Program on Education Policy and Governance Working Papers Series

Evolution vs. Creationism in the Classroom: The Lasting Effects of Science Education

Benjamin W. Arold [1]

PEPG 22-01

Harvard Kennedy School
79 JFK Street, Taubman 220
Cambridge, MA 02138
Tel: 617-495-7976
www.hks.harvard.edu/pepg/

[1] Arold: ETH Zurich.

Evolution vs. Creationism in the Classroom: The Lasting Effects of Science Education*

Benjamin W. Arold

September 5, 2022

Abstract

Anti-scientific attitudes can impose substantial costs on societies. Can schools be an important agent in mitigating the propagation of such attitudes? This paper investigates the effect of the content of science education on anti-scientific attitudes, knowledge, and choices. The analysis exploits staggered reforms that reduce or expand the coverage of evolution theory in US state science education standards. I compare adjacent cohorts in models with state and cohort fixed effects and conduct fine-grained placebo tests to rule out scientific, religious and political confounders. There are three main results. First, expanded evolution coverage increases students' knowledge about evolution. Second, the reforms translate into greater evolution belief in adulthood, but do not crowd out religiosity or affect political attitudes. Third, the reforms affect high-stakes life decisions, namely the probability of working in life sciences.

JEL-Codes: I28; J24; P16; Z12
Keywords: Evolution, Religion, Science Education, Human Capital, Beliefs, Occupational Choice

---

* **Arold**: ETH Zurich (benjamin.arold@gess.ethz.ch). I am grateful to Karun Adusumilli, Elliott Ash, Robert Barro, Luca Braghieri, Davide Cantoni, Sarah Eichmeyer, Eric Hanushek, Martina Magli, Rachel McCleary, Markus Nagler, Paul Peterson, Matteo Sandi, M. Danish Shakeel, Claudia Steinwender, Joseph Stiglitz, Gregory Veramendi, Martin West, Brian Wheaton, Scott Williamson, Ludger Woessmann, Larissa Zierow, as well as the participants of the seminar in Religion and Political Economy (Harvard), the Graduate Workshop on Political Economy (Harvard), seminars at ETH Zurich, UZH Zurich, Mannheim, briq Institute Bonn, Luxembourg, Nuremberg, and the Lindau Nobel Laureate Meeting 2022, ASSA/AEA Annual Meeting 2022, ASREC Conference 2022, ifo-WZB EffEE Conference 2022, ESPE Conference 2021, EALE/SOLE/AASLE World Conference 2020, German Economic Association (VfS) Meeting 2020, AEFP Conference 2020, EffEE PhD Workshop 2021, LMU Empirical PhD Seminar 2020, and the LMU Munich Young Economist Meeting 2019. Caroline Michler and Sophia Stutzmann provided excellent research assistance. I thank Paul Peterson, Antonio Wendland and the Program on Education Policy and Governance at Harvard University for their hospitality while writing parts of this paper. I gratefully acknowledge financial support by the DAAD through a one-year scholarship for doctoral students funding my stay at Harvard University. Financial support by the Leibniz Competition (SAW 2019) is also gratefully acknowledged. Some of the data used in this analysis are derived from Sensitive Data Files of the GSS, obtained under special contractual arrangements designed to protect the anonymity of respondents. These data are not available from the author. Persons interested in obtaining GSS Sensitive Data Files should contact the GSS at GSS@NORC.org. Similar requirements apply to NAEP. All errors are my own.

1 INTRODUCTION

Anti-scientific attitudes can impose substantial costs on public health, the environment, and the economy. Misinformation about the danger of Covid-19 and a lack of trust in scientists have undermined compliance with social distancing measures and vaccination recommendations, prolonging the pandemic (Bursztyn et al., 2020; Algan et al., 2021; Brzezinski et al., 2021; Jin et al., 2021). Climate change denial has reduced the support for policies cutting greenhouse gas emissions, contributing to its environmental and economic damage (Akter et al., 2012; Linden et al., 2015). The rejection of evolution theory has been used to justify white supremacy and racism in the US (Marks, 2012), and has contributed to anti-scientific agricultural policies and associated food shortages in the Soviet Union (Graham, 2016).[1] While there is broad understanding of the societal costs of anti-scientific attitudes, evidence on its determinants is surprisingly scant despite the relevance for effective policy responses.

This paper isolates the content of science education in high school as one determinant of anti-scientific attitudes that is directly subject to policy makers.[2] To study whether the content of science education has a lasting impact on individuals beyond attitudinal outcomes, the paper also analyzes how it affects scientific knowledge and life decisions. Specifically, I estimate the causal effect of students' exposure to the teaching of evolution theory in science education on (i) their knowledge about evolution at the end of high school, (ii) their belief in evolution in adulthood, and (iii) the probability that they work in life sciences.

The focus of this paper is on evolution theory because of its fundamental role in

[1]The pseudoscientific theories of Trofim Lysenko, then-president of the Academy of Agricultural Sciences of the USSR and leading agricultural advisor to Joseph Stalin, have been made responsible for prolonging Soviet foot shortages in the 1930s ("Lysenkoism") (Joravsky, 1962).

[2]In general, attitudes are shaped by a multitude of factors many of which are rather shielded in the private domain. An extensive literature on the formation of attitudes and beliefs has emphasized the impact of inter-generational transmission in families (Bisin and Verdier, 2001; Guiso et al., 2008; Tabellini, 2008). Other determinants include peers and social networks (Sacerdote, 2001; Bailey et al., 2020), the media (Martin et al., 2017), political systems (Alesina and Fuchs-Schündeln, 2007), and macroeconomic conditions (Giuliano and Spilimbergo, 2014).

science, and its controversy in the population and the education system. Evolution can scientifically explain the existence of all species including our own. The American Association for the Advancement of Science (2021) states that "the foundation of all life sciences is biological evolution". 98 percent of its members express support for the statement that humans have evolved over time (Pew Research Center, 2015). In contrast, evolution is a highly charged topic among the US population with only 65 percent agreeing that humans have evolved over time. Prior to the First World War and up to the present day, this controversy has been reflected in heated debates and legal battles on whether evolution is supposed to be taught in schools.[3] Teachers and school districts have been convicted for not following the education standards' stance on evolution. Even today, there is substantial variation across US states and time in the way how evolution is covered in education standards.

To isolate exogenous variation in students' exposure to the teaching of evolution, this paper exploits staggered state-level reforms of the coverage of evolution in US State Science Education Standards (Science Standards). In the study period from 2000 until 2009, 22 states expanded the coverage of evolution in their education standards, while 15 states reduced it. I argue that the political and institutional processes leading to these reforms, in particular the predetermined timing of gubernatorial elections and the tenure of members of State Boards of Education, create idiosyncrasies in the determination of the precise reform years. This setting allows for the estimation of causal effects in two-way fixed effects models with state and cohort fixed effects, overcoming the identification problem that the content of science education is generally correlated with scientific, religious and political attitudes of the students' environment which independently affect student outcomes.

Beyond the theoretical argument that the reform timing is determined by institutional idiosyncrasies, my empirical setup explicitly accounts for a range of endogeneity concerns by comparing adjacent cohorts around sharp reforms of the

[3] For example, the New York Times published a report on recent controversies with the headline "Questioning Evolution: The Push to Change Science Class" (Haberman, 2017).

Science Standards. Specifically, the performed two-way fixed effects estimations can rule out as confounding factors (i) state-specific differences (such as education levels), (ii) cohort-specific differences (such as national changes in attitudes across time), (iii) time-varying state-specific shocks that affect adjacent cohorts similarly (such as natural disasters or state-level political or religious shocks that do not differentially affect children of different cohorts), and (iv) time-varying state-specific shocks that affect adjacent cohorts differentially, but smoothly (such as state-specific trends in science skepticism), in a robustness test that includes state-specific time trends. To conduct the set of analyses, I link state-level data on the evolution coverage in Science Standards with three individual-level datasets.

First, this paper shows that the evolution coverage in Science Standards affects what students learn about evolution in school. Specifically, I use the National Assessment for Educational Progress (NAEP) to demonstrate that students being exposed to a more comprehensive evolution coverage in high school are more likely to correctly answer knowledge questions on evolution by the end of high school. This finding exemplifies how the content of education standards can foster scientific knowledge, an outcome of direct economic importance given its effects on earnings and economic growth in the long run (Lucas, 1988; Barro, 2001; Hanushek and Woessmann, 2008, 2012).

Second, this paper demonstrates that the evaluated reforms have lasting effects on attitudes. To that end, I make use of the General Social Survey (GSS) to show that evolution teaching affects the probability of believing in the concept of evolution in adulthood. Being exposed to a comprehensive evolution coverage in the education standards in high school compared to no evolution coverage increases evolution belief in adulthood by 57 percent of the sample mean, corresponding to a persuasion rate of 79 percent (DellaVigna and Gentzkow, 2010). This analysis underscores that reform effects persist long after students have left high school. This result exemplifies how science education can promote scientific attitudes, which can be directly relevant for improving public health, the environment, and the economy (Brzezinski et al., 2021;

Martinez-Bravo and Stegmann, 2021).

Third, this paper shows that the evaluated reforms affect high-stakes choices, namely occupational choice. I hypothesize that learning about evolution, the fundamental theory of life sciences, affects the probability of working in life sciences in adulthood. Using the American Community Survey (ACS), I demonstrate that high school exposure to a comprehensive evolution coverage in the education standards compared to no evolution coverage increases the probability of working in life sciences in adulthood by 23 percent of the sample mean. This effect mostly comes from the subgroup of biology, the subject in which evolution is typically being taught. This finding exemplifies how science education can attract future STEM workers, which not only raises wages at the individual level (Hastings et al., 2013; Kirkeboen et al., 2016; Deming and Noray, 2020), but also has wider economic consequences through fostering innovation, technological change, labor productivity and economic growth (Griliches, 1992; Jones, 1995; Kerr and Lincoln, 2010; Peri et al., 2015).

A particularly useful feature of focusing on one topic such as evolution is the possibility of constructing fine-grained placebo tests. Testing whether reforms affect non-evolution outcomes constitutes falsification tests. Specifically, I show null effects for (i) non-evolution scientific knowledge by the end of high school, (ii) non-evolution scientific, religious, and political attitudes in adulthood, and (iii) the probabilities of working in non-scientific occupational fields. These results provide empirical support for the interpretation that it is in fact institutional idiosyncrasies which determine the exact reform timing rather than scientific, religious, and political trends or shocks. I further demonstrate that the reform effect on evolution knowledge is specific to students in public schools, while there is no effect for a placebo sample of private school students for whom Science Standards have never been binding. Another robustness check replicates the main results on a subsample using only states with closely elected governors ruling at the time of the reform. In addition, the results are immune to potential biases in staggered two-way fixed effects designs from time-varying treatment effects (Callaway and Sant'Anna, 2021).

This paper contributes to the literature on the political economy of schooling (Lott, 1999; Cantoni and Yuchtman, 2013). I provide the first quasi-experimental evidence that attitudinal changes induced by reforms of the content of education translate into high-stakes choices of individuals. Cantoni et al. (2017) exploit a Chinese textbook reform to show that the content of education affects students' political and economic attitudes. Other seminal papers study the effects of the content of education on cultural identity (Clots-Figueras and Masella, 2013), civic values (Bandiera et al., 2019), and religiosity (Bazzi et al., 2020). While these papers show effects on attitudes, I go beyond this by demonstrating that high-stakes occupational choice is also affected.

This finding also enhances our understanding of how to increase the share of STEM graduates, which is a policy goal with widespread support in many societies.[4] Occupational sorting is influenced by demand side factors such as expected earnings and non-pecuniary job benefits (Wiswall and Zafar, 2018; Arcidiacono et al., 2020), perceived ability (Stinebrickner and Stinebrickner, 2014; Arcidiacono et al., 2016a), and heterogeneous tastes (Wiswall and Zafar, 2015). Supply side factors such as grading policies (Butcher et al., 2014), admissions systems (Bordon and Fu, 2015), affirmative action policies (Arcidiacono et al., 2016b), and the provision of role models (Porter and Serra, 2020) can also play a role (for an overview, see also (Altonji et al., 2016)). I demonstrate that the content of science education in high school can be an effective policy tool to attract STEM graduates.

This paper also speaks to the emerging literature on the determinants of religiosity (Iannaccone, 1998; Iyer, 2016; McCleary and Barro, 2019). Finding null effects on religious outcomes demonstrates that expanding the scientific content of science education neither reduces the belief in nor the belonging to a religion. This is true despite the fact that being raised as Evangelical is a large negative predictor of evolution belief. While a number of

[4]In the US, increasing the number of STEM graduates is a central policy goal of the Federal Government's strategic plan for STEM education 2018-2023 (National Science and Technology Council, 2018). Similarly, the EU aims to increase the number of STEM graduates as one of its twelve policy goals of the European Skills Agenda 2020-2025 (European Commission, 2020).

studies have found a positive relationship between education and religiosity (McCleary and Barro, 2006; Glaeser and Sacerdote, 2008; Meyersson, 2014), other research suggests that education can decrease religiosity (Hungerman, 2014; Becker et al., 2017). In the specific setting of evolution teaching in the US, religiosity is not crowded out.

Finally, this paper contributes to the literature on the effects of the content of education on students' knowledge. While there is broad understanding about the effects of topic-specific instruction time (Cortes and Goodman, 2014), minimum high school course requirements (Goodman, 2019), advanced placement courses (Conger et al., 2021), vocational school curricula (Schultheiss and Backes-Gellner, 2021), and the interaction of curricula and internet penetration (Sen and Tucker, 2022), this paper can show that the content of education standards affects the knowledge of students on the topic in question in the intended direction.[5] What is more, the effects of the content of education standards last until adulthood.

The paper proceeds as follows. Section 2 outlines the historical and institutional background of the teaching of evolution. Section 3 provides information on the data measuring the coverage of evolution in Science Standards and the microeconometric datasets. Section 4 describes the identification strategy. Section 5 presents the results. Section 6 discusses robustness tests. Section 7 concludes.

2 INSTITUTIONAL BACKGROUND

2.1 The Battle for Teaching Evolution in US Public Schools

For at least a century, the teaching of evolution in public schools has been a contested issue in the US. Although the scientific community reached a consensus on the validity of evolution relatively soon after Charles Darwin's publication of "On the Origin of Species" (Darwin, 1859),[6] the public did not share the consensus. This was and still is reflected

[5] Arold and Shakeel (2021) show that the adoption of centralized education standards in the US in math and ELA had *unintended* effects on students' *overall* science achievement.

[6] Thomas Henry Huxley (1880, p.1) stated that "there is no field of biological inquiry in which the influence of the 'Origin of Species' is not traceable [. . .] and the general doctrine of evolution [. . .] may

in the educational system. For the decades prior to the First World War, Beale (1941) describes how teachers wanting to teach evolution in an average American school had difficulties in doing so. Only one quarter of the biology textbooks published between 1900 and 1919 contained information about evolution (Skoog, 2005). No book covered human evolution. In the 1920s, about one third of biology textbooks covered human evolution, documenting an early phase of a gradual and non-linear development throughout the 20^{th} century towards more evolution coverage in US high school biology textbooks.

However, the 1920s also marked the start of a series of legal disputes regarding teaching evolution in US schools throughout the 20^{th} century. At least 20 states considered bills to ban the coverage of evolution in public schools in the 1920s (Numbers, 1982). Among other states, such a bill became law in Tennessee, known as the Butler Act, resulting in the famous Scopes trial in 1925. John T. Scopes, a biology teacher from Tennessee, was convicted in Tennessee v. Scopes for having taught evolution in the classroom. Although the Tennessee Supreme Court overturned the decision on a technicality, it decided that the law banning evolution from schools was not unconstitutional (Larson, 1999).

In the second half of the 20^{th} century, legislative and adjudicative decisions became more favorable towards the coverage of evolution in public schools (Moore et al., 2003b). In 1967, the Butler Act was repealed by the Tennessee legislature. One year later, the Supreme Court of the US ruled that a law banning the teaching of evolution in schools in Arkansas was unconstitutional in Epperson v. Arkansas. As a reaction, creationists lobbied for laws requiring that equal time must be spent on teaching evolution and creation. In 1987, this was ruled unconstitutional by the US Supreme Court in Edwards v. Aguillard. In sum, the legal decisions of the 20^{th} century have paved the way for evolution to be taught in public schools. In the 21^{st} century, creationism and intelligent design are no longer permitted to be taught in US public schools. Still, there continues to be substantial variation in evolution teaching across states and years, as the

conduct its conquest of the whole realm of Nature". Ernst Mayr (1991, p.25), a leading evolutionary biologist of the 20^{th} century, wrote that "within fifteen years of the publication of the Origin hardly a qualified biologist was left who had not become an evolutionist". This influence was also reflected in the cultural discourse at the time (Giorcelli et al., 2022).

subsequent analysis of the evolution coverage in Science Standards demonstrates.

2.2 US State Science Standards

US State Science Standards serve as state-wide school curriculum frameworks in science. The content of US education has historically been determined at the local level. However, concerns about a decline in achievement among US students in the 1960s and 1970s and resulting economic costs (Hanushek, 1986; Bishop, 1989) gave rise to calls to establish rigorous and comparable education standards. In 1983, the report "A Nation at Risk" (National Commission on Excellence in Education, 1983) proposed the introduction of centralized education standards.[7] Several organizations have proposed guidelines for centralized educational standards for the different school subjects. Regarding science, the American Association for the Advancement of Science developed the Science Standards guidelines "Science for All Americans" (1990) and "Benchmarks for Science Literacy" (1994), and the National Research Council published the "National Science Education Standards" (1996). By 2000, all states except for Iowa had adopted Science Standards (Lerner, 2000a).

Science Standards define the scientific knowledge and skills that students are supposed to master in a given grade in public schools. Scientific teaching which a student is ultimately exposed to in class does not solely depend on the Science Standard of her state, but also on local school curricula, the selection of textbooks, the knowledge, ability and ideology of teachers, testing formats, and other factors. However, Science Standards form the basis of many of these factors and thus, indirectly, affect science teaching in schools. For instance, they affect how local curricula and teachers' lesson plans are written (Lerner, 2000b). Furthermore, science textbooks are arranged to match the content laid out in Science Standards, reflecting the standards from larger

[7] Theoretically, centralized education standards can be more rigorous as they overcome a free-riding problem induced by the mobility of high school graduates across school districts and their pooling in the local labor markets (Costrell, 1994, 1997). At the same time, centralization can also reduce the incentive to develop rigorous and innovative education standards by abolishing competition between school districts (Tiebout, 1956; Oates, 1999).

states in particular. Moreover, state-wide standardized exams often directly test the content set out in the Science Standards. Lerner (2000b, p.ix) summarizes that "the knowledge and skills set forth in state standards are supposed to form the core of "standard based" education reform. They are meant to serve as the frame to which everything else is attached, the desired outcome that drives countless other decisions about how best to attain it." With regards to evolution, 88 percent of a nation-wide representative sample of US public high school biology teachers state that they focus heavily on what students need to know to meet Science Standards when teaching evolution, see Figure A.1.

2.3 The Adoption Process of Reforms of Science Standards

Understanding the political process leading to reforms of the evolution coverage in Science Standards is of particular interest for assessing whether they create exogenous variation in students' exposure to the teaching of evolution. In this section, I argue that the fact that such reforms happen *at some point* is not as-good-as-random, but that the *specific timing* of such reforms is as-good-as-random due to substantial institutional idiosyncrasies.

Reforms of Science Standards are decided by majority vote of the members of the State Boards of Education. The selection process of the members of the State Boards of Education differs across states. In some states, members are appointed by the governor, sometimes with the advice and consent of the senate (for example in California and Florida). In other states, members are elected by the public, typically in a staggered election across districts (for example in the District of Columbia and Texas). A few states combine the two selection mechanisms by appointing some members and electing others (for example Louisiana and Ohio). Student representatives or external experts are also appointed or elected in some states (for example in Alaska and Massachusetts).

Before the members of the State Board of Education vote on a reform of Science Standards, standards are typically drafted by advisory committees. Again, the composition of these advisory committees depends on the state. In general, advisory committees consist of a panel of teachers and other stakeholders which sometimes includes scientists. In addition to the input of the advisory board, most State Boards of Education hold hearings or testimonies of stakeholders such as parents, scientists, religious representatives, among others. At this point, it typically becomes clear which interest groups lobby in favor or against the proposed reform. For example, the National Center for Science Education has lobbied for a more comprehensive coverage of evolution in multiple cases, while the Discovery Institute has spoken out against it. Following the period of public comment, the State Board of Education has the final vote.

On the one hand, the political process described above implies that these reforms happening at some point in a given state is not random. Instead, they reflect changing political views, either expressed by the election of a governor who subsequently appoints members of the State Boards of Education, or by direct election of members of the State Boards of Education.

On the other hand, the exact reform year in a given state can be regarded as-good-as random. If beliefs in evolution or science in general change among the population in a certain year, it will take an arbitrary number of years until this results in a reform of Science Standards due to institutional idiosyncrasies. In states where members of the Board of Education are appointed by the governor, the year of a reform crucially depends on the governors' year of election, as determined by the legislation period lasting four years in general. In states where members of the Board of Education are directly elected, the reform year depends on the elections, which typically take place with different districts in different years in a staggered manner. Further state-specific idiosyncrasies are induced by the fact that the tenure of members of the Boards of Education differs across states, which can last up to nine years such as in West Virginia. Even after a new majority in the Board of Education is in power, the drafting, hearing, and voting on new standards causes further delay, as this can take months or years. In some cases, there are also spillovers in the sense that Science Standards reforms of one state affect the teaching in other states. This occurs, for example, because textbooks used in smaller states may follow Science

Standards reforms of larger states. In sum, there may be a great number of years between a scientific, religious, or political shock and a reform of the evolution coverage in Science Standards. However, it can also be small if election dates and the tenure expiration of the marginal board member occur shortly after a given shock. Hence, the precise timing of such reforms is arguably exogenous. Appendix A.1 provides anecdotal evidence on the political processes leading to reforms in Florida and Texas. While Florida expanded the evolution coverage in 2008, Texas reduced it in 2009; with neither reform following a partisan change in government.

In the empirical analysis, placebo tests showing null effects on non-evolution scientific, religious and political outcomes test this narrative empirically. The same is true for regressions conditioning on the party of the governor.

2.4 The Implementation of Reforms of Science Standards

After new Science Standards are adopted, their implementation in the classroom tends to be rather swift. In general, widely publicized lawsuits convicting school districts for not implementing the teaching of evolution as outlined in Science Standards contribute towards a fast implementation of such reforms.[8] In Florida in 2008, for example, school districts were supposed to adjust their lessons by comprehensively including evolution as outlined in the newly adopted Science Standard within one year. Furthermore, evolution was required to become part of standardized testing in Florida from 2012 onward. In the 2009 Texas reform, the evolution coverage of the new Science Standard had to be in textbooks from 2011 onward.

[8]For example, a lawsuit that received national attention was Kitzmiller v. Dover Area School District in 2005. The Dover Area School District had required biology teachers to teach intelligent design (a form of creationism attributing the creation of the world to an intelligent designer) as an alternative to evolution. This requirement contradicted the content of the Science Standard in force at the time, and was ruled unconstitutional in Kitzmiller v. Dover Area School District. Specifically, the verdict prohibited the Dover Area School District from requiring teachers to "denigrate or disparage the scientific theory of evolution, and from requiring teachers to refer to a religious, alternative theory known as intelligent design." (Kitzmiller v. Dover Area School District, 400 F. Supp. 2d 707 (M.D. Pa. 2005)).

3 DATA

3.1 Coding of Reforms of Science Standards

To measure the coverage of evolution in Science Standards, I make use of the "evolution score" provided by Lerner (2000a) and Mead and Mates (2009). The evolution score is a composite index based on an evaluation of whether the word "evolution" appears in a Science Standard, of the respective coverages of biological, human, geological, and cosmological evolution, and of the connection of the different aspects of evolution. Moreover, the absence of creationist jargon and creationist disclaimers in textbooks is taken into account. The evolution score is defined between 0 and 1, with 0.01 increments. An evolution score of 0 indicates no or a non-scientific/creationist coverage of evolution, and a score of 1 a very comprehensive coverage of evolution. Notably, the creationist jargon in all Standards evaluated in this paper is never openly religious, which would be unconstitutional. However, there is large variation in the emphasis of (alleged) weaknesses and critique of evolution theory, creating or removing scope for teachers wishing to teach creationist content.[9]

The evolution score is available for all states for the years 2000 and 2009, provided by Lerner (2000a) and Mead and Mates (2009), respectively. They also provide information on the evolution score's year of reform for each state between 2000 and 2009 (if there was any reform). If more than one reform took place between 2000 and 2009 in a given state, there is information on the last reform.[10] The evolution score serves as a treatment

[9]In 2000, Kansas received an out-of-range score of -0.18, as "it is a special case, unique in the extremity of its exclusion of evolution from statewide science standards" (Lerner, 2000b, p.16). For example, it did not cover Darwin, biological evolution and any reference to the age of the earth. In this paper, I change this evolution score from -0.18 to 0 for ease of interpretability of regression results. All results using the original score of -0.18 for Kansas instead of 0 do not differ meaningfully (results available upon request). Iowa had no Science Standards in 2000 which is coded as missing. The District of Columbia is treated as a state throughout this paper. The evolution score was originally defined between 0 and 100, but I re-scale it by dividing it by 100, again for ease of interpretability. More information about the details of the scoring scheme are provided in Lerner (2000b, pp.10-17).

[10]This implies that reforms before the respective last reform are not taken into account in the analyses. In theory, ignoring these prior reforms merely creates attenuation bias as long as these prior reforms are uncorrelated with the timing of the last reform in a given state. To explicitly test for this, I perform a robustness check restricting the sample to students from states for which careful examination of academic

variable in all analyses presented in this paper. When merging it with individual-level datasets, each individual is defined as being exposed to the evolution score from 2000 if she started high school before the reform year in her state, and being exposed to the evolution score from 2009 if she started high school in the year of the reform in her state or later. The high school entry year is the pertinent year, as most of the teaching on evolution takes place at the beginning of high school.[11]

To illustrate the identifying variation, Figure 1 depicts the state-level evolution score *difference* between 2000 and 2009.[12] The evolution score increased in 22 states (implying a positive evolution score difference) and decreased in 15 states (implying a negative evolution score difference) between 2000 and 2009. In the remaining 13 states, it remained unchanged. The states with the largest evolution score increases are Kansas, Mississippi, and Florida. The largest evolution score decreases are found in Connecticut, Louisiana, and Texas. By construction, the changes partly depend on the baseline level, in the sense that Science Standards covering evolution very comprehensively in 2000 cannot expand the coverage by much until 2009, and vice versa. However, by identifying from changes within states I control for fixed differences between states. Overall, the evolution score changes are fairly well spread over the US, with each census region having at least one state in which the evolution coverage became more comprehensive, less comprehensive, and remained unchanged, respectively.

3.2 Micro Data

The following subsection describes the three micro-level datasets used in this paper. These repeated cross-sectional datasets are standardized and hence comparable across US states and cohorts, making them suitable for analyses with state and cohort fixed effects.

In all three datasets, I keep students in the sample who have no missing basic controls

articles, legal documents, and state education websites indicates that they only had one reform between 2000 and 2009, see Section A.4.2 and Table A.1 for more details.

[11]The standard high school curriculum typically features biology (the subject in which evolution is being taught) in the first year of high school.

[12]Figure A.2 also depicts the evolution score levels in 2000 and 2009.

variables and who start high school after 1990 and before 2010 in my preferred sample cut. Thereby, I balance temporal proximity to the reform years and having sufficient years to estimate pre-trends and fixed effects credibly (and with statistical power in general). This approach also prevents identification from the adoption of the Next Generation Science Standards which started in 2013. The results of this paper do not depend on this specific sample cut, as shown in robustness tests in Section A.4.2.

3.2.1 NAEP: Evolution Knowledge in School

To estimate the effect of students' exposure to the teaching of evolution in high school on their knowledge about evolution by the end of high school, I link the evolution score with the restricted-use individual-level National Assessment of Educational Progress (NAEP). NAEP is a standardized student achievement test, measuring the knowledge of US students in various subjects since 1990. For this study, I use the NAEP test for science in grade 12 as it contains questions on evolution. Students are coded as exposed to the Science Standard in place in the year and state of their high school entry, assuming that they started high school three years prior to taking the test in grade 12 in the same state.

The main outcome variable "evolution knowledge" is defined as the share of correctly answered questions on evolution. The nine categories of scientific knowledge on topics other than evolution are defined analogously. They serve as placebo outcomes in subsequent analyses and include topics such as "reproduction", "climate" or the "universe". In addition, the NAEP student surveys provide rich student-level control variables. They include, inter alia, variables measuring the socio-economic status such as subsidized lunch status, parental education and home possessions.

The main sample only contains public school students, as Science Standards have never been binding for private schools. However, the latter serve as a placebo sample in robustness checks. The main sample consists of more than 15,000 public school students who were asked at least one question on evolution. The descriptive statistics show that the average evolution score equals 0.65, implying that sampled students were on average exposed to a "satisfactory" evolution coverage.[13] The mean of the main outcome variable "evolution knowledge" equals 0.32. The fact that, on average, not even one third of the questions on evolution is being answered correctly underscores the questions' difficulty. Appendix A.2.1 provides detailed tables of the descriptive statistics and raw correlations. It also presents sample questions, explains how the science questions are grouped into topical categories, and how missing observations are dealt with.

3.2.2 GSS: Evolution Belief in Adulthood

To estimate the effect of students' exposure to the teaching of evolution in high school on their belief in evolution in adulthood, I link the evolution score with the restricted-use individual-level General Social Survey (GSS). The GSS is a biennial cross-sectional survey which monitors societal change by interviewing a nationally representative sample of adults in the US since 1972. Since 2006, respondents have been asked about their belief in evolution. The GSS also provides the state of residence at age 16 and the birth year. I assume that respondents started high school in this state at age 14 and merge the evolution score for this state-year combination accordingly. Hence, I can link individuals' belief in evolution in adulthood to the evolution coverage of the Science Standard they were exposed to as students, even if they migrated to other states after finishing school.

The main outcome variable "evolution belief" is based on the question "Human beings, as we know them today, developed from earlier species of animals. Is that true or false?".[14] The corresponding indicator variable is set to one if the answer "true" was given, and set to zero if any other answer option was reported such as "false", "don't know", or "no answer". The GSS also asks a broad range of questions on scientific topics besides evolution, and on religious, political and partisan attitudes. Other variables capturing different dimensions of the childhood environment serve as control variables, including

[13] Lerner (2000b) classifies evolution scores between 0.60 and 0.79 as "satisfactory".

[14] The words "human beings" are replaced by the word "elephants" for 10 percent of the questions on evolution belief in the sample. Table A.14 shows that the results are robust to dropping these 10 percent from the sample.

the religion a respondent was raised in.

The GSS is sampled from the entire US adult population irrespective of type of school attendance. It does not make it possible to differentiate between public and private school attendance as the NAEP. Instead, one can estimate effects net of endogenous sorting across school types, including homeschooling. The estimation sample of individuals who were asked the question on evolution belief contains more than 1,800 individuals. The descriptive statistics show that 58 percent of the sample believe in evolution which largely represents the average evolution belief in the US population at the time (Pew Research Center, 2009). More details on descriptive statistics, raw correlations and data background is provided in Appendix A.2.2.

3.2.3 ACS: Occupational Choice

To estimate the effect of students' exposure to the teaching of evolution in high school on their probability of working in life sciences during adulthood, I link the evolution score with the individual-level IPUMS American Community Survey (ACS) (Ruggles et al., 2020). The ACS is a large-scale demographic survey which draws from a national random sample of the US population. Responding and providing correct information is required by US law. The ACS contains detailed information on the respondents' occupational field. It also elicits the state and year of birth. I assume that students start high school in this state at age 14, and accordingly merge the evolution score for this state-year combination.

Given that evolution is the fundamental theory of life sciences, the occupational field of primary interest in this study is life sciences. The main outcome variable "working in life sciences" is coded as an indicator variable equal to one if the respondent works in life sciences, and equal to zero otherwise. All other occupational fields are coded analogously. The ACS also enables the division of occupational fields into more fine-grained occupational subfields. The occupational field "life sciences" can be divided into the subfields "biology", "agriculture and food", "conservation and forestry" and "medical and other" for the purpose of subgroup analyses. Beyond sciences, I also analyze other occupational fields such as management, engineering and education. There are 25 non-scientific occupational fields in total, including one category for unemployed/not in the labor market which serve as placebo outcomes in robustness checks.

Like in the GSS, the ACS is sampled from the entire US population which also includes individuals who went to private school and homeschoolers. The estimation sample of individuals who are older than 18 years (i.e., who typically completed secondary education) consists of more than 6 million individuals. Further information, including descriptive statistics, is provided in Appendix A.2.3.

4 IDENTIFICATION STRATEGY

The analyses presented in this paper are based on the following two-way fixed effects (TWFE) model. The TWFE model exploits the different timing of reforms of the evolution coverage in Science Standards across states, and the fact that some of the reforms extended the coverage of evolution, while others reduced it, and a third group of states did not reform the evolution coverage. It compares outcomes of cohorts who went to high school in states where the evolution coverage was reformed with previous cohorts from the same states prior to reforms, relative to how outcomes of these cohorts changed in states that did not reform at the time, after accounting for fixed differences between states and birth cohorts. The baseline parametric TWFE model is specified as follows:

$$Y_{istu} = \beta \cdot Evolution\_Score_{st} + \gamma \cdot \mathbf{X_i} + \delta_s + \lambda_t + \theta_u + \epsilon_{istu} \tag{1}$$

where Y_{istu} is the outcome of interest of individual i, who started high school in state s and year t, and completed the test or survey in year u. The treatment variable $Evolution\_Score_{st}$ measures the intensity of the evolution coverage in the Science Standard in state s and year t. β is the parameter of interest capturing the effect on the outcome of being exposed to a very comprehensive coverage of evolution ($Evolution\_Score_{st}$=1) as compared to being exposed to no or a non-scientific/creationist coverage of evolution ($Evolution\_Score_{st}$=0). The vector $\mathbf{X_i}$

contains the individual-level control variables. State fixed effects δ_s, birth cohort/high school entry cohort fixed effects λ_t, test/survey year fixed effects θ_u, and an error term complete the model. Standard errors are clustered at the state level which is conservative in this setting and accounts for the potential correlation of error terms across cohorts within states (Abadie et al., 2017; Athey and Imbens, 2021).

The baseline model addresses a range of concerns on the ability to estimate causal effects of the evolution coverage in Science Standards. One might be concerned that state-level differences in scientific, religious or political attitudes are correlated with the evolution coverage in Science Standards and affect scientific knowledge, beliefs as well as occupational choice. The state fixed effects absorb all differences in outcomes that are constant between states. In addition, one might be worried that national trends, such as attitudinal trends on scientific, religious or political topics, might erroneously appear as reform effects. To counter this concern, the cohort fixed effects eliminate all national differences between cohorts.

A remaining concern are time-varying state-specific trends or shocks. For example, state-specific trends in human capital levels, or regional religiosity shocks induced by, for instance, church scandals may affect attitudes towards evolution differentially in different states. However, such factors only threaten the ability to estimate causal effects if they affect different high school entry cohorts differently. Many state-specific factors may be time-varying, but still have a similar effect on adjacent cohorts. This is the case, for example, if a church scandal occurring in a given year and state provokes similar reactions in adjacent cohorts. However, my empirical setup exploits cross-cohort variation within a narrow time window around the reforms, and identifies from reforms of the evolution coverage in Science Standards that affect adjacent cohorts in different ways. Although reforms of Science Standards are generally applicable to all cohorts from the year of adoption onward, the change in evolution coverage typically only affects the high school entry cohort (and younger cohorts in the following years when they start high school). This is true as the high school entry year is the year in which evolution is typically being taught.[15] The state fixed effects capture such time-varying state-specific factors that affect students of different cohorts equally.

Moreover, I address concerns about time-varying state-specific trends or shocks that affect adjacent cohorts differently, but smoothly, by conducting robustness checks with state-specific time trends. For example, the trust in science among students could develop differently in the various states, but change smoothly across cohorts. The presented specification with linear and quadratic state-specific time trends is particularly demanding in terms of statistical power, as reform effects are only detectable as "jumps" from the cross-cohort trend. Showing that (at least the point estimates of) the main results hold in this specification reaffirms a causal interpretation of the findings presented.

In addition, the individual-level control variables take out observable differences between individuals that vary non-smoothly across states and cohorts. For example, controlling for the religion in which an individual was raised ensures that outcomes across individuals are compared while holding constant their religion of the childhood.

To validate my econometric approach, I conduct a series of placebo tests. In the first analysis, I test whether the coverage of evolution affects evolution knowledge (main outcome), but not knowledge on scientific topics other than evolution (placebo outcomes). In the second analysis, I test whether the coverage of evolution affects evolution belief in adulthood (main outcome), but not other scientific, religious or political attitudes (placebo outcomes). In the third analysis, I test whether the coverage of evolution affects the probability of working in life sciences (and in particular biology), but not the probabilities of working in non-scientific occupational fields. Null effects on placebo outcomes suggest that no previously uncontrolled scientific, religious or political shock coincides with the timing of the reforms. They also demonstrate that the effects reported in this paper are strongly linked to the topic of evolution, providing empirical support to the claim that the exact timing of reforms is driven by institutional idiosyncrasies in lieu of political

[15]To the extent that evolution is also being taught in higher grade levels, the difference in exposure to the teaching of evolution between pre- and post-reform cohorts is overstated in my coding. Hence, I interpret the results as lower-bound estimates as parts of the cohorts coded as exposed to pre-reform Science Standards may be partially treated by post-reform Science Standards.

changes. At the same time, they show that these outcomes themselves are not affected by the reforms.

Another placebo test makes use of a placebo sample of private school students for whom education standards have never been binding. One can test whether the reform effect on evolution knowledge is specific to public school students (main sample), but not detectable for private school students (placebo sample).

Even in the absence of confounding trends or shocks, consistent estimation of reform effects requires homogeneity in treatment effects (Chaisemartin and D'Haultfœuille, 2020; Borusyak et al., 2021; Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Sun and Abraham, 2021). The treatment effect from the baseline TWFE model is a weighted average of all possible 2x2 difference-in-differences comparisons between treated and untreated groups as well as groups treated at different points in time (Goodman-Bacon, 2021). In settings with staggered treatment timing such as in this study, time-varying treatment effects can bias results away from the true effect if already-treated students act as controls for later-treated students (negative weighting).

To show that my TWFE estimator is immune to this bias, I run the estimator by Callaway and Sant'Anna (2021) (CS estimator; in the implementation of Rios-Avila et al. (2022)), which excludes those 2x2 difference-in-differences comparisons in which already-treated students act as controls from the sample. Out of the set of newly developed estimators that account for this issue of negative weighting (Chaisemartin and D'Haultfœuille, 2020; Borusyak et al., 2021; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021), the CS estimator is preferred in my setting, as it includes never-treated and not-yet-treated units as controls, and runs on cross-sectional datasets, see also the discussion in De Chaisemartin and D'Haultfoeuille (2022).[16]

The continuous nature of the evolution score treatment variable requires a strong parallel trends assumption that is likely met in my setting. Specifically, it requires that

[16]Another approach would be to only focus on the CS estimator and abandon the TWFE estimator altogether. However, this is not necessarily advisable: The TWFE estimator estimates a convex combination of effects, and often has a lower variance compared to the newly developed TWFE estimators (De Chaisemartin and D'Haultfoeuille, 2022).

the average potential outcomes for individuals that are exposed to a reform of the evolution coverage are the same at each level of evolution coverage, or that there is no selection into a particular level of evolution coverage (Callaway et al., 2021). Given that the evolution coverage varies by state, selection into exposure to a different State Science Standards requires moving across states. At the same time, a costly relocation across states, presumably even involving families, seems unnecessary in most cases as students can simply switch to a private school (or do homeschooling).[17] Furthermore, the institutional idiosyncrasies determining the exact reform timing, as discussed in Section 2, make it difficult to pro-actively select into certain evolution coverages.

To explicitly address any remaining concerns about selection into evolution coverage, I conduct a range of robustness checks. First, the CS event-study graphs provide empirical support for the standard parallel trends assumption. Second, I show that reform effects also hold when transforming the continuous evolution score variable into binary variables, which does not require the strong parallel trend assumption. Third, I present CS estimators on a smaller sample without individuals belonging to the top and bottom 20 percent of the evolution score distribution. Following an idea by Marie and Zwiers (2022), this approach alleviates concerns about selection into evolution coverage, as individuals from the extremes of the evolution coverage distribution, with arguably the strongest incentive to move, are excluded from the sample. In sum, the analyses on time-varying treatment effects indicate that the conditions that bias TWFE estimations do not seem to be satisfied in this setting.

5 RESULTS

In three steps, this section shows that the evolution coverage in Science Standards affects the knowledge about evolution of students, the belief in evolution in adulthood, and the probability of working in life sciences.

[17]The main regression estimates are based on a sample of students from public and private schools (and homeschoolers in case of the GSS and ACS), and are hence net of spurious selection across school types.

5.1 Evolution Knowledge in School

The first analysis demonstrates that the evaluated reforms affect what students learn about evolution in school. To that end, I regress the share of questions on evolution answered correctly in the 12^{th} grade NAEP science test on the evolution coverage in Science Standards and different sets of control variables. Column (1) of Table 1 displays the raw correlation without any control variables. The positive raw correlation could imply that being exposed to a comprehensive coverage of evolution increases students' knowledge about evolution (reform effect). However, it could also reflect that comparatively high average levels of evolution knowledge raise the probability of states adopting Science Standards that cover evolution more comprehensively, for example because students might be less willing to accept creationist teaching (reverse causality). The positive raw correlation could also be driven by third variables such as parental education affecting both the probability of states adopting comprehensive Science Standards and the probability of students having knowledge about evolution (omitted variable bias).

To isolate the effect of the coverage of evolution in Science Standards on evolution knowledge, I add different sets of control variables in columns (2)-(4). When adding student-level control variables in column (2), or state and cohort fixed effects in column (3), or both the student controls and the fixed effects in column (4), the positive correlation persists and becomes even larger compared to the raw correlation. The full model in column (4) is the preferred specification since it exploits the reforms of Science Standards as a source of arguably exogenous variation by controlling for time-invariant differences between states, national differences between cohorts, time-varying state-specific shocks that affect adjacent cohorts similarly, as well as student level characteristics. It corresponds to the TWFE approach as specified in equation 1.

Regarding the main variable of interest, I find that being exposed to an evolution score of one, i.e., to a very comprehensive coverage of evolution, as compared to an evolution score of zero, i.e., to no or a non-scientific/creationist coverage of evolution, increases the share of questions on evolution answered correctly by 5.8 percentage points (p-value = 0.004). Given that, on average, students answer 32 percent of the questions on evolution correctly, the reported effect equals 18 percent of the sample mean.

Next, I hypothesize that the reform effect on evolution knowledge is disproportionately large for underprivileged students as they might rely more on schools to compensate for the lack of science exposure they receive from their parents and private environments. To begin, I note that variables typically associated with a lack of privilege such as being Black (relative to being White) tend to predict knowledge about evolution negatively, see Table 1. Conversely, variables typically reflecting privilege such as having a computer at home tend to predict knowledge about evolution positively.

Subgroup analysis by student characteristics reveals that the point estimates of reform effects on evolution knowledge are larger for underprivileged students, i.e. respectively for students who are Black, receive subsidized lunch, or do not have a computer at home, see Figure 2. However, these differences are not significantly different, in contrast to analogous subgroup results on evolution belief in adulthood shown below.

Furthermore, I present suggestive evidence from a teacher survey that the evaluated reforms indeed affect the evolution instruction in the classroom. Specifically, I show in this supplementary analysis that high school biology teachers who are exposed to a more comprehensive coverage of evolution in the Science Standards spend more time on teaching evolution (see Appendix A.3 and Table A.2). Other teaching strategies including the expression of teachers' personal opinions on the validity of evolution remain unaffected.

5.2 Evolution Belief in Adulthood

The second analysis shows that the teaching of evolution has a lasting impact on attitudes in adulthood, shedding light on the persistence of effects of scientific educational content. At the same time, it examines whether the effect on evolution knowledge translates into neutral settings in adulthood during which the scientifically correct answer is not encouraged. It could well be that students exposed to evolution content are both willing and able to answer science exam questions correctly to gain points in an exam as the NAEP, but that they are not convinced of the correctness of evolution theory.

Table 2 presents the GSS results from regressions of evolution belief in adulthood on the evolution score in high school, conditional on different sets of control variables. The raw correlation in column (1) is positive and significant. When subsequently adding student-level controls and fixed effects, the effect becomes even larger. The estimate in the full model presented in column (4) shows that individuals who were exposed to an evolution score of one, as compared to an evolution score of zero, are 33.3 percentage points more likely to believe in evolution in adulthood (p-value = 0.003). This effect amounts to 57 percent of the sample mean, making it larger than the corresponding effect on evolution knowledge reported in section 5.1.

To benchmark the effect size relative to other determinants of attitudes, I calculate persuasion rates (DellaVigna and Gentzkow, 2010). I define the persuasion rate induced by a reform changing the evolution score from zero to one as the average treatment effect on evolution belief divided by the share of students who do not believe in evolution in the entire sample.[18] The corresponding persuasion rate equals 79 percent. This is larger than the persuasion rates Cantoni et al. (2017) report for a Chinese school textbook reform on a range of outcomes.[19] It is also on the upper end of the persuasion rate distribution of media which includes rates from 3-8 percent (DellaVigna and Kaplan, 2007) to 65 percent (Enikolopov et al., 2011) for different media, settings and outcomes.

Regarding subgroups, the religion in which an individual was raised gives rise to a particularly interesting heterogeneity analysis, given the large differences in attitudes on

[18] Another definition of the persuasion rate would require dividing the treatment effect of the average reform by the share of individuals who do not believe in evolution and who studied before the evolution coverage was reformed. However, compositional differences by states and cohorts between individuals who studied before and after the reforms would bias results. Similarly, calculating the persuasion rate based on predicting treated and untreated students' beliefs and subtracting the treatment effect from the treated students' beliefs as in Cantoni et al. (2017) is not feasible as most students are treated to some extent even prior to the reforms, which then go in different directions with different intensities.

[19] They find the largest persuasion rates for the outcomes "Not investing in a bond" (50 percent persuasion rate) and "Trusting the local government" (47 percent persuasion rate).

evolution and creationism between religious groups. I first document that individuals raised as Evangelicals are 29 percentage points less likely to believe in evolution in adulthood compared to individuals being raised without a religion, conditional on the other regressors, see Table 2. The predictive power for individuals raised as Mainline Protestants is substantially weaker. For Catholics, it cannot be distinguished from those raised without a religion.

The subgroup analysis depicted in Figure 3 shows large reform effects for individuals raised as Mainline Protestants who hold moderate views on evolution on average. In contrast, students are less susceptible to the effects of evolution teaching if they were raised in a religion with strong average anti-evolution views like Evangelicals, or with strong average pro-evolution views like those raised as without a religion. The difference in reform effects between those raised as Mainline Protestants and those without a religion is statistically significant. Furthermore, reform effects are significantly larger for individuals who grew up in urban areas instead of rural ones, and for Blacks relative to Whites.

5.3 Occupational Choice

The third analysis reveals that the teaching of evolution translates into real-world high-stakes outcomes beyond attitudinal outcomes. Specifically, I focus on occupational choice as one high-stakes life decision in which an individuals' attitudes, values and beliefs may be revealed. I hypothesize that exposure to evolution theory (and hence to the fundamental scientific theory about the existence of life) affects individuals' probability of choosing to work in life sciences.

Using the ACS, this analysis shows that being exposed to a more comprehensive teaching of evolution in school increases the probability of working in life sciences during adulthood, as presented in Table 3. The point estimate is significant and stable across specifications. The full model presented in column (4) shows that individuals who were exposed to an evolution score of one, as compared to an evolution score of zero, are 0.035 percentage points more likely to work in life sciences as adults (p-value = 0.016).

This effect is small in absolute terms since few people work in life sciences in relation to the total US labor force. However, if expressed relative to the sample mean, the effect amounts to 23 percent.

The corresponding subgroup results by individual-level characteristics are in line with those from the previous subsections. Table A.3 shows that the point estimate on the probability of working in life sciences is larger for females than for males, and for Blacks than for other racial/ethnic groups, if expressed relative to the respective subsample mean (although the respective differences are insignificant). The ACS does not provide more individual-level covariates. However, one can conduct insightful subgroup analyses by the outcome variable, namely by the four subfields of life sciences. Figure 4 depicts the reform's positive and highly significant effect on the probability of working in biology. It is large in relative size, amounting to more than 39 percent of the sample mean. For all other subfields of life sciences, the reform effects that are much smaller in size and not statistically different from zero (the effect on biology is significantly different from the effects on agriculture and food as well as conservation and forestry). This subgroup pattern underpins that it is indeed the evolution teaching which drives reform effects, in line with the fundamental relevance of evolution for biology,[20] and given that evolution is being taught in biology.

6 ROBUSTNESS

The presented TWFE estimations can be interpreted causally if the identifying assumptions on parallel trends and the homogeneity of treatment effects hold, as described in section 4. To assess the validity of these assumptions, the following subsections show placebo tests, robustness checks on time-varying treatment effects including event-study graphs, and a large range of further checks. These include specifications that control for state-specific trends, and estimations that run on a

[20] This can be illustrated by the well-known assertion by Dobzhansky (2013) that "nothing in biology makes sense except in the light of evolution".

subsample of states with closely elected governors.

6.1 Placebo Tests

A key threat to internal validity are state-specific shocks, events, or trends that affect adjacent cohorts differently and coincided with reforms of the evolution coverage in Science Standards and affect the respective outcomes. The following placebo tests are designed to assess this threat. I show below that neither changes in knowledge on non-evolution scientific topics at the end of high school, nor changes in non-evolution scientific, religious and political attitudes in adulthood, nor changes in the probabilities of working in non-scientific occupations appear as reform effects. These findings support (i) that reform coefficients do not reflect underlying shocks or trends and (ii) that the reforms themselves do not affect these outcomes. The fact that the reform effects are neatly tied to the topic of evolution in all three independent datasets and outcomes therein supports a causal interpretation of the results of this paper. In sum, the placebo tests provide empirical support for the theoretical assessment that the exact reform timing is determined by institutional idiosyncrasies and not by scientific, religious or political confounders.

Evolution Knowledge in School: As is visible in Table 4, there is no effect of the evolution coverage in Science Standards on student knowledge in any of the non-evolution scientific topics such as reproduction or climate. I also conduct a regression in which the outcome variable is the share of questions on any of the non-evolution scientific topics answered correctly. This averaging allows for an overall assessment of the reform effect on non-evolution scientific knowledge. In theory, it could be significantly different from zero even if none of the individual effects are significant, for example through the reduction of measurement error. Averaging over all non-evolution scientific questions also alleviates concerns about multiple hypothesis testing (by reducing the number of tested hypotheses, see (Anderson, 2008)). As can be seen in Column (11), I find that this average effect on non-evolution scientific knowledge is also insignificant and close to zero.

To rule out that shocks or events specific to evolution but not related to the Science

Standards drive the main effect, I perform the main analysis on a placebo sample of students from private schools for whom the reforms were never compulsory. As shown in Table 5, the point estimate measuring the effect of the evolution coverage on evolution knowledge of private school students is very close to zero (although imprecisely estimated and therefore not significantly different from the point estimate of public school students). From this result I conclude that there are unlikely to be shocks or events related to evolution coincident with the reform of evolution coverage in Science Standards, at least as long as they affect both public and private school students. This result also suggests that there are no spillovers from public school curricula to private school curricula. In addition, the main effect holds on a joint sample of both public and private school students. This addresses the concern that spurious selection of students or school curricula into (or out of) private schools coincidental to the reform drives the results.

In another placebo analysis, I randomly reshuffle the reform years across the different reforming states. The density plot of the placebo coefficients based on 1000 permutations shows that the baseline estimated reform effect on evolution knowledge is larger than the 95^{th} percentile of the distribution of the 1000 placebo reform effects, see Figure A.3.

Evolution Belief in Adulthood: Table 6 demonstrates that the evolution coverage does not affect non-evolution scientific outcomes in adulthood on topics such as radioactivity or antibiotics. This is true for each of the nine non-evolution scientific outcomes, and for the average of all nine outcomes. This finding can be interpreted as the adulthood equivalent of the placebo tests on non-evolution scientific outcomes measured at the end of high school shown above.

Table 7 shows that the evolution coverage has no effect on religious outcomes in adulthood. Religious outcomes include variables capturing (i) religious beliefs such as belief in God, (ii) religious belonging such as religious affiliation or churchgoing, and (iii) general religiosity.[21] There is no effect that is statistically different from zero on any

[21] The distinction between believing and belonging follows Barro and McCleary (2003) and McCleary and Barro (2019) who find in cross-country analyses that believing stimulates economic growth, while belonging tends to reduce economic growth at given levels of religious beliefs.

religious outcome. The same is true for the average of all religious outcomes.

Table 8 demonstrates null effects of the reform on political outcomes. These outcomes encompass general political attitudes such as thinking of oneself as a Republican (as opposed to Democrat, Independent, or something else), political attitudes on specific topics typically regarded as controversial or partisan such as same-sex marriage, and preferences for governmental spending increases in areas such as alternative energy sources. There is no effect that is statistically different from zero on any political outcome including the average of all political outcomes.

Had there been, say, a negative coefficient on religiosity or political conservatism, it would be difficult to distinguish whether this result was driven by confounding shocks or by the reforms. However, a null finding implies that neither confounding shocks nor reform effects are in effect, because they plausibly operate in the same direction and do not offset each other. For example, it would be implausible to assume that negative confounders, for instance, state-specific church scandals coincident with the reforms, reduce the coverage of evolution in Science Standards causing a negative effect on religiosity; while at the same time offsetting this negative effect by increasing religiosity through other channels.

As before, I also show that the estimated baseline reform effect on evolution belief is larger than the 95^{th} percentile of the distribution of 1000 placebo reform effects based on random reshuffling of reform years across reforming states, see Figure A.4.

Occupational Choice: To begin, this analysis contrasts the reform effect on life sciences with effects on other scientific occupational fields. Table 9 shows that the positive and significant reform effect on the probability of working in life sciences is the most significant and largest (if measured relative its sample mean). It is followed by a positive effect on physical sciences. The point estimates on social sciences and science technicians are smaller in size and not significant, implying that effects are strongest for natural sciences.

An additional analysis about reform effects on non-scientific occupational fields does not yield significant results for any of the 25 non-scientific occupational fields, see Table

10. This finding suggests that there are no confounding shocks along specific topical dimensions. Had there been a specific negative effect on the probability of working one field, say, finance, one could be concerned about shocks simultaneous to the evolution reforms that deter individuals from working in finance. To increase statistical power and reduce concerns about multiple hypothesis testing, I also run a regression on the probability of working in any non-scientific field. Conversely to the positive effects on the natural sciences, I here find a negative effect, see Column (26). This finding implies that the reform caused a net increase in scientists, rather than a shift within different scientific subfields.

In addition, I demonstrate that the estimated reform effect on the probability of working in life sciences is larger than the 90^{th} percentile of the distribution of 1000 placebo reform effects based on random reshuffling of reform years across reforming states, see Figure A.5.

6.2 Time-Varying Treatment Effects, and Further Robustness

To address concerns about heterogeneous treatment effects, I implement the estimator by Callaway and Sant'Anna (2021). The CS estimator has to be estimated separately for subsets of states that reduce and expand the evolution coverage, respectively, because joint CS estimations would cancel out effects from opposing reforms. Table A.4 presents the parametric CS estimators for the subset of states reducing the evolution coverage. As expected, this reduction decreases evolution knowledge, belief, and the probability of working in life sciences. Column (1) shows the estimates for the full sample of the states reducing the evolution coverage. For the first outcome evolution knowledge by the end of high school, the CS estimate shows a reform effect of (an absolute value of) 5.6 percentage points. It is highly significant and very similar to the overall TWFE effect of 5.8 percentage points. The coefficients on evolution belief in adulthood and the probability of working in life sciences are also significant and close to the overall TWFE estimates, respectively. In column (2), I also present regressions in which all individuals belonging to the top and bottom 20 percent of the evolution score distribution are dropped from the sample. Here, the CS estimator remains significant and becomes even larger in absolute terms for all three outcomes. This finding is not only policy-relevant, but also alleviates concerns about the continuous nature of the evolution score measure.

For the subset of states expanding the evolution coverage, the CS estimates are positive for all three outcomes, as expected, see Table A.5. The effects are relatively small in size and insignificant for the full subsample. However, the effects become larger and mostly significant when estimated on the reduced sample dropping individuals belonging to the top and bottom 20 percent of the evolution score distribution.

Another, more direct way to assess pre-trends and account for heterogeneous treatment effects in the presence of staggered treatment timing are CS event study graphs. For example, evolution belief trending in the direction of estimated reform effects prior to the reform could indicate a bias from underlying trends in the data. Analogous to the parametric approach described above, CS event-study models are estimated separately for the subsets of states that reduce and expand the evolution coverage, respectively.

Figure A.6 displays CS event-study graphs for the subset of states where the reform reduces the evolution coverage in Science Standards.[22] They are depicted one below the other for the three main outcomes: evolution knowledge by the end of high school, evolution belief in adulthood, and the probability of working in life sciences. For all three outcomes, there is no indication of differential pre-trends between reforming and non-reforming states, supporting the parallel trends assumption.[23] Furthermore, it holds for all three outcomes that reform effects set in shortly after reform adoption. As expected, evolution knowledge, belief, and the probability of working in life sciences decrease significantly after removal of evolution from the Science Standards.

[22]In all CS event-study graphs, three years are grouped together to one bin to smooth the number of observations across bins as not all microdata are collected in every year (see section 3). The bins at the beginning (end) of the domain additionally include the years prior to (following) the domain's starting (ending) year. Black (red) 95% confidence intervals indicate pre (post) reform years. Longer post-reform time horizons are not available in the microdata.

[23]The only significant pre-reform coefficient appears in the first analysis on evolution knowledge at year -4.

For the subset of states where the reform expands the evolution coverage in Science Standards, the CS event-study graphs, again, provide empirical support for the parallel trends assumption. Moreover, the expansions in evolution content led to gradual increases in evolution knowledge, belief, and the probability to work in life sciences, see Figure A.7.

Appendix A.4.1 presents another robustness check on time-varying treatment effects that estimates the main regressions on subsamples of individuals coming from states with the same reform year following Cengiz et al. (2019). In sum, the findings of the entire section on time-varying treatment effects indicate that the conditions under which the main TWFE estimator is biased due to negative weights do not seem to be satisfied in my setting.

Appendix A.4.2 covers a range of further robustness checks including specifications that control for state-specific trends, are estimated on a subsample of closely elected governors, account for multiple hypothesis testing, and define treatment as a binary variable.

7 CONCLUSION

This paper shows that school curricula have lasting effects on students. To demonstrate this, the paper focuses on the teaching of evolution theory in the US. Exploiting institutional idiosyncrasies in the timing of reforms of the evolution coverage in US State Science Education Standards, I first document that the teaching of evolution causally affects students' knowledge about evolution at the end of high school. Second, I show that the teaching of evolution shapes attitudes on evolution of exposed students in adulthood. Third, I demonstrate that the teaching of evolution impacts high-stakes life decisions, namely occupational choice. In sum, the three sets of findings exemplify that science education can have lasting effects on students by affecting their knowledge, attitudes and choices.

To illustrate the effect sizes, I calculate changes in outcomes that one would expect to observe if all states adopted Science Standards with a highly comprehensive evolution coverage relative to the average coverage in the sample. Linear extrapolation of the presented estimation results suggests that the evolution belief in the US population would increase by 20 percent of the sample mean in such a scenario. Analogously, the number of adults working in life sciences would increase by 8 percent of the sample mean, and in the subfield of biology by 13 percent of the sample mean.

The three sets of results provide empirical support to important arguments raised in the policy debate about evolution teaching. As suggested by proponents of evolution teaching, the results indicate that teaching evolution has wider economic and societal benefits given the positive effects of scientific knowledge (Hanushek and Woessmann, 2008), scientific attitudes (Brzezinski et al., 2021), and working in STEM occupations (Peri et al., 2015) on individual and societal outcomes. Consensus on topics such as evolution could also reduce societal polarization and its associated costs (Alesina et al., 2020). Furthermore, the results speak against a major concern brought forward by some skeptics of evolution teaching, namely that teaching evolution might undermine students' religiosity. The null findings on various religious outcomes imply that neither believing in nor belonging to a religion (Barro and McCleary, 2003; McCleary and Barro, 2019) is crowded out by teaching evolution. The same is true for political attitudes.

This paper shows that the content of education standards is relevant for individuals in the short- and long-run. This conclusion challenges the notion that education standards have no meaningful impact on students as prevalent in the academic and political debate. It has been argued that, in reality, there is limited scope for education standards to affect teaching due to the dominance of other factors such as the teachers' own ideology for curriculum design in school (Moore et al., 2003a; Loveless, 2021). Still, legal pressures on school districts to follow education standards, the reflection of the content of education standards in textbooks, as well as the gradual expansion of standardized testing covering the content of education standards have arguably incentivized teachers to follow education standards. The analyses presented in this paper empirically demonstrate that they do in fact affect what students learn.

More broadly, this paper shows that the content of school curricula and instruction shapes students over the long-term. This is true even for a topic like evolution that is highly charged in political and societal debates. Despite its fundamental relevance for and overwhelming acceptance in science, people have strong partisan views on it. These views are likely to be determined by a multitude of factors. Still, what schools teach has long lasting effects on individuals' fundamental views and translates into high-stakes choices.

Beyond the evolution content of US State Science Education Standards evaluated in this paper, the findings indicate potential relevance of other education policies that increase the time teachers spend on teaching evolution.[24] Examples of such policies include the adoption of the Next Generation Science Standards and improvements in pre-service teacher education (Plutzer et al., 2020).

Beyond the US, the findings may also have a bearing for other countries where the teaching of evolution is controversial.[25]

Beyond the topic of evolution, the findings of this paper might also be relevant more broadly for further topics of science teaching, such as vaccinations, climate change or the trust in science in general. It is up to future research to study this explicitly.

[24] Between 2007 and 2019, the average number of hours a high school biology teachers in U.S. public schools spend on teaching human evolution almost doubled from 4.1 to 7.7 class hours (Plutzer et al., 2020).

[25] Examples include Israel, Turkey and India, as illustrated by the news headlines "Israeli schools largely avoid teaching evolution" by the Times of Israel (Staff, 2018), "Turkey's new school year: Jihad in, evolution out" by the BBC (Altunas, 2017), and "Indian education minister dismisses theory of evolution" by the Guardian (Safi, 2018).

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge (2017). "When should you adjust standard errors for clustering?" *National Bureau of Economic Research Working Paper Series.*

Akter, Sonia, Jeff Bennett, and Michael B. Ward (2012). "Climate Change Scepticism and Public Support for Mitigation: Evidence from an Australian Choice Experiment". *Global Environmental Change* 22(3), pp. 736–745.

Alesina, Alberto and Nicola Fuchs-Schündeln (2007). "Good-bye Lenin (or Not?): The Effect of Communism on People's Preferences". *American Economic Review* 97(4), pp. 1507–1528.

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva (2020). "The Polarization of Reality". *AEA Papers and Proceedings* 110, pp. 324–28.

Algan, Yann, Daniel Cohen, Eva Davoine, Martial Foucault, and Stefanie Stantcheva (2021). "Trust in scientists in times of pandemic: Panel evidence from 12 countries". *Proceedings of the National Academy of Sciences* 118(40).

Altonji, Joseph G., Peter Arcidiacono, and Arnaud Maurel (2016). "Chapter 7 - The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects". Ed. by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann. Vol. 5. Handbook of the Economics of Education. Elsevier, pp. 305–396.

Altunas, Öykü (2017). "Turkey's new school year: Jihad in, evolution out". *BBC.* URL: https://www.bbc.com/news/world-europe-41296714 (visited on 07/09/2021).

American Association for the Advancement of Science (2021). "Dialogue Science, Ethics, and Religion". *Thematic Areas.* URL: https://www.aaas.org/programs/dialogue-science-ethics-and-religion/thematic-areas (visited on 05/08/2021).

Anderson, Michael L. (2008). "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects". *Journal of the American Statistical Association* 103(484), pp. 1481–1495.

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom (2016a). "College Attrition and the Dynamics of Information Revelation". *National Bureau of Economic Research Working Paper Series, No. 22325.*

Arcidiacono, Peter, Esteban M. Aucejo, and V. Joseph Hotz (2016b). "University Differences in the Graduation of Minorities in STEM Fields: Evidence from California". *American Economic Review* 106(3), pp. 525–62.

Arcidiacono, Peter, V. Joseph Hotz, Arnaud Maurel, and Teresa Romano (2020). "Ex Ante Returns and Occupational Choice". *Journal of Political Economy* 128(12), pp. 4475–4522.

Arold, Benjamin W. and M. Danish Shakeel (2021). "The Unintended Effects of Common Core State Standards on Non-Targeted Subjects". *PEPG 21-03. Program on Education Policy and Governance Working Papers Series, Harvard University.*

Athey, Susan and Guido W. Imbens (2021). "Design-Based Analysis in Difference-In-Differences Settings with Staggered Adoption". *Journal of Econometrics, forthcoming.*

Bailey, Michael, Drew M. Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel, and Johannes Stroebel (2020). *Social Networks Shape Beliefs and Behavior: Evidence from Social Distancing During the COVID-19 Pandemic.* Tech. rep.

Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo (2019). "Nation-building Through Compulsory Schooling during the Age of Mass Migration". *Economic Journal* 129(617), pp. 62–109.

Barro, Robert J. (2001). "Human Capital and Growth". *American Economic Review* 91(2), pp. 12–17.

Barro, Robert J. and Rachel M. McCleary (2003). "Religion and Economic Growth across Countries". *American Sociological Review* 68(5), pp. 760–781.

Bazzi, Samuel, Masyhur Hilmy, and Benjamin Marx (2020). "Islam and the State: Religious Education in the Age of Mass Schooling". *National Bureau of Economic Research Working Paper Series No. 27073.*

Beale, H. K. (1941). *A History of the Freedom of Teaching in American Schools in Report of the Commission on the Social Studies, Part 16.* Chicago, IL: Charles Scribner & Sons.

Becker, Sascha O., Markus Nagler, and Ludger Woessmann (2017). "Education and Religious Participation: City-Level Evidence from Germany's Secularization Period 1890–1930". *Journal of Economic Growth* 22(3), pp. 273–311.

Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli (2006). "Adaptive linear step-up procedures that control the false discovery rate". *Biometrika* 93(3), pp. 491–507.

Berkman, Michael and Eric Plutzer (2011). "Defeating Creationism in the Courtroom, But Not in the Classroom". *Science* 331(6016), pp. 404–405.

Berkman, Michael B., Julianna Sandell Pacheco, and Eric Plutzer (2008). "Evolution and Creationism in America's Classrooms: A National Portrait". *PLoS Biology* 6(5), pp. 920–924.

Bishop, John H. (1989). "Is the Test Score Decline Responsible for the Productivity Growth Decline?" *American Economic Review* 79(1), pp. 178–197.

Bisin, Alberto and Thierry Verdier (2001). "The Economics of Cultural Transmission and the Dynamics of Preferences". *Journal of Economic Theory* 97(2), pp. 298–319.

Bordon, Paola and Chao Fu (2015). "College-Major Choice to College-Then-Major Choice". *Review of Economic Studies* 82(4), pp. 1247–1288.

Borusyak, Kirill, Xavier Jaravel, and Jan Spiess (2021). "Revisiting Event Study Designs: Robust and Efficient Estimation". *Working Paper.*

Brzezinski, Adam, Valentin Kecht, David Van Dijcke, and Austin L Wright (2021). "Science skepticism reduced compliance with COVID-19 shelter-in-place policies in the United States". *Nature Human Behaviour*, pp. 1–9.

Bursztyn, Leonardo, Aakaash Rao, Christopher P. Roth, and David H. Yanagizawa-Drott (2020). "Misinformation During a Pandemic". *National Bureau of Economic Research Working Paper Series, No. 27417.*

Butcher, Kristin F., Patrick J. McEwan, and Akila Weerapana (2014). “The Effects of an Anti-Grade-Inflation Policy at Wellesley College”. *Journal of Economic Perspectives* 28(3), pp. 189–204.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro HC Sant’Anna (2021). “Difference-in-differences with a continuous treatment”. *arXiv preprint arXiv:2107.02637*.

Callaway, Brantly and Pedro H.C. Sant’Anna (2021). “Difference-in-Differences with multiple time periods”. *Journal of Econometrics* 225(2), pp. 200–230.

Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang (2017). “Curriculum and Ideology”. *Journal of Political Economy* 125(2), pp. 338–392.

Cantoni, Davide and Noam Yuchtman (2013). “The Political Economy of Educational Content and Development: Lessons from History”. *Journal of Development Economics* 104, pp. 233–244.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019). “The Effect of Minimum Wages on Low-Wage Jobs”. *Quarterly Journal of Economics* 134(3), pp. 1405–1454.

Chaisemartin, Clément de and Xavier D’Haultfœuille (2020). “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects”. *American Economic Review* 110(9), pp. 2964–96.

Clots-Figueras, Irma and Paolo Masella (2013). “Education, Language and Identity”. *Economic Journal* 123(570), pp. 332–357.

Conger, Dylan, Alec I. Kennedy, Mark C. Long, and Raymond McGhee (2021). “The Effect of Advanced Placement Science on Students’ Skills, Confidence, and Stress”. *Journal of Human Resources* 56(1), pp. 93–124.

Cortes, Kalena and Joshua Goodman (2014). “Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement”. *American Economic Review* 104(5), pp. 400–405.

Costrell, Robert M. (1994). "A Simple Model of Educational Standards". *American Economic Review* 84(4), pp. 956–971.

— (1997). "Can Centralized Educational Standards Raise Welfare?" *Journal of Public Economics* 65(3), pp. 271–293.

Darwin, Charles R. (1859). *On the Origin of Species*. London: John Murray.

De Chaisemartin, Clément and Xavier D'Haultfoeuille (2022). "Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey". *National Bureau of Economic Research Working Paper Series* No. 29691.

DellaVigna, Stefano and Matthew Gentzkow (2010). "Persuasion: Empirical Evidence". *Annual Review of Economics* 2(1), pp. 643–669.

DellaVigna, Stefano and Ethan Kaplan (2007). "The Fox News Effect: Media Bias and Voting". *Quarterly Journal of Economics* 122(3), pp. 1187–1234.

Deming, David J and Kadeem Noray (2020). "Earnings Dynamics, Changing Job Skills, and STEM Careers". *Quarterly Journal of Economics* 135(4), pp. 1965–2005.

Dobzhansky, Theodosius (2013). "Nothing in Biology Makes Sense Except in the Light of Evolution". *American Biology Teacher* 75(2), pp. 87–91.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya (2011). "Media and Political Persuasion: Evidence from Russia". *American Economic Review* 101(7), pp. 3253–85.

European Commission (2020). "European skills agenda for sustainable competitiveness, social fairness and resilience". *Publications Office of the European Union Luxembourg*.

Giorcelli, Michela, Nicola Lacetera, and Astrid Marinoni (2022). "How Does Scientific Progress Affect Cultural Changes? A Digital Text Analysis". *Journal of Economic Growth*.

Giuliano, Paola and Antonio Spilimbergo (2014). "Growing up in a Recession". *Review of Economic Studies* 81(2), pp. 787–817.

Glaeser, Edward L. and Bruce I. Sacerdote (2008). "Education and Religion". *Journal of Human Capital* 2(2), pp. 188–215.

Goodman, Joshua (2019). "The Labor of Division: Returns to Compulsory High School Math Coursework". *Journal of Labor Economics* 37(4), pp. 1141–1182.

Goodman-Bacon, Andrew (2021). "Difference-in-Differences with Variation in Treatment Timing". *Journal of Econometrics, forthcoming.*

Graham, Loren (2016). *Lysenko's Ghost: Epigenetics and Russia.* Harvard University Press.

Griliches, Zvi (1992). "The Search for R&D Spillovers". *Scandinavian Journal of Economics* 94, pp. 29–47.

Gross, Paul R (2005). "The State of State Science Standards". *Thomas B. Fordham Foundation & Institute.*

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2008). "Social Capital as Good Culture". *Journal of the European Economic Association* 6, pp. 295–320.

Haberman, Clyde (2017). "Questioning Evolution: The Push to Change Science Class". *New York Times.* URL: https://www.nytimes.com/2017/11/19/us/retro-report-evolution-science.html (visited on 07/09/2021).

Hanushek, Eric A. (1986). "The Economics of Schooling: Production and Efficiency in Public Schools". *Journal of Economic Literature* 24(3), pp. 1141–1177.

Hanushek, Eric A. and Ludger Woessmann (2008). "The Role of Cognitive Skills in Economic Development". *Journal of Economic Literature* 46(3), pp. 607–668.

— (2012). "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation". *Journal of Economic Growth* 17(4), pp. 267–321.

Hastings, Justine S., Christopher A. Neilson, and Seth D. Zimmerman (2013). "Are Some Degrees Worth More than Others? Evidence from College Admission Cutoffs in Chile". *National Bureau of Economic Research Working Paper Series, No. 19241.*

Hungerman, Daniel (2014). "The effect of education on religion: Evidence from compulsory schooling laws". *Journal of Economic Behavior and Organization* 104(C), pp. 52–63.

Huxley, T. H (1880). "The Coming of Age of the Origin of Species". *Nature* 22(549), pp. 1–4.

Iannaccone, Laurence R. (1998). “Introduction to the Economics of Religion”. *Journal of Economic Literature* 36(3), pp. 1465–1495.

Iyer, Sriya (2016). “The New Economics of Religion”. *Journal of Economic Literature* 54(2), pp. 395–441.

Jin, Qiang, Syed Hassan Raza, Muhammad Yousaf, Umer Zaman, and Jenny Marisa Lim Dao Siang (2021). “Can Communication Strategies Combat COVID-19 Vaccine Hesitancy with Trade-Off between Public Service Messages and Public Skepticism? Experimental Evidence from Pakistan”. *Vaccines* 9(7).

Jones, Charles I. (1995). “R & D-Based Models of Economic Growth”. *Journal of Political Economy* 103(4), pp. 759–784.

Joravsky, David (1962). “The Lysenko Affair”. *Scientific American* 207(5), pp. 41–49.

Kerr, William R. and William F. Lincoln (2010). “The Supply Side of Innovation: H-1B Visa Reforms and U.S. Ethnic Invention”. *Journal of Labor Economics* 28(3), pp. 473–508.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad (2016). “Field of Study, Earnings, and Self-Selection”. *Quarterly Journal of Economics* 131(3), pp. 1057–1111.

Larson, Edward J. (1999). “The Scopes Trial and the Evolving Concept of Freedom”. *Virginia Law Review* 85(3), pp. 503–529.

Lerner, Lawrence S (2000a). “Good and Bad Science in US Schools”. *Nature* 407(6802), pp. 287–290.

— (2000b). “Good Science, Bad Science: Teaching Evolution in the States”. *Washington, DC: Thomas B. Fordham Foundation.*

Linden, Sander van der, Anthony Leiserowitz, Geoffrey Feinberg, and Edward Maibach (2015). “The Scientific Consensus on Climate Change as a Gateway Belief: Experimental Evidence”. *PloS ONE* 10, e0118489.

Lott, John R. Jr. (1999). “Public Schooling, Indoctrination, and Totalitarianism”. *Journal of Political Economy* 107(6), pp. 127–157.

Loveless, Tom (2021). *Between the State and the Schoolhouse: Understanding the Failure of Common Core.* Harvard Education Press.

Lucas, Robert E. (1988). "On the Mechanics of Economic Development". *Journal of Monetary Economics* 22(1), pp. 3–42.

Marie, Olivier and Esmee Zwiers (2022). "Religious Barriers to Birth Control Access". *CEPR Discussion Paper* No. 17427.

Marks, Jonathan (2012). "Why be Against Darwin? Creationism, Racism, and the Roots of Anthropology". *American Journal of Physical Anthropology* 149(55), pp. 95–104.

Martin, Alicia R, Christopher R Gignoux, Raymond K Walters, Genevieve L Wojcik, Benjamin M Neale, Simon Gravel, Mark J Daly, Carlos D Bustamante, and Eimear E Kenny (2017). "Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations". *American Journal of Human Genetics* 100(4), pp. 635–649.

Martinez-Bravo, Monica and Andreas Stegmann (2021). "In Vaccines We Trust? The Effects of the CIA's Vaccine Ruse on Immunization in Pakistan". *Journal of the European Economic Association.*

Mayr, Ernst (1991). *One Long Argument: Charles Darwin and the Genesis of Modern Evolutionary Thought.* Vol. 2. Harvard University Press.

McCleary, Rachel M. and Robert J. Barro (2006). "Religion and Economy". *Journal of Economic Perspectives* 20(2), pp. 49–72.

— (2019). *The Wealth of Religions: The Political Economy of Believing and Belonging.* Princeton University Press.

Mead, Louise and Anton Mates (2009). "Why Science Standards are Important to a Strong Science Curriculum and How States Measure Up". *Evolution: Education and Outreach* 2, pp. 359–371.

Meyersson, Erik (2014). "Islamic Rule and the Empowerment of the Poor and Pious". *Econometrica* 82(1), pp. 229–269.

Moore, Randy, Murray Jensen, and Jay Hatch (2003a). "The Problems with State Educational Standards." *Science Education Review* 2(3).

Moore, Randy, Murray Jensen, and Jay Hatch (2003b). "Twenty Questions: What Have the Courts Said about the Teaching of Evolution and Creationism in Public Schools?" *BioScience* 53(8), pp. 766–771.

National Commission on Excellence in Education (1983). "A Nation at Risk: The Imperative for Educational Reform". *Elementary School Journal* 84(2), pp. 113–130.

National Science and Technology Council (2018). "Charting a Course for Success: America's Strategy for STEM Education". *Report by the Committee on STEM Education of the National Science and Technology Council.*

Numbers, Ronald L. (1982). "Creationism in 20th-Century America". *Science* 218(4572), pp. 538–544.

Oates, Wallace E. (1999). "An Essay on Fiscal Federalism". *Journal of Economic Literature* 37(3), pp. 1120–1149.

Peri, Giovanni, Kevin Shih, and Chad Sparber (2015). "STEM Workers, H-1B Visas, and Productivity in US Cities". *Journal of Labor Economics* 33(1), pp. 225–255.

Pew Research Center (2009). *Public Praises Science; Scientists Fault Public, Media - Scientific Achievements Less Prominent Than a Decade Ago.*

— (2015). *Public and Scientists' View on Science and Society.*

Plutzer, Eric, Glenn Branch, and Ann Reid (2020). "Teaching Evolution in US Public Schools: A Continuing Challenge". *Evolution: Education and Outreach* 13(1), pp. 1–15.

Porter, Catherine and Danila Serra (2020). "Gender Differences in the Choice of Major: The Importance of Female Role Models". *American Economic Journal: Applied Economics* 12(3), pp. 226–54.

Rios-Avila, Fernando, Pedro Sant'Anna, and Brantly Callaway (2022). "CSDID: Stata module for the estimation of Difference-in-Difference models with multiple time periods". *Working Paper.*

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek (2020). "Integrated Public Use Microdata Series USA: Version 10.0 [dataset]". *Minneapolis, MN: IPUMS.*

Sacerdote, Bruce (2001). “Peer Effects with Random Assignment: Results for Dartmouth Roommates”. *Quarterly Journal of Economics* 116(2), pp. 681–704.

Safi, Michael (2018). “Indian education minister dismisses theory of evolution”. *The Guardian*. URL: https://www.theguardian.com/world/2018/jan/23/indian-education-minister-dismisses-theory-of-evolution-satyapal-singh (visited on 07/09/2021).

Schultheiss, Tobias and Uschi Backes-Gellner (2021). “Updated education curricula and accelerated technology diffusion in the workplace: Micro-evidence on the race between education and technology”. *Swiss Leading House Economics of Education Working Paper Series* No. 173.

Sen, Ananya and Catherine Tucker (2022). “Product Quality and Performance in the Internet Age: Evidence from Creationist-Friendly Curriculum”. *Journal of Marketing Research* 59(1), pp. 211–229.

Skoog, Gerald (2005). “The Coverage of Human Evolution in High School Biology Textbooks in the 20th Century and in Current State Science Standards”. *Science & Education* 14(3), pp. 395–422.

Staff, Toi (2018). “Israeli schools largely avoid teaching evolution”. *Times of Israel*. URL: https://www.timesofisrael.com/israeli-schools-largely-avoid-teaching-evolution-report/ (visited on 07/09/2021).

Stinebrickner, Ralph and Todd R. Stinebrickner (2014). “A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout”. *Review of Economic Studies* 81(1 (286)), pp. 426–472.

Sun, Liyang and Sarah Abraham (2021). “Estimating dynamic treatment effects in event studies with heterogeneous treatment effects”. *Journal of Econometrics* 225(2), pp. 175–199.

Swanson, Christopher B (2005). “Evolution in State Science Education Standards”. *Editorial Projects in Education Research Center*.

Tabellini, Guido (2008). “The Scope of Cooperation: Values and Incentives”. *Quarterly Journal of Economics* 123(3), pp. 905–950.

Tiebout, Charles M. (1956). “A Pure Theory of Local Expenditures”. *Journal of Political Economy* 64(5), pp. 416–424.

Wiswall, Matthew and Basit Zafar (2015). “Determinants of College Major Choice: Identification using an Information Experiment”. *Review of Economic Studies* 82(2), pp. 791–824.

— (2018). “Preference for the Workplace, Investment in Human Capital, and Gender”. *Quarterly Journal of Economics* 133(1), pp. 457–507.

MAIN TABLES AND FIGURES

Table 1 – Effect of evolution coverage in Science Standards on evolution knowledge in school

	Evolution Knowledge			
	(1)	(2)	(3)	(4)
Evolution Score	0.036*	0.039***	0.069*	0.058**
	(0.018)	(0.011)	(0.028)	(0.019)
Female		-0.003		-0.004
		(0.008)		(0.008)
Race/Ethnicity: Black		-0.084***		-0.082***
		(0.007)		(0.007)
Race/Ethnicity: Hispanic		-0.051***		-0.048***
		(0.009)		(0.009)
Subsidized Lunch		-0.012		-0.011
		(0.006)		(0.006)
Parental Education: Graduated High School		-0.009		-0.010
		(0.011)		(0.012)
Parental Education: Some education after High School		0.002		0.002
		(0.011)		(0.011)
Parental Education: Graduated College		0.023*		0.021
		(0.010)		(0.011)
Computer at Home		0.011		0.022**
		(0.007)		(0.007)
State FEs	NO	NO	YES	YES
Birth Year FEs	NO	NO	YES	YES
Other Controls	NO	YES	NO	YES
Mean of Dep. Var.	0.32	0.32	0.32	0.32
Std. Dev. of Dep. Var.	0.42	0.42	0.42	0.42
Adj. R-squared	0.001	0.043	0.015	0.049
Observations	15,530	15,520	15,530	15,520

Note: Dependent variable: Share of questions about evolution answered correctly. Other Controls: Indicator variables for asian (race/ethnicity) other (race/ethnicity), English language learner status, disability status, parental education, home possessions (books), and test year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Table 2 – Effect of evolution coverage in Science Standards on evolution belief in adulthood

	Evolution Belief			
	(1)	(2)	(3)	(4)
Evolution Score	0.108**	0.089**	0.205	0.333**
	(0.040)	(0.033)	(0.115)	(0.107)
Female		-0.053*		-0.050*
		(0.022)		(0.022)
Race/Ethnicity: Black		-0.158***		-0.149***
		(0.038)		(0.040)
Race/Ethnicity: Hispanic		-0.100*		-0.091
		(0.044)		(0.056)
Raised in Rural Area		-0.014		-0.003
		(0.024)		(0.025)
Raised as Protestant: Mainline		-0.141***		-0.121**
		(0.035)		(0.035)
Raised as Protestant: Evangelical		-0.302***		-0.290***
		(0.046)		(0.047)
Raised as Catholic		0.018		0.019
		(0.037)		(0.040)
State FEs	NO	NO	YES	YES
Birth Year FE	NO	NO	YES	YES
Other Controls	NO	YES	NO	YES
Mean of Dep. Var.	0.58	0.58	0.58	0.58
Std. Dev. of Dep. Var.	0.49	0.49	0.49	0.49
Adj. R-squared	0.005	0.088	0.038	0.107
Observations	1,812	1,801	1,812	1,801

Note: Dependent variable: Belief in Evolution ("Human beings, as we know them today, developed from earlier species of animals - Is that true or false?", Indicator variable, 1=true, 0=false; don't know). Other Controls: Indicator variables for white (race/ethnicity; omitted category) other (race/ethnicity), parents born abroad, parental education, having lived with parents in adolescence, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism (all reported here), no religion (omitted category), judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: General Social Survey.

Table 3 – Effect of evolution coverage in Science Standards on probability of working in life sciences

	Life Sciences			
	(1)	(2)	(3)	(4)
Evolution Score	0.039*	0.035*	0.035*	0.035*
	(0.018)	(0.013)	(0.014)	(0.014)
Female		0.014*		0.013*
		(0.006)		(0.006)
Race/Ethnicity: Black		-0.127***		-0.115***
		(0.007)		(0.006)
Race/Ethnicity: Hispanic		-0.106***		-0.085***
		(0.008)		(0.008)
State FEs	NO	NO	YES	YES
Birth Year FEs	NO	NO	YES	YES
Other Controls	NO	YES	NO	YES
Mean of Dep. Var.	0.15	0.15	0.15	0.15
Std. Dev. of Dep. Var.	3.84	3.84	3.84	3.84
Adj. R-squared	0.000	0.000	0.000	0.001
Observations	6,460,650	6,460,650	6,460,650	6,460,650

Note: Dependent variable: Probability of working in life sciences (multiplied by 100 for interpretability). Other controls: Indicator variables for asian (race/ethnicities), other (race/ethnicities), multiple (race/ethnicities), and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: American Community Survey.

Table 4 – Placebo tests: Effect of evolution coverage in Science Standards on knowledge about non-evolution scientific topics

	Main Outcome:	Placebo Outcomes (Knowledge on the following Non-Evolution Scientific Topic):									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Evolution	Motion	Matter and Mass	Energy	Reproduction	Climate	Pollution	Earth	Tectonics	Universe	Average
Evolution Score	0.058**	-0.006	0.013	-0.014	0.024	-0.026	-0.014	0.039	0.009	0.003	0.008
	(0.019)	(0.045)	(0.047)	(0.024)	(0.028)	(0.046)	(0.059)	(0.023)	(0.020)	(0.051)	(0.013)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.32	0.51	0.30	0.38	0.38	0.39	0.15	0.41	0.17	0.31	0.35
Std. Dev. of Dep. Var.	0.42	0.43	0.43	0.43	0.42	0.39	0.27	0.42	0.27	0.42	0.27
Adj. R-squared	0.049	0.090	0.140	0.110	0.100	0.073	0.077	0.140	0.312	0.049	0.180
Observations	15,520	9,510	17,000	22,910	18,610	19,080	4,770	13,710	6,730	6,260	32,850

Note: Dependent variables: Shares of questions answered correctly about scientific topics indicated in the column headers. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, English language learner status, disability status, parental education, home possessions (separate indicator variables for computer and books), and test year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Table 5 – Placebo tests: Effect of evolution coverage in Science Standards on evolution knowledge in private schools

	Evolution Knowledge		
	(1) Only Public School Students	(2) Only Private School Students	(3) Overall
Evolution Score	0.058**	0.003	0.046*
	(0.019)	(0.062)	(0.018)
State FEs	YES	YES	YES
Birth Year FEs	YES	YES	YES
Controls	YES	YES	YES
Mean of Dep. Var.	0.32	0.43	0.34
Std. Dev. of Dep. Var.	0.42	0.38	0.41
Adj. R-squared	0.049	0.045	0.056
Observations	15,520	3,160	18,680

Note: Regressions by students' school type as indicated in the column headers. Dependent variable: Share of questions about evolution answered correctly. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, English language learner status, disability status, parental education, home possessions (separate indicator variables for computer and books), and test year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Table 6 – Placebo tests: Effect of evolution coverage in Science Standards on non-evolution scientific topics

	Main Outcome:	Placebo Outcomes: Non-Evolution Scientific Topics									
	(1) Evolution	(2) Earth	(3) Radioactivity	(4) Reproduction	(5) Lasers	(6) Electrons	(7) Antibiotics	(8) Universe	(9) Tectonics	(10) Sun	(11) Average
Evolution Score	0.333**	0.000	-0.125	0.192	0.042	-0.133	0.175	-0.191	-0.179	-0.266	-0.053
	(0.107)	(0.091)	(0.138)	(0.107)	(0.181)	(0.144)	(0.158)	(0.113)	(0.091)	(0.164)	(0.057)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.58	0.88	0.66	0.62	0.47	0.57	0.51	0.45	0.85	0.79	0.65
Std. Dev. of Dep. Var.	0.49	0.32	0.47	0.48	0.50	0.49	0.50	0.50	0.36	0.41	0.22
Adj. R-squared	0.107	0.051	0.074	0.035	0.091	0.038	0.092	0.113	0.054	0.090	0.158
Observations	1,801	1,800	1,797	1,747	1,799	1,800	1,801	1,796	1,801	1,801	1,801

Note: Dependent variables: Shares of questions answered correctly about scientific topics indicated in the column headers. Controls: Indicator variables for gender, races/ethnicities, parents born abroad, parental education, having lived with parents in adolescence, raised in rural area, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism, no religion, judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: General Social Survey.

Table 7 – Placebo tests: Effect of evolution coverage in Science Standards on religious outcomes

	Main Outcome:	Religious Placebo Outcomes: Believing					Religious Placebo Outcomes: Belonging and Activities					Religious Placebo Outcomes: Overall			Religious Placebo Outcomes: Average
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Evolution	God	Bible	Afterlife	Rebirth	Strong Believer	Religious Affiliation	Church-going	Church Activities	Personal Prayer	Missionize	Spiritual Person	Religious Person	Fundamentalist	Average
Evolution Score	0.333**	-0.021	-0.009	-0.023	-0.097	-0.091	0.099	0.102	-0.072	-0.108	-0.028	0.060	-0.206	0.015	-0.029
	(0.107)	(0.096)	(0.114)	(0.128)	(0.123)	(0.150)	(0.109)	(0.120)	(0.091)	(0.113)	(0.073)	(0.101)	(0.162)	(0.129)	(0.068)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.58	0.87	0.72	0.73	0.34	0.32	0.70	0.35	0.17	0.65	0.41	0.56	0.43	0.24	0.50
Std. Dev. of Dep. Var.	0.49	0.33	0.45	0.45	0.47	0.47	0.46	0.48	0.38	0.48	0.49	0.50	0.50	0.43	0.28
Adj. R-squared	0.107	0.104	0.086	0.029	0.134	0.086	0.163	0.081	0.040	0.149	0.130	0.061	0.077	0.244	0.201
Observations	1,801	1,797	1,794	1,797	1,796	1,783	1,799	1,801	1,801	1,798	1,801	1,801	1,799	1,718	1,801

Note: Dependent variables: Main outcome and religious outcomes indicated in the column headers. Controls: Indicator variables for gender, races/ethnicities, parents born abroad, parental education, having lived with parents in adolescence, raised in rural area, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism, no religion, judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: General Social Survey.

Table 8 – Placebo tests: Effect of evolution coverage in Science Standards on political outcomes

	Main Outcome:	Political Placebo Outcomes: General		Political Placebo Outcomes: Conservative attitude on:							Political Placebo Outcomes: Conservative attitude on governmental spending for:								Political Placebo Outcomes: Average
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
	Evolution	Repub-lican	Conser-vative	Prayer in Public Schools	Sex Education in Public Schools	Same-Sex Marriage	Abor-tion	Marijuana Legali-zation	Capital Punish-ment	Gun Control	Immi-gration	Environ-ment	Alternative Energy Sources	Educa-tion	Scientific Research	Reducing Income Differences	Assistance to the Poor	Conditions of Blacks	Average
Evolution Score	0.333**	0.028	-0.018	-0.061	-0.085	-0.084	0.257	0.006	-0.101	0.013	0.082	0.062	-0.045	0.047	0.116	-0.137	-0.148	0.044	-0.002
	(0.107)	(0.114)	(0.132)	(0.104)	(0.065)	(0.225)	(0.188)	(0.153)	(0.162)	(0.135)	(0.141)	(0.105)	(0.146)	(0.104)	(0.138)	(0.100)	(0.101)	(0.116)	(0.053)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.58	0.28	0.26	0.45	0.06	0.36	0.55	0.45	0.61	0.30	0.84	0.29	0.41	0.19	0.63	0.46	0.51	0.64	0.43
Std. Dev. of Dep. Var.	0.49	0.45	0.44	0.50	0.25	0.48	0.50	0.50	0.49	0.46	0.36	0.45	0.49	0.39	0.48	0.50	0.50	0.48	0.19
Adj. R-squared	0.107	0.058	0.024	0.087	0.048	0.110	0.028	0.079	0.063	0.060	-0.001	0.018	0.011	0.027	0.012	0.001	0.060	0.095	0.072
Observations	1,801	1,792	1,791	1,200	1,200	1,056	1,059	1,336	1,788	1,063	1,174	1,801	1,195	1,801	1,799	1,337	1,798	1,788	1,801

Note: Dependent variables: Main outcome and political outcomes indicated in the column headers. All political outcomes are coded such that an increase in the variable implies an increase in political conservatism, see Appendix A.2.2 for details. Controls: Indicator variables for gender, races/ethnicities, parents born abroad, parental education, having lived with parents in adolescence, raised in rural area, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism, no religion, judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: General Social Survey.

Table 9 – Effect of evolution coverage in Science Standards on probability of working in different scientific occupational fields

	(1) Life Sciences	(2) Physical Sciences	(3) Social Sciences	(4) Science Technicians
Evolution Score	0.035*	0.042*	0.031	-0.027
	(0.014)	(0.018)	(0.028)	(0.053)
State FEs	YES	YES	YES	YES
Birth Year FEs	YES	YES	YES	YES
Controls	YES	YES	YES	YES
Mean of Dep. Var.	0.15	0.22	0.16	0.32
Std. Dev. of Dep. Var.	3.84	4.68	4.03	5.62
Adj. R-squared	0.00064	0.00083	0.00096	0.00073
Observations	6,460,650	6,460,650	6,460,650	6,460,650

Note: Dependent variable: Probability of working in occupational field indicated in the column header (multiplied by 100 for interpretability). Controls: Indicator variables for gender, races/ethnicities, and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: American Community Survey.

Table 10 – Effect of evolution coverage in Science Standards on probability of working in specific non-scientific occupational fields

	Non-Scientific Occupational Fields												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Management	Business Operations	Finance	IT	Engineering	Social	Legal	Education	Arts	Health-care	Health-care Support	Protective Services	Food
Evolution Score	0.281	0.127	0.120	0.115	0.023	0.133	0.145	-0.015	0.074	-0.053	-0.045	0.035	0.125
	(0.191)	(0.099)	(0.079)	(0.080)	(0.068)	(0.097)	(0.078)	(0.241)	(0.111)	(0.189)	(0.182)	(0.201)	(0.386)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	5.44	1.87	1.63	1.97	1.39	1.39	0.82	5.58	2.04	4.33	2.63	2.13	7.88
Std. Dev. of Dep. Var.	22.68	13.54	12.67	13.89	11.69	11.70	9.01	22.96	14.15	20.36	15.99	14.43	26.94
Adj. R-squared	0.01926	0.00489	0.00421	0.01092	0.00774	0.00417	0.00382	0.02060	0.00198	0.02620	0.01644	0.00655	0.01962
Observations	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650

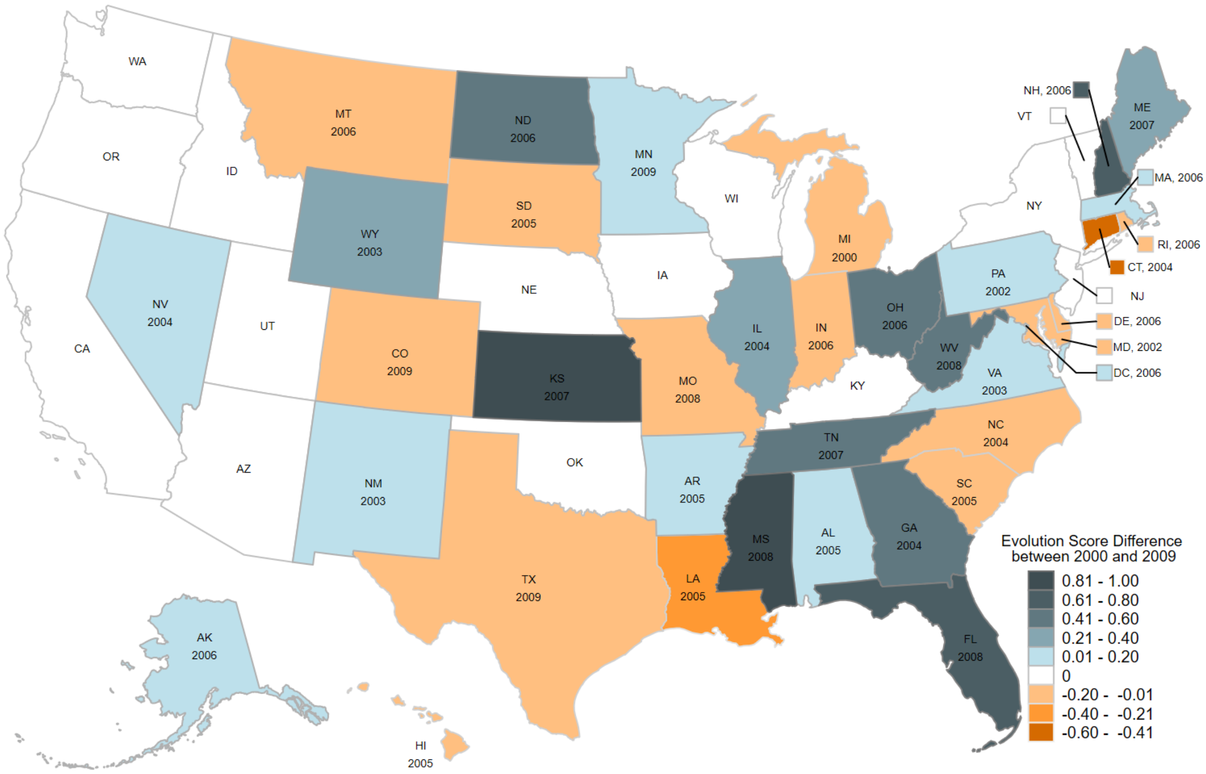
Note: Dependent variable: Probability of working in occupational field indicated in the column header (multiplied by 100 for interpretability). Controls: Indicator variables for gender, races/ethnicities, and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: American Community Survey.

Table 10 (continued) – Placebo test: Effect of evolution coverage in Science Standards on probability of working in specific non-scientific occupational fields

	Non-Scientific Occupational Fields												Any Non-Scientific Occupational Field
	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
	Buildings	Personal Care	Sales	Office	Farming	Construc-tion	Extrac-tion	Installa-tion	Produc-tion	Transporta-tion	Armed Forces	Unemployed / Not in Labor Market	
Evolution Score	0.073	-0.078	-0.502	-0.059	-0.011	-0.227	0.058	0.060	-0.171	0.190	0.046	-0.525	-0.081*
	(0.283)	(0.284)	(0.267)	(0.488)	(0.111)	(0.249)	(0.047)	(0.117)	(0.220)	(0.329)	(0.080)	(1.778)	(0.037)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	2.84	3.96	11.42	13.25	0.66	4.45	0.21	2.80	4.87	5.35	0.75	9.48	99.15
Std. Dev. of Dep. Var.	16.61	19.51	31.81	33.90	8.07	20.63	4.60	16.50	21.53	22.50	8.62	29.30	9.16
Adj. R-squared	0.00622	0.01477	0.01027	0.02388	0.00547	0.04532	0.00463	0.02608	0.01894	0.02964	0.00522	0.03058	0.00166
Observations	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650

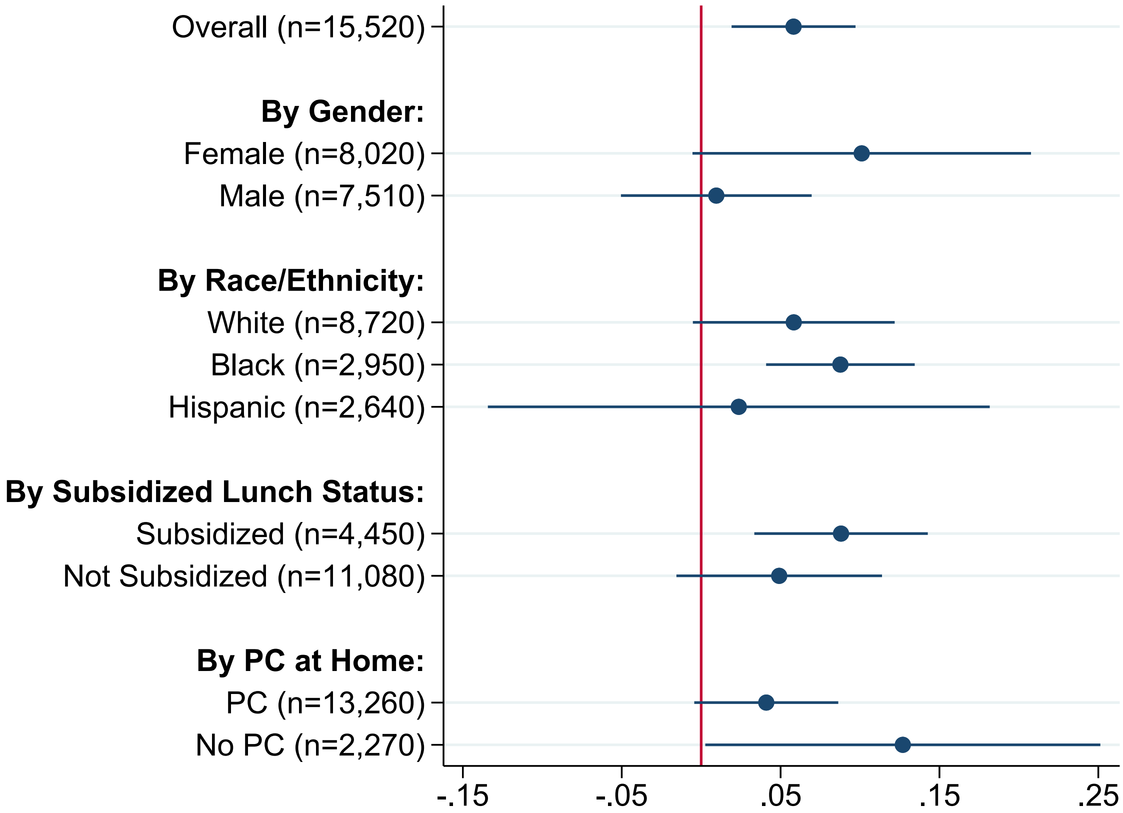
Note: Dependent variable: Probability of working in occupational field indicated in the column header (multiplied by 100 for interpretability). Controls: Indicator variables for gender, races/ethnicities, and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: American Community Survey.

Figure 1 – US map of evolution score difference between 2000 and 2009



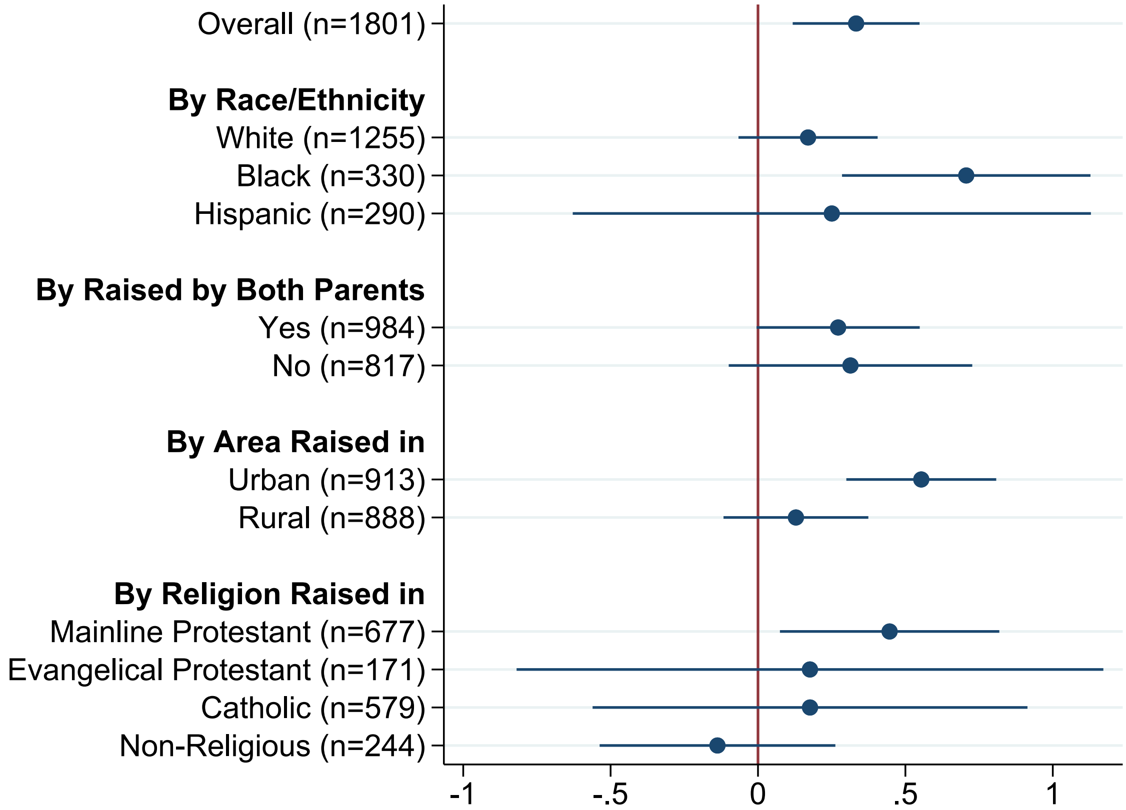
Note: Map depicts the evolution score difference, which I define as the evolution score of 2009 minus the evolution score of 2000. A positive (negative) difference implies an increase (decrease) in the evolution score between 2000 and 2009, as indicated by blue (orange) coloring. White coloring indicates no change of the evolution score between 2000 and 2009. The years reported below the two-letter state codes mark the respective reform years. A list of the evolution score differences and reform years underlying this map is provided in Table A.1. Data source: Lerner (2000b), Mead and Mates (2009)

Figure 2 – Effect of evolution coverage in Science Standards on evolution knowledge in school, by subgroups



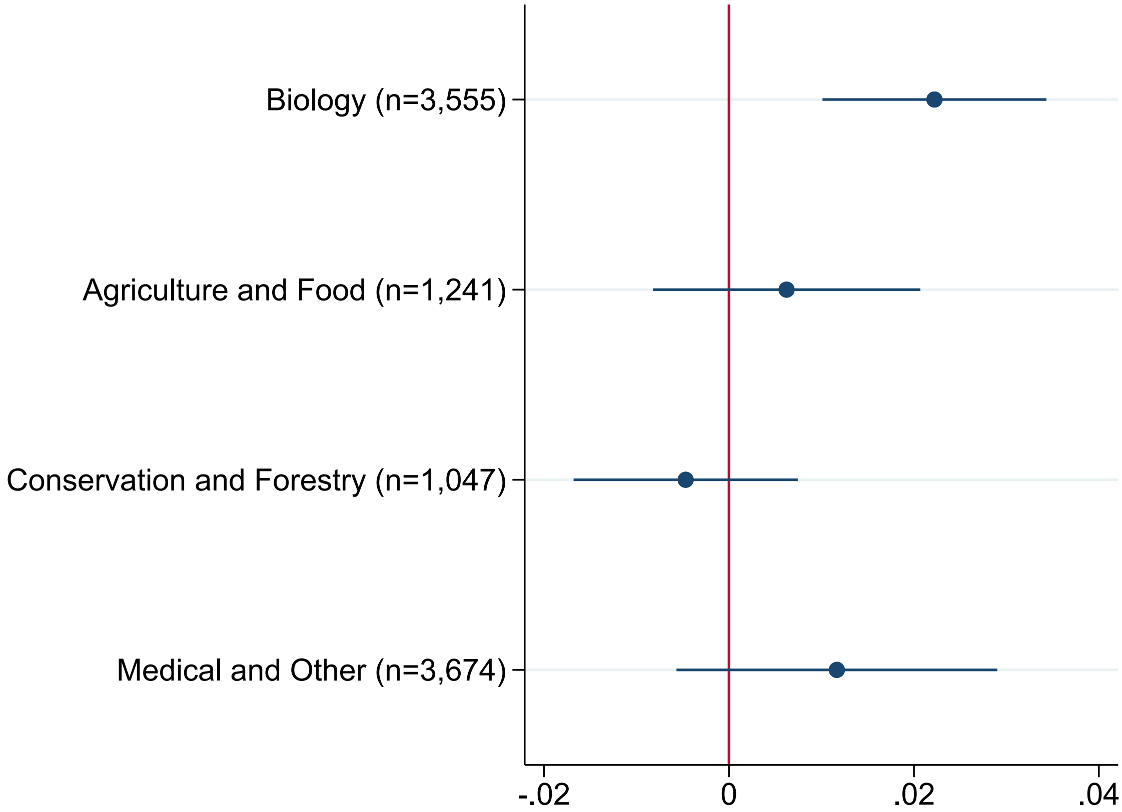
Note: Figure displays effect of evolution coverage in Science Standards on share of questions about evolution answered correctly, by individual subgroup as indicated in rows. Sample sizes of subgroups in parenthesis. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, English language learner status, disability status, parental education, home possessions (separate indicator variables for computer and books), and fixed effects for state, birth year, and test year. Standard errors clustered at the state level. 95% confidence intervals displayed. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Figure 3 – Effect of evolution coverage in Science Standards on evolution belief in adulthood, by subgroups



Note: Figure displays effect of evolution coverage in Science Standards on belief in evolution in adulthood ("Human beings, as we know them today, developed from earlier species of animals - Is that true or false?", Indicator variable, 1=true, 0=false; don't know), by individual subgroup as indicated in rows. Sample sizes of subgroups in parenthesis. Controls: Indicator variables for gender, races/ethnicities, parents born abroad, parental education, having lived with parents in adolescence, raised in rural area, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism, no religion, judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), and fixed effects for state, birth year, and test year. Standard errors clustered at the state level. 95% confidence intervals displayed. Data source: General Social Survey

Figure 4 – Effect of evolution coverage in Science Standards on probability of working in life sciences, by subfields of life sciences



Note: Figure displays effect of evolution coverage in Science Standards on probability of working in life sciences, by subfields of life sciences as indicated in rows (multiplied by 100 for interpretability). Sample sizes of subfields in parenthesis (raw value). Controls: Indicator variables for gender, races/ethnicities and fixed effects for state, birth year, and test year. Standard errors clustered at the state level. 95% confidence intervals displayed. Data source: American Community Survey

A FOR ONLINE PUBLICATION

A.1 Reform Examples from Florida and Texas

Reforms of the evolution coverage in Science Standards form the basis of the two-way fixed effects design performed in this paper. The following two reform examples illustrate how such reforms come into existence. While Florida expanded the evolution coverage in 2008, Texas reduced it in 2009. The Science Standard in power in Florida before 2008 did not mention the word "evolution", and its discussion of evolutionary processes (under a different wording) were minimal.[26] In February 2008, the Florida Board of Education voted 4:3 in favor of a new Science Standard that comprehensively included evolution. This close majority emerged following years of debating and drafting the Standard. In fact, the Standard was re-drafted yet again just hours before the final vote. Replacing the term "evolution" by "the scientific theory of evolution" ultimately secured the majority. The new Standard comprehensively captured biological, geological, cosmological and even human evolution (Mead and Mates, 2009).

In contrast to Florida, Texas reduced the evolution coverage in 2009. The evolution coverage in the Science Standard in place in 2000 was described as "brief but satisfactory" (Lerner, 2000b, p.15). It encompassed all areas of evolution except for human evolution. In 2003, Don McLeroy, the then-chairman of the Texas Board of Education, advocated a far more limited evolution coverage. He stated that he personally does not believe in Darwin's evolution theory and in the earth being older than a couple of thousand years, which was in part reflected in the Science Standard proposal. In 2003, his reform proposal found no majority in the Board of Education, and years of debate followed. In 2009, he proposed another Science Standard which required that "strengths and weaknesses" of evolution should be taught. Some regarded this as an attempt to facilitating teaching of creationism at the teachers' discretion, without explicitly mentioning creationism in

[26] Lerner (2000b, p.14) describes the Science Standard as "Extensive standards that skim lightly over biological and geological evolution without ever mentioning the word. Not satisfactory."

the Science Standard. It was voted down 8-7. A second version required students to study the "sufficiency or insufficiency" of key principles of evolution. It was also voted down 8-7. A third attempt which contained more subtle creationist jargon was ultimately approved by 13-2 votes. This new Science Standard omitted some areas of the teaching of evolution and added "pieces of creationist jargon" (Mead and Mates, 2009, p.366). For example, the phrase that "the estimated age of the universe was 14 billion years" was removed. Notably, the reforms in Florida and Texas did not follow a partisan change since all governors in 21^{st} Century Florida and Texas have been Republican. Both reform examples shed light on the political process behind such reforms, and show that they do not simply result from a change of government.

A.2 Data Appendix

A.2.1 NAEP: Evolution Knowledge in School

The NAEP is a congressionally mandated project also known as the Nation's Report Card. It is administered by the National Center for Education Statistics (NCES), a body within the Institute of Education Sciences (IES) and the US Department of Education. Throughout the paper I use data from the Main-NAEP and not the Long-Term Trend NAEP, as the Main-NAEP has much larger sample sizes, is state-representative and, particularly relevant for this analysis, also covers science.

I categorize a question as addressing evolution if it contains the words "evolution" or "natural selection", or if it contains words that are based on the same word stem, such as "evolutionary".[27] I transform each question into a binary variable that is set equal to one if the correct answer was given, and equal to zero for any other answer, whether it is incorrect, partially correct, off task, etc. (the specific available categories depend on the question type). Figure A.8 presents two sample questions, one on general Darwinian theory, and one on evolutionary trees. For each student, I calculate the share of questions on evolution that the student answered correctly. This share serves as the main outcome variable measuring a student's knowledge on evolution. I analogously group questions into nine categories of scientific topics other than evolution.[28] An increase in such a variable always implies an increase in scientific knowledge on the topic in question. Hence, the average knowledge of non-evolution scientific topics is calculated as the non-

[27]Sometimes, the dataset does not contain the full wording of the questions but question keywords due to data protection reasons. I code such cases analogously, i.e. as addressing evolution if their keywords contain the words "evolution" or "natural selection", or if they contain words that are based on the same word stem.

[28]Notably, the number of questions available for each scientific topic in the pool of NAEP questions differs across scientific topics. Furthermore, each student receives only a subset of the pool of questions during the test. This test design explains why the number of questions answered on a given scientific topic differs across students. To address this issue, I calculate the share of questions answered correctly on a given scientific topic instead of the number of questions answered correctly. Moreover, this test design also explains why the number of students answering questions on a given scientific topic differs across scientific topics, resulting in varying sample sizes across scientific topics. These sample size differences are not a result of spurious selection, but are induced by the test design.

missing average of all questions from these nine non-evolution scientific topics. Table A.6 shows that knowledge on evolution is in general positively correlated with knowledge on non-evolution scientific topics.

In the preferred sample cut of keeping individuals who enter high school after 1990 and before 2010, I use the NAEP tests for science in grade 12 from 1996, 2000, 2005, and 2009. Regarding missings, I keep all students without missings on basic controls such as gender, and who come from birth cohorts of at least 10 observations. I set missings of other control variables to zero and add separate explanatory binary variables to account for these missings.[29]

The descriptive statistics for the main treatment, outcome, and control variables are presented in Table A.7. The treatment variable "evolution score" captures the score of the evolution coverage of the Science Standard in power in the state and year of a student's high school entry. The average evolution score equals 0.65, implying that students were on average exposed to a "satisfactory" evolution coverage.[30] The main outcome variable "evolution knowledge" is defined as the share of questions on evolution a student answers correctly. The fact that only 32 percent of questions on evolution are answered correctly on average underscores the difficulty of the test. For instance, the shares of students giving correct answers to the sample questions reported in Figure A.8 equal 54 percent and 28 percent, respectively. Regarding non-evolution scientific topics, the average share of questions answered correctly amounts to 35 percent, indicating that the average difficulty of questions on evolution is largely similar to the overall difficulty. With regards to control variables, about half of the sample are female (51 percent). The shares of Whites, Blacks, Hispanics and Asians amount to 57 percent, 19 percent, 16 percent, and 6 percent, respectively. The various variables on the socio-economic status indicate that a non-negligible share of students from grade 12 lives in underprivileged circumstances as measured by subsidized lunch status (30 percent), having no PC at home (16 percent), or disability status (11 percent).

[29]The results are robust to not imputing the missings, as shown in Table A.13.

[30]Lerner (2000b) classifies evolution scores between 0.60 and 0.79 as "satisfactory".

A.2.2 GSS: Evolution Belief in Adulthood

The GSS data in the main sample comes from the waves from 2006, 2008, 2010, 2012, 2014, and 2016. For all scientific (religious) outcome variables, an increase in the variable always implies an increase in scientific knowledge (religiosity). Hence, the average of scientific (religious) variables is calculated as the average of the non-missing scientific (religious) outcomes. For the political outcomes, an increase in a variable does not always imply an increase in the same political direction. For example, being in favor of prayer in public schools is positively correlated with being politically conservative in the sample, while being in favor of sex education in public schools is negatively correlated with being politically conservative. To facilitate the interpretation, I recode all political variables such that an increase in the variable implies an increase in political conservatism. Specifically, I define a political attitude as politically conservative (progressive) if the share of respondents self-identifying as politically conservative who believe in/approve of the attitude is larger (smaller) than the share of respondents not self-identifying as politically conservative.[31]

Regarding correlations, the belief in evolution is almost only positively correlated with the other scientific outcomes, see Table A.8. For all religious variables, I find a negative raw correlation with evolution belief as is visible in Table A.9. The correlations between politically conservative attitudes on different topics and evolution belief also tend to be negative, see Table A.10.

Table A.11 shows the descriptive statistics for the main treatment, outcome, and control variables. The individuals in the sample were exposed to an evolution score of 0.63 on average which is very similar to corresponding sample average in NAEP, as

[31]This definition allows for an unequivocal assignment of attitudes to political conservatism/progressivism, but does not reflect absolute belief/approval rates. For example, the share of conservatives being in favor of increasing governmental spending for education is larger than 50 percent, but smaller than the corresponding share of non-conservatives. Thus, being in favor of increasing governmental spending for education is classified a progressive attitude. The raw variable is then recoded such that an increase in the variable implies an increase in conservatism. In general, the recoding is undertaken to facilitate the interpretation of results and allow for meaningful averaging across political outcomes.

expected given the comparable sample cut. Regarding the main outcome variable evolution belief, I find that 58 percent of sample say that the aforementioned statement about evolution is true. Regarding non-evolution scientific topics, six of the nine non-evolution scientific topics display higher rates of correct answers than evolution, with an average of 64 percent across these nine topics. Looking at religious outcomes, I note that 87 percent of respondents believe in God, and 70 percent are affiliated with a church. To give examples on conservative political attitudes, 45 percent come out in favor of a conservative attitude on prayer on public schools (implying being in favor of prayer in schools), while only 6 percent come out in favor of a conservative attitude on sex education in public schools (implying being against sex education). With regard to the religious upbringing of these individuals, I observe that the most common religion/denomination an individual was raised in is Mainline Protestantism (37 percent), followed by Catholicism (32 percent), Non-Religious/Agnosticism/Atheism (14 percent), and Evangelicalism (9 percent).

A.2.3 ACS: Occupational Choice

The estimation sample combines ACS waves from 2000-2017. The descriptive statistics are presented in Table A.12. For the treatment variable, I find that the average evolution score exposure equals 0.67, which is similar to the corresponding averages from the analyses using the NAEP and the GSS. Regarding the outcome variables, all indicator variables for occupational fields are multiplied by 100 to ease the readability of descriptive statistics and reform effects. Hence, the descriptive statistics including mean and standard deviation are multiplied by 100 as well. For example, the sample mean of respondents working in life sciences equals 0.15, which implies that 0.15 percent of the sample work in this field. 0.85 percent of the sample work in any scientific field. Out of all 26 occupational fields, the largest sample shares are found for respondents working in office (13.2 percent) and in sales (11.5 percent).

A.3 Evidence from High School Biology Teachers

This supplementary analysis provides suggestive evidence that teachers base their evolution teaching on the evolution coverage of the Science Standard in power in their state. To show this, I draw on the National Survey of High School Biology Teachers conducted by the Survey Research Center of Penn State University in 2007. The survey focuses on the biology teachers' approach to teaching evolution (and creationism) in the classroom, as well as their educational background and personal attitudes on evolution. It contains a nationally representative sample of high school biology teachers who are teaching in a public school where grades 9 and 10 are offered, who taught a high school biology class in at least the previous year, and who had not recently retired (see Berkman et al. (2008) and Berkman and Plutzer (2011) for more information).

First, I report that the large majority of biology teachers states that they align their evolution teaching with the evolution coverage of their Science Standard. Specifically, 88 percent of high school biology teachers strongly agree or agree with the statement "When I do teach evolution, I focus heavily on what students need to know to meet state science standards" (see Figure A.1).

Second, I show that high school biology teachers who are exposed to a more comprehensive evolution coverage in their Science Standard spend more time on teaching evolution. To demonstrate this, I link information on the time spent on teaching evolution (and various other pro-evolution and pro-creationism teaching strategies) to the evolution score measuring the evolution coverage of the Science Standard in power in the state of the teacher in 2007, the year of the survey. The between-states model is specified as follows:

$$Y_{is} = \beta \cdot Evolution\_Score_s + \gamma \cdot \mathbf{X_i} + \eta_c + \epsilon_{is} \tag{2}$$

where Y_{is} is the outcome of interest of teacher i, who teaches in state s and is surveyed in 2007. The treatment variable $Evolution\_Score_s$ measures the evolution coverage in the Science Standard in state s in 2007. β is the parameter of interest capturing the conditional association of the outcome with being exposed to a very comprehensive coverage of evolution ($Evolution\_Score_s$=1) as compared to being exposed to no or a non-scientific/creationist coverage of evolution ($Evolution\_Score_s$=0). The vector $\mathbf{X_i}$ contains control variables, η_c captures census division fixed effects, and an error term completes the model. The standard errors are clustered at the state level.

In contrast to the main effects shown in this paper, these conditional associations should be interpreted as suggestive rather than causal evidence. Due to the fact that the teacher data is only available for one year, there is no within-state variation over time that would allow for identification of effects from reforms of Science Standards. Instead, the variation here stems from differences between states at one point in time. The main concern for a causal interpretation is that not only the evolution coverage in Science Standards differs between states, but many other factors including teachers' own attitude on evolution. To partially account for that, I control for detailed teacher characteristics including information on their education about biology and evolution specifically, and their personal attitude and knowledge about evolution. Second, I control for census division fixed effects which ensures that the identifying variation stems from between-state comparisons within relatively homogeneous subgroups of states.

The conditional associations show that teachers with similar characteristics in the same census division in different states who are exposed to an evolution score of one, i.e. to a very comprehensive coverage of evolution, as compared to an evolution score of zero, i.e. to no or a non-scientific/creationist coverage of evolution, are 33 percentage points more likely to spend at least 5 class hours per year on teaching evolution (Table A.2). This positive, large, and significant association is specific to teaching hours spent on evolution. Other strategies regarding the teaching of evolution (and creationism) do not significantly differ by the evolution score. Taken together, the results presented in this supplementary analysis suggest that biology teachers (i) focus their evolution teaching on what students need to know to meet Science Standards, and (ii) adjust the time spent on teaching evolution accordingly, while other teaching strategies such as the expression of personal opinions on the validity of evolution do not differ.

A.4 Robustness Appendix

A.4.1 Additional Robustness on Time-Varying Treatment Effects

Beyond the presented CS analyses, the issue of negative weights can be assessed in this setting by estimating the main regressions on subsamples of individuals coming from states with the same reform year (plus individuals from non-reforming states as "clean control states") following Cengiz et al. (2019). Specifically, I run seven subsample regressions for the reform years 2003-2009 separately.[32] These regressions are reform-year-specific and do not exploit any staggered reform timing.

Figure A.9 depicts the seven reform-year-specific estimates of the effect of the evolution coverage on evolution knowledge, ordered by effect size. Reassuringly, six of the seven estimates are positive, and four of them are significant. The only negative point estimate is small in size and imprecisely estimated. The analogous analyses for evolution belief in adulthood, shown in Figure A.10, and for the probability of working in life sciences, presented in Figure A.11, yield a similar picture at slightly lower levels of significance. For all three analyses, the respective shares of significant effects are larger than analogous share of 25 percent in the corresponding analysis by Cengiz et al. (2019).

A.4.2 Further Robustness Checks

This subsection covers a range of further robustness checks. The first test replicates the main analysis on a subset of reforms which themselves can arguably be regarded as-good-as-random (and not only their specific timing). This subset contains reforms in states where the governor decides on the members of the State Board of Education, and where the governor ruling at the time of the reform adoption won the previous election by a small margin. In these states, the outcome of the election, and hence the political direction of the Boards of Education and their reforms, is somewhat arbitrary. Although the set of states with close pre-reform gubernatorial elections reduces the sample size by

[32] The reform years 2000-2002 are dropped from this analysis due to too few reforms and, hence, too small sample sizes.

around two thirds, the reform effects on evolution knowledge are robust (see column (1) of Table A.13). The same is true for analogous analyses on evolution belief in adulthood (see column (1) of Table A.14) and on the probability of working in life sciences (with the latter being estimated less precisely, see column (1) of Table A.15). These findings lend empirical support to a causal interpretation of the presented estimation results, even if it was not true that institutional idiosyncrasies were quasi-randomizing the reform timing.

Another, more direct, way to control for political changes is the inclusion of state-by-year controls for the political affiliation of the governor ruling in the state and year of the respective individuals' high school entry. As reported in column (2) of Tables A.13, A.14, and A.15, respectively, this test yields robust results throughout the three analyses both in terms of size and significance.

Adding state-specific time trends as control variables to the baseline TWFE model constitutes another way of assessing robustness. These trends explicitly account for time-varying state-specific shocks that affect adjacent cohorts differentially, but smoothly. As is visible in column (3) of the three Tables listed above, the levels of significance tend to decrease in this demanding specification, while the point estimates largely hold and partly become even larger.

Another robustness check reduces the sample to states that had only one reform event between 2000 and 2009 based on careful examination of academic articles, legal documents, and state education websites. As shown in column (4) of the three Tables listed above, the results are largely robust and partly even more pronounced.

In addition, the results hold if the observation period of the main sample is defined differently. As reported in columns (5) and (6) of the Tables listed above, the results are largely robust to sample definitions with fewer pre-reform cohorts, with the earliest cohorts starting high school in 1995 and 2000, respectively. Moreover, the results do not depend on the precise coding of the outcome variables (for this test, the column numbers depends on the analysis, see footnotes of the three Tables listed above for more information). For example, the results are robust to coding those individuals who do not know how to answer the question on evolution belief as a missing observation instead of non-believing. There are also corresponding results for the analysis on evolution knowledge, but not for the probability of working in life sciences as the latter has no such outcome category. The remaining columns of the three Tables listed above show that the results are largely robust to conducting logit and probit specifications, and to dropping missing observations of control variables instead of imputing them.

Although the tested hypotheses about effects on evolution knowledge, evolution belief, and the probability of working in life sciences are inherently linked to evolution teaching through the common focus on evolution, I still conduct robustness checks on multiple hypothesis testing. I implement two multiple hypotheses corrections, first by presenting p-values that are adjusted using the false discovery rate (FDR) procedure (Benjamini et al., 2006; Anderson, 2008), and second by implementing the particularly conservative Bonferroni correction. The treatment effects on the three main outcomes remain statistically significant when correcting for multiple hypothesis testing. Tables A.16, A.17, and A.18 show the adjusted p-values for the three analyses on evolution knowledge in school, evolution belief in adulthood, and the probability of working in life sciences, respectively. The adjusted p-values for these three outcomes range from 0.009 for the effect on evolution knowledge using the FDR procedure (Anderson, 2008) to 0.078 for the effect on the probability of working in the life sciences using the Bonferroni correction.

Lastly, the interpretation of the results does not change meaningfully when transforming the treatment variable to indicator variables. Specifically, the first (second) indicator variable is set to one if the evolution score is larger than 0.1 (0.2), and zero otherwise. The seven other indicator variables are coded accordingly. This coding eliminates a substantial amount of treatment variation, but allows to assess which domain of the evolution score distribution is particularly important for the production of evolution knowledge, evolution belief, and the probability of working in life sciences. Tables A.19, A.20, and A.21 show that most domains of the evolution score distribution are important for the production of outcomes with the exception of the highest value. This finding also implies that results do not hinge on the continuous nature of the treatment, alleviating concerns about the related strong parallel trends assumption (Callaway et al., 2021).

A.5 Supplementary Tables and Figures

Table A.1 – Evolution scores and reform year, by state

State	Evolution Score: 2009	Evolution Score: 2000	Evolution Score Difference 2009 - 2000	Reform Year	Only One Reform Event
Alabama	0.21	0.09	0.12	2005	NO
Alaska	0.59	0.48	0.11	2006	NO
Arkansas	0.66	0.55	0.11	2005	YES
DC	0.96	0.80	0.16	2006	YES
Florida	0.91	0.16	0.75	2008	YES
Georgia	0.66	0.07	0.59	2004	YES
Illinois	0.82	0.45	0.37	2004	YES
Kansas	0.96	0.00	0.96	2007	NO
Maine	0.68	0.30	0.38	2007	YES
Massachusetts	0.84	0.82	0.02	2006	NO
Minnesota	0.89	0.86	0.03	2009	NO
Mississippi	0.86	0.05	0.81	2008	NO
Nevada	0.77	0.70	0.07	2004	YES
New Hampshire	0.91	0.23	0.68	2006	YES
New Mexico	0.91	0.73	0.18	2003	YES
North Dakota	0.64	0.09	0.55	2006	NO
Ohio	0.86	0.28	0.58	2006	NO
Pennsylvania	0.96	0.91	0.05	2002	YES
Tennessee	0.55	0.02	0.53	2007	NO
Virginia	0.68	0.50	0.18	2003	YES
West Virginia	0.46	0.03	0.43	2008	NO
Wyoming	0.61	0.36	0.25	2003	YES
Colorado	0.82	0.86	-0.04	2009	NO
Connecticut	0.59	1.00	-0.41	2004	YES
Delaware	0.80	0.91	-0.11	2006	YES
Hawaii	0.75	0.91	-0.16	2005	YES
Indiana	0.96	1.00	-0.04	2006	NO
Louisiana	0.27	0.64	-0.37	2005	NO
Maryland	0.73	0.77	-0.04	2002	NO
Michigan	0.80	0.84	-0.04	2000	YES
Missouri	0.78	0.82	-0.04	2008	NO
Montana	0.75	0.82	-0.07	2006	YES
North Carolina	0.82	1.00	-0.18	2004	YES
Rhode Island	0.82	1.00	-0.18	2006	YES
South Carolina	0.91	0.95	-0.04	2005	NO
South Dakota	0.77	0.82	-0.05	2005	YES
Texas	0.46	0.64	-0.18	2009	YES
Arizona	0.82	0.82	0.00	-	-
California	1.00	1.00	0.00	-	-
Idaho	0.82	0.82	0.00	-	-
Iowa	0.77	No Standard	-	-	-
Kentucky	0.55	0.55	0.00	-	-
Nebraska	0.66	0.66	0.00	-	-
New Jersey	1.00	1.00	0.00	-	-
New York	0.68	0.68	0.00	-	-
Oklahoma	0.25	0.25	0.00	-	-
Oregon	0.82	0.82	0.00	-	-
Utah	0.82	0.82	0.00	-	-
Vermont	0.86	0.86	0.00	-	-
Washington	0.86	0.86	0.00	-	-
Wisconsin	0.55	0.55	0.00	-	-

Note: Table reports the evolution score from 2009 based on Mead and Mates (2009), the evolution score from 2000 based on Lerner (2000b), and the difference of the evolution scores (evolution score from 2009 minus evolution score from 2000). States are listed in three panels, positive, negative, and zero evolution score change. For states that changed their evolution score, the respective year of the (last) reform as noted in Mead and Mates (2009) is also provided, and whether this reform is the only reform event between 2000 and 2009. The latter information on the only reform event is based on Gross (2005), Swanson (2005) as well as my own examination of state education websites.

Table A.2 – Conditional associations of evolution coverage in Science Standards with pro-evolution and pro-creationism teaching strategies

	Pro-Evolution Teaching				Pro-Creationism Teaching			
	(1) Teaching Hours On Evolution	(2) Emphasize Consensus about Evolution	(3) Agree: Evolution Is Unifying Theme	(4) Emphasize Scientists Reject Creationism	(5) Teaching Hours On Creationism	(6) Emphasize Evolution May Be Wrong	(7) Emphasize Creationism As Valid Alternative	(8) Believe Evolution Not Needed For Good Course
Evolution Score	0.333**	0.041	0.116	0.193	0.008	-0.091	-0.051	-0.027
	(0.120)	(0.087)	(0.148)	(0.129)	(0.023)	(0.094)	(0.169)	(0.072)
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.67	0.79	0.65	0.52	0.01	0.71	0.30	0.13
Std. Dev. of Dep. Var.	0.47	0.41	0.48	0.50	0.11	0.45	0.46	0.33
Adj. R-squared	0.134	0.105	0.163	0.117	0.013	0.091	0.191	0.127
Observations	814	802	794	368	808	804	390	806

Note: Dependent variables (indicator variables) indicated in the column headers as follows : (1) Teacher typically spends at least 5 class hours in biology course for the year on general evolutionary processes; (2) When teaching evolution, teacher emphasizes the broad consensus that evolution is fact even as scientists disagree about the specific mechanisms through which evolution occurred; (3) Teacher agrees that evolution serves as the unifying theme for the content of the course; (4) When teaching creationism or intelligent design, teacher emphasizes that almost all scientists reject these as valid accounts of the origin of species; (5) Teacher typically spends at least 5 class hours in biology course for the year on intelligent design or creationism; (6) When teaching evolution, teacher emphasizes the possibility that portions of evolutionary theory may be proven wrong; (7) When teaching creationism or intelligent design, teacher emphasizes that this is a valid, scientific alternative to Darwinian explanations for the origin of species; (8) Teacher believes it is possible to offer an excellent general biology course for high school students that includes no mention of Darwin or evolutionary theory. Controls: Teacher's gender, age, years of teaching experience, undergraduate and graduate courses (separate variables for undergraduate and graduate credit hours in biology; specific college-level course in evolution; major, minor, or special emphasis in science education, biology, other science, statistics, or education), college degrees (separate variables for associate degree; Bachelor of Arts; Bachelor of Science; Master's degree in education; Master's degree in science; PhD in education; PhD in science), type of teaching certificate, teacher's continuing education about scientific debates of last years (separate variables for textbooks; science journalism; science education websites; scientific journals; or taking science courses), teacher's own view about evolution regardless of classroom activities (separate variables for creationist view; acknowledgment of evolution as guided by God; non-religious view on evolution), teacher's self-assessed knowledge about evolution (separate variables for excellent, very good, typical, or not good), high school biology assessment test in place, and census division fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: National Survey for High School Biology Teachers, 2007.

Table A.3 – Effect of evolution coverage in Science Standards on probability of working in life sciences, by subgroups

	By Gender		By Race/Ethnicity		
	(1) Females	(2) Males	(3) Whites	(4) Blacks	(5) Hispanics
Evolution Score	0.052* (0.020)	0.018 (0.020)	0.038* (0.016)	0.012 (0.016)	0.004 (0.034)
State FEs	YES	YES	YES	YES	YES
Birth Year FEs	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.15	0.14	0.16	0.04	0.06
Std. Dev. of Dep. Var.	3.92	3.75	4.05	2.06	2.44
Adj. R-squared	0.00068	0.00063	0.00047	0.00022	0.00030
Observations	3,220,042	3,240,608	5,023,449	789,587	765,295

Note: Regressions by selected subgroups, as indicated in the columns headers. Dependent variable: Probability of working in life sciences (multiplied by 100 for interpretability). Controls: Indicator variables for gender, races/ethnicities, and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: American Community Survey.

Table A.4 – CS parametric estimator: Only states reducing evolution coverage

	Full sample	Excluding individuals with 20% highest and lowest evolution coverage in Science Standards
	(1)	(2)
	(a) Evolution knowledge in school	
Evolution Score	-0.056***	-0.120***
	(0.017)	(0.016)
	(b) Evolution belief in adulthood	
Evolution Score	-0.274***	-0.384***
	(0.083)	(0.121)
	(c) Working in life sciences	
Evolution Score	-0.036*	-0.055**
	(0.016)	(0.021)

Note: Each entry is from separate regression model. CS estimator (Callaway and Sant'Anna, 2021), accounting for heterogeneous treatment effects and staggered treatment timing. Simple aggregation of all post treatment effects, using doubly robust inverse probability weighting. Controls: Not-yet-treated and never-treated observations. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data sources: (a) U.S. Department of Education, National Center for Education Statistics, 1996-2009 National Assessment of Educational Progress; (b) General Social Survey; (c) American Community Survey.

Table A.5 – CS parametric estimator: Only states expanding evolution coverage

	Full sample (1)	Excluding individuals with 20% highest and lowest evolution coverage in Science Standards (2)
	(a) Evolution knowledge in school	
Evolution Score	0.028	0.058**
	(0.022)	(0.022)
	(b) Evolution belief in adulthood	
Evolution Score	0.198	0.432**
	(0.127)	(0.173)
	(c) Working in life sciences	
Evolution Score	0.007	0.010
	(0.011)	(0.029)

Note: Each entry is from separate regression model. CS estimator (Callaway and Sant'Anna, 2021), accounting for heterogeneous treatment effects and staggered treatment timing. Simple aggregation of all post treatment effects, using doubly robust inverse probability weighting. Controls: Not-yet-treated and never-treated observations. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data sources: (a) U.S. Department of Education, National Center for Education Statistics, 1996-2009 National Assessment of Educational Progress; (b) General Social Survey; (c) American Community Survey.

Table A.6 – Correlation coefficients of knowledge about evolution and other scientific areas

	Evolution Knowledge
Motion	0.0894***
Matter and Mass	0.0836***
Energy	0.129***
Reproduction	0.283***
Climate	0.0524***
Pollution	0.150***
Earth	0.0924***
Tectonics	0.0183
Universe	0.117***
Non-Evolution Scientific Topics: Average	0.252***

Note: Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Table A.7 – Descriptive statistics of NAEP data

	Mean	Std. Dev.	Min.	Max.
Treatment Variable:				
Evolution Score	0.65	0.31	0.00	1.00
Main Outcome:				
Evolution Knowledge	0.32	0.42	0.00	1.00
Placebo Outcomes - Non-Evolution Scientific Topics:				
Motion	0.51	0.43	0.00	1.00
Matter and Mass	0.30	0.43	0.00	1.00
Energy	0.38	0.43	0.00	1.00
Reproduction	0.38	0.42	0.00	1.00
Climate	0.39	0.39	0.00	1.00
Pollution	0.15	0.28	0.00	1.00
Earth	0.41	0.42	0.00	1.00
Tectonics	0.17	0.27	0.00	1.00
Universe	0.32	0.42	0.00	1.00
Non-Evolution Scientific Topics: Average	0.35	0.27	0.00	1.00
Controls:				
Female	0.51	0.50	0.00	1.00
Race/Ethnicity: White	0.57	0.49	0.00	1.00
Race/Ethnicity: Black	0.19	0.39	0.00	1.00
Race/Ethnicity: Hispanic	0.16	0.37	0.00	1.00
Race/Ethnicity: Asian	0.06	0.23	0.00	1.00
Race/Ethnicity: Other	0.01	0.11	0.00	1.00
English Language Learner	0.05	0.22	0.00	1.00
Disabled	0.11	0.32	0.00	1.00
Subsidized Lunch	0.30	0.46	0.00	1.00
Parental Education: Did not finish High School	0.09	0.29	0.00	1.00
Parental Education: Graduated High School	0.20	0.40	0.00	1.00
Parental Education: Some education after High School	0.26	0.44	0.00	1.00
Parental Education: Graduated College	0.44	0.50	0.00	1.00
Computer at Home	0.84	0.37	0.00	1.00
Books at Home: 0−10	0.23	0.42	0.00	1.00
Books at Home: 11−25	0.27	0.44	0.00	1.00
Books at Home: 26−100	0.33	0.47	0.00	1.00
Books at Home: >100	0.17	0.38	0.00	1.00

Note: Descriptive statistics (mean, standard deviation, minimum, maximum) for treatment, outcome, and control variables. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Table A.8 – Correlation coefficients of belief in evolution and other scientific areas

	Evolution Belief
Earth	0.120***
Radioactivity	0.145***
Reproduction	-0.0222
Lasers	0.106***
Electrons	0.169***
Antibiotics	0.107***
Universe	0.415***
Tectonics	0.248***
Sun	0.109***
Non-Evolution Scientific Topics: Average	0.314***

Note: Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: General Social Survey.

Table A.9 – Correlation coefficients of evolution belief and religious outcomes

	Evolution Belief
God	-0.194***
Bible	-0.272***
Afterlife	-0.106***
Rebirth	-0.313***
Strong Believer	-0.284***
Religious Affiliation	-0.212***
Church-going	-0.267***
Church Activities	-0.203***
Personal Prayer	-0.282***
Missionize	-0.275***
Spiritual Person	-0.158***
Religious Person	-0.241***
Fundamentalist	-0.248***
Religious Outcomes: Average	-0.374***

Note: Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: General Social Survey.

Table A.10 – Correlation coefficients of evolution belief and political outcomes

	Evolution Belief
Republican	-0.120***
Conservative	-0.126***
Prayer in Public Schools	-0.236***
Sex Education in Public Schools	-0.198***
Same-Sex Marriage	-0.287***
Abortion	-0.240***
Marijuana Legalization	-0.128***
Capital Punishment	-0.0136
Gun Control	-0.0279
Immigration	-0.00435
Environment	-0.0890***
Alternative Energy Sources	-0.0880**
Education	-0.0560*
Scientific Research	-0.163***
Reducing Income Differences	-0.0912***
Assistance to the Poor	0.00113
Conditions of Blacks	-0.0594*
Political Outcomes: Average	-0.249***

Note: All political outcomes are coded such that an increase in the variable implies an increase in political conservatism, see Appendix A.2.2 for details. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: General Social Survey.

Table A.11 – Descriptive statistics of GSS data

	Mean	Std. Dev.	Min.	Max.
Treatment Variable:				
Evolution Score	0.63	0.33	0.00	1.00
Main Outcome:				
Evolution Belief	0.58	0.49	0.00	1.00
Placebo Outcomes - Non-Evolution Scientific Topics:				
Earth	0.88	0.33	0.00	1.00
Radioactivity	0.65	0.48	0.00	1.00
Reproduction	0.62	0.48	0.00	1.00
Lasers	0.47	0.50	0.00	1.00
Electrons	0.57	0.50	0.00	1.00
Antibiotics	0.51	0.50	0.00	1.00
Universe	0.45	0.50	0.00	1.00
Tectonics	0.85	0.36	0.00	1.00
Sun	0.79	0.41	0.00	1.00
Non-Evolution Scientific Topics: Average	0.64	0.22	0.00	1.00
Placebo Outcomes - Religious Attitudes:				
God	0.87	0.33	0.00	1.00
Bible	0.72	0.45	0.00	1.00
Afterlife	0.72	0.45	0.00	1.00
Rebirth	0.34	0.47	0.00	1.00
Strong Believer	0.32	0.47	0.00	1.00
Religious Affiliation	0.70	0.46	0.00	1.00
Church-going	0.35	0.48	0.00	1.00
Church Activities	0.17	0.38	0.00	1.00
Personal Prayer	0.65	0.48	0.00	1.00
Missionize	0.41	0.49	0.00	1.00
Spiritual Person	0.56	0.50	0.00	1.00
Religious Person	0.43	0.49	0.00	1.00
Fundamentalist	0.24	0.43	0.00	1.00
Religious Outcomes: Average	0.50	0.28	0.00	1.00
Placebo Outcomes - Political Attitudes:				
Republican	0.28	0.45	0.00	1.00
Conservative	0.26	0.44	0.00	1.00
Prayer in Public Schools	0.45	0.50	0.00	1.00
Sex Education in Public Schools	0.06	0.24	0.00	1.00
Same-Sex Marriage	0.36	0.48	0.00	1.00
Abortion	0.55	0.50	0.00	1.00
Marijuana Legalization	0.45	0.50	0.00	1.00
Capital Punishment	0.61	0.49	0.00	1.00
Gun Control	0.30	0.46	0.00	1.00
Immigration	0.85	0.36	0.00	1.00

Table A.11 (continued) – Descriptive statistics of GSS data

	Mean	Std. Dev.	Min.	Max.
Placebo Outcomes - Political Attitudes (continued):				
Environment	0.29	0.45	0.00	1.00
Alternative Energy Sources	0.41	0.49	0.00	1.00
Education	0.19	0.40	0.00	1.00
Scientific Research	0.63	0.48	0.00	1.00
Reducing Income Differences	0.47	0.50	0.00	1.00
Assistance to the Poor	0.51	0.50	0.00	1.00
Conditions of Blacks	0.64	0.48	0.00	1.00
Political Outcomes: Average	0.43	0.19	0.00	1.00
Controls:				
Female	0.57	0.50	0.00	1.00
Race/Ethnicity: White	0.70	0.46	0.00	1.00
Race/Ethnicity: Black	0.19	0.39	0.00	1.00
Race/Ethnicity: Other	0.12	0.32	0.00	1.00
Race/Ethnicity: Hispanic	0.16	0.37	0.00	1.00
Raised in Rural Area	0.49	0.50	0.00	1.00
Parents born in US	0.19	0.39	0.00	1.00
Parents born abroad	0.81	0.39	0.00	1.00
Parental Education: No Highschool	0.11	0.31	0.00	1.00
Parental Education: Highschool	0.50	0.50	0.00	1.00
Parental Education: More than Highschool	0.39	0.49	0.00	1.00
Growing up: Both Parents	0.55	0.50	0.00	1.00
Growing up: One Parent, one Stepparent	0.12	0.33	0.00	1.00
Growing up: Single Parent	0.25	0.43	0.00	1.00
Growing up: Other	0.05	0.22	0.00	1.00
Raised as Protestant: Mainline	0.37	0.48	0.00	1.00
Raised as Protestant: Evangelical	0.09	0.29	0.00	1.00
Raised as Catholic	0.32	0.47	0.00	1.00
Raised as Jew	0.01	0.10	0.00	1.00
Raised as Non-Religious	0.14	0.34	0.00	1.00
Raised as Other	0.01	0.08	0.00	1.00
Raised as Buddhist	0.00	0.06	0.00	1.00
Raised as Hindu	0.00	0.05	0.00	1.00
Raised as Other Eastern Rel.	0.00	0.03	0.00	1.00
Raised as Muslim	0.00	0.06	0.00	1.00
Raised as Orthodox-Christian	0.00	0.05	0.00	1.00
Raised as Christian	0.04	0.20	0.00	1.00
Raised as Native American	0.00	0.03	0.00	1.00
Raised as Inter-Nondenominational	0.00	0.02	0.00	1.00

Note: Descriptive statistics (mean, standard deviation, minimum, maximum) for treatment, outcome, and controls variables. Data source: General Social Survey.

Table A.12 – Descriptive statistics of ACS data

	Mean	Std. Dev.	Min.	Max.
Treatment Variable:				
Evolution Score	0.67	0.30	0.00	1.00
Main Outcomes - Working in scientific fields:				
Life Sciences	0.15	3.84	0.00	100.00
Physical Sciences	0.22	4.68	0.00	100.00
Social Sciences	0.16	4.03	0.00	100.00
Science Technicians	0.32	5.62	0.00	100.00
Additional Outcomes - Working in non-scientific fields:				
Management	5.44	22.68	0.00	100.00
Analysts	1.87	13.54	0.00	100.00
Finance	1.63	12.67	0.00	100.00
IT	1.97	13.89	0.00	100.00
Engineering	1.39	11.69	0.00	100.00
Social	1.39	11.70	0.00	100.00
Legal	0.82	9.01	0.00	100.00
Education	5.58	22.96	0.00	100.00
Arts	2.04	14.15	0.00	100.00
Health Care	4.33	20.36	0.00	100.00
Health Care Support	2.63	15.99	0.00	100.00
Protective Services	2.13	14.43	0.00	100.00
Food	7.88	26.94	0.00	100.00
Buildings	2.84	16.61	0.00	100.00
Personal Care	3.96	19.51	0.00	100.00
Sales	11.42	31.81	0.00	100.00
Office	13.25	33.90	0.00	100.00
Farming	0.66	8.07	0.00	100.00
Construction	4.45	20.63	0.00	100.00
Extraction	0.21	4.60	0.00	100.00
Installation	2.80	16.50	0.00	100.00
Production	4.87	21.53	0.00	100.00
Transportation	5.35	22.50	0.00	100.00
Armed Forces	0.75	8.62	0.00	100.00
Unemployed / Not in Labor Market	9.48	29.30	0.00	100.00
All Non-Scientific Occupations	99.15	9.16	0.00	100.00

Table A.12 (continued) – Descriptive statistics of ACS data

	Mean	Std. Dev.	Min.	Max.
Controls:				
Female	0.50	0.50	0.00	1.00
Race/Ethnicity: White	0.78	0.42	0.00	1.00
Race/Ethnicity: Black	0.12	0.33	0.00	1.00
Race/Ethnicity: Asian	0.02	0.16	0.00	1.00
Race/Ethnicity: Native	0.01	0.11	0.00	1.00
Race/Ethnicity: Other	0.03	0.18	0.00	1.00
Race/Ethnicity: Multiple	0.03	0.17	0.00	1.00
Race/Ethnicity: Hispanic	0.12	0.32	0.00	1.00

Note: Descriptive statistics (mean, standard deviation, minimum, maximum) for treatment, outcome (multiplied by 100 for interpretability), and controls variables. Data source: American Community Survey.

Table A.13 – Effect of evolution coverage in Science Standards on evolution knowledge, further robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Close Elections	Control: Governor's Party	State Specific Time Trends	Only One Reform Event	Sample Start: 1995	Sample Start: 2000	Outcome Coding: Indicator Variation	Logit	Probit	Drop Missings
Evolution Score	0.083***	0.059**	0.042	0.079**	0.042*	0.033	0.051*	0.065*	0.063*	0.063**
	(0.021)	(0.020)	(0.068)	(0.022)	(0.016)	(0.023)	(0.022)	(0.032)	(0.031)	(0.022)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.33	0.32	0.32	0.32	0.31	0.30	0.35	0.32	0.32	0.33
Std. Dev. of Dep. Var.	0.42	0.42	0.42	0.42	0.42	0.44	0.43	0.42	0.42	0.42
Adj. R-squared	0.046	0.049	0.048	0.049	0.045	0.038	0.041	0.083	0.083	0.043
Observations	5,200	15,520	15,520	7,000	14,030	11,390	14,470	15,510	15,510	13,550

Note: Robustness checks indicated in the column headers as follows: (1) Sample only includes states where the members of the State Board of Education are appointed by the governor, and where the governor in office at the time of the reform was voted into office with a margin of less than 10 percentage points compared to the runner-up; (2) Regressions control for political affiliation of governor ruling in the state and year of the student's high school entry; (3) Regressions include state-specific linear and quadratic time trends; (4) Sample only includes individuals from states that had only one reform event between 2000 and 2009, see Table A.1 for more details; (5) Sample only includes individuals who started high school after 1994; (6) Sample only includes individuals who started high school after 1999; (7) Re-coding of dependent variable: Share of questions about evolution answered correctly. Indicator variable, 1=true, 0=false, missing=omitted/not reached/off-task/etc. (dependent on the question type); (8) Re-coding of dependent variable: Individuals are dropped from the sample if the question on evolution belief contains the word "elephants" instead of "human beings" (9) Coefficient reports average marginal treatment effect of logit specification; (10) Coefficient reports average marginal treatment effect of probit specification; (11) Sample excludes individuals who have missing values on controls instead of imputing them. Dependent variable: Share of questions about evolution answered correctly. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, English language learner status, disability status, parental education, home possessions (separate indicator variables for computer and books), and test year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Table A.14 – Effect of evolution coverage in Science Standards on evolution belief in adulthood, further robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Close Elections	Control: Governor's Party	State Specific Time Trends	Only One Reform event	Sample Start: 1995	Sample Start: 2000	Outcome Coding Variation 1	Outcome Coding Variation 2	Logit	Probit	Drop Missings
Evolution Score	0.605**	0.332**	0.625**	0.394*	0.257*	0.313	0.288*	0.426**	0.347**	0.329*	0.304*
	(0.188)	(0.111)	(0.218)	(0.163)	(0.116)	(0.171)	(0.138)	(0.145)	(0.129)	(0.130)	(0.136)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.58	0.58	0.58	0.57	0.59	0.61	0.66	0.55	0.58	0.58	0.58
Std. Dev. of Dep. Var.	0.49	0.49	0.49	0.49	0.49	0.49	0.47	0.50	0.49	0.49	0.49
Adj. R-squared	0.102	0.107	0.096	0.115	0.092	0.077	0.127	0.102	0.117	0.117	0.110
Observations	589	1,801	1,801	709	1,299	654	1,571	1,617	1,780	1,780	1,751

Note: Robustness checks indicated in the column headers as follows: (1) Sample only includes states where the members of the State Board of Education are appointed by the governor, and where the governor in office at the time of the reform was voted into office with a margin of less than 10 percentage points compared to the runner-up; (2) Regressions control for political affiliation of governor ruling in the state and year of the student's high school entry; (3) Regressions include state-specific linear and quadratic time trends; (4) Sample only includes individuals who started high school after 1994; (5) Sample only includes individuals from states that had only one reform event between 2000 and 2009, see Table A.1 for more details; (6) Sample only includes individuals who started high school after 1999; (7) Sample excludes individuals whose dependent variable question on evolution replaces the words "human beings" with the word "elephants"; (8) Recoding of dependent variable: Belief in Evolution ("Human beings, as we know them today, developed from earlier species of animals - Is that true or false?", Indicator variable, 1=true, 0=false; missing=don't know); (9) Coefficient reports average marginal treatment effect of logit specification; (10) Coefficient reports average marginal treatment effect of probit specification; (11) Sample excludes individuals who have missing values on parental controls (parents born abroad, parental education) instead of imputing these values. Dependent variable: Belief in Evolution ("Human beings, as we know them today, developed from earlier species of animals - Is that true or false?", Indicator variable, 1=true, 0=false; don't know; unless noted otherwise). Controls: Indicator variables for gender, races/ethnicities, parents born abroad, parental education, having lived with parents in adolescence, raised in rural area, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism, no religion, judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: General Social Survey.

Table A.15 – Effect of evolution coverage in Science Standards on probability of working in life sciences, further robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Close Elections	Control: Governor's Party	State Specific Time Trends	Only One Reform Event	Sample Start: 1995	Sample Start: 2000	Logit	Probit
Evolution Score	0.039	0.036*	0.025	0.031	0.036**	0.029*	0.033	0.035
	(0.025)	(0.014)	(0.025)	(0.021)	(0.013)	(0.012)	(0.028)	(0.026)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FEs	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.14	0.15	0.15	0.14	0.13	0.10	0.15	0.15
Std. Dev. of Dep. Var.	3.80	3.84	3.84	3.77	3.58	3.10	3.84	3.84
Adj. R-squared	0.001	0.001	0.001	0.001	0.001	0.001		
Pseudo R-squared							0.032	0.032
Observations	2,022,927	6,460,650	6,460,650	2,522,283	4,821,487	2,762,694	6,460,650	6,460,650

Note: Robustness checks indicated in the column headers as follows: (1) Sample only includes states where the members of the State Board of Education are appointed by the governor, and where the governor in office at the time of the reform was voted into office with a margin of less than 10 percentage points compared to the runner-up; (2) Regressions control for political affiliation of governor ruling in the state and year of the student's high school entry; (3) Regressions include state-specific linear and quadratic time trends; (4) Sample only includes individuals from states that had only one reform event between 2000 and 2009, see Table A.1 for more details; (5) Sample only includes individuals who started high school after 1994; (6) Sample only includes individuals who started high school after 1999; (7) Coefficient reports average marginal treatment effect of logit specification; (8) Coefficient reports average marginal treatment effect of probit specification. Dependent variable: Probability of working in life sciences (multiplied by 100 for interpretability). Controls: Indicator variables for gender, races/ethnicities, and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: American Community Survey.

Table A.16 – Effect of evolution coverage in Science Standards on evolution knowledge and on average of non-evolution scientific knowledge: Correction for multiple hypothesis testing

	Outcomes:	
	Evolution knowledge	Non-evolution scientific knowledge (average)
	(1)	(2)
Effects	0.058	0.008
Standard p-values	0.004	0.516
Anderson-adjusted p-values	0.009	0.348
Bonferroni-adjusted p-values	0.009	1.000

Note: Table displays effects sizes and different p-values robust to multiple hypothesis testing of effects of evolution coverage in Science Standards on outcomes indicated in the column headers. P-values: Standard p-values based on clustering at the state level; sharpened False Discovery Rate (FDR) p-values as introduced by Benjamini et al. (2006), using the stata code provided by Anderson (2008); Bonferroni-adjusted p-values (standard p-value multiplied by number of tested hypothesis; capped at 1). Dependent variables multiplied by 100 for interpretability. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, English language learner status, disability status, parental education, home possessions (separate indicator variables for computer and books), as well as state, birth year, and test year fixed effects. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12.

Table A.17 – Effect of evolution coverage in Science Standards on evolution belief in adulthood and on respective averages of adulthood non-evolution scientific outcomes, religious outcomes, and political outcomes: Correction for multiple hypothesis testing

	Outcomes:			
	Evolution Belief	Non-evolution scientific outcomes (average)	Religious outcomes (average)	Political outcomes (average)
	(1)	(2)	(3)	(4)
Effects	0.333	-0.053	-0.029	-0.002
Standard p-values	0.003	0.354	0.668	0.971
Anderson-adjusted p-values	0.013	1.000	1.000	1.000
Bonferroni-adjusted p-values	0.013	1.000	1.000	1.000

Note: Table displays effects sizes and different p-values robust to multiple hypothesis testing of effects of evolution coverage in Science Standards on outcomes indicated in the column headers. P-values: Standard p-values based on clustering at the state level; sharpened False Discovery Rate (FDR) p-values as introduced by Benjamini et al. (2006), using the stata code provided by Anderson (2008); Bonferroni-adjusted p-values (standard p-value multiplied by number of tested hypothesis; capped at 1). Controls: Indicator variables for gender, races/ethnicities, parents born abroad, parental education, having lived with parents in adolescence, raised in rural area, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism, no religion, judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), as well as state, birth year, and survey year fixed effects. Data source: General Social Survey.

Table A.18 – Effect of evolution coverage in Science Standards on probability of working in life sciences, physical sciences, social sciences, as science technicians, and in non-scientific occupations: Correction for multiple hypothesis testing

	Outcomes: Probability of working in:				
	Life Sciences (1)	Physical Sciences (2)	Social Sciences (3)	Science Technicians (4)	Non-Scientific Occupations (5)
Effects	0.035	0.042	0.031	-0.027	-0.081
Standard p-values	0.016	0.025	0.277	0.603	0.032
Anderson-adjusted p-values	0.057	0.057	0.161	0.318	0.057
Bonferroni-adjusted p-values	0.078	0.127	1.000	1.000	0.159

Note: Table displays effects sizes and different p-values robust to multiple hypothesis testing of effects of evolution coverage in Science Standards on outcomes indicated in the column headers. P-values: Standard p-values based on clustering at the state level; sharpened False Discovery Rate (FDR) p-values as introduced by Benjamini et al. (2006), using the stata code provided by Anderson (2008); Bonferroni-adjusted p-values (standard p-value multiplied by number of tested hypothesis; capped at 1). Controls: Indicator variables for gender, races/ethnicities, as well as state, birth year, and survey year fixed effects. Data source: Anmerican Community Survey.

Table A.19 – Effect of evolution coverage in Science Standards on evolution knowledge in school, by evolution score indicator variables

	Evolution Knowledge								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Evolution Score > 0.90	0.028								
	(0.018)								
Evolution Score > 0.80		0.013							
		(0.014)							
Evolution Score > 0.70			0.018						
			(0.013)						
Evolution Score > 0.60				0.023*					
				(0.009)					
Evolution Score > 0.50					0.023*				
					(0.009)				
Evolution Score > 0.40						0.018			
						(0.011)			
Evolution Score > 0.30							0.025*		
							(0.012)		
Evolution Score > 0.20								0.032**	
								(0.010)	
Evolution Score > 0.10									0.032**
									(0.010)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32
Std. Dev. of Dep. Var.	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
Adj. R-squared	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049
Observations	15,520	15,520	15,520	15,520	15,520	15,520	15,520	15,520	15,520

Note: Dependent variable: Share of questions about evolution answered correctly. Explanatory variables: Evolution score indicator variables (equals one if evolution score is larger than indicated level, and zero otherwise). Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, English language learner status, disability status, parental education, home possessions (separate indicator variables for computer and books), and test year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Table A.20 – Effect of evolution coverage in Science Standards on evolution belief in adulthood, by evolution score indicator variables

	Evolution Belief								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Evolution Score > 0.90	0.117								
	(0.116)								
Evolution Score > 0.80		0.170*							
		(0.072)							
Evolution Score > 0.70			0.197*						
			(0.092)						
Evolution Score > 0.60				0.126					
				(0.069)					
Evolution Score > 0.50					0.139				
					(0.072)				
Evolution Score > 0.40						0.245***			
						(0.058)			
Evolution Score > 0.30							0.222**		
							(0.070)		
Evolution Score > 0.20								0.152	
								(0.109)	
Evolution Score > 0.10									0.073
									(0.108)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58
Std. Dev. of Dep. Var.	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49
Adj. R-squared	0.104	0.106	0.106	0.105	0.105	0.107	0.106	0.105	0.104
Observations	1,801	1,801	1,801	1,801	1,801	1,801	1,801	1,801	1,801

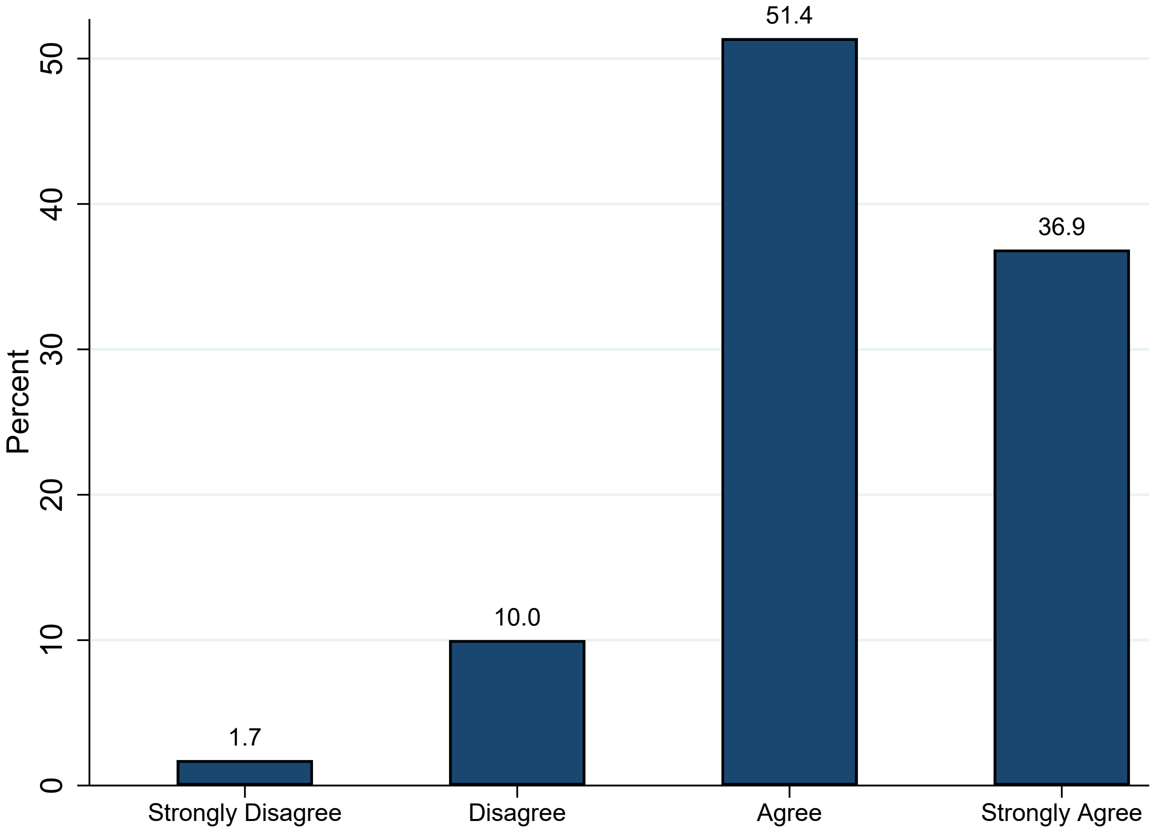
Note: Dependent variable: Belief in Evolution ("Human beings, as we know them today, developed from earlier species of animals - Is that true or false?", Indicator variable, 1=true, 0=false; don't know). Controls: Indicator variables for gender, races/ethnicities, parents born abroad, parental education, having lived with parents in adolescence, raised in rural area, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism, no religion, judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: General Social Survey.

Table A.21 – Effect of evolution coverage in Science Standards on probability of working in life sciences, by evolution score indicator variables

	Life Sciences								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Evolution Score > 0.90	0.003 (0.019)								
Evolution Score > 0.80		0.022* (0.010)							
Evolution Score > 0.70			0.023* (0.009)						
Evolution Score > 0.60				0.012 (0.007)					
Evolution Score > 0.50					0.013 (0.008)				
Evolution Score > 0.40						0.019 (0.010)			
Evolution Score > 0.30							0.020 (0.012)		
Evolution Score > 0.20								0.036** (0.012)	
Evolution Score > 0.10									0.027** (0.010)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Birth Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Std. Dev. of Dep. Var.	3.84	3.84	3.84	3.84	3.84	3.84	3.84	3.84	3.84
Adj. R-squared	0.00064	0.00064	0.00064	0.00064	0.00064	0.00064	0.00064	0.00064	0.00064
Observations	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650	6,460,650

Note: Dependent variable: Probability of working in life sciences (multiplied by 100 for interpretability). Explanatory variables: Evolution score indicator variables (equals one if evolution score is larger than indicated level, and zero otherwise). Controls: Indicator variables for gender, races/ethnicities, and survey year fixed effects. Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 5%, 1%, and 0.1% levels, respectively. Data source: American Community Survey.

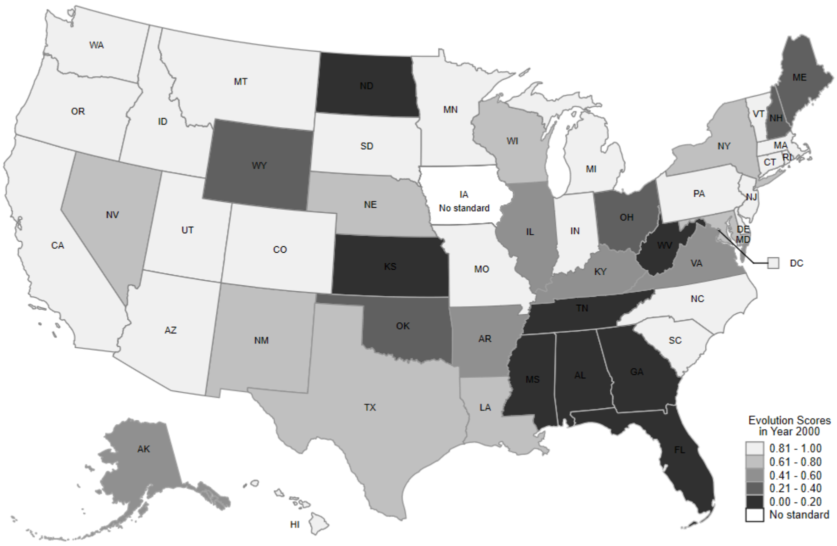
Figure A.1 – Teachers' focus on Science Standards when teaching evolution



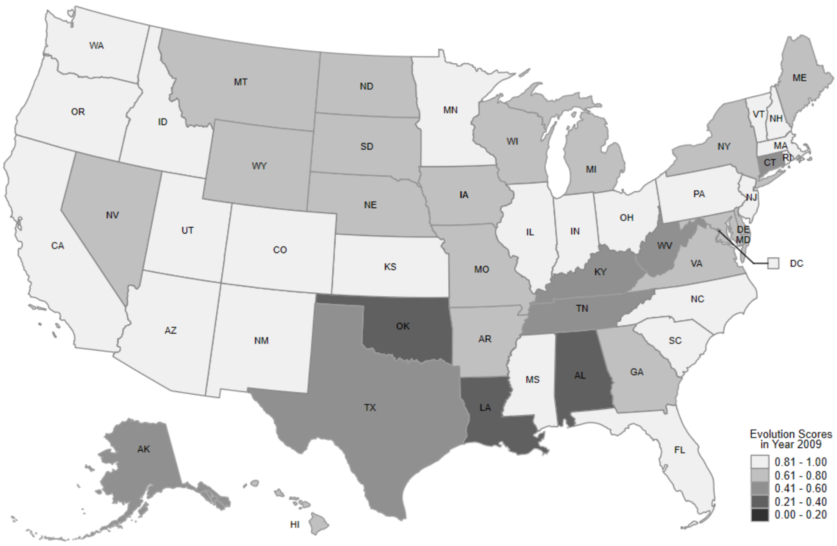
Note: Histogram depicts answer categories on agreement with the statement "When I do teach evolution, I focus heavily on what students need to know to meet state science standards". Data Source: National Survey for High School Biology Teachers, 2007.

Figure A.2 – US map of evolution scores

(a) Year 2000

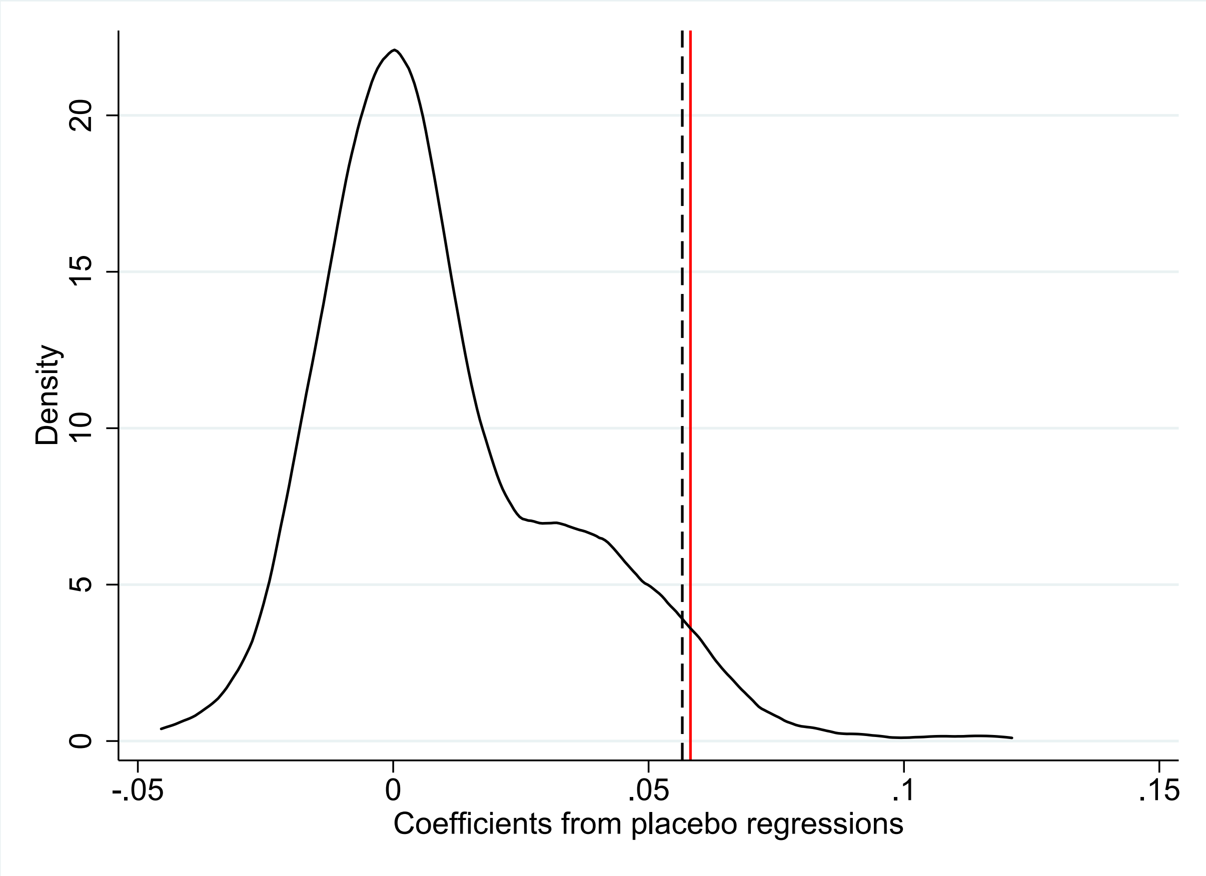


(b) Year 2009



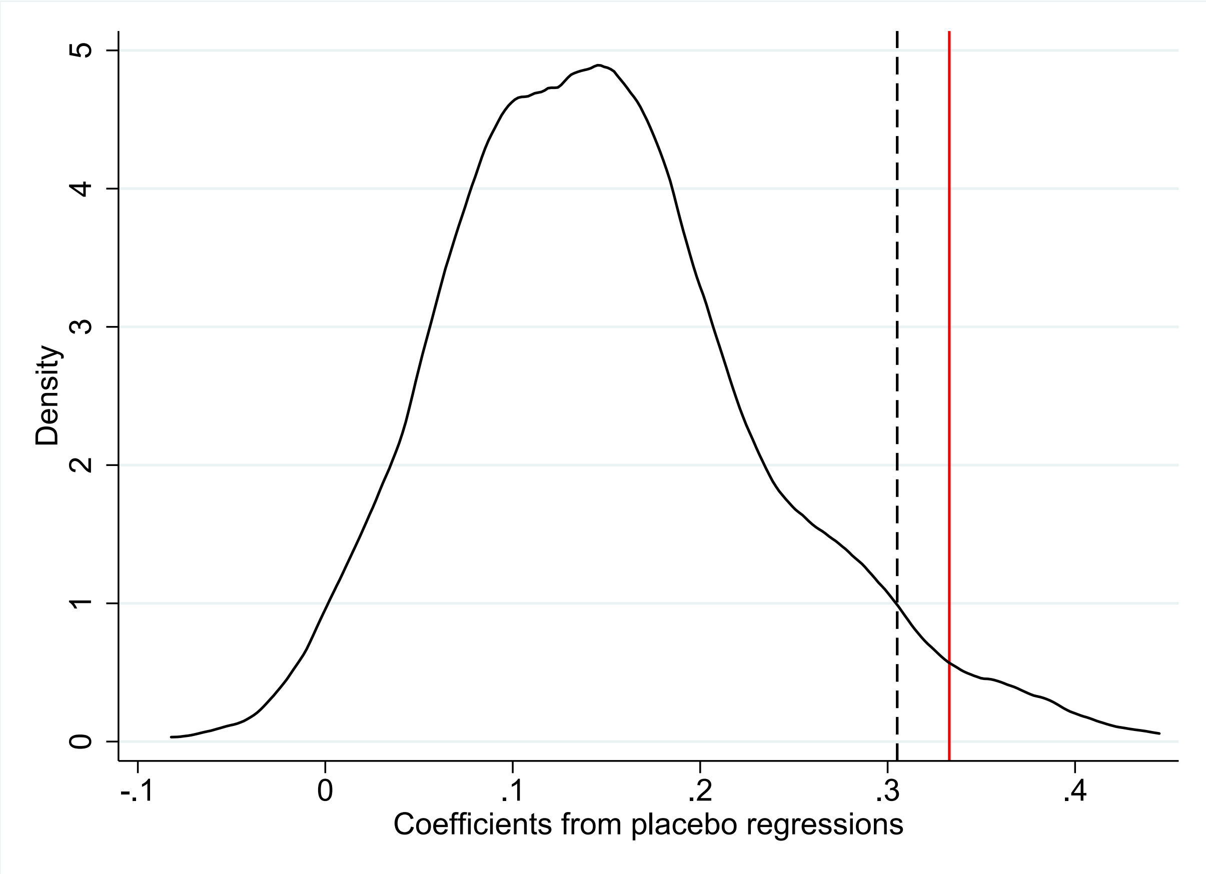
Note: Map depicts the evolution score of US States in 2000 and 2009, respectively. The evolution score measures the coverage of evolution in Science Standards, as reported in Lerner (2000b) and Mead and Mates (2009). An evolution score of 0 indicates no or a non-scientific/creationist coverage of evolution, and a score of 1 a very comprehensive coverage of evolution. A list of the evolution scores underlying this map is provided in Table A.1. Data source: Lerner (2000b) and Mead and Mates (2009)

Figure A.3 – Placebo tests: Density plot of placebo estimates of effect of evolution coverage in Science Standards on knowledge about evolution, based on 1000 permutations



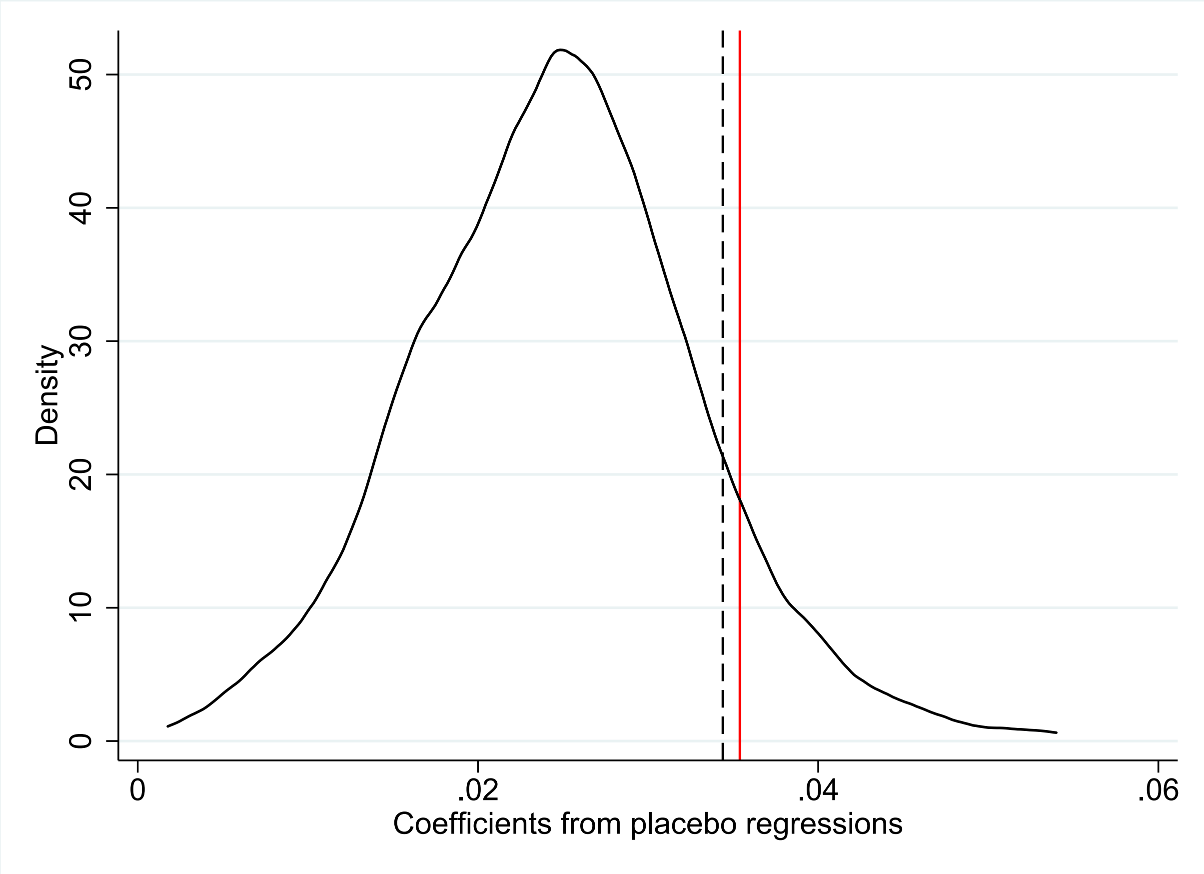
Note: Kernel density plot of coefficients of effect of evolution coverage in Science Standards on knowledge about evolution with randomly reshuffled reform years across reforming states (1000 permutations). Red solid vertical line indicates coefficient of reform effect from baseline model (0.058). Black dashed vertical line indicates 95^{th} percentile of the distribution of the 1000 placebo coefficients (0.057). Dependent variable: Share of questions about evolution answered correctly. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, English language learner status, disability status, parental education, home possessions (separate indicator variables for computer and books), and fixed effects for state, birth year, and test year. Standard errors clustered at the state level. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Figure A.4 – Placebo tests: Density plot of placebo estimates of effect of evolution coverage in Science Standards on evolution belief in adulthood, based on 1000 permutations



Note: Kernel density plot of coefficients of effect of evolution coverage in Science Standards on evolution belief with randomly reshuffled reform years across reforming states (1000 permutations). Red solid vertical line indicates coefficient of reform effect from baseline model (0.333). Black dashed vertical line indicates 95^{th} percentile of the distribution of the 1000 placebo coefficients (0.305). Dependent variable: Belief in Evolution ("Human beings, as we know them today, developed from earlier species of animals - Is that true or false?", Indicator variable, 1=true, 0=false; don't know). Controls: Indicator variables for gender, races/ethnicities, parents born abroad, parental education, having lived with parents in adolescence, raised in rural area, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism, no religion, judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), and fixed effects for state, birth year, and test year. Standard errors clustered at the state level. Data source: General Social Survey

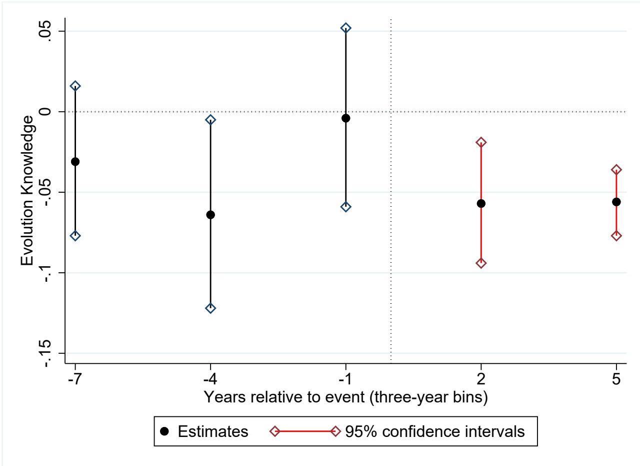
Figure A.5 – Placebo tests: Density plot of placebo estimates of effect of evolution coverage in Science Standards on probability of working in life sciences, based on 1000 permutations



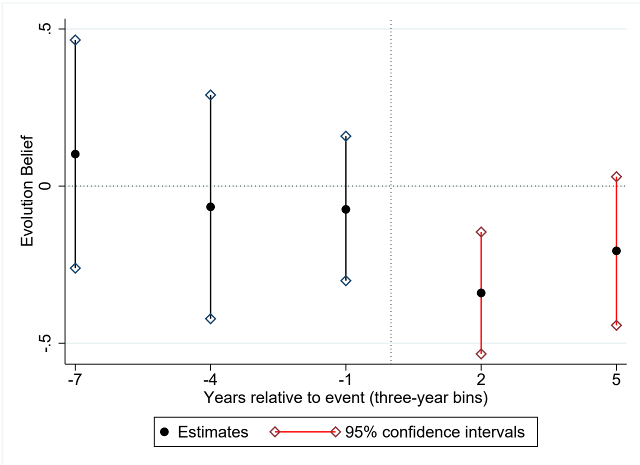
Note: Kernel density plot of coefficients of effect of evolution coverage in Science Standards on probability of working in life sciences with randomly reshuffled reform years across reforming states (1000 permutations). Red solid vertical line indicates coefficient of reform effect from baseline model (0.035). Black dashed vertical line indicates 90^{th} percentile of the distribution of the 1000 placebo coefficients (0.034). Dependent variable: Probability of working in life sciences (multiplied by 100 for interpretability). Controls: Indicator variables for gender, races/ethnicities and fixed effects for state, birth year, and test year. Standard errors clustered at the state level in parenthesis. Data source: American Community Survey.

Figure A.6 – CS event-study graphs: Only states reducing evolution coverage

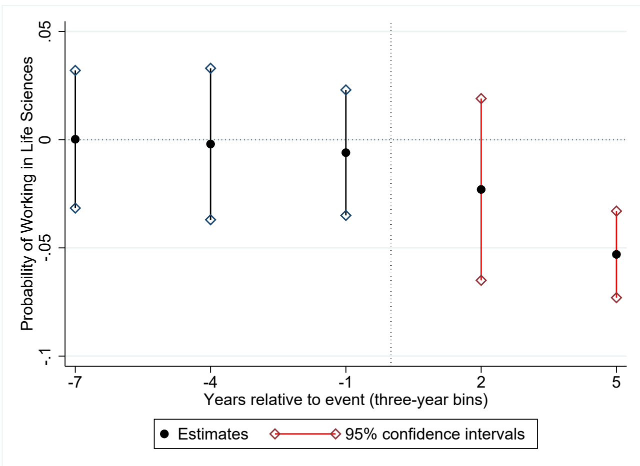
(a) Evolution knowledge in school



(b) Evolution belief in adulthood



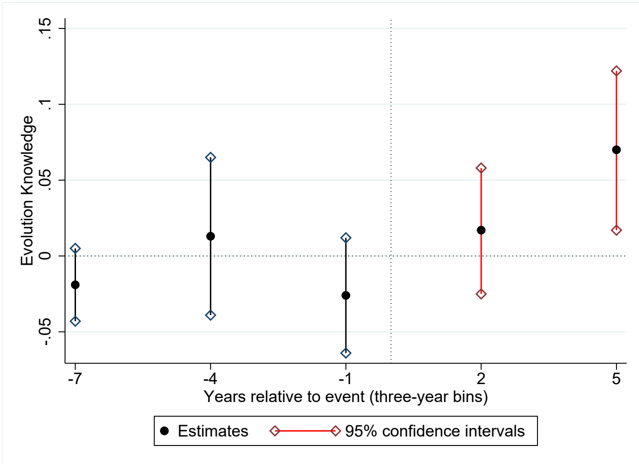
(c) Working in life sciences



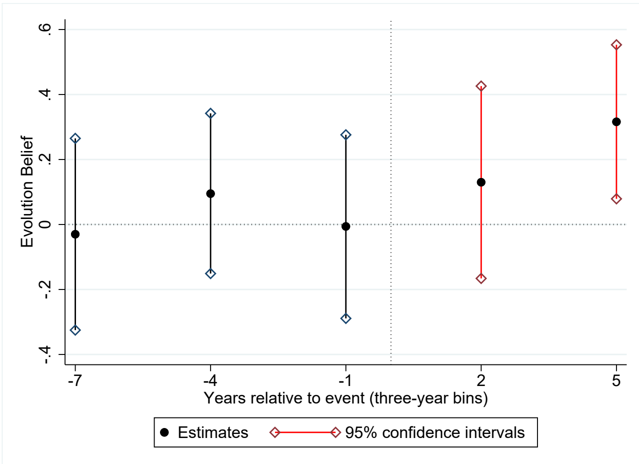
Note: Coefficients and 95% confidence intervals from CS estimator (Callaway and Sant'Anna, 2021), accounting for heterogeneous treatment effects and staggered treatment timing. Dynamic aggregation/event study effects, using doubly robust inverse probability weighting. Controls: Not-yet-treated and never-treated observations. Numbers on horizontal axis refer to final year of respective three-year bins; i.e., -1 = last three years prior to treatment. Inference: Clustering at state level. Data sources: (a) U.S. Department of Education, National Center for Education Statistics, 1996-2009 National Assessment of Educational Progress; (b) General Social Survey; (c) American Community Survey.

Figure A.7 – CS event-study graphs: Only states expanding evolution coverage

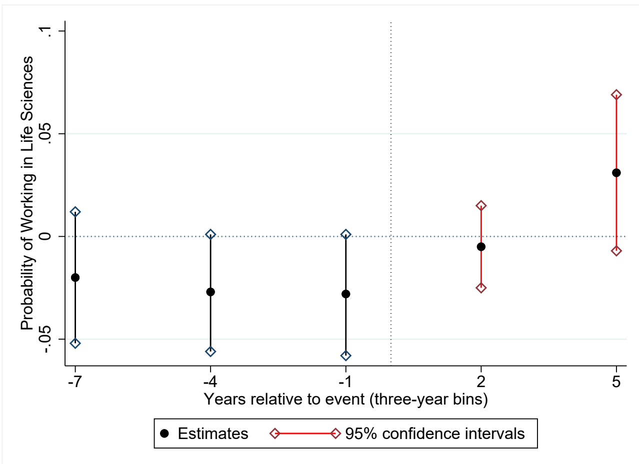
(a) Evolution knowledge in school



(b) Evolution belief in adulthood



(c) Working in life sciences



Note: Coefficients and 95% confidence intervals from CS estimator (Callaway and Sant'Anna, 2021), accounting for heterogeneous treatment effects and staggered treatment timing. Dynamic aggregation/event study effects, using doubly robust inverse probability weighting. Controls: Not-yet-treated and never-treated observations. Numbers on horizontal axis refer to final year of respective three-year bins; i.e., -1 = last three years prior to treatment. Inference: Clustering at state level. Data sources: (a) U.S. Department of Education, National Center for Education Statistics, 1996-2009 National Assessment of Educational Progress; (b) General Social Survey; (c) American Community Survey.

Figure A.8 – Two NAEP sample questions on evolution knowledge

(a) Sample Question 1

Which of the following is NOT a part of Darwin's theory of evolution by natural selection?

A. Individuals in a population vary in many ways.

B. Some individuals possess features that enable them to survive better than individuals lacking those features.

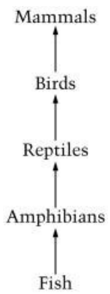
C. More offspring are produced than can generally survive.

D. Changes in an individual's genetic material are usually harmful.

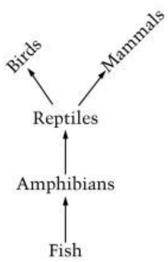
(b) Sample Question 2

According to evolutionary theory, which of the following evolutionary trees best describes the relationship between groups of vertebrates?

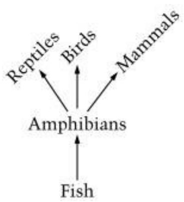
A.



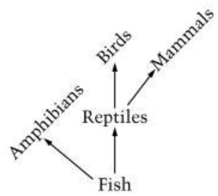
B.



C.

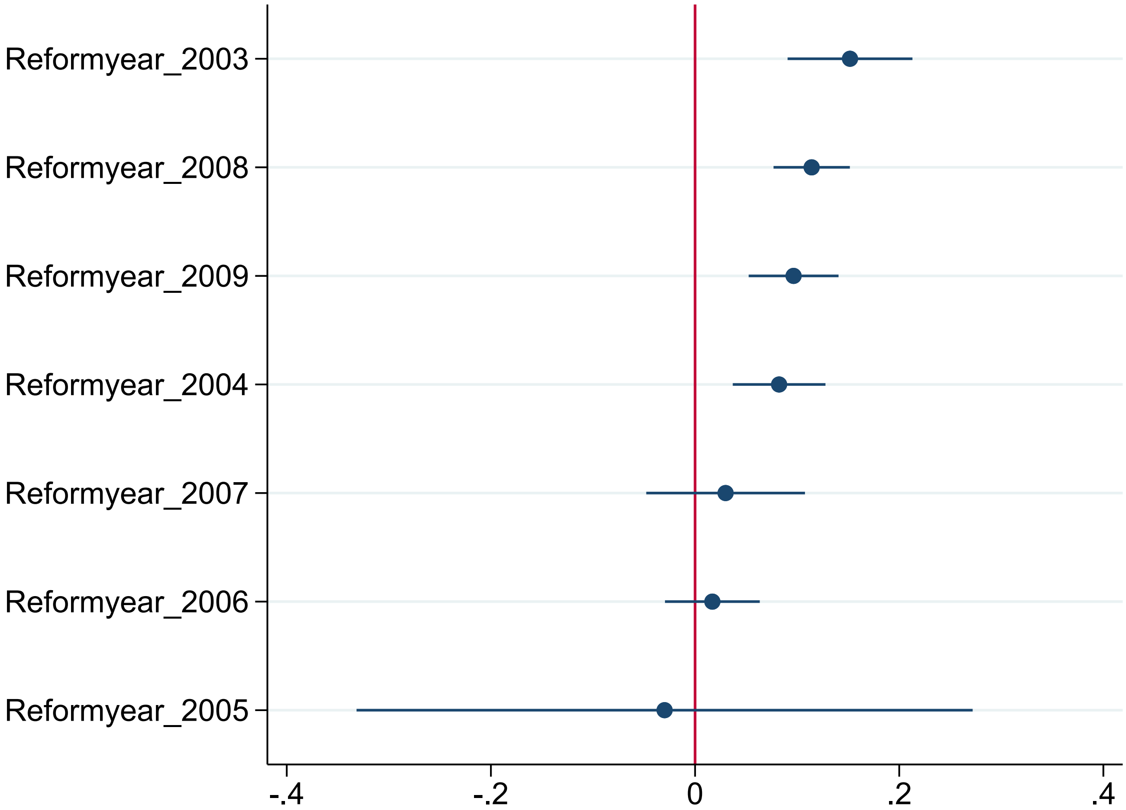


D.



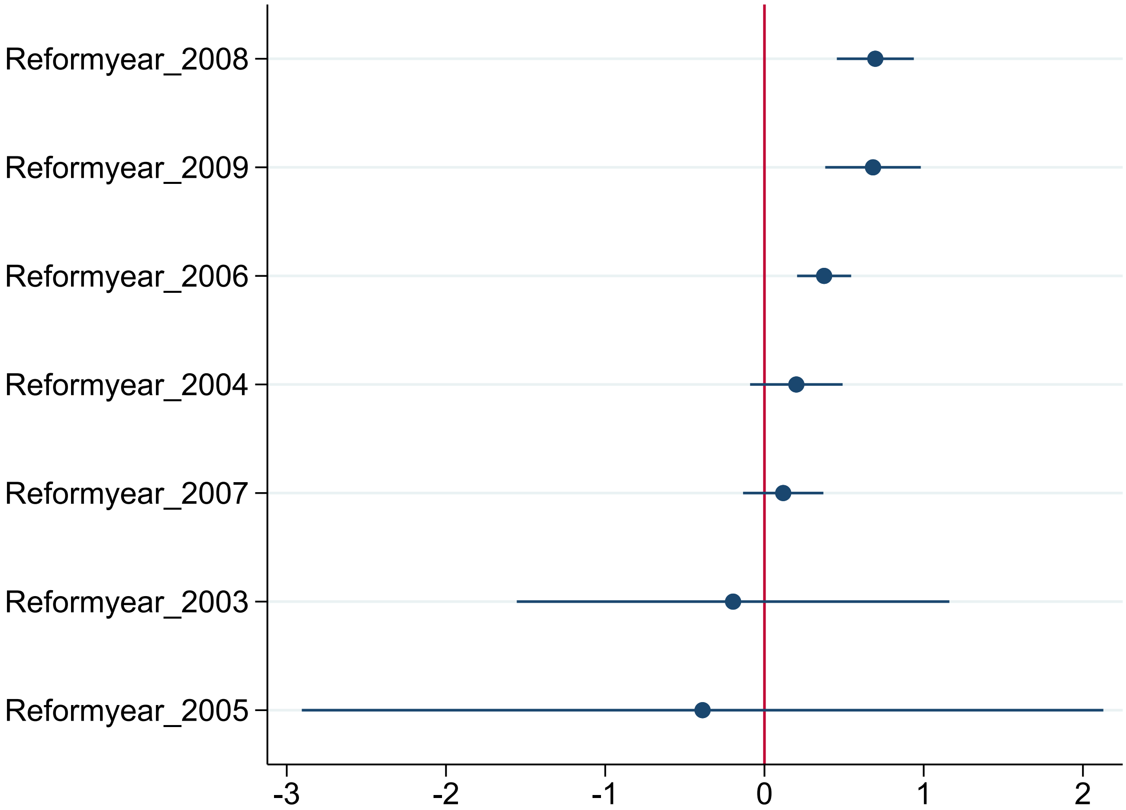
Note: Sample question on evolution knowledge from NAEP Science Test, Grade 12, Year 2000. Question also accessible online at NAEP question tool. Question 1: Answer D is correct. Question 2: Answer B is correct. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2000 Science Assessment for Grade 12

Figure A.9 – Reform-year-specific estimates of effect of evolution coverage in Science Standards on knowledge about evolution



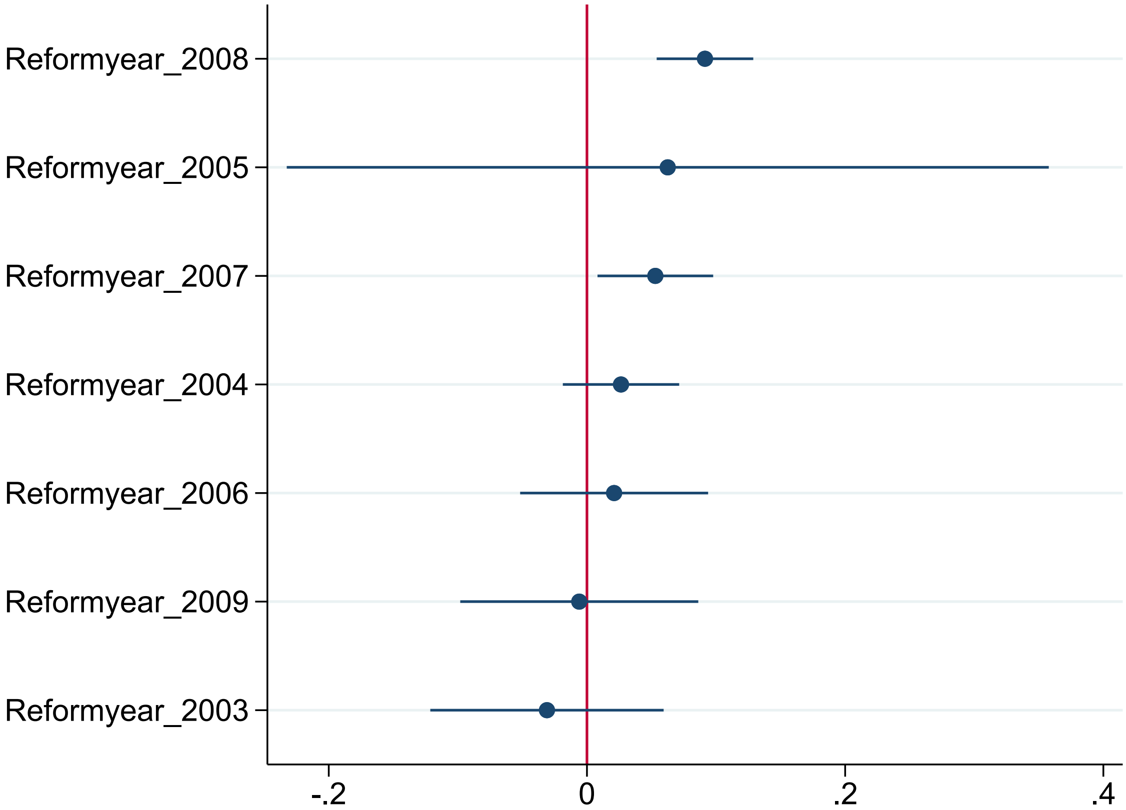
Note: Figure displays reform-year-specific estimates of effect of evolution coverage in Science Standards on share of questions about evolution answered correctly. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, English language learner status, disability status, parental education, home possessions (separate indicator variables for computer and books), and fixed effects for state, birth year, and test year. Standard errors clustered at the state level. 95% confidence intervals displayed. Respective reform-year samples contain individuals from all non-reforming states and from all states with a reform of the evolution coverage in the respective reform year. Years 2000-2002 are dropped due to too few reforms. Data source: U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1996-2009 Science Assessments for Grade 12

Figure A.10 – Reform-year-specific estimates of effect of evolution coverage in Science Standards on evolution belief in adulthood



Note: Figure displays reform-year-specific estimates of effect of evolution coverage in Science Standards on evolution belief in adulthood ("Human beings, as we know them today, developed from earlier species of animals - Is that true or false?", Indicator variable, 1=true, 0=false; don't know). Controls: Indicator variables for gender, races/ethnicities, parents born abroad, parental education, having lived with parents in adolescence, raised in rural area, religion raised in (indicator variables for mainline protestantism, evangelical protestantism, catholicism, no religion, judaism, buddhism, hinduism, other eastern, islam, orthodox-christian, christian, native american, inter-nondenominational, other religion), and fixed effects for state, birth year, and test year. Standard errors clustered at the state level. 95% confidence intervals displayed. Respective reform-year samples contain individuals from all non-reforming states and from all states with a reform of the evolution coverage in the respective reform year. Years 2000-2002 are dropped due to too few reforms. Data source: General Social Survey

Figure A.11 – Reform-year-specific estimates of effect of evolution coverage in Science Standards on probability of working in life sciences



Note: Figure displays reform-year-specific estimates of effect of evolution coverage in Science Standards on probability of working in life sciences (multiplied by 100 for interpretability). Controls: Indicator variables for gender, races/ethnicities and fixed effects for state, birth year, and test year. Standard errors clustered at the state level. 95% confidence intervals displayed. Respective reform-year samples contain individuals from all non-reforming states and from all states with a reform of the evolution coverage in the respective reform year. Years 2000-2002 are dropped due to too few reforms. Data source: American Community Survey