EVOLUTION VS. CREATIONISM IN THE CLASSROOM: THE LASTING EFFECTS OF SCIENCE EDUCATION*

BENJAMIN W. AROLD

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 student cohorts in models with state and cohort fixed effects. 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*: Z12, I28, J24, P16

*Arold: Center for Law & Economics, ETH Zurich, Haldeneggsteig 4, 8092 Zurich, Switzerland (aroldb@ethz.ch; +41(0)44/26919). I am very grateful to Elliott Ash, Ludger Woessmann, and Davide Cantoni for their support throughout this project. I also wish to thank the editors Robert Barro and Lawrence Katz, four anonymous referees, and Karun Adusumilli, Luca Braghieri, Sarah Eichmeyer, Eleonora Guarnieri, Eric Hanushek, Martina Magli, Rachel McCleary, Markus Nagler, Nathan Nunn, Paul Peterson, Matteo Sandi, M. Danish Shakeel, Claudia Steinwender, Joseph Stiglitz, Gregory Veramendi, Martin West, Brian Wheaton, Scott Williamson, 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 University of Cambridge, UC Riverside, ETH Zurich, UZH Zurich, LMU Munich, 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, and the AEFP Conference 2020. Clementine Abed Meraim, Shufan Ma, 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. Word count: 11,963.

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, Bennett, and Ward, 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 science, and its controversy in the population and the education system. Evolution can

[1]The rejection of evolution theory by Trofim Lysenko, the then-president of the Academy of Agricultural Sciences of the USSR and leading agricultural advisor to Joseph Stalin, has 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, Sapienza, and Zingales, 2008; Tabellini, 2008). Other determinants include peers and social networks (Sacerdote, 2001; Bailey et al., 2020), the media (Martin et al., 2017), and political systems (Alesina and Fuchs-Schündeln, 2007).

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, 15 states reduced the coverage of evolution in their education standards, while 22 states expanded it. I argue that the political and institutional processes leading to these reforms, in particular the predetermined timing of gubernatorial elections in combination with 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 Science Standards. Specifically, the performed two-way fixed effects estimations can rule out as confounding factors (i) time-invariant state-specific differences (such as education

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

levels), (ii) cohort-specific national 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 or science prestige), in a robustness test that includes state-specific time trends. In addition, I account for potential biases in staggered two-way fixed effects designs from time-varying treatment effects (Callaway and Sant'Anna, 2021). 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 of 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 innovation, earnings, and economic growth in the long run (Lucas, 1988; Barro, 2001; Hanushek and Woessmann, 2008, 2012).[4]

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. There are no effects on religiosity and political attitudes. 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). Effects are largest for Mainline Protestants. This analysis underscores that reform effects persist long after students have left high school. This result

[4] At the correlational level, Biasi and Ma (2022) show a direct link between higher education curricula and innovation.

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. It seems plausible that if people know more about and believe more in a given theory, they are also more interested in working in the field that was founded on it. After all, skills and interest are important determinants of occupational choice (Speer, 2017). Specifically, 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 subfield 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, Neilson, and Zimmerman, 2013; Kirkeboen, Leuven, and Mogstad, 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, Shih, and Sparber, 2015). In all three analyses presented in this paper, I consistently find that effects are larger in absolute terms for the subgroup of reforms that reduce the evolution coverage compared to the subgroup that expands it.

This study provides evidence on how learning about evolution theory – the fundamental theory of the evolution of life that explains the existence of humans – actually shapes humans. Specifically, I show that evolution teaching does not only affect the related knowledge of students, but also translates into evolution belief in adulthood. This implies that scientific beliefs can be "taught" in school, and persist into adulthood. This finding complements seminal research on the effects of non-science-related school curricula reforms on financial decision-making (Bernheim, Garrett, and Maki, 2001), identity (Clots-Figueras and Masella, 2013), labor market participation and employment (Fuchs-Schündeln and Masella, 2016; Costa-Font, García-Hombrados, and Nicińska, 2024), political and economic preferences (Cantoni and Yuchtman, 2013; Cantoni et al., 2017), civic values (Bandiera et al., 2019), and religiosity (Bazzi, Hilmy, and Marx, 2020).[5]

This paper further demonstrates that attitudinal changes induced by science school curricula reforms translate into high-stakes choices of individuals. Specifically, the finding on occupational choice enhances our understanding of how to increase the share of STEM graduates, which is a policy goal with widespread support in many societies.[6] 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., 2016), and heterogeneous tastes (Wiswall and Zafar, 2015). Supply side factors such as grading policies (Butcher, McEwan, and Weerapana, 2014), admissions systems (Bordon and Fu, 2015), affirmative action policies (Arcidiacono, Aucejo, and Hotz, 2016), and the provision of role models (Porter and Serra, 2020) can also play a role; see Altonji, Arcidiacono, and Maurel (2016) for an overview. 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.[7] This is true

[5]Scientific attitudes are arguably more challenging to change, as they pertain to the acceptance of knowledge rather than opinions. Note that the only scientific attitude addressed in the previous papers, environmental attitudes in Cantoni et al. (2017), did not change due to the curriculum reform (in contrast to political and economic attitudes).

[6]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).

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

despite the fact that being raised as Evangelical is a large negative predictor of evolution belief. While a number of studies have found a positive relationship between education and religiosity (McCleary and Barro, 2006a,b; Glaeser and Sacerdote, 2008; Meyersson, 2014), other research suggests that education can decrease religiosity (Hungerman, 2014; Becker, Nagler, and Woessmann, 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), 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. What is more, the effects of the content of education standards last until adulthood. In sum, this paper demonstrates that high school curricula exert a lifetime influence on students.

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

2 INSTITUTIONAL BACKGROUND

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. Before the First World War, evolution teaching was rare (Beale, 1941). In the 1920s, more than 20 states even considered bills to ban the teaching of evolution. Among other states, such a bill became law in Tennessee, resulting in the famous "Monkey trial" where a biology teacher was convicted for having taught evolution.[8] This law from Tennessee was

[8] This decision was later overturned on a technicality, see Larson (1999).

overruled in 1967, and similar decisions followed in other states in the following years. In 1987, another law requiring that equal time must be spent on teaching evolution and creation was ruled unconstitutional by the US Supreme Court. In short, the legislative and adjudicative decisions of the second half of 20^{th} century have paved the way for evolution to be taught in public schools (Moore, Jensen, and Hatch, 2003b). Still, there continues to be substantial variation across states and years in the 21^{st} century, as the subsequent analysis of the evolution coverage in Science Standards demonstrates.

US State Science Education Standards: In general, Science Standards define the scientific knowledge and skills that students are supposed to master in a given grade in public schools. The scientific teaching which a student is ultimately exposed to also depends on local school curricula, the selection of textbooks (Adukia et al., 2023), the knowledge, ability and ideology of teachers, testing formats, and other factors. However, Science Standards form the basis of many of these factors. 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. Moreover, state-wide standardized exams often directly test the content set out in the Science Standards. Lerner (2000b, p.ix) summarizes that Science Standards "[...] 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 Online Appendix Figure A.I.

Science Standards cover many topics. However, the reforms evaluated in this paper primarily concentrate on evolution. In Online Appendix A.1, I present evidence from a text analysis of the Science Standards, demonstrating that while the reforms modify the coverage of many topics, the treatment of evolution is altered to a significantly greater extent.

The Adoption Process of Reforms of Science Standards: How do reforms of Science Standards come into existence? In each state, they are decided by majority vote of the members of the State Boards of Education. The selection process of these members differs across states. In some states, members are appointed by the governor, sometimes with the consent of the senate (for example in California and Florida). In other states, members are elected by the public (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). Before the final vote of the members of the Board of Education, Science Standards are typically drafted by advisory committees. These consist of a panel of teachers and other stakeholders such as parents, scientists, and religious representatives. Moreover, the Boards of Education hold public hearings.

This process implies that these reforms happening at some point is not random. Instead, it reflects changing views, either expressed by the election of a governor, or by direct election of members of the Boards of Education.

However, the exact reform year in a given state can be regarded as-good-as random due to institutional idiosyncrasies. If beliefs in evolution 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. 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 in a staggered manner across districts and years. Further 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 can take months or years.[9] In sum, there may be a great number of years between a shock and a reform of the 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. I identify from this arguably exogenous reform timing.

[9]In some cases, there are also idiosyncrasies induced by 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. Building on this point, I show that reform effects also hold in a subsample of large states only, see column (6) in Table II.

Online Appendix A.2 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.

Online Appendix A.3 provides quantitative evidence on the exogenous timing of the reforms. I regress state-by-year characteristics such as (i) unemployment rate, partisan composition, and school resources, and (ii) Google search frequencies of keywords specific to evolution and creationism, on the evolution coverage in the Science Standards, in models with state and year fixed effects. All estimates are insignificant and close to zero. This suggests that the timing of the reforms is independent of (i) economic, political, and educational conditions, and (ii) the interest of the population in evolution and creationism.

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.[10] Newspaper articles and policy reports suggest that the content of textbooks, lesson plans, and standardized testing questions was changed due to the reforms, while there is no indication of new teachers being hired.[11] 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. 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. These dates are likely the upper end of the implementation timeline, as teachers in reality have to implement

[10]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).

[11]The likelihood of discussing evolution vs creationism in the household could be affected as well, and would be a mechanism of the main results to the extent that it is specific to the treated cohorts.

many changes earlier to allow for a smooth transition of classroom activities before the deadline. Still, some individuals coded as exposed to the post-reform treatment may have been exposed to some pre-reform treatment. This dilutes the treatment-control contrast and implies that any effects should be interpreted as lower bound estimates.

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 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.[12]

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

[12]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). In this paper, I change this evolution score from -0.18 to 0 for ease of interpretability. 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 scoring scheme is provided in Lerner (2000b, pp.10-17). I also estimate the effects of the different components of the evolution score separately on the relevant outcomes, but find that no component stands out in particular (results available upon request).

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.[13] The evolution score serves as a treatment variable 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 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.[14]

[Figure I about here.]

To illustrate the identifying variation, Figure I depicts the state-level evolution score *difference* between 2000 and 2009.[15] The evolution score decreased in 15 states (implying a negative evolution score difference) and increased in 22 states (implying a positive evolution score difference) between 2000 and 2009. In the remaining 13 states, it remained unchanged. The states with the largest evolution score decreases are Connecticut, Louisiana, and Texas. The largest evolution score increases are found in Kansas, Mississippi, and Florida. 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,

[13]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 articles, legal documents, and state education websites indicates that they only had one reform between 2000 and 2009, see column (5) in Table II.

[14]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 cohorts. This is as the standard high school curriculum typically features biology (the subject in which evolution is being taught) in the first year of high school. To account for the possibility that evolution could also be taught in other years, I also run a dosage specification. Here, the main evolution score is replaced by the weighted average of the pre- and post-reform evolution scores, with the weights corresponding to the number of pre- and post-reform high school years, see column (4) in Table II.

[15]Online Appendix Figure A.II also depicts the evolution score levels in 2000 and 2009.

with each census region having at least one state in which the evolution coverage became less comprehensive, more comprehensive, and remained unchanged, respectively.

3.2 Micro Data

The following subsection describes the three micro-level datasets used in this paper. Each of these repeated cross-sectional datasets is standardized and hence comparable across US states and cohorts, making it suitable for analyses with state and cohort fixed effects. In all three datasets, I keep students in the sample who start high school after 1990 and before 2010. 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).[16] This approach also prevents identification from the adoption of the Next Generation Science Standards which started in 2013.

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 include topics such as "reproduction", "climate change" or the "universe" and are defined analogously. In addition, the NAEP student surveys provide rich student-level control variables, such as subsidized lunch status or home possessions to approximate the socio-economic status.

[16] The results of this paper do not depend on this specific sample cut, see columns (8) and (9) in Table II.

The main sample consists of more than 14,000 students who were asked at least one question on evolution. It contains only public school students, as Science Standards have never been binding for private schools. The average evolution score equals 0.65, implying that the sampled students were on average exposed to a "satisfactory" evolution coverage.[17] 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. Online Appendix A.4.1 provides descriptive statistics, correlations, sample questions, and further information.

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?".[18] 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 other scientific topics, and on religious and political attitudes. Variables capturing dimensions of the childhood environment serve as controls, including the religion a respondent was raised in.

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

[18] The words "human beings" are replaced by the word "elephants" for 10 percent of the questions on evolution belief. Column (11), Panel B, of Table II shows that the results are robust to dropping those from the sample.

The GSS is sampled from the entire US adult population irrespective of type of school attendance. This makes it impossible 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. 58 percent of this sample believe in evolution which largely corresponds to other surveys at the time (Pew Research Center, 2009). Further descriptive statistics, raw correlations and data background is provided in Online Appendix A.4.2.[19]

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. It can be divided into the subfields "biology", "agriculture and food", "conservation and forestry" and "medical and other".

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 Online Appendix A.4.3.

[19]Building on Barro (2022), I show for example that state-level averages of evolution belief and the evolution score correlate positively with COVID vaccination rates and negatively with Trump voting, see Online Appendix Tables A.I and A.II.

4 IDENTIFICATION STRATEGY

The analyses presented in this paper are based on the following two-way fixed effects (TWFE) model. The 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 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. Hence, the treatment status is based on the evolution score in force in the state and year in which an individual enters high school. 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 creationist coverage of evolution ($Evolution\_Score_{st}$=0). The vector $\mathbf{X_i}$ contains 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.[20] Standard errors are clustered at the state level to account for the potential correlation of error terms across cohorts within states (Abadie et al., 2022; Athey and Imbens, 2022).

The key identifying assumptions are that, in the absence of the evolution coverage

[20] In the analyses using NAEP, cohort and year fixed effects generally coincide as each cohort was examined in grade 12.

reforms, the outcomes of students attending high school in different states would have evolved along parallel trends, and that treatment effects are homogeneous over time.

The TWFE model allows to rule out various concerns on the ability to estimate causal effects of the evolution coverage in Science Standards. First, 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.

Moreover, the state fixed effects rule out time-varying state-specific shocks as long as they affect adjacent cohorts equally. This is as the empirical setup exploits cross-cohort variation within a narrow time window around the evolution curriculum reforms, and identifies from reforms that affect adjacent cohorts in different ways. Specifically, a reform adopted in a given state and calendar year primarily affects the new high school entry cohort (and younger cohorts in the following years when they start high school), see Section 3.1.[21] I argue that many potential shocks that one might typically be concerned about (such as state-specific church scandals or shocks that affect the prestige of science) do not discontinuously affect different high school cohorts.[22] Furthermore, a robustness specification with state-specific linear and quadratic time trends accounts for time-varying state-specific shocks that affect adjacent cohorts differentially, but smoothly. This specification is particularly demanding in terms of statistical power, as reform effects are only detectable as "jumps" from the cross-cohort trend. For example,

[21] 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 my results as lower-bound estimates because parts of the cohorts coded as exposed to pre-reform Science Standards would be partially treated by post-reform Science Standards in this scenario.

[22] See also Cantoni et al. (2017, p.363), who also exploit a cohort-specific introduction of a curriculum reform in models with state and cohort fixed effect and note: “This method of introducing the new curriculum considerably reduces concerns about omitted variables, as many time-varying, province-specific shocks seem unlikely to have very different effects across adjacent cohorts of students”.

this specification accounts for changes in trust in science that could develop differently in the various states, and change smoothly across cohorts.

As the evolution score treatment variable is of continuous nature, the parallel trends assumption has to hold in its strong form: The average potential outcomes for individuals that are exposed to a reform of the evolution coverage have to be the same at each level of evolution coverage. This excludes selection into a particular level of dosage (evolution coverage) (Callaway, Goodman-Bacon, and Sant'Anna, 2021). Following Cook et al. (2022), I probe the plausibility of this assumption by testing for correlations between the evolution score and covariates. Specifically, I regress the evolution score on pre-determined observables and find little evidence of systematic correlations.[23] I also show in robustness analyses that reform effects hold when transforming the continuous evolution score variable into binary variables. Moreover, I note (i) that active selection into exposure to a different State Science Standards requires moving across states, which seems unnecessarily costly in most cases as students can simply switch to a private school (or do homeschooling),[24] and (ii) that the institutional idiosyncrasies determining the exact reform timing, as discussed in Section 2, make it difficult to pro-actively select into certain evolution coverages. Last, I probe robustness 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 most diverse potential outcomes and the strongest incentive to move, are excluded from the sample.

To further probe the plausibility of the parallel trends assumption, I estimate an event study version of equation (1), in which the reform of the evolution coverage in Science

[23]Each row of Online Appendix Figure A.III displays the estimate from a separate regression of the evolution score on the covariate defined along the vertical axis and survey year fixed effects. The estimates show little correlation overall. The one notable exception is Mainline Protestants, which, however, has the opposite sign as compared to other religious groups such as Evangelicals and Catholics.

[24]The main regression estimates using the GSS and ACS are based on a sample of students from public school, private schools, and homeschoolers that are net of spurious selection across school types. For NAEP, the results hold on a sample of public and private schools, see column (4) of Online Appendix Table A.III, while there is no NAEP test for homeschoolers.

Standards in a given state and year is referred to as the "event". This approach now views every reform as a discrete event and ignores differences in the intensity of the reform as indicated by the evolution score used in the baseline TWFE model, as follows:

$$Y_{istu} = \sum_{k=-7}^{5} 1(t_{is} = t_s + k)\beta_k + \gamma * X_{istu} + \delta_s + \lambda_t + \theta_u + \epsilon_{istu} \tag{2}$$

I estimate the effect of exposure to a reform of the evolution coverage in Science Standards in year t_s on outcomes of students starting high school k years before and after the evolution coverage reform, as captured by the parameter vector β_k. All event-study regressions are estimated by grouping two years together to one bin to smooth the number of observations across bins as not all micro data are collected in every year (see Section 3).[25] This model allows to examine non-linear pre-reform trends in outcome variables, and disentangle effects which directly occur at the time of the reforms from those that phase-in gradually after the reform. Note that the event-study estimations yield changes in outcomes induced by the average reform. This requires running the event-study models separately for the sets of states that reduce and expand the evolution coverage, respectively, because joint event-study models would cancel out effects from opposing reforms. Within each set of states, the regression coefficients can be interpreted as changes in outcomes induced by the average reform of the evolution coverage in that set.

Even in the absence of confounding trends or shocks, consistent estimation of reform effects requires homogeneity in treatment effects. 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, time-varying treatment effects can bias results away from the true effect if already-treated students act as controls for later-treated students, which also applies to

[25] Longer post-reform time horizons are not available in the micro data. Effects are estimated relative to the bin directly before the reform, i.e. the bin that covers the years [-2,-1]. The bins at the beginning (end) of the domain additionally include the years prior to (following) the domain's starting (ending) year. In event studies with GSS data, the individual-level covariates are dropped as the number of observations is particularly small relative to the number of covariates and estimated coefficients.

the event studies (Sun and Abraham, 2021).

To test whether my TWFE OLS event study regressions are immune to this bias, I run the CS estimator (Callaway and Sant'Anna, 2021), which excludes those 2x2 difference-in-differences comparisons in which already-treated students act as controls from the sample. As I run the event-study models separately for the sets of states that reduce and expand the evolution coverage without never treated states, the CS estimator uses not-yet-treated students as controls. To alleviate concerns about the multi-collinearity of dynamic treatment effects and time fixed effects (Borusyak, Jaravel, and Spiess, 2023), I bin both endpoints which allows for separate identification even when using not-yet-treated units only (Schmidheiny and Siegloch, 2023). In robustness analyses I also add a set of never-treated states to the sample.[26]

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.

5.1 Evolution Knowledge in School

Baseline Estimates: To assess reform effects on students' knowledge about evolution, 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, following equation 1.

[Table I about here.]

As shown in Column (1) of Panel A, Table I, the raw correlation is positive. This could imply that a comprehensive coverage of evolution increases students' evolution knowledge (reform effect), that evolution knowledge raises the probability of states adopting Science

[26] I add the four never-treated states Nebraska, Oregon, Utah, Wisconsin to the sample of states that reduce the evolution coverage. These four states are selected to be largely representative for the US, and to be large enough to allow for separate identification of dynamic treatment effects and time fixed effects.

Standards that cover evolution comprehensively (for example because students might not accept creationist content; reverse causality), or that third variables such as socio-economic status affect both (omitted variable bias). To isolate reform effects, I add control variables in columns (2)-(4). The full model in column (4) is the preferred specification since it exploits the idiosyncratic timing of reforms of Science Standards as a source of arguably exogenous variation. 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 creationist coverage of evolution, increases the share of questions on evolution answered correctly by 6.5 percentage points (p-value = 0.005). Given that, on average, students answer 32 percent of the questions on evolution correctly, the reported effect amounts to 20 percent of the sample mean.

Event Study Graphs: To test for parallel trends and study the dynamics of treatment effects, I conduct event study regressions following equation 2.

[Figure II about here.]

Figure II, Panel A, depicts the event-study graph of evolution knowledge in school for the set of states where the reform reduces the evolution coverage in Science Standards. The TWFE OLS coefficients of pre-reform years are all close to zero: The largest pre-reform point estimate is smaller than 0.018, which is less than one third of the smallest post-reform coefficient. All pre-reform coefficients are individually and jointly insignificant, even at the 10 percent level (F-test of joint significance: F = 0.27, p = 0.844). These findings are consistent with the idiosyncratic timing of the reforms and the parallel trends assumption. After the reform, we observe immediate, stable, and significant changes in evolution knowledge. The second post-reform coefficient is the largest and amounts to 9 percentage points (p-value = 0.025).

To account for heterogeneous treatment effects in the presence of staggered treatment timing, I complement the OLS estimates with CS estimates (Callaway and Sant'Anna, 2021). The CS pre-reform and post-reform coefficients are similar to the OLS coefficients in terms of size and significance, which implies that the OLS results are immune to bias from time-varying treatment effects. Specifically, all CS pre-reform coefficients are individually and jointly insignificant, and all post-reform coefficient are individually and jointly significant (the last two individual coefficients at the 1 percent level). The corresponding simple aggregate CS estimates based on the discrete reforms are also mostly significant (see Online Appendix Table A.IV).[27]

Figure II, Panel B, depicts the event-study graph of evolution knowledge in school for the set of states where the reform expands the evolution coverage in Science Standards. I can formally reject the significance of pre-reform coefficients and document some significant post-reform coefficients. However, the reform effects are more pronounced for the states that reduce the evolution coverage (this pattern consistently re-emerges in the analyses on evolution belief and occupational choice). Why is this the case? Plausible reasons include the disproportionate threat of lawsuits against teachers who do not adhere to Science Standards after evolution was removed from the standards, given the long history of anti-evolution advocacy groups litigating against teachers who "illegally" teach evolution. Another plausible reason is the disproportionate media attention that such reforms typically garner, which may heighten the awareness of teachers and other stakeholders, and contribute to a swift implementation.[28]

Further Classroom Outcomes: Beyond reform effects on evolution knowledge, I present supporting analyses on a range of other classroom outcomes related to evolution and biology. Using teacher survey data, I show suggestive evidence 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 Online Appendix A.5). Other teaching strategies including the expression of teachers' personal opinions on the validity of evolution remain unaffected.

[27]They are stronger for the set of reforms that reduce the evolution coverage as compared those that expand it, and stronger for the smaller sample without individuals belonging to the top 20 percent or bottom 20 percent of the evolution score (Marie and Zwiers, 2022).

[28]For example, the reforms that reduce the evolution coverage in Texas and Louisiana received heightened media attention. For Texas, journalists published numerous local and national newspaper articles, for example in the NYT (McKinley, 2009), and even the film "The Revisionaries" (Gold, 2012). For Louisiana, journalists published numerous newspaper articles as well, for example in the NYT (Nossiter, 2009), and teacher organizations (including the National Association of Biology Teachers) alongside 78 Nobel Laureates endorsed the repeal of this reform, further contributing to its visibility.

Moreover, I demonstrate that the reforms do not affect the probability of students' selection of high school biology courses, see Online Appendix Table A.V.

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 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.

Baseline Estimates: Panel B of Table I 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 evolution score estimate of 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.

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.[29] 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

[29] 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.

on a range of outcomes.[30] 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, Petrova, and Zhuravskaya, 2011) for different media, settings and outcomes. Reasons for the large persuasion rate of evolution teaching may include the large amount of time dedicated to evolution teaching,[31] the credibility of the persuader,[32] and the difficulty to avoid exposure.[33]

In terms of effect heterogeneities, I observe significantly larger reform impacts on evolution beliefs among individuals raised as Mainline Protestants compared to those raised non-religiously, as well as among Blacks compared to Whites, and among those raised in urban areas compared to rural areas. These heterogeneity results, along with others from the analyses on evolution knowledge and occupational choice, are discussed in more detail in Online Appendix A.6.

Event Study Graphs: The event study graphs for reform effects on evolution belief are presented in Figure III.

[Figure III about here.]

For the set of states reducing the evolution coverage shown in Panel A of Figure III, the pre-reform TWFE OLS coefficients are close to zero as well as individually and jointly insignificant, even at the 10 percent level (F-test of joint significance: $F = 0.75$, $p = 0.540$). After the reform, I observe immediate and significant changes in evolution belief. The first post-reform coefficient is the largest; it amounts to 31 percentage points (p-value = 0.017). The CS results are similar to the OLS results, both in terms of size and significance, for the pre- and post-reform coefficients. These findings are consistent with the idiosyncratic timing of the reforms and the parallel trends assumption, and with the

[30] 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).

[31] The number of hours a high school biology teacher in U.S. public schools spends on teaching evolution amounts to 14-20 hours on average, see Plutzer, Branch, and Reid (2020).

[32] Trust in public schools has exceeded trust in other persuaders such as newspapers and television news consistently throughout the last decades, see Gallup (2022).

[33] I.e., the inability of students to “switch off” the teacher like a TV program conditional on being in class.

absence of contamination by time-varying treatment effects of the OLS. As before, reform effects are more pronounced for the states that reduce the evolution coverage (Panel A) compared to the states that expand it (Panel B).

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. We know from the literature that skills and interest are important determinants of occupational choice (Speer, 2017). If education standards are able to change students' knowledge about and belief in a given theory, they could also change the desire to work in the area that is built on this theory. 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.

Baseline Estimates: Panel C of Table I shows that being exposed to a more comprehensive teaching of evolution in school increases the probability of working in life sciences during adulthood. The full model presented in column (4) demonstrates 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 numerically small, however, it amounts to 23 percent of the sample mean. In absolute terms, it corresponds to more than 90,000 workers (approx. 260,000,000 adults in the US * 0.00035). As the life science industry faces a severe shortage of skilled workers, an addition of 90,000 workers would yield a substantial contribution.[34] This is also of wider economic relevance, given the significance of the life science industry for growth, innovation, employment, and trade.[35]

[34]In recent years, the number of unique vacancies has exceeded the number of new hires by about 40 percent in the life science industry, indicating a tight labor market (Cushman and Wakefield, 2023). And the report "How to Ensure That America's Life-Sciences Sector Remains Globally Competitive" by the Information Technology & Innovation Foundation concludes "Overall, we need more qualified STEM workers" (Kennedy, 2020, p.5).

[35]The life science industry accounts for 2.9 percent of the US GDP (or 7.3 percent when indirect impacts are included) and has grown by an average of 7.8 percent per year over the last decade (Biotechnology

Not least the COVID-19 pandemic has underscored the importance of the life science sector and its innovations (Barro, 2022).

Subfield analysis reveals that the overall effect on the life sciences is mostly coming from the subfield of biology, see Figure IV.

[Figure IV about here.]

The effect on biology is large in relative size, amounting to more than 39 percent of the sample mean (p-value = 0.001). It is significantly different from the effects on agriculture and food as well as conservation and forestry (with reform effects on all non-biology subfields themselves being indistinguishable from zero). 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,[36] and given that evolution is being taught in biology.

Event Study Graphs: Figure V depicts the event study graph of the probability to work in life sciences.

[Figure V about here.]

For the states reducing the evolution coverage depicted in Panel A of Figure V, the pre-reform TWFE OLS coefficients are close to zero as well as individually and jointly insignificant, even at the 10 percent level (F-test of joint significance: F = 0.17, p = 0.912). After the reform, I observe changes in the probability of working in life sciences that peak at the second post-reform coefficient of 0.052 percentage points (p-value = 0.008). We repeatedly observe in this paper that reform effects do not reach their peak immediately after reform adoption (with the OLS event study on evolution belief for the set of states reducing the evolution coverage shown in Panel A of Figure III being an exception to this

Innovation Organization, 2022). Furthermore, it plays a particularly crucial role in driving innovation in the US, with about 20 percent of US patents originating from the life sciences (Cipher, 2022). Moreover, the employment multiplier of the life science industry is estimated to be 4.82, implying that for every job in the life sciences, an additional 3.82 jobs are supported across the rest of the US economy (Biotechnology Innovation Organization, 2022). The export volume of US biopharmaceutical goods has quadrupled in the period between 2002 and 2022, amounting to almost 90 billion US dollars in 2022 (US Census Bureau).

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

pattern). In general, this pattern aligns with reports of slight delays in the implementation of reforms, likely attributable to adjustments in lesson plans, curricula, textbooks, and standardized testing, as discussed in Section 2. Also, note that in Panel A of Figure V, the CS results are similar to the OLS results both in terms of size and significance. As before, reform effects are more pronounced for the states that reduce the evolution coverage (Panel A) compared to the states that expand it (Panel B).

6 ROBUSTNESS

This section shows additional analyses on placebo reforms, other outcome variables, event study figures, and a range of further checks including estimations that control for state-specific trends, and that run on a set of states with closely elected governors.

Placebo Tests: As an alternative way of inference, I randomly reshuffle the reform years across the different reforming states. For the analysis on evolution knowledge, the density plot of the baseline TWFE OLS 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 Online Appendix Figure A.IV. The baseline effect on evolution belief is also larger than the 95^{th} percentile of the placebo reform effect distribution (see Online Appendix Figure A.V), and the baseline effect on the probability to work in life sciences is larger than the 90^{th} percentile of the placebo reform effect distribution (see Online Appendix Figure A.VI). These findings suggest that all three main effects of this paper are unlikely to be spurious.

Other Outcomes: Beyond evolution knowledge and belief, I also estimate reform effects on knowledge of non-evolution scientific topics at the end of high school, and attitudes about non-evolution scientific, religious, and political topics in adulthood. In principle, these outcomes could interact with each other, and react to the reforms.

As is visible in column (2) of Online Appendix Table A.VI, there is no effect of the evolution coverage in Science Standards on the average student knowledge in the non-evolution scientific topics (I reject the equality of coefficients with evolution knowledge at the 10 percent level; p-value = 0.058). The coefficients for the nine individual non-evolution topics are all numerically smaller than that of evolution, but I only reject the equality of coefficients of evolution and "motion" at the 10 percent level (p-value = 0.086), and "energy" at the 5 percent level (p-value = 0.034). As is visible in Online Appendix Tables A.VII, A.VIII, A.IX, and A.X, respectively, there is no effect of the reforms on non-evolution scientific, religious, and political outcomes.[37] In sum, the reform effects are neatly tied to the topic of evolution in independent datasets and outcomes therein.

How could the null finding on religious outcomes arise? For some subgroups, such as those raised as Evangelicals and Non-religious, there is no reform effect on evolution belief. For those, we do not expect to observe a "crowding out" of religious beliefs. But what about the other subgroups? For some of those, evolution may not be contradictory to their religious beliefs, for example, if they follow "Theistic Evolution" (the belief that God does not interfere with the laws of nature after creation). For some others where evolution and religion contradict, compartmentalization, a mechanism where a person separates conflicting beliefs from each other, could be at work. Furthermore, religious beliefs may be harder to change by the school curriculum compared to evolution beliefs, for example due to the relatively early exposure to religiosity during childhood (Huber, 2007), the connection of religious activities to social groups (Levy and Razin, 2012), and the social pressure to maintain religious beliefs (Edgell, Gerteis, and Hartmann, 2006; Wiles, 2014). Finally, if the impact of evolution teaching on religiosity manifests at a later stage in life compared to its effect on evolution belief, my data structure may not be capable of detecting religiosity effects. (Note that I do not have data on older post-reform cohorts, as the evaluated reforms occurred between 2000 and 2009).

Event Study Graphs: I also assess the sensitivity of the event study estimates to violations of parallel trends. Following Rambachan and Roth (2023), I compare the

[37]Between the main outcome "evolution belief" and the average of the scientific topical outcomes, I reject the equality of coefficients at the 1 percent level (p-value = 0.0001). The analogous equality for the average of the broader science attitudes (interpreted as science prestige) is rejected at the 1 percent level as well (p-value = 0.004). The analogous equality for the average of the political and religious outcomes is rejected at the 1 percent level (p-value = 0.003), and at the 5 percent level (p-value = 0.022), respectively. Regarding the individual outcomes, I reject the equality of coefficients between "evolution belief" and the individual scientific topical outcomes for 6 out of 9 comparisons, and between "evolution belief" and the individual broader science attitudes for 3 out of 3 comparisons. For the individual religious and political outcomes, I reject the analogous equality for 8 out of 13 comparisons, and for 7 out of 17 comparisons, respectively.

confidence intervals of my main event study estimates from Figures II, III, and V with confidence intervals that account for precision of pre-trends and allow for per-period deviations from linear trends up to parameter M. These confidence sets correct for the unintuitive property of standard pre-trend testing that zero pre-trends are less likely to be rejected if they are estimated less precisely (Roth, 2022; Roth et al., 2023).

For the states reducing the evolution coverage shown in Panel A of Online Appendix Figures A.VII, A.VIII, and A.IX, the 95% confidence intervals become smaller and exclude the zero when assuming linear trends (M=0) in all three analyses. Confidence intervals then gradually increase with larger deviations from linearity, and exclude the zero for values of M up until 0.002, 0.11, and 0.005, respectively. These thresholds correspond to 8 percent, 94 percent, and 33 percent of the standard error of the respective coefficient of interest. This indicates that the finding on belief in evolution is most robust against low power of pre-trends and non-linear violations of parallel trends, followed by the result on the probability of working in life sciences, while the finding on knowledge of evolution only demonstrates robustness against relatively smooth non-linear trends. Note that these estimates are conservative as I do not impose any restriction on the sign or monotonicity of the differential trends. In contrast, reform effects for the states expanding the evolution coverage are not robust to non-linear violations of parallel trends, see Panel B of Online Appendix Figures A.VII, A.VIII, and A.IX.

Moreover, I probe the robustness of the main event study graphs to adding of a set of never-treated states to the sample. The event study graphs do not change meaningfully, see Online Appendix Figures A.X, A.XI, and A.XII.[38]

Further Robustness: Table II covers a range of further robustness checks including specifications that control for state-specific trends, and that are estimated on a subsample of closely elected governors.

[Table II about here.]

More information about these robustness checks, as well as other checks that account

[38]For the finding on the probability of working in the life sciences, I observe a positive pre-trend for the OLS estimator. This pre-trend completely disappears for the CS estimator that accounts for heterogeneous treatment effects.

for multiple hypothesis testing or define treatment as a binary variable, is presented in Online Appendix A.7.1. Note that some of the point estimates are quite sensitive to the inclusion of control variables, which I mostly attribute to the relatively small number of observations within each state-cohort cell relative to the number of control variables added to some specifications (for example, compare the main results in columns (3) and (4) of Table I, or the robustness check with state-specific time trends in column (3) of Table II to the main results in column (4) of Table I). The robustness test on the set of states with closely elected governors is particularly demanding as it reduces the sample size by around two thirds, but reform effects are still robust (with the effects on the probability of working in life science being estimated less precisely, but yielding a similar point estimate).

Online Appendix A.7.2 assesses student movements between public schools, private schools, and homeschooling due to the reforms, and finds no evidence of meaningful changes to the sample composition by school type.

7 CONCLUSION

This paper shows that school curricula have lasting effects on students by affecting their knowledge, attitudes, and choices. 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 show that the teaching of evolution affects (i) students' knowledge about evolution, (ii) their belief in evolution in adulthood, and (iii) the probability of working in life sciences. To illustrate 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. Naive linear extrapolation of the 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 (biology) would increase by 8 percent (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, Shih, and Sparber, 2015) on individual and societal outcomes (although ultimate welfare implications depend also on other factors such as substitution patterns). 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. 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 (Moore, Jensen, and Hatch, 2003a; Loveless, 2021). Still, legal pressures on school districts to follow education standards, and the reflection of education standards in textbooks and standardized testing questions have arguably incentivized teachers to follow education standards. The analyses presented in this paper provide empirical evidence that education standards indeed 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 reforms evaluated in this paper, the findings indicate potential relevance of other education policies that increase the time teachers spend on teaching evolution, such as the adoption of the Next Generation Science Standards. Beyond the US, the findings may also have a bearing for other countries where the teaching of evolution is controversial.[39] 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.

ETH ZURICH, SWITZERLAND

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

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge, "When Should You Adjust Standard Errors for Clustering?" *Quarterly Journal of Economics*, 138 (2022), 1–35.

Adukia, Anjali, Alex Eble, Emileigh Harrison, Hakizumwami Birali Runesha, and Teodora Szasz, "What We Teach About Race and Gender: Representation in Images and Text of Children's Books," *Quarterly Journal of Economics*, 138 (2023), 2225–2285.

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

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

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

Altonji, Joseph G., Peter Arcidiacono, and Arnaud Maurel, "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, 2016, 305–396.

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

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

Arcidiacono, Peter, Esteban Aucejo, Arnaud Maurel, and Tyler Ransom, "College Attrition and the Dynamics of Information Revelation," *NBER Working Paper No. 22325* (2016).

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

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

Athey, Susan and Guido W. Imbens, "Design-Based Analysis in Difference-In-Differences Settings with Staggered Adoption," *Journal of Econometrics*, 226 (2022), Annals Issue in Honor of Gary Chamberlain, 62–79.

Bailey, Michael, Drew M. Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel, and Johannes Stroebel, "Social Networks Shape Beliefs and Behavior: Evidence from Social Distancing During the COVID-19 Pandemic," *NBER Working Paper, No. 28234* (2020).

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

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

— "Vaccination rates and COVID outcomes across U.S. states," *Economics & Human Biology*, 47 (2022), 101–201.

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

Bazzi, Samuel, Masyhur Hilmy, and Benjamin Marx, "Religion, Education, and Development," *NBER Working Paper No. 27073* (2020).

Beale, H. K., *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, 1941.

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

Bernheim, B.Douglas, Daniel M. Garrett, and Dean M. Maki, "Education and saving: The long-term effects of high school financial curriculum mandates," *Journal of Public Economics*, 80 (2001), 435–465.

Biasi, Barbara and Song Ma, "The education-innovation gap," *NBER Working Paper No. 29853* (2022).

Biotechnology Innovation Organization, "The U.S. Bioscience Industry: Fostering Innovation and Driving America's Economy Forward," *Biotechnology Innovation Organization Report* (2022).

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

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

Borusyak, Kirill, Xavier Jaravel, and Jan Spiess, "Revisiting Event Study Designs: Robust and Efficient Estimation," *Review of Economic Studies, accepted* (2023).

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

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

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

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

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

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

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

Cipher, "Global Innovation Report: A breakdown of the world's inventions and activity by technology," *Cipher Report* (2022).

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

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

Cook, Lisa D, Maggie E C Jones, Trevon D Logan, and David Rosé, "The Evolution of Access to Public Accommodations in the United States," *Quarterly Journal of Economics*, 138 (2022), 37–102.

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

Costa-Font, Joan, Jorge García-Hombrados, and Anna Nicińska, "Long-lasting effects of indoctrination in school: Evidence from the People's Republic of Poland," *European Economic Review*, 161 (2024), 104641.

Cushman and Wakefield, "Life Sciences: Update. March 2023, United States," *Cushman and Wakefield Report* (2023).

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

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

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

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

Edgell, Penny, Joseph Gerteis, and Douglas Hartmann, "Atheists As "Other": Moral Boundaries and Cultural Membership in American Society," *American Sociological Review*, 71 (2006), 211–234.

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

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

Fuchs-Schündeln, Nicola and Paulo Masella, "Long-Lasting Effects of Socialist Education," *Review of Economics and Statistics*, 98 (2016), 428–441.

Gallup, *Gallup News Service: June 1-20, 2022 - Final Topline*, 2022.

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

Gold, Daniel, "Culture Wars in the School Board; Movie Review of The Revisionaries". *New York Times*, 2012, URL: https://www.nytimes.com/2012/10/26/movies/the-revisionaries-looks-at-textbooks-in-texas.html.

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

Goodman-Bacon, Andrew, "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 225 (2021), Themed Issue: Treatment Effect 1, 254–277.

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

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

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

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

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

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

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

Huber, Stefan, "Are religious beliefs relevant in daily life?" *Religion inside and outside traditional institutions*, Brill, 2007, 209–230.

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

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

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

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

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

Joravsky, David, "The Lysenko Affair," *Scientific American*, 207 (1962), 41–49.

Kennedy, Joe, "How to Ensure That America's Life-Sciences Sector Remains Globally Competitive," *Information Technology and Innovation Foundation, Update (Original Report from 2018)*, 15 (2020).

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

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

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

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

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

Levy, Gilat and Ronny Razin, “Religious Beliefs, Religious Participation, and Cooperation,” *American Economic Journal: Microeconomics*, 4 (2012), 121–51.

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

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

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

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

Marks, Jonathan, “Why be Against Darwin? Creationism, Racism, and the Roots of Anthropology,” *American Journal of Physical Anthropology*, 149 (2012), 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, “Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations,” *American Journal of Human Genetics*, 100 (2017), 635–649.

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

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

— “Religion and Political Economy in an International Panel,” *Journal for the Scientific Study of Religion*, 45 (2006), 149–175.

McCleary, Rachel M. and Robert J. Barro, *The Wealth of Religions: The Political Economy of Believing and Belonging*, Princeton University Press, 2019.

McKinley, James C., "Split Outcome in Texas Battle on Teaching of Evolution". *New York Times*, 2009, URL: https://www.nytimes.com/2009/01/24/education/24texas.html.

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

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

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

— "Twenty Questions: What Have the Courts Said about the Teaching of Evolution and Creationism in Public Schools?" *BioScience*, 53 (2003), 766–771.

National Science and Technology Council, "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* (2018).

Nossiter, Adam, "Boycott by Science Group Over Louisiana Law Seen as Door to Teaching Creationism ". *New York Times*, 2009, URL: https://www.nytimes.com/2009/02/17/us/17boycott.html.

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

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

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

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

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

Rambachan, Ashesh and Jonathan Roth, “A More Credible Approach to Parallel Trends,” *Review of Economic Studies*, 90 (2023), 2555–2591.

Roth, Jonathan, “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends,” *American Economic Review: Insights*, 4 (2022), 305–22.

Roth, Jonathan, Pedro H.C. Sant’Anna, Alyssa Bilinski, and John Poe, “What’s trending in difference-in-differences? A synthesis of the recent econometrics literature,” *Journal of Econometrics*, 235 (2023), 2218–2244.

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

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

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

Schmidheiny, Kurt and Sebastian Siegloch, “On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization,” *Journal of Applied Econometrics*, 38 (2023), 695–713.

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

Speer, Jamin D, “Pre-market skills, occupational choice, and career progression,” *Journal of Human Resources*, 52 (2017), 187–246.

Staff, Toi, “Israeli schools largely avoid teaching evolution”. *Times of Israel*, 2018, URL: https://www.timesofisrael.com/israeli-schools-largely-avoid-teaching-evolution-report/.

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

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

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

Wiles, Jason R, “Gifted students’ perceptions of their acceptance of evolution, changes in acceptance, and factors involved therein,” *Evolution: Education and Outreach*, 7 (2014), 1–19.

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

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

MAIN TABLES AND FIGURES

Table I
Effect of Evolution Coverage in Science Standards on Evolution Knowledge in School (Panel A), on Evolution Belief in Adulthood (Panel B), and on Probability of Working in Life Sciences (Panel C)

	(1) Controls : NO State FE: NO Cohort FE: NO	(2) Controls : YES State FE: NO Cohort FE: NO	(3) Controls : NO State FE: YES Cohort FE: YES	(4) Controls : YES State FE: YES Cohort FE: YES
Panel A: Outcome: Evolution Knowledge in School				
Evolution Score	0.039**	0.035***	0.074**	0.065***
	(0.018)	(0.012)	(0.028)	(0.022)
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	0.035	0.015	0.041
Observations	14,080	14,080	14,080	14,080
Panel B: Outcome: Evolution Belief in Adulthood				
Evolution Score	0.122***	0.084**	0.235*	0.333***
	(0.042)	(0.036)	(0.125)	(0.107)
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	0.006	0.088	0.038	0.107
Observations	1,801	1,801	1,801	1,801
Panel C: Outcome: Working in Life Sciences				
Evolution Score	0.039**	0.035**	0.035**	0.035**
	(0.018)	(0.013)	(0.014)	(0.014)
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: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1, for different sets of control variables and fixed effects as indicated in the column header. Each entry is from separate regression model. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. **Panel A:** Dependent variable: Share of questions about evolution answered correctly. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, home possessions (separate indicator variables for computer and books), birth month, test session, and test year. 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. **Panel B:** 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. Data source: General Social Survey. **Panel C:** Dependent variable: Probability of working in life sciences (multiplied by 100 for interpretability). Controls: Indicator variables for gender, races/ethnicities, and survey year. Data source: American Community Survey.

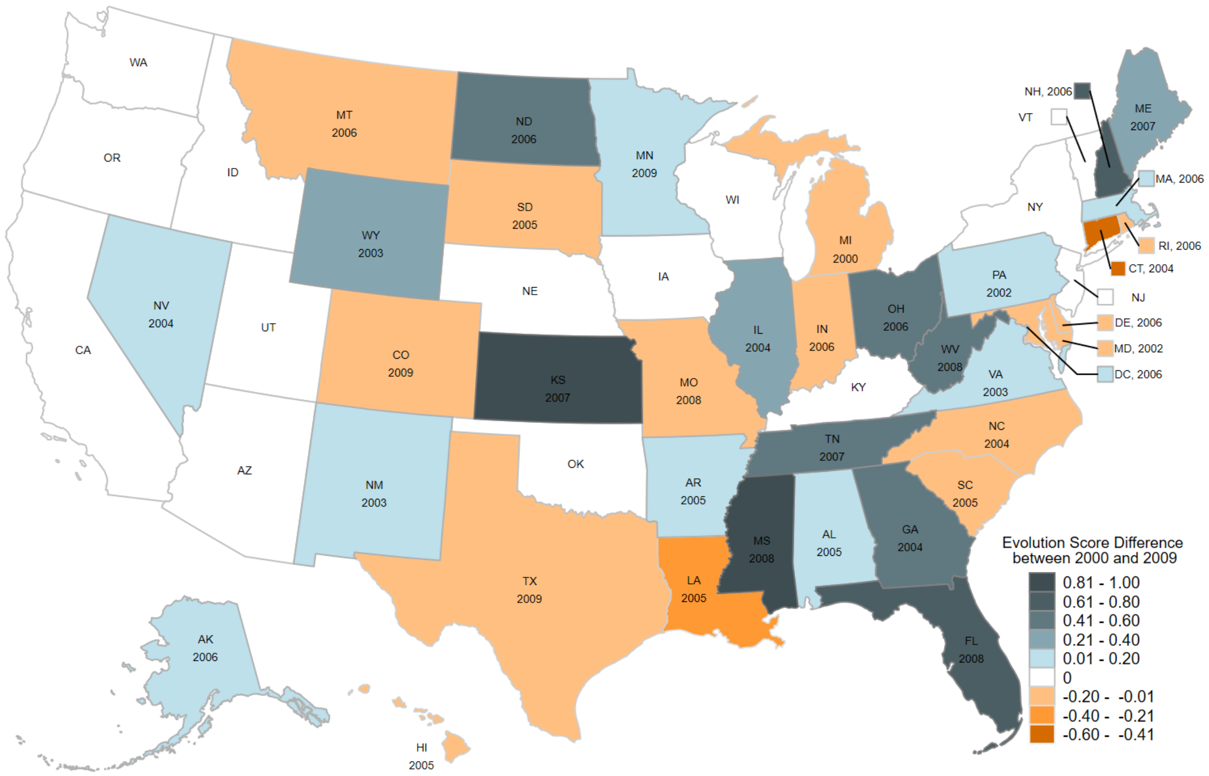
Table II

Robustness on Effect of Evolution Coverage in Science Standards on Evolution Knowledge in School (Panel A), on Evolution Belief in Adulthood (Panel B), and on Probability of Working in Life Sciences (Panel C)

	(1) Close Elections	(2) Control: Governor's Party	(3) State Specific Time Trends	(4) Dosage Treatment	(5) Only One Reform Event	(6) Only Large States	(7) Only States With Std. Text	(8) Sample Start: 1995	(9) Sample Start: 2000	(10) Probit	(11) Outcome Coding Variation 1	(12) Outcome Coding Variation 2
Panel A: Outcome: Evolution Knowledge in School												
Evolution Score	0.093***	0.066***	0.043	0.042	0.089***	0.056**	0.078***	0.048***	0.037	0.076**	0.058**	n/a
	(0.022)	(0.022)	(0.065)	(0.034)	(0.029)	(0.022)	(0.025)	(0.017)	(0.026)	(0.032)	(0.022)	
Equality Test w/ Base. Coef. (P-value)	0.265	0.798	0.533	0.220	0.314	0.621	0.340	0.052	0.014	0.494	0.153	
Panel B: Outcome: Evolution Belief in Adulthood												
Evolution Score	0.605***	0.332***	0.625***	0.177**	0.394**	0.433***	0.322**	0.257**	0.313*	0.329**	0.288**	0.426***
	(0.188)	(0.111)	(0.218)	(0.070)	(0.163)	(0.117)	(0.131)	(0.116)	(0.171)	(0.130)	(0.138)	(0.145)
Equality Test w/ Base. Coef. (P-value)	0.072	0.954	0.042	0.461	0.582	0.158	0.893	0.029	0.852	0.792	0.323	0.107
Panel C: Outcome: Working in Life Sciences												
Evolution Score	0.039	0.036**	0.025	0.034**	0.031	0.040**	0.057***	0.036***	0.029**	0.035	n/a	n/a
	(0.025)	(0.014)	(0.025)	(0.016)	(0.021)	(0.017)	(0.016)	(0.013)	(0.012)	(0.026)		
Equality Test w/ Base. Coef. (P-value)	0.848	0.731	0.439	0.963	0.724	0.663	0.080	0.982	0.385	0.984		

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1, for different 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) Main evolution score is replaced by weighted average of pre- and post-reform evolution scores, with weights corresponding to number of pre- and post-reform high school years; (5) Sample only includes individuals from states that had only one reform event between 2000 and 2009, see Table A.XI for more details; (6) Sample only includes individuals from 20 states with the largest population; (7) Sample only includes individuals from states for which text of Science Standards is available for text analysis implemented in Online Appendix A.1; (8) Sample only includes individuals who started high school after 1994; (9) Sample only includes individuals who started high school after 1999; (10) Coefficient reports average marginal treatment effect of probit specification; (11) Panel A: 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); Panel B: Sample excludes individuals whose dependent variable question on evolution replaces the words "human beings" with the word "elephants"; Panel C: n/a; (12) Panel A: n/a; Panel B: 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); Panel C: n/a. Each entry is from separate regression model. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. "Equality Test w/ Base. Coef. (P-value)" refers to p-value of test for equality between coefficient reported in given column and corresponding baseline coefficient presented in Table I, Column 4. **Panel A:** Dependent variable: Share of questions about evolution answered correctly. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, home possessions (separate indicator variables for computer and books), birth month, test session, and fixed effects for state, cohort, and test year. 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. **Panel B:** 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, cohort, and survey year. Data source: General Social Survey. **Panel C:** 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, cohort, and survey year. Data source: American Community Survey. Online Appendix Tables A.XII, A.XIII, and A.XIV, provide further statistics of each regression including mean and standard deviation of dependent variable, adjusted R-Squared, and number of observations.

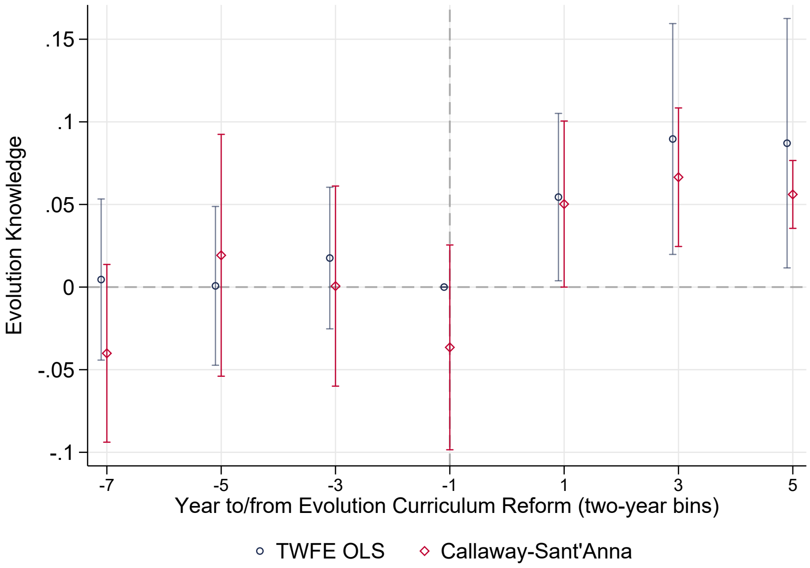
Figure I
US Map of Evolution Score Difference Between 2000 and 2009



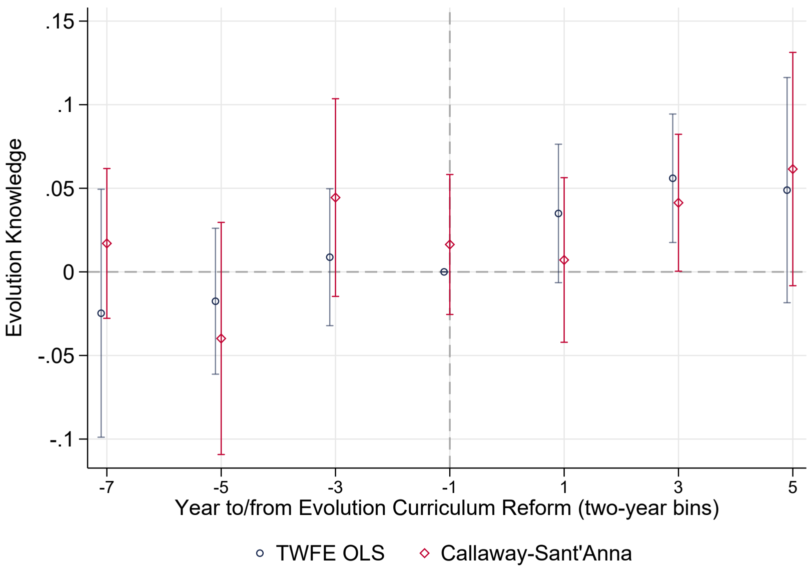
Note: Map depicts the evolution score difference, which I define as the evolution score of 2009 minus the evolution score of 2000. A positive (negative) difference implies an increase (decrease) in the evolution score between 2000 and 2009, as indicated by blue (orange) coloring. White coloring indicates 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.XI. Data source: Lerner (2000b), Mead and Mates (2009)

Figure II
Event-Study Graphs for Effect of Evolution Coverage in Science Standards on Evolution Knowledge in School

(A) States reducing evolution coverage



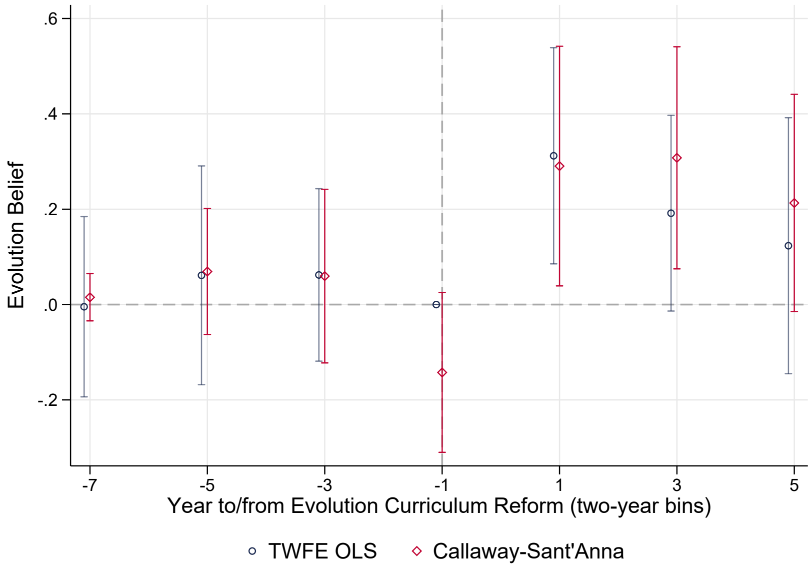
(B) States expanding evolution coverage



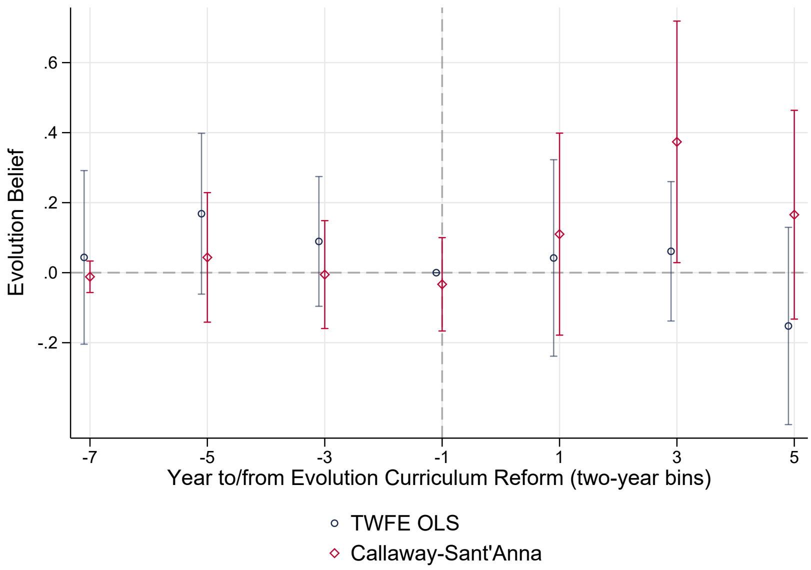
Note: Figure shows point estimates and 95% confidence intervals from estimating equation 2. Dependent variable: Share of questions about evolution answered correctly. In Panel A, outcome variable is multiplied by minus 1 (inverted outcome due to inverted reform, to allow for comparability across results). TWFE OLS depicted in blue with circle markers. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, home possessions (separate indicator variables for computer and books), birth month, test session, and fixed effects for state, cohort, and test year. CS estimators (Callaway and Sant'Anna, 2021) depicted in red with diamond markers, using doubly robust inverse probability weighting and not-yet-treated observations as controls. Numbers on horizontal axis refer to final year of respective two-year bins; i.e., -1 = last two years prior to treatment. Inference: Clustering at state level. Data source: National Assessment of Educational Progress

Figure III
Event-Study Graphs for Effect of Evolution Coverage in Science Standards on Evolution Belief in Adulthood

(A) States reducing evolution coverage

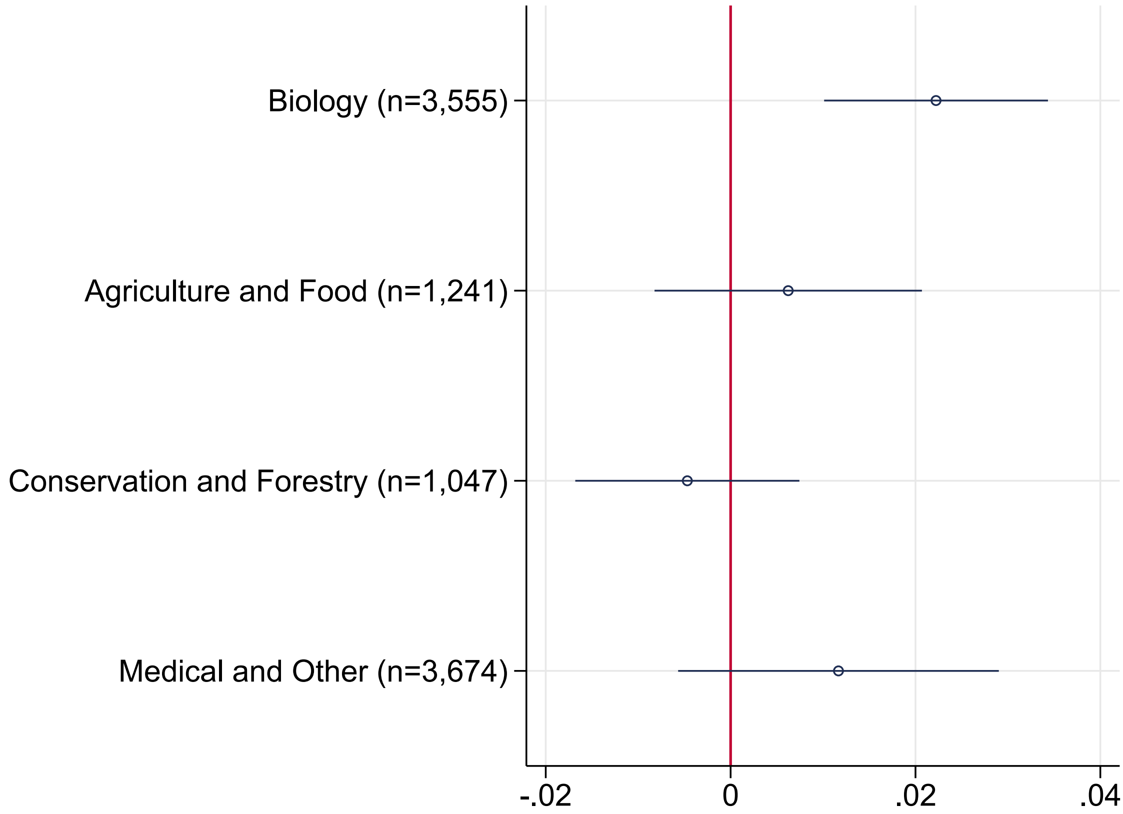


(B) States expanding evolution coverage



Note: Figure shows point estimates and 95% confidence intervals from estimating equation 2. 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). In Panel A, outcome variable is multiplied by minus 1 (inverted outcome due to inverted reform, to allow for comparability across results). TWFE OLS depicted in blue with circle markers. Controls: Fixed effects for state, cohort, and survey year. CS estimators (Callaway and Sant'Anna, 2021) depicted in red with diamond markers, using doubly robust inverse probability weighting and not-yet-treated observations as controls. Numbers on horizontal axis refer to final year of respective two-year bins; i.e., -1 = last two years prior to treatment. Inference: Clustering at state level. Data source: General Social Survey

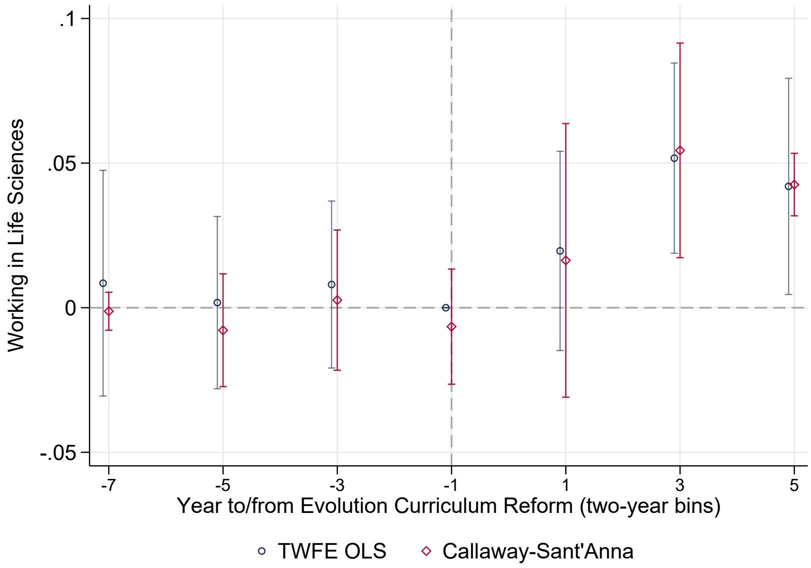
Figure IV
Effect of Evolution Coverage in Science Standards on Probability of Working in Life Sciences, by Subfields of Life Sciences



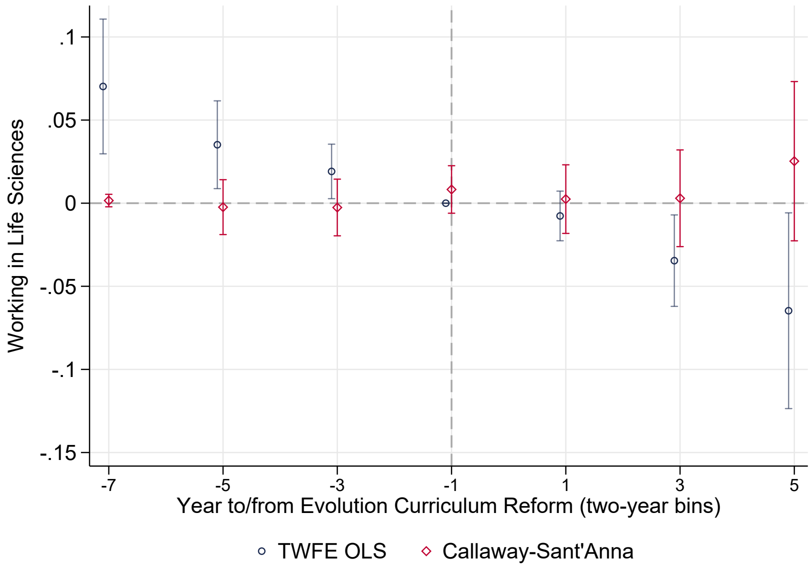
Note: Figure shows TWFE OLS coefficients and 95% confidence intervals of effect of evolution coverage in Science Standards on probability of working in life sciences (multiplied by 100 for interpretability), by subfields of life sciences as indicated along the vertical axis from estimating equation 1. Sample sizes of subfields in parenthesis (raw value). Controls: Indicator variables for gender, races/ethnicities, and fixed effects for state, cohort, and survey year. Inference: Clustering at state level. Data source: American Community Survey

Figure V
Event-Study Graphs for Effect of Evolution Coverage in Science Standards on Probability of Working in Life Sciences

(A) States reducing evolution coverage



(B) States expanding evolution coverage



Note: Figure shows point estimates and 95% confidence intervals from estimating equation 2. Dependent variable: Probability of working in life sciences (multiplied by 100 for interpretability). In Panel A, outcome variable is multiplied by minus 1 (inverted outcome due to inverted reform, to allow for comparability across results). TWFE OLS depicted in blue with circle markers. Controls: Indicator variables for gender, races/ethnicities, and fixed effects for state, cohort, and survey year. CS estimators (Callaway and Sant'Anna, 2021) depicted in red with diamond markers, using doubly robust inverse probability weighting and not-yet-treated observations as controls. Numbers on horizontal axis refer to final year of respective two-year bins; i.e., -1 = last two years prior to treatment. Inference: Clustering at state level. Data source: American Community Survey

EVOLUTION VS. CREATIONISM IN THE CLASSROOM: THE LASTING EFFECTS OF SCIENCE EDUCATION

BENJAMIN W. AROLD

Online Appendix

A ONLINE APPENDIX

A.1 Text Analysis on Science Standards

This analysis provides evidence from the text of the Science Standards that the evaluated reforms alter the coverage of evolution in the Science Standards to a much larger extent compared to the coverage of non-evolution scientific topics. The following paragraphs describe the data, empirical approach, and results.

Data: To implement the analysis, I require the original texts of the evaluated Science Standards. In total, I managed to collect 73 Science Standards (27 standards from 2000, and 46 standards from 2009, with all 27 standards from 2000 also being available for 2009) out of 100 standards (50 standards from 2000 and 2009, respectively).[1] The subset of states for which I have Science Standards from 2000 and 2009 covers all census regions and 8 out of 9 census divisions. It is also largely representative of the US along scientific, religious, and political dimensions, exhibiting moderately higher belief in evolution and slightly lower levels of religiosity and political conservatism, see Online Appendix Table A.XV. The main results of the paper all replicate in this subset of states, as shown in column (7) of Table II.

To construct measures of the coverage of different scientific topics in the Science Standards, I create word count measures for a range of scientific topics, including evolution. To maintain consistency, I use the exact same 10 scientific topics as those used to classify the NAEP questions (see Section 3.2). These topics include evolution, motion, matter and mass, energy, reproduction, climate change, pollution, the earth, tectonics, and the universe. To begin, I extract the relevant parts of the education standards that focus on science.[2] Then, I convert the PDFs to machine-readable TXT files using an OCR engine. Next, I retrieve the counts for the 10 scientific words listed

[1] I am grateful to Lawrence S. Lerner, the author of Lerner (2000), Louise Mead, the author of Mead and Mates (2009), the Thomas B. Fordham Foundation, the National Center for Science Education, and different State Departments of Education that helped me to collected as many of the historic Science Standards as possible. The Wayback Machine (https://archive.org/web/) proved to be particularly useful in this endeavor too.

[2] I aimed to closely align with the chapters/documents referenced in Lerner (2000) and Mead and Mates (2009).

above. Specifically, I count all words with the applicable word stem, for example, I categorized "evolved" under the "evolution" word category (using Python's Regular Expression Operations Module). Last, I generate a dataset containing the different word counts, the evolution score assigned to the standards by Lerner (2000) and Mead and Mates (2009), as well as state and year indicators.

Empirical Strategy: Using this dataset, I regress the different word count measures on the evolution score as well as state and year fixed effects, as follows:

$$Y_{st} = \beta \cdot Evolution\_Score_{st} + \delta_s + \lambda_t + \epsilon_{st} \tag{3}$$

where Y_{st} is the logarithmized count of a given word in a Science Standard in state s and year t. The treatment variable $Evolution\_Score_{st}$ measures the intensity of the evolution coverage in the Science Standard in state s and year t. State fixed effects δ_s, year fixed effects λ_t, and an error term complete the model. The standard errors are clustered at the state level.

Results: As shown in column (1) of Online Appendix Table A.XVI, a larger evolution score is positively associated with the word count for evolution. Specifically, Science Standards with an evolution score of 1 use the word "evolution" more than 7 times ($e^{1.986}$ = 7.286) more often than Science Standards with an evolution score of 0 (p-value = 0.010), controlling for time-invariant differences between states and national differences between years. This positive significant association is specific to evolution. All coefficients for non-evolution scientific words are smaller and not statistically different from zero (although estimated with some imprecision and not always statistically different from the coefficient on evolution). The sum of all evaluated non-evolution scientific words appears twice as often in Science Standards with an evolution score of 1 compared to those with an evolution score of 0, as reported in column (11) ($e^{0.798}$ = 2.221). This coefficient is not significantly different from zero (p-value = 0.463), and is much smaller in size compared to and significantly different from the corresponding coefficient on evolution reported in column (1) (p = 0.0404). Similarly, the sum of all words (scientific words and all other words including headlines and stopwords) appears about 30 percent as often in Science

Standards with an evolution score of 1 compared to those with an evolution score of 0, as reported in column (12) ($e^{0.300} = 1.349$). This coefficient is not significantly different from zero (p-value = 0.835), and is much smaller in size compared to and significantly different from the corresponding coefficient on evolution reported in column (1) (p = 0.0402).

In sum, these results demonstrate that while the Science Standards cover various scientific dimensions, the non-evolution dimensions are not changing to the same extent as the evolution dimension due to the evaluated reforms.

A.2 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.[3] After years of debating and drafting, in February 2008 the Florida Board of Education voted 4:3 in favor of a new Science Standard that comprehensively included evolution. 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, 2000, 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 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

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

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.3 Tests on the Timing of Reforms

This analysis provides evidence on the exogenous timing of the evaluated reforms, using state-by-year level measures of economic, political, educational, evolution-related, and creationism-related conditions. I regress these characteristics on the evolution score, in models with state and year fixed effects. In contrast to the main analysis that exploits the cohort-specific introduction of the reforms to estimate cohort-specific reform effects, this analysis compares state-by-year-level outcomes in the years around the introduction of the reform. The following paragraphs describe the data, empirical approach, and results.

Data: The first set of outcome variables is chosen to represent the economic, political, and educational conditions in a given state and year. They come from public sources including the Bureau of Economic Analysis and the National Center for Education Statistics. They are collected for the preferred sample cut of the main analysis, i.e. for all years after 1990 and before 2010.

The second set of outcome variables is selected to capture the interest in evolution and creationism of the population in a given state and year. I approximate the population's interest in a topic by the average Google search frequencies of keywords specific to the topic. This data comes from google trends and is available from 2004 onward at the state-by-year level. The search frequencies are provided by Google as normalized, i.e. with a range from 0 to 1, where 0 indicates no search interest in a given keyword (or the absence of sufficient data), and 1 indicates the highest search interest in that keyword in the history of Google Trends for a given state. Hence, the values are directly comparable within state over time, but not across states. However, scaling differences between states are constant over time and therefore accounted for by state fixed effects in the regressions below. In total, I select 5 keywords related to evolution, and 2 keywords related to creationism. I also calculate the average of the 7 keywords, in two ways. In the first average, the words related to evolution enter positively, and the words related to creationism enter negatively. I interpret this as a measure of pro-evolution interest. In the second average, the words related to evolution enter positively, and the words related to creationism enter also positively. I interpret this as a general measure of interest in the topic of evolution and creationism.

Empirical Strategy: I regress the different state-by-year characteristics on the evolution score as well as state and year fixed effects, as follows:

$$Y_{st} = \beta \cdot Evolution\_Score_{st} + \delta_s + \lambda_t + \epsilon_{st} \tag{4}$$

where Y_{st} is the outcome in state s and year t. The treatment variable $Evolution\_Score_{st}$ measures the intensity of the evolution coverage in the Science Standard in state s and year t. State fixed effects δ_s, year fixed effects λ_t, and an error term complete the model. The standard errors are clustered at the state level.

Results: As shown in Online Appendix Table A.XVII, all economic, political, and educational estimates are insignificant and close to zero. This implies that the timing of the evaluated reforms is independent of economic, political, and educational conditions in the years around the reforms. Furthermore, the estimates are also insignificant and close to zero for the outcomes measuring Google search frequencies of keywords specific to evolution and creationism, see Online Appendix Table A.XVIII. This suggests that the interest in evolution and creationism of the population did not change systematically in the years around the reform. Taken together, these findings support the notion that the specific reform timing is as-good-as random.

A.4 *Data Appendix*

A.4.1 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".[4] 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). Online Appendix Figure A.XIII 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.[5] 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-missing average of all questions from these nine

[4]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.

[5]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.

non-evolution scientific topics. Online Appendix Table A.XIX 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. As the CS estimator (Callaway and Sant'Anna, 2021) in the implementation of Rios-Avila, Sant'Anna, and Callaway (2022) requires equal gaps between test/survey years, I reassign the test year of 1996 as 1997, and 2000 as 2001 for all CS estimations. 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.

The descriptive statistics for the main treatment, outcome, and control variables are presented in Online Appendix Table A.XX. 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, as classified by Lerner (2000). 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 Online Appendix Figure A.XIII 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), or by having no PC at home (16 percent).

A.4.2 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 re-code 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.[6]

Regarding correlations, the belief in evolution is almost only positively correlated with the other scientific topical outcomes and science attitudes, see Online Appendix Tables A.XXI and A.XXII. For all religious variables, I find a negative raw correlation with evolution belief as is visible in Online Appendix Table A.XXIII. The correlations between politically conservative attitudes on different topics and evolution belief also tend to be negative, see Online Appendix Table A.XXIV.

Online Appendix Table A.XXV 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

[6]This definition allows for an unequivocal assignment of attitudes to political conservatism/progressiveness, 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 as a progressive attitude. The raw variable is then re-coded such that an increase in the variable implies an increase in conservatism. The re-coding is undertaken to facilitate the interpretation of results and allow for meaningful averaging across political outcomes.

average in NAEP, as 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 88 percent of respondents believe in God, and 70 percent are affiliated with a church. To give examples on conservative political attitudes, 46 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 (38 percent), followed by Catholicism (32 percent), Non-Religious/Agnosticism/Atheism (13 percent), and Evangelicalism (9 percent).

A.4.3 ACS: Occupational Choice

The estimation sample combines ACS waves from 2000-2017. The descriptive statistics are presented in Online Appendix Table A.XXVI. 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.

A.5 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 (Berkman, Pacheco, and Plutzer, 2008; Berkman and Plutzer, 2011).

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 Online Appendix Figure A.I).

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:

$$Y_{is} = \beta \cdot Evolution\_Score_s + \gamma \cdot \mathbf{X_i} + \eta_c + \epsilon_{is} \qquad (5)$$

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 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. Third, I note that the raw correlations (and correlations conditional on individual-level control variables) of the evolution score with evolution knowledge, evolution belief, and the probability of working in the life sciences are identical in terms of direction, and very similar in terms of significance compared to the main specifications with fixed effects, compare columns (1) and (2) with column (4) in Table I. This suggests that the conditional correlations of the evolution score and teacher outcomes presented here might not be too far off from the causal effect in terms of direction and significance.

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 creationist coverage of evolution, are 33 percentage points more likely to spend at least 5 class hours per year on teaching evolution (Online Appendix Table

A.XXVII). 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.6 Heterogeneity Analysis

In the following, I first provide a short summary on the subgroup effects, by outcome variable. Second, I discuss potential reasons that might have given rise to the observed pattern of subgroup effects.

Evolution Knowledge in School: Subgroup analysis by NAEP student characteristics reveals that the point estimates of reform effects on evolution knowledge are numerically larger for students that may be considered underprivileged, i.e. respectively for students who do not have a computer at home, receive subsidized lunch, are female and non-white, see Online Appendix Figure A.XIV. However, these differences are not significant, in contrast to some subgroup results on evolution belief in adulthood discussed next.

Evolution Belief in Adulthood: The subgroup analysis by GSS characteristics shows large reform effects for individuals raised as Mainline Protestants, as depicted in Online Appendix Figure A.XV. In contrast, students are less susceptible to the effects of evolution teaching if they were raised as Evangelicals, Catholics, or Non-religious. Only the difference in reform effects between those raised as Mainline Protestants and as non-religious is statistically significant (p-value = 0.003). Furthermore, reform effects are significantly larger for Blacks relative to Whites (p-value = 0.008) and for those raised in urban areas rather than rural areas (p-value = 0.018). There are no differences by gender.

Occupational Choice: The corresponding subgroup results by ACS characteristics are largely in line with those from the previous subsections. Online Appendix Table A.XXVIII 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.

Discussion of Subgroup Differences by Gender: Despite a lack of statistical significance, reform effects are numerically larger for females than for males on evolution knowledge in school and on the probability of working in life sciences. There is no relevant numerical gender difference in reform effects on evolution belief in adulthood.

In the following, I first provide a list of theoretical reasons for why the effect on evolution knowledge may plausibly be larger for females than for males. Then, I discuss how this can translate into occupational choice. Lastly, I conclude with a discussion of the implications of evolution belief reacting similarly between genders.

Why do females react more strongly to the reforms in terms of their learning about evolution? First, there is evidence that, on average, female students behave differently in the classroom compared to males. If these differences are conducive to learning, then the quality of evolution instruction might be more relevant for females when learning about evolution. For example, there is evidence that female students pay more attention in class, complete their homework more diligently, and interrupt less (Gershenson and Holt, 2015; Farsani et al., 2021).

Second, female high school students are more likely to attend Advanced Placement (AP)/IB biology courses, such as physiology, anatomy, and genetics (Cunningham, Hoyer, and Sparks, 2015). Some of these courses may cover evolution, if the Science Standards allow, and thereby improve the NAEP results on evolution questions.[7] In 2009, the difference in attendance amounted to 10 percentage points (49.9% for females, 39.4% for males; p-value of difference <0.05). While this difference is unlikely to be the main driver of the overall gender gap in reform effects on NAEP evolution knowledge, it may contribute to it.

Third, females enter high school with lower knowledge of life sciences compared to their male counterparts. For example, they consistently perform worse in the life science component of the NAEP Science Test in grade 8 (U.S. Department of Education, 2023). Hence, they have more potential to improve their knowledge about life sciences, including evolution, in high school. This may imply that differences in evolution coverage in the curriculum have a greater impact on the production of evolution knowledge for females

[7]As the Science Standards typically apply to all science courses, the evolution coverage of AP biology courses plausibly reacts to the evaluated reforms in a similar way as the main biology courses. Hence, AP courses may contribute to the overall reform effect for the subset of students that select AP courses. This implies that the gender difference in AP course attendance can be a mechanism of the gender differences in reform effects as long as these courses are relevant for evolution learning. At the same time, the probability of selecting an AP course itself is not affected by the reforms and hence not a mechanism of the overall reform effect, see Online Appendix Table A.V.

relative to males. (It is not possible to compare pre-high school differences in evolution knowledge, as evolution knowledge is only elicited in the NAEP Science Test for grade 12.)

Is it plausible that these gender-specific effects on learning translate into gender-specific effects on occupational choice? It appears plausible that a group that earns more knowledge and skills in a particular area is more likely to work in this area, given the importance of knowledge and skills for occupational choice (Speer, 2017). In addition, there is one gender-specific reason that might be of particular relevance for females and STEM, namely self-confidence. We know from the literature that a lack of confidence and a "sense of belongingness" to science is a primary obstacle for females to choose to work in STEM (Litzler, Samuelson, and Lorah, 2014; Kahn and Ginther, 2017). Enhanced knowledge about evolution as demonstrated through the NAEP performance effect may boost girls' confidence in their skills related to evolution and the life sciences, and therefore translate into a larger probability of choosing to work in life sciences.

We do not observe a corresponding gender difference in effects on evolution belief; what does this imply? The effect on evolution knowledge and occupational choice for females is also there for evolution belief. However, the latter effect is of similar size for males, in contrast to the two former effects, resulting in different gender gaps across analyses. I interpret this pattern as suggestive evidence that females pick up the "intellectual" and "ideological" aspect of evolution teaching. In contrast to females, males pick up only the ideological component: They are more likely to believe in evolution theory without being able to answer in-depth scientific questions about evolution to a greater extent.[8] Since the effect on occupational choice is only large for females, it suggests that merely raising belief in evolution theory is not enough to significantly impact the probability of working in the life sciences. The intellectual component appears to be a key factor.

[8] In this interpretation, I regard the effect on evolution knowledge as evidence of the intellectual component, i.e., understanding, and the effect on evolution belief as the ideological component. A key point here is that both components are conceptually distinct as one can score well in a science exam question on evolution without believing in the correctness of evolution, and one can express belief in evolution in a survey without understanding it.

Discussion of Subgroup Differences by Race, Urbanicity, and Religious Upbringing:

The shares of females and males do not differ between different racial and socioeconomic groups. In contrast, many of the racial and socioeconomic variables are correlated with each other. Therefore, I discuss the effect heterogeneities by these factors in the context of the GSS analysis, as it contains information on race, socioeconomic status, and a range of further characteristics of an individual's upbringing. Specifically, I focus on heterogeneities by race, urbanicity, and religious upbringing for which I find significant differences.

The reform effects for Blacks/Hispanics are numerically larger compared to Whites (if expressed relative to the respective subsample mean) across all three main analyses of the paper. Are these differences by race to be expected? It should be noted that throughout the paper, effects tend to be numerically larger for groups typically regarded as less privileged or of lower socioeconomic status (such as those without a PC at home in the NAEP analysis). Blacks have been over-represented in underprivileged groups in society for a long time, both in general economic terms (Chetty et al., 2020) and specifically in education contexts (Bertocchi and Dimico, 2012). Beyond socioeconomic differences, Blacks are over-represented in other groups for which there are large reform effects, such as being raised in urban areas or as Mainline Protestants (see Online Appendix Table A.XXIX). Hence, the large reform effects for Blacks are in line with the general pattern of results. Moreover, the effects for Blacks are in line with the findings of Goodman (2019) who documents that states' high school math requirements increase math coursework and later earnings primarily by Blacks.

Are effect differences by growing up in urban as opposed to rural areas to be expected? As shown in Online Appendix Table A.XXIX, urban individuals are more likely to exhibit characteristics associated with large reform effects, such as being Black/Hispanic and being raised by parents born abroad (i.e., in family circumstances where parental involvement might be lower on average, and schools might be relatively more important). Furthermore, there are other fundamental differences between urban and rural schooling contexts that might contribute to a more efficient reform implementation in urban areas. These include differences in schooling resources, including internet access (Pew Research Center, 2018) and access to mental health support (Andrilla et al., 2018), differences in school choice, school competition, and, relatedly, commuting time (Gibbons and Silva, 2008), as well as differences in information, network effects, and economies of scale more generally (Hoxby and Avery, 2012).

Are effect differences by the childhood religion to be expected? While Mainline Protestants with moderate views on evolution, on average, display the largest reform effects, religious groups at the ends of the evolution belief spectrum (Evangelicals and non-religious, respectively) are less influenced by the school. This may happen, arguably, as they have strong views on the topic to begin with. Of course, there are also differences in correlations with covariates that may contribute to the overall results pattern, such as Mainline Protestants being over-proportionally likely to be Black.

In general, this study does not attempt to separate different subgroup effects, such as the effect of being Mainline Protestant from the effect of being Black, given the empirical setting and corresponding data limitations. This implies that the discussed reasons for subgroup effect differences should be interpreted as plausible suggestions rather than definitive answers.

A.7 Robustness Appendix

A.7.1 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 around two thirds, the reform effects are robust (see column (1) of Table II, with the effect on the probability of working in life sciences being estimated less precisely). 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 Table II, 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 Table II, 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 relaxes the assumption that evolution is taught in the first year of high school. The main treatment coding assigns an individual the evolution score in force in the year of her high school entry. Although plausible, there might be cases where evolution is also taught in other grades, for example for individuals choosing an advanced placement course in biology. To account for such possibilities, I go to the other extreme and create a treatment variable that assumes equal exposure to teaching of evolution or creationism in all high school grades. For students who had a reform while being in high school, this approach effectively replaces the main evolution score by the weighted average of the pre- and post-reform evolution scores, with the weights corresponding to the number of pre- and post-reform high school years. As shown in column (4) of Table II, results are largely robust to this dosage treatment specification.

In another robustness check, I reduce 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 (5) of Table II, the results are largely robust and partly even more pronounced. In a similar vein, excluding small states who may adopt textbooks from larger states from the sample yields robust results, see column (6). As reported in column (7), the results are also robust to dropping states for which the original text of the evaluated Science Standards is not available for text analysis implemented in Online Appendix A.1.

In addition, the results hold if the observation period of the main sample is defined differently. As reported in columns (8) and (9) of Table II, 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 are largely robust to conducting a probit specification, see column (10) of Table II. Lastly, the results do not depend on the precise coding of the outcome variables, see columns (11) and (12) of Table II. 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.

Although the tested hypotheses about effects on evolution knowledge and evolution belief are inherently linked to evolution teaching through the common focus on evolution, I still conduct robustness checks on multiple hypothesis testing. First, I note that averaging over all non-evolution scientific questions, and over non-evolution scientific, political, and religious attitudes, respectively, alleviates concerns about multiple hypothesis testing by reducing the number of tested hypotheses, see Anderson (2008). Second, I implement the particularly conservative Bonferroni correction. The treatment effects on the main outcomes remain statistically significant when correcting for multiple hypothesis testing, see Online Appendix Tables A.XXX, and A.XXXI.

Lastly, the interpretation of the results does not change meaningfully when transforming the treatment variable to different indicator variables, i.e. to discrete reforms. Specifically, the first (second) indicator variable is set to one if the evolution score is larger than 0.1 (0.2), and to 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. Online Appendix Tables A.XXXII, A.XXXIII, and A.XXXIV 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 coding of the treatment, alleviating concerns about the related strong parallel trends assumption (Callaway, Goodman-Bacon, and Sant'Anna, 2021).

A.7.2 Robustness: Student Movement Between School Types

I assess reform effects on student movements between public and private schools (and homeschooling), as well as subgroup effects by school type. One can imagine that the inclusion of evolution in Science Standards changes the sample composition of students in different school types, as Science Standards are only binding for public schools.

To this end, I add the set of private school students covered by NAEP to the main sample of public school students. Specifically, I estimate equation 1 on a joint sample, using an indicator variable as outcome that equals one if the student attends a private school, and zero if the student attends a public school. As can be seen in column (1) of Online Appendix Table A.III, students are 6 percentage points more likely to attend a private school if exposed to an evolution score of one relative to zero. This amounts to approximately one third of the sample mean. However, it is unlikely to be a key mechanism or pose a threat to identification for the following three reasons. First, the effect is estimated imprecisely (p-value = 0.409) and should be interpreted with caution. Second, the effect is not meaningfully larger for females and students without a PC at home, see columns (2) and (3), respectively. These are the two subgroups for which I find the largest reform effects (that are statistically different from zero) on evolution knowledge (see Online Appendix Figure A.XIV). If school type selection was an important mechanism for the overall results pattern on evolution knowledge, one could expect to see differentially large effects on these groups. Third, the reform effect on evolution knowledge also holds in a sample of students from both public and private schools, see column (4). This coefficient is net of selection effects between public and private schools.

Furthermore, I decompose the overall effect into subgroup effects by school type, see columns (5) and (6). The difference between reform effects for public and private schools is not statistically significant. However, the point estimate for public school students is numerically larger than that for private school students, which is in line with the fact that the evaluated Science Standards have never been binding for private schools. This finding also mirrors the result by Goodman (2019) who documents that states' high school math requirements increase math coursework and later earnings by Blacks from public schools, while the corresponding private school effect is indistinguishable from zero.

I complement this main analysis on movements of students between school types of cohorts just exposed and just not exposed to the evolution reforms with a second analysis at the calendar year level. This second analysis demonstrates that overall private school enrollment, private school graduation, the numbers of teachers in private schools, the number of private schools as well as the number of students that are homeschooled did not significantly change in the years after the evolution curriculum reform. While the first analysis above uses individual-level data (and hence allows to run cohort-specific analyses), the second analysis uses state-level data (and hence allows to conduct calendar-year-specific analyses, analogous to the analyses of Online Appendix A.3). The second analysis is of interest if the mechanism in mind is that the evolution reforms work not through cohort-specific exposure to evolution theory, but through a general exposure to the topic to all cohorts in a given calendar year, for example through the media.

To conduct this second analysis, I set up a regression model using state-by-year level characteristics of private schools and homeschooling as outcome variables from the National Center for Education Statistics (NCES) and the International Center for Home Education Research (ICHER), respectively. The NCES data is collected for the preferred sample cut of the main analysis, i.e. for all available years after 1990 and before 2010, with the years 1993, 1995, 1997, 1999, 2001, 2003, 2005, 2007, and 2009 being available in this case. The ICHER data on homeschooling is available for a set of 30 states from 2000 until 2009 (with occasional missings).

I regress the different state-by-year characteristics of private schools on the evolution score as well as state and year fixed effects, as follows:

$$Y_{st} = \beta \cdot Evolution\_Score_{st} + \delta_s + \lambda_t + \epsilon_{st} \tag{6}$$

where Y_{st} is the outcome in state s and year t. The treatment variable $Evolution\_Score_{st}$ measures the intensity of the evolution coverage in the Science Standard in state s and year t. State fixed effects δ_s, year fixed effects λ_t, and an error term complete the model. The standard errors are clustered at the state level.

As shown in Online Appendix Table A.XXXV, all private school estimates are positive but insignificant and relatively small. This implies that private school enrollment, graduation as well as numbers of teachers and schools do not change in the years around the evolution curriculum reforms. For homeschooling, the point estimate is a bit larger, but also insignificant (and still only a relatively small fraction of the subsample mean). One can interpret these findings as indication that the timing of the evaluated reforms is largely independent of characteristics of different school types in the years around the reforms. This result complements the cohort-specific NAEP finding that the evaluated reforms have no relevant association with students' school choices.

While the presented analyses using data from NAEP, NCES, and ICHER can alleviate concerns about the sample composition of public and private school students and homeschoolers directly, the GSS and ACS analyses of the main paper address this topic indirectly. Both are sampled from the entire population including private school students and homeschoolers. Hence, the reform effects on evolution belief in adulthood and on the probability of working in life sciences presented in the result section of the main paper are net of selection of students into public schools, private schools, and homeschooling.

B TABLES AND FIGURES OF ONLINE APPENDIX

Table A.I
Correlation Coefficients

	Vaccination Rates	Trump Vote	Evolution Belief	Evolution Score (Curriculum)
Vaccination Rates	1			
Trump Vote	-0.785***	1		
Evolution Belief	0.576***	-0.517***	1	
Evolution Score (Curriculum)	0.398**	-0.466***	0.288*	1

Note: Sample: 50 U.S. states plus District of Columbia. Vaccination Rates defined as fraction of the population fully vaccinated against COVID-19 (not counting booster shots), averaged over following period: 8/23/21–11/17/21 (this equals "Period 3" in Barro (2022)). Data sources: CDC, NYT, Johns Hopkins Coronavirus Resource Center, U.S. Department of Health & Human Services, all provided by Opportunity Insights, Economic Tracker (Chetty, Friedman, Stepner, et al. (2023)). Trump vote: Republican share of the 2020 Presidential vote. Data source: Federal Election Commission. Evolution belief: Approval of the following statement, averaged across survey respondents within states across survey years of the main sample of the main paper (2006 – 2016): "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. Data source: General Social Survey. Evolution score (curriculum): Measure of coverage of evolution in State Science Education Standards in 2009. Data source: Mead and Mates (2009). Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.II
Regressions for Vaccination Rates, Inspired by Barro (2022)

	Outcome: Vaccination Rates					
	(1)	(2)	(3)	(4)	(5)	(6)
Trump Vote	-0.530***		-0.449***		-0.517***	-0.441***
	(0.115)		(0.114)		(0.127)	(0.124)
Evolution Belief		0.245***	0.099**			0.098**
		(0.056)	(0.040)			(0.041)
Evolution Score (Curriculum)				0.174***	0.018	0.012
				(0.041)	(0.039)	(0.037)
Adj. R-squared	0.609	0.318	0.642	0.141	0.602	0.635
Observations	51	51	51	51	51	51

Note: OLS estimation with heteroskedasticity robust standard errors in parentheses. Sample: 50 U.S. states plus District of Columbia. Dependent variable: Vaccination Rates defined as fraction of the population fully vaccinated against COVID-19 (not counting booster shots), averaged over following period: 8/23/21–11/17/21 (this equals "Period 3" in Barro (2022)). Data sources: CDC, NYT, Johns Hopkins Coronavirus Resource Center, U.S. Department of Health & Human Services, all provided by Opportunity Insights, Economic Tracker (Chetty, Friedman, Stepner, et al. (2023)). Explanatory variables: Trump vote: Republican share of the 2020 Presidential vote. Data source: Federal Election Commission. Evolution belief: Approval of the following statement, averaged across survey respondents within states across survey years of the main sample of the main paper (2006 – 2016): "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. Data source: General Social Survey. Evolution score (curriculum): Measure of coverage of evolution in State Science Education Standards in 2009. Data source: Mead and Mates (2009). Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table A.III
Effect of Evolution Coverage in Science Standards on Private School Attendance, and on Evolution Knowledge by School Type

	Attending Private School			Evolution Knowledge		
	(1) All Students	(2) Only Female	(3) Only Without PC At Home	(4) All Students	(5) Only Public School Students	(6) Only Private School Students
Evolution Score	0.064	0.066	-0.025	0.051**	0.065***	0.023
	(0.077)	(0.076)	(0.119)	(0.019)	(0.022)	(0.066)
State FEs	YES	YES	YES	YES	YES	YES
Cohort FEs	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.18	0.17	0.10	0.34	0.32	0.43
Std. Dev. of Dep. Var.	0.38	0.38	0.30	0.42	0.42	0.38
Adj. R-squared	0.336	0.353	0.201	0.049	0.041	0.043
Observations	17,150	8,790	1,770	17,150	14,080	3,070

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. Dependent variable: Columns (1) - (3): Indicator variable for private school attendance, for all students and selected subgroups as indicated in the column headers; columns (4) - (6): Share of questions about evolution answered correctly, for all students and selected subgroups as indicated in the column headers. Controls: Indicator variables for gender, races/ethnicities, home possessions (separate indicator variables for computer and books), birth month, test session, and test year. Subsidized lunch status not in set of controls variables, as missing for most private school students. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 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.IV
CS Baseline Panel Estimator

	States that reduce evolution coverage		States that expand evolution coverage	
	All Individuals	Excluding individuals with 20% highest and 20% lowest evolution coverage in Science Standards	All Individuals	Excluding individuals with 20% highest and 20% lowest evolution coverage in Science Standards
	(1)	(2)	(3)	(4)
	Panel A: Outcome: Evolution Knowledge in School			
Evolution Reform	0.056***	0.082***	0.028	0.061***
	(0.017)	(0.031)	(0.022)	(0.020)
	Panel B: Outcome: Evolution Belief in Adulthood			
Evolution Reform	0.274***	0.384***	0.198	0.434***
	(0.083)	(0.121)	(0.127)	(0.168)
	Panel C: Outcome: Working in Life Sciences			
Evolution Reform	0.036**	0.055***	0.007	0.005
	(0.016)	(0.021)	(0.011)	(0.014)

Note: Table shows CS estimator (Callaway and Sant'Anna, 2021), accounting for heterogeneous treatment effects and staggered treatment timing. Simple aggregation of all post treatment effects (ATT), using doubly robust inverse probability weighting and not-yet-treated observations as controls. Each entry is from a separate regression model. Sample of states and individuals indicated in column headers and subheaders, respectively. Dependent variables as indicated in panel header, times minus 1 for states that reduce evolution coverage (inverted outcome due to inverted reform, to allow for comparability across results). Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 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.V
Effect of Evolution Coverage in Science Standards on Selection into Biology Courses

	(1) First Year Biology	(2) Second Year Biology	(3) Advanced Placement Biology
Evolution Score	0.028	0.007	0.012
	(0.025)	(0.041)	(0.020)
State FEs	YES	YES	YES
Cohort FEs	YES	YES	YES
Controls	YES	YES	YES
Mean of Dep. Var.	0.94	0.26	0.11
Std. Dev. of Dep. Var.	0.25	0.44	0.31
Adj. R-squared	0.041	0.029	0.021
Observations	32,650	31,700	26,520

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. Dependent variables: Indicator variables indicating whether student ever took high school course indicated in column header. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, home possessions (separate indicator variables for computer and books), birth month, test session, and test year fixed effects. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 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.VI
Effect of Evolution Coverage in Science Standards on Knowledge about Non-Evolution Scientific Topics

	Main Outcome:	Non-Evolution Scientific Topics:									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Evolution	Average	Motion	Matter and Mass	Energy	Reproduction	Climate Change	Pollution	Earth	Tectonics	Universe
Evolution Score	0.065***	0.016	0.010	0.024	-0.011	0.024	-0.008	-0.010	0.049*	0.020	0.003
	(0.022)	(0.011)	(0.035)	(0.042)	(0.023)	(0.031)	(0.042)	(0.061)	(0.026)	(0.023)	(0.030)
Mean of Dep. Var.	0.32	0.36	0.51	0.30	0.37	0.38	0.39	0.16	0.42	0.18	0.32
Std. Dev. of Dep. Var.	0.42	0.26	0.43	0.43	0.43	0.41	0.40	0.28	0.41	0.27	0.43
Adj. R-squared	0.041	0.159	0.086	0.126	0.108	0.088	0.066	0.068	0.122	0.310	0.047
Observations	14,080	28,980	9,190	14,930	20,380	16,760	17,450	4,270	12,580	6,130	7,260

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. Dependent variables: Shares of questions answered correctly about scientific topics indicated in the column headers (on average in column (2), and individually in columns (2) -(11)). Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, home possessions (separate indicator variables for computer and books), birth month, test session, and fixed effects for state, cohort, and test year. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 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.VII

Effect of Evolution Coverage in Science Standards on Non-Evolution Scientific Topics

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

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. 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 fixed effects for state, cohort, and survey year. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: General Social Survey.

Table A.VIII
Effect of Evolution Coverage in Science Standards on Science Attitudes

	Main Outcome:	Non-Evolution Science Attitudes:			
	(1)	(2)	(3)	(4)	(5)
	Evolution	Average	Scientific Research is Necessary	Confidence in Science	Interest in Science
Evolution Score	0.333***	-0.072	-0.067	0.045	-0.080
	(0.107)	(0.070)	(0.127)	(0.094)	(0.080)
Mean of Dep. Var.	0.58	0.87	0.86	0.90	0.86
Std. Dev. of Dep. Var.	0.49	0.25	0.35	0.30	0.35
Adj. R-squared	0.107	0.043	0.027	0.015	0.013
Observations	1,801	1,801	1,739	1,337	1,433

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. Dependent variables: Approval of the science attitude indicated in the column header. All science attitude outcomes are coded such that an increase in the variable implies an increase in pro-science attitude. Average of pro-science attitudes, presented in column (2) interpret as proxy for science prestige. 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, cohort, and survey year. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: General Social Survey.

Table A.IX
Effect of Evolution Coverage in Science Standards on Religious Outcomes

	Main Outcome:	Religious Outcomes: Average	Religious Outcomes: Believing				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Evolution	Average	God	Bible	Afterlife	Rebirth	Strong Believer
Panel A Outcomes: Main; Religious Average; Religious Believing							
Evolution Score	0.333***	-0.029	-0.021	-0.009	-0.023	-0.097	-0.091
	(0.107)	(0.068)	(0.096)	(0.114)	(0.128)	(0.123)	(0.150)
Mean of Dep. Var.	0.58	0.50	0.87	0.72	0.73	0.34	0.32
Std. Dev. of Dep. Var.	0.49	0.28	0.33	0.45	0.45	0.47	0.47
Adj. R-squared	0.107	0.201	0.104	0.086	0.029	0.134	0.086
Observations	1,801	1,801	1,797	1,794	1,797	1,796	1,783

	Religious Outcomes: Belonging and Activities					Religious Outcomes: Overall		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Religious Affiliation	Church-going	Church Activities	Personal Prayer	Missionize	Spiritual Person	Religious Person	Fundamentalist
Panel B Outcomes: Religious Belonging; Religious Overall								
Evolution Score	0.099	0.102	-0.072	-0.108	-0.028	0.060	-0.206	0.015
	(0.109)	(0.120)	(0.091)	(0.113)	(0.073)	(0.101)	(0.162)	(0.129)
Mean of Dep. Var.	0.70	0.35	0.17	0.65	0.41	0.56	0.43	0.24
Std. Dev. of Dep. Var.	0.46	0.48	0.38	0.48	0.49	0.50	0.50	0.43
Adj. R-squared	0.163	0.081	0.040	0.149	0.130	0.061	0.077	0.244
Observations	1,799	1,801	1,801	1,798	1,801	1,801	1,799	1,718

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. Dependent variables: Main outcome and religious outcomes indicated in the column headers. All religious outcomes are coded such that an increase in the variable implies an increase in religiosity. 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, cohort, and survey year. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: General Social Survey.

Table A.X
Effect of Evolution Coverage in Science Standards on Political Outcomes

	Main Outcome:	Political Outcomes: Average	Political Outcomes: Conservative attitude on:							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Evolution	Average	Prayer in Public Schools	Sex Education in Public Schools	Same-Sex Marriage	Abortion	Marijuana Legalization	Capital Punishment	Gun Control	Immigration
Panel A Outcomes: Main; Political Average; Political Attitudes										
Evolution Score	0.333***	-0.002	-0.061	-0.085	-0.084	0.257	0.006	-0.101	0.013	0.082
	(0.107)	(0.053)	(0.104)	(0.065)	(0.225)	(0.188)	(0.153)	(0.162)	(0.135)	(0.141)
Mean of Dep. Var.	0.58	0.43	0.45	0.06	0.36	0.55	0.45	0.61	0.30	0.84
Std. Dev. of Dep. Var.	0.49	0.19	0.50	0.25	0.48	0.50	0.50	0.49	0.46	0.36
Adj. R-squared	0.107	0.072	0.087	0.048	0.110	0.028	0.079	0.063	0.060	-0.001
Observations	1,801	1,801	1,200	1,200	1,056	1,059	1,336	1,788	1,063	1,174

	Political Outcomes: Conservative attitude on governmental spending for:							Political Outcomes: General	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Environment	Alternative Energy Sources	Education	Scientific Research	Reducing Income Differences	Assistance to the Poor	Conditions of Blacks	Republican	Conservative
Panel B Outcomes: Political Spending; Political General									
Evolution Score	0.062	-0.045	0.047	0.116	-0.137	-0.148	0.044	0.028	-0.018
	(0.105)	(0.146)	(0.104)	(0.138)	(0.100)	(0.101)	(0.116)	(0.114)	(0.132)
Mean of Dep. Var.	0.29	0.41	0.19	0.63	0.46	0.51	0.64	0.28	0.26
Std. Dev. of Dep. Var.	0.45	0.49	0.39	0.48	0.50	0.50	0.48	0.45	0.44
Adj. R-squared	0.018	0.011	0.027	0.012	0.001	0.060	0.095	0.058	0.024
Observations	1,801	1,195	1,801	1,799	1,337	1,798	1,788	1,792	1,791

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. 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 Online Appendix A.4.2 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 fixed effects for state, cohort, and survey year. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: General Social Survey.

Table A.XI
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
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
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
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 (2000), and the difference of the evolution scores (evolution score from 2009 minus evolution score from 2000). States are listed in three panels: Negative, positive, 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 the author's examination of state education websites.

Table A.XII
Effect of Evolution Coverage in Science Standards on Evolution Knowledge: Further Robustness Checks

	(1) Close Elections	(2) Control: Governor's Party	(3) State Specific Time Trends	(4) Dosage Treatment	(5) Only One Reform Event	(6) Only Large States	(7) Only States With Std. Text	(8) Sample Start: 1995	(9) Sample Start: 2000	(10) Probit	(11) Outcome Coding: Indicator Variation
Evolution Score	0.093***	0.066***	0.043	0.042	0.089***	0.056**	0.078***	0.048***	0.037	0.076**	0.058**
	(0.022)	(0.022)	(0.065)	(0.034)	(0.029)	(0.022)	(0.025)	(0.017)	(0.026)	(0.032)	(0.022)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Cohort FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Equal. Test w/ Base. Coef. (P-val.)	0.265	0.798	0.533	0.220	0.314	0.621	0.340	0.052	0.014	0.494	0.153
Mean of Dep. Var.	0.33	0.32	0.32	0.32	0.32	0.32	0.33	0.32	0.31	0.32	0.35
Std. Dev. of Dep. Var.	0.43	0.42	0.42	0.42	0.42	0.42	0.43	0.43	0.44	0.42	0.43
Adj. R-squared	0.040	0.041	0.029	0.041	0.042	0.043	0.043	0.037	0.032	0.076	0.035
Observations	4,700	14,080	15,530	14,080	6,480	10,560	8,860	12,950	10,970	14,070	13,140

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1, for different 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) Main evolution score is replaced by weighted average of pre- and post-reform evolution scores, with weights corresponding to number of pre- and post-reform high school years; (5) Sample only includes individuals from states that had only one reform event between 2000 and 2009, see Table A.XI for more details; (6) Sample only includes individuals from 20 states with the largest population; (7) Sample only includes individuals from states for which text of Science Standards is available for text analysis implemented in Online Appendix A.1; (8) Sample only includes individuals who started high school after 1994; (9) Sample only includes individuals who started high school after 1999; (10) Coefficient reports average marginal treatment effect of probit specification; (11) 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). Dependent variable: Share of questions about evolution answered correctly. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, home possessions (separate indicator variables for computer and books), birth month, test session, and test year (specification with state-specific time trends only controls for gender, race/ethnicity, and test year fixed effects due to power constraints). Standard errors clustered at the state level in parenthesis. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. "Equal. Test w/ Base. Coef. (P-val.)" refers to p-value of test for equality between coefficient reported in given column and corresponding baseline coefficient presented in Table I, Panel A, Column 4. 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.XIII
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)	(12)
	Close Elections	Control: Governor's Party	State Specific Time Trends	Dosage Treatment	Only One Reform Event	Only Large States	Only States With Std. Text	Sample Start: 1995	Sample Start: 2000	Probit	Outcome Coding Variation 1	Outcome Coding Variation 2
Evolution Score	0.605***	0.332***	0.625***	0.177**	0.394**	0.433***	0.322**	0.257**	0.313*	0.329**	0.288**	0.426***
	(0.188)	(0.111)	(0.218)	(0.070)	(0.163)	(0.117)	(0.131)	(0.116)	(0.171)	(0.130)	(0.138)	(0.145)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Cohort FE	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
Equal. Test w/ Base. Coef. (P-val.)	0.072	0.954	0.042	0.461	0.582	0.158	0.893	0.029	0.852	0.792	0.323	0.107
Mean of Dep. Var.	0.58	0.58	0.58	0.58	0.57	0.59	0.60	0.59	0.61	0.58	0.66	0.55
Std. Dev. of Dep. Var.	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.49	0.47	0.50
Adj. R-squared	0.102	0.107	0.097	0.105	0.115	0.100	0.119	0.092	0.077	0.117	0.127	0.102
Observations	589	1,801	1,801	1,801	709	1,359	1,213	1,299	654	1,780	1,571	1,617

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1, for different 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) Main evolution score is replaced by weighted average of pre- and post-reform evolution scores, with weights corresponding to number of pre- and post-reform high school years; (5) Sample only includes individuals from states that had only one reform event between 2000 and 2009, see Table A.XI for more details; (6) Sample only includes individuals from 20 states with the largest population; (7) Sample only includes individuals from states for which text of Science Standards is available for text analysis implemented in Online Appendix A.1; (8) Sample only includes individuals who started high school after 1994; (9) Sample only includes individuals who started high school after 1999; (10) Coefficient reports average marginal treatment effect of probit specification; (11) Sample excludes individuals whose dependent variable question on evolution replaces the words "human beings" with the word "elephants"; (12) 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). 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 10%, 5%, and 1% levels, respectively. "Equal. Test w/ Base. Coef. (P-val.)" refers to p-value of test for equality between coefficient reported in given column and corresponding baseline coefficient presented in Table I, Panel B, Column 4. Data source: General Social Survey.

Table A.XIV

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)	(9)	(10)
	Close Elections	Control: Governor's Party	State Specific Time Trends	Dosage Treatment	Only One Reform Event	Only Large States	Only States With Std. Text	Sample Start: 1995	Sample Start: 2000	Probit
Evolution Score	0.039	0.036**	0.025	0.034**	0.031	0.040**	0.057***	0.036***	0.029**	0.035
	(0.025)	(0.014)	(0.025)	(0.016)	(0.021)	(0.017)	(0.016)	(0.013)	(0.012)	(0.026)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Cohort FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Equal. Test w/ Base. Coef. (P-val.)	0.848	0.731	0.439	0.963	0.724	0.663	0.080	0.982	0.385	0.984
Mean of Dep. Var.	0.14	0.15	0.15	0.15	0.14	0.14	0.15	0.13	0.10	0.15
Std. Dev. of Dep. Var.	3.80	3.84	3.84	3.84	3.77	3.76	3.91	3.58	3.10	3.84
Adj. R-squared	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.032
Observations	2,022,927	6,460,650	6,460,650	6,460,650	2,522,283	4,828,471	4,139,919	4,821,487	2,762,694	6,460,650

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1, for different 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) Main evolution score is replaced by weighted average of pre- and post-reform evolution scores, with weights corresponding to number of pre- and post-reform high school years; (5) Sample only includes individuals from states that had only one reform event between 2000 and 2009, see Table A.XI for more details; (6) Sample only includes individuals from 20 states with the largest population; (7) Sample only includes individuals from states for which text of Science Standards is available for text analysis implemented in Online Appendix A.1; (8) Sample only includes individuals who started high school after 1994; (9) Sample only includes individuals who started high school after 1999; (10) 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 10%, 5%, and 1% levels, respectively. "Equal. Test w/ Base. Coef. (P-val.)" refers to p-value of test for equality between coefficient reported in given column and corresponding baseline coefficient presented in Table I, Panel C, Column 4. Data source: American Community Survey.

Table A.XV
Text Analysis: Mean Comparison (Sample Representativity)

	Estimation Sample		Population	
	Std. avail.	Std. n/a	Std. avail.	Std. n/a
Main Outcome:				
Evolution Belief	0.60	0.54	0.52	0.46
Non-Evolution Scientific Outcomes:				
Earth	0.90	0.85	0.84	0.80
Radioactivity	0.66	0.64	0.73	0.68
Reproduction	0.62	0.63	0.62	0.60
Lasers	0.49	0.43	0.49	0.43
Electrons	0.59	0.53	0.53	0.49
Antibiotics	0.52	0.48	0.55	0.50
Universe	0.47	0.41	0.41	0.38
Tectonics	0.86	0.84	0.81	0.79
Sun	0.80	0.76	0.74	0.70
Non-Evolution Scientific Topics: Average	0.66	0.62	0.63	0.60
Religious Outcomes:				
God	0.87	0.89	0.91	0.92
Bible	0.70	0.77	0.76	0.82
Afterlife	0.72	0.74	0.73	0.76
Rebirth	0.31	0.40	0.36	0.47
Strong Believer	0.30	0.36	0.42	0.46
Religious Affiliation	0.68	0.75	0.80	0.83
Church-going	0.33	0.41	0.42	0.48
Church Activities	0.16	0.19	0.24	0.26
Personal Prayer	0.63	0.69	0.73	0.79
Missionize	0.39	0.47	0.42	0.51
Spiritual Person	0.56	0.57	0.67	0.70
Religious Person	0.42	0.45	0.55	0.62
Fundamentalist	0.21	0.29	0.26	0.34
Religious Outcomes: Average	0.48	0.54	0.56	0.61
Political Outcomes:				
Republican	0.27	0.29	0.35	0.34
Conservative	0.25	0.29	0.33	0.36
Prayer in Public Schools	0.42	0.52	0.54	0.62
Sex Education in Public Schools	0.07	0.06	0.10	0.11
Same-Sex Marriage	0.35	0.39	0.50	0.58
Abortion	0.52	0.60	0.53	0.63
Marijuana Legalization	0.44	0.46	0.54	0.57
Capital Punishment	0.60	0.62	0.64	0.63
Gun Control	0.29	0.32	0.26	0.30
Immigration	0.85	0.83	0.88	0.87
Environment	0.28	0.30	0.36	0.39
Alternative Energy Sources	0.41	0.42	0.40	0.44
Education	0.19	0.20	0.25	0.26
Scientific Research	0.63	0.64	0.61	0.61
Reducing Income Differences	0.46	0.48	0.53	0.54
Assistance to the Poor	0.52	0.51	0.54	0.53
Conditions of Blacks	0.64	0.65	0.69	0.68
Political Outcomes: Average	0.42	0.44	0.48	0.49

Note: Means of scientific, religious, and political outcome variables, presented by population and estimation sample, and by availability of text of Science Standard for text analysis. Estimation sample refers to main estimation sample (with cohorts entering high school between 1990 and 2010); population refers to sample without cohort restrictions. All political outcomes are coded such that an increase in the variable implies an increase in political conservatism, see Online Appendix A.4.2 for details. Data source: General Social Survey.

Table A.XVI
Text Analysis: Effect of Evolution Coverage in Science Standards (as Measured by the Evolution Score) on Counts of Different Scientific Words in Science Standards

	Word Count of:	Counts of the following Non-Evolution Scientific Words:										Counts of all Words:
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Evolution	Motion	Matter and Mass	Energy	Reproduction	Climate Change	Pollution	Earth	Tectonics	Universe	Sum	Sum
Evolution Score	1.986***	0.505	1.137	0.725	0.494	0.933	0.357	1.186	1.195	1.202	0.798	0.300
	(0.736)	(1.061)	(1.031)	(1.022)	(0.665)	(0.745)	(1.863)	(1.566)	(1.294)	(1.801)	(1.078)	(1.431)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	2.38	3.09	3.64	4.02	2.29	0.33	0.93	3.99	1.07	2.62	5.31	9.50
Std. Dev. of Dep. Var.	0.99	0.88	0.90	0.93	0.81	0.52	0.91	0.97	0.76	0.91	0.90	0.93
Adj. R-squared	0.359	0.387	0.403	0.321	0.362	0.278	0.146	0.474	0.157	0.307	0.400	0.517
Observations	73	73	73	73	73	73	73	73	73	73	73	73

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 3. Dependent variables: Count of words in Science Standards (logarithmized), as indicated in the column headers. Sum in column 11 refers to (logarithmized) sum of all non-evolution scienfic words. Sum in column 12 refers to logarithmized sum of all words (total word count). Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: Author's collection of US State Science Education Standards (73 Standards available in total: 27 Standards in power in 2000 as referenced in Lerner (2000), 46 Standards in power in 2009 as referenced in Mead and Mates (2009)).

Table A.XVII
Effect of Evolution Coverage in Science Standards on Different State-by-Year Characteristics

	(1) Unemployment Rate	(2) Share of Democrat Seats (State House)	(3) Share of Democrat Seats (State Senate)	(4) School Expenditure
Evolution Score	-0.000	-0.038	-0.043	0.022
	(0.005)	(0.050)	(0.054)	(0.029)
State FEs	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES
Mean of Dep. Var.	0.05	0.54	0.53	7.17
Std. Dev. of Dep. Var.	0.02	0.17	0.17	0.30
Adj. R-squared	0.770	0.874	0.826	0.950
Observations	953	934	915	953

Note: Table shows TWFE OLS coefficients with state and year fixed effects, and standard errors clustered at the state level in parenthesis. Dependent variables indicated in the column headers as follows: (1) State annual average of number of unemployed as a percentage of the labor force; (2) Number of Democrat seats over the sum of Democrat seats and Republican seats in State House of Representatives; (3) Number of Democrat seats over the sum of Democrat seats and Republican seats in State Senate; (4) Direct expenditures per capita of state and local governments for elementary and secondary education (logarithmized). Nebraska dropped from sample in Columns (2) and (3) as it does not officially recognize a representative's party affiliation. Washington DC dropped from Column (3) as it has unicameral system. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data sources: Column (1): Bureau of Economic Analysis; Columns (2) and (3): Websites of United States House of Representatives/Senate, State Elections Offices, wikipedia.org; Column (4): National Center for Education Statistics.

Table A.XVIII
Effect of Evolution Coverage in Science Standards on Different Keyword Search Frequencies on Google

	Google Searches of the Following Words:							Average:	
	(1) Evolution	(2) Evolution Theory	(3) Evolutionary Process	(4) Darwin	(5) Natural Selection	(6) Creationism	(7) Intelligent Design	(8) Sign Adjusted	(9) No Sign Adjusted
Evolution Score	0.002	-0.020	0.000	-0.002	0.006	-0.010	-0.000	-0.001	-0.003
	(0.021)	(0.027)	(0.028)	(0.011)	(0.008)	(0.010)	(0.008)	(0.008)	(0.007)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean of Dep. Var.	0.31	0.10	0.02	0.11	0.03	0.03	0.03	0.07	0.09
Std. Dev. of Dep. Var.	0.07	0.09	0.04	0.04	0.02	0.03	0.03	0.02	0.03
Adj. R-squared	0.737	0.674	0.348	0.632	0.316	0.376	0.518	0.749	0.790
Observations	303	303	303	303	303	303	303	303	303

Note: Table shows TWFE OLS coefficients with state and year fixed effects, and standard errors clustered at the state level in parenthesis. Dependent variables: Columns (1) - (7): State annual average frequency of Google search of keyword indicated in the column header; Column (8): Average of state annual average frequencies of Google search of keywords from Column (1) to (7), where keywords "Creationism" and "Intelligent Design" enter negatively (such that an increase in each component of the average implies an increase in evolution search interest or a decrease in creationism search interest); Column (9): Average of state annual average frequencies of Google search of keywords from Column (1) to (7), where keywords "Creationism" and "Intelligent Design" enter positively (such that an increase in each component of the average implies an increase in evolution search interest or an increase in creationism search interest). All dependent variables are normalized to range from 0 to 1, where 0 indicates no search interest in given keyword (or the absence of sufficient data), and 1 indicates highest search interest in that keyword in the history of Google Trends for a given state. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Sampling period: 2004-2009. Data source: Google trends.

Table A.XIX
Correlation Coefficients of Knowledge about Evolution and Other Scientific Areas

	Evolution Knowledge
Non-Evolution Scientific Topics: Average	0.257***
Motion	0.0894***
Matter and Mass	0.0836***
Energy	0.129***
Reproduction	0.283***
Climate Change	0.0524***
Pollution	0.150***
Earth	0.0924***
Tectonics	0.0183
Universe	0.152***

Note: Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 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.XX
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
Non-Evolution Scientific Outcomes:				
Non-Evolution Scientific Topics: Average	0.35	0.27	0.00	1.00
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 Change	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	0.31	0.43	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
Subsidized Lunch	0.30	0.46	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
Computer at Home	0.84	0.37	0.00	1.00
Test Session	1.75	0.97	1.00	8.00
Birth Month	6.60	3.40	1.00	12.00

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.XXI
Correlation Coefficients of Belief in Evolution and Other Scientific Areas

	Evolution Belief
Non-Evolution Scientific Topics: Average	0.316***
Earth	0.119***
Radioactivity	0.146***
Reproduction	-0.0220
Lasers	0.109***
Electrons	0.168***
Antibiotics	0.111***
Universe	0.414***
Tectonics	0.248***
Sun	0.109***

Note: Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: General Social Survey.

Table A.XXII
Correlation Coefficients of Belief in Evolution and Other Science Attitudes

	Evolution Belief
Science Attitudes: Average	0.152***
Science is Necessary	0.123***
Confidence in Science	0.116***
Interest in Science	0.0931***

Note: Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: General Social Survey.

Table A.XXIII
Correlation Coefficients of Evolution Belief and Religious Outcomes

	Evolution Belief
Religious Outcomes: Average	-0.376***
God	-0.193***
Bible	-0.276***
Afterlife	-0.108***
Rebirth	-0.309***
Strong Believer	-0.289***
Religious Affiliation	-0.212***
Church-going	-0.273***
Church Activities	-0.207***
Personal Prayer	-0.285***
Missionize	-0.277***
Spiritual Person	-0.156***
Religious Person	-0.244***
Fundamentalist	-0.245***

Note: Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: General Social Survey.

Table A.XXIV
Correlation Coefficients of Evolution Belief and Political Outcomes

	Evolution Belief
Political Outcomes: Average	-0.248***
Prayer in Public Schools	-0.238***
Sex Education in Public Schools	-0.195***
Same-Sex Marriage	-0.283***
Abortion	-0.237***
Marijuana Legalization	-0.128***
Capital Punishment	-0.0132
Gun Control	-0.0249
Immigration	-0.00674
Environment	-0.0917***
Alternative Energy Sources	-0.0888***
Education	-0.0549**
Scientific Research	-0.162***
Reducing Income Differences	-0.0878***
Assistance to the Poor	0.00132
Conditions of Blacks	-0.0600**
Republican	-0.119***
Conservative	-0.128***

Note: All political outcomes are coded such that an increase in the variable implies an increase in political conservatism, see Online Appendix A.4.2 for details. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: General Social Survey.

Table A.XXV
Descriptive Statistics of GSS Data

	Mean	Std. Dev.	Min.	Max.
Treatment Variable:				
Evolution Score	0.64	0.32	0.00	1.00
Main Outcome:				
Evolution Belief	0.58	0.49	0.00	1.00
Non-Evolution Scientific Topics:				
Non-Evolution Scientific Topics: Average	0.64	0.22	0.00	1.00
Earth	0.88	0.32	0.00	1.00
Radioactivity	0.66	0.47	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	0.49	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 Science Attitudes:				
Science Attitudes: Average	0.87	0.25	0.00	1.00
Science is Necessary	0.86	0.35	0.00	1.00
Confidence in Science	0.90	0.30	0.00	1.00
Interest in Science	0.86	0.35	0.00	1.00
Religious Attitudes:				
Religious Outcomes: Average	0.50	0.28	0.00	1.00
God	0.88	0.33	0.00	1.00
Bible	0.72	0.45	0.00	1.00
Afterlife	0.73	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	0.36	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	0.50	0.00	1.00
Fundamentalist	0.24	0.43	0.00	1.00
Political Attitudes:				
Political Outcomes: Average	0.43	0.19	0.00	1.00
Prayer in Public Schools	0.46	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
Environment	0.29	0.45	0.00	1.00

- Continued -

	Mean	Std. Dev.	Min.	Max.
Political Attitudes (continued):				
Alternative Energy Sources	0.42	0.49	0.00	1.00
Education	0.19	0.39	0.00	1.00
Scientific Research	0.63	0.48	0.00	1.00
Reducing Income Differences	0.46	0.50	0.00	1.00
Assistance to the Poor	0.51	0.50	0.00	1.00
Conditions of Blacks	0.65	0.48	0.00	1.00
Republican	0.28	0.45	0.00	1.00
Conservative	0.26	0.44	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	0.18	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	0.24	0.43	0.00	1.00
Growing up: Other	0.05	0.21	0.00	1.00
Raised as Protestant: Mainline	0.38	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	0.13	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	0.19	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.XXVI
Descriptive Statistics of ACS Data

	Mean	Std. Dev.	Min.	Max.
Treatment Variable:				
Evolution Score	0.67	0.30	0.00	1.00
Main Outcome - Working in life sciences:				
Life Sciences	0.15	3.84	0.00	100.00
Additional Outcomes - Working in subfields of life sciences:				
Biology	0.06	2.35	0.00	100.00
Agriculture and Food	0.02	1.39	0.00	100.00
Conservation and Forestry	0.02	1.27	0.00	100.00
Medical and Other	0.06	2.38	0.00	100.00
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.XXVII
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: Table shows multivariate OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 5. 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. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: National Survey for High School Biology Teachers, 2007.

Table A.XXVIII
Effect of Evolution Coverage in Science Standards on Probability of Working in Life Sciences, by Subgroups

	By Gender		By Race/Ethnicity		
	(1)	(2)	(3)	(4)	(5)
	Females	Males	Whites	Blacks	Hispanics
Evolution Score	0.052**	0.018	0.038**	0.012	0.004
	(0.020)	(0.020)	(0.016)	(0.016)	(0.034)
State FEs	YES	YES	YES	YES	YES
Cohort 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: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1, by selected subgroups, as indicated in the column 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 10%, 5%, and 1% levels, respectively. Data source: American Community Survey.

Table A.XXIX
Descriptive Statistics: Selected Subsamples

	Black	White	Rural	Urban	Mainline Protestant	Evangelical	Catholic	Non-Religious
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Growing up: Single Parent	0.44	0.20	0.21	0.28	0.25	0.20	0.21	0.32
Growing up: One Parent, one Stepparent	0.13	0.13	0.12	0.12	0.12	0.19	0.11	0.11
Growing up: Both Parents	0.30	0.60	0.58	0.52	0.53	0.53	0.62	0.46
Growing up: Other	0.10	0.03	0.06	0.04	0.06	0.02	0.04	0.05
Parents born in US	0.87	0.87	0.86	0.76	0.92	0.87	0.66	0.86
Parents born abroad	0.12	0.13	0.13	0.24	0.07	0.13	0.34	0.14
Race/Ethnicity: Black	1.00	0.00	0.12	0.24	0.28	0.19	0.05	0.17
Race/Ethnicity: White	0.00	1.00	0.77	0.62	0.68	0.71	0.74	0.71
Race/Ethnicity: Other	0.00	0.00	0.10	0.13	0.03	0.10	0.21	0.12
Race/Ethnicity: Hispanic	0.04	0.11	0.11	0.21	0.03	0.13	0.36	0.10
Raised in Rural Area	0.33	0.55	1.00	0.00	0.53	0.52	0.45	0.55
Raised as Protestant: Mainline	0.58	0.37	0.40	0.35	1.00	0.00	0.00	0.00
Raised as Protestant: Evangelical	0.10	0.10	0.10	0.09	0.00	1.00	0.00	0.00
Raised as Catholic	0.10	0.34	0.29	0.35	0.00	0.00	1.00	0.00
Raised as Non-Religious	0.13	0.14	0.15	0.12	0.00	0.00	0.00	1.00

Note: Table shows mean of selected characteristics for different subsamples indicated in the column headers. Numbers do not have to add up to one for categorical dummies of given variable due to rounding and not all categories being shown. Data source: General Social Survey.

Table A.XXX
Effect of Evolution Coverage in Science Standards on Evolution Knowledge and Average of Non-Evolution Scientific Knowledge: Correction for Multiple Hypothesis Testing

	Outcomes:	
	Evolution knowledge	Non-evolution scientific knowledge (average)
	(1)	(2)
Effects	0.065	0.016
Standard p-values	0.005	0.156
Bonferroni-adjusted p-values	0.009	0.312

Note: Table shows TWFE OLS coefficients and p-values robust to multiple hypothesis testing of effects of evolution coverage in Science Standards on outcomes indicated in the column headers, from estimating equation 1. P-values: Standard p-values based on clustering at the state level; Bonferroni-adjusted p-values (standard p-value multiplied by number of tested hypothesis; capped at 1). Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, home possessions (separate indicator variables for computer and books), birth month, test session, as well as state, cohort, 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.XXXI
Effect of Evolution Coverage in Science Standards on Evolution Belief in Adulthood and 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)	Non-evolution science attitudes (average)	Religious outcomes (average)	Political outcomes (average)
	(1)	(2)	(3)	(4)	(5)
Effects	0.333	-0.053	-0.072	-0.029	-0.002
Standard p-values	0.003	0.354	0.306	0.668	0.971
Bonferroni-adjusted p-values	0.016	1.000	1.000	1.000	1.000

Note: Table shows TWFE OLS coefficients and p-values robust to multiple hypothesis testing of effects of evolution coverage in Science Standards on outcomes indicated in the column headers, from estimating equation 1. P-values: Standard p-values based on clustering at the state level; 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, cohort, and survey year fixed effects. Data source: General Social Survey.

Table A.XXXII
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	0.033								
	(0.020)								
Evolution Score > 0.80		0.024							
		(0.015)							
Evolution Score > 0.70			0.029*						
			(0.016)						
Evolution Score > 0.60				0.027**					
				(0.011)					
Evolution Score > 0.50					0.026***				
					(0.009)				
Evolution Score > 0.40						0.018			
						(0.011)			
Evolution Score > 0.30							0.023*		
							(0.012)		
Evolution Score > 0.20								0.029**	
								(0.012)	
Evolution Score > 0.10									0.029**
									(0.012)
State FEs	YES	YES	YES	YES	YES	YES	YES	YES	YES
Cohort 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	0.041	0.041	0.041	0.041	0.041	0.041	0.041	0.041	0.041
Observations	14,080	14,080	14,080	14,080	14,080	14,080	14,080	14,080	14,080

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. 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, home possessions (separate indicator variables for computer and books), birth month, test session, and test year fixed effects. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 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.XXXIII
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
Cohort 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: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. 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). 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, 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. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: General Social Survey.

Table A.XXXIV
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
Cohort 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

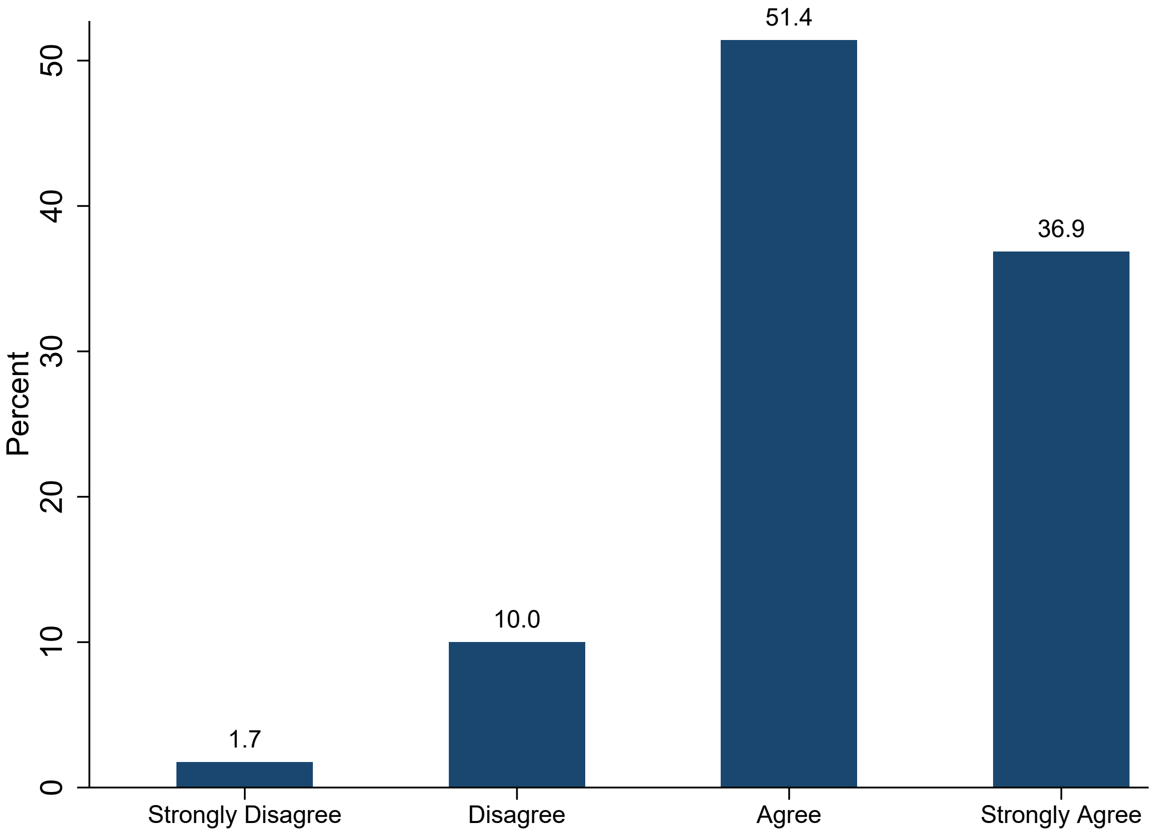
Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 1. Dependent variable: Probability of working in life sciences (multiplied by 100 for interpretability). Explanatory variables: Evolution score indicator variables (equals one if evolution score is larger than indicated level, and zero otherwise). Controls: Indicator variables for gender, races/ethnicities, and survey year fixed effects. Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data source: American Community Survey.

Table A.XXXV
Effect of Evolution Coverage in Science Standards on Different State-by-Year Characteristics of Private Schools and Homeschooling

	Private Schools:				Homeschooling:
	(1) Students	(2) Graduates	(3) Teachers	(4) Schools	(5) Students
Evolution Score	0.034	0.027	0.036	0.038	0.102
	(0.056)	(0.064)	(0.048)	(0.098)	(0.095)
State FEs	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES
Mean of Dep. Var.	10.93	7.98	8.39	5.87	9.07
Std. Dev. of Dep. Var.	1.23	1.27	1.18	1.06	1.09
Adj. R-squared	0.990	0.976	0.991	0.979	0.991
Observations	452	449	452	452	244

Note: Table shows TWFE OLS coefficients and standard errors clustered at the state level in parenthesis from estimating equation 6. Dependent variables: Column (1): Number of students enrolled in private elementary and secondary schools (logarithmized); column (2): Number of high school graduates in private secondary schools (logarithmized); column (3): Number of teachers in private elementary and secondary schools (logarithmized); column (4): Number of private schools (logarithmized); column (5): Number of students who are homeschooled (logarithmized). Outcome variables in columns (1) - (4) available biennially from 1993 until (including) 2009. Data on graduates is missing for Wyoming for 2005 and 2007, and for North Dakota for 2007 as "reporting standards not met" according to NCES. Before 2003, teacher data did not include prekindergarten teachers; since 2003, teacher data includes prekindergarten teachers who do not teach only prekindergarten (teachers who teach only prekindergarten are excluded). Outcome variable in column (5) available from 2000 until 2009, for 30 states (with missings). Single, double, and triple asterisks indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Data sources: Columns (1) - (4): National Center for Education Statistics; column (5): International Center for Home Education Research.

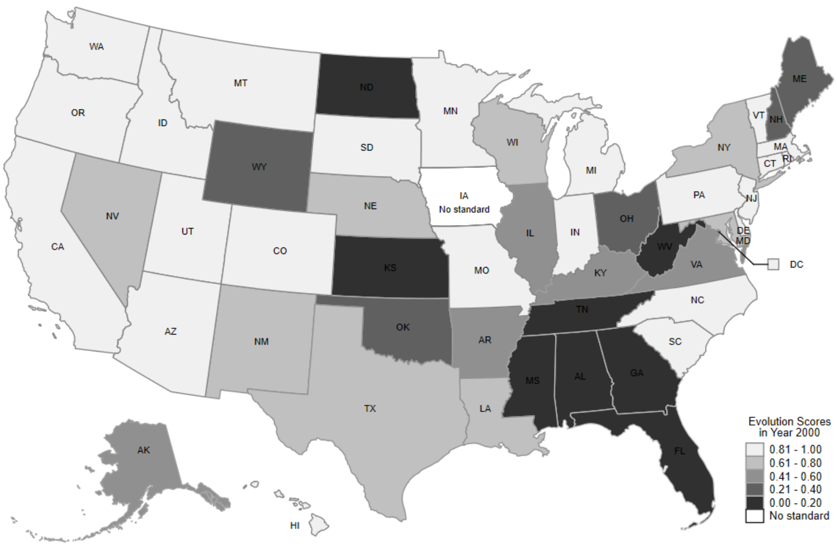
Figure A.I
Teachers' Focus on Science Standards When Teaching Evolution



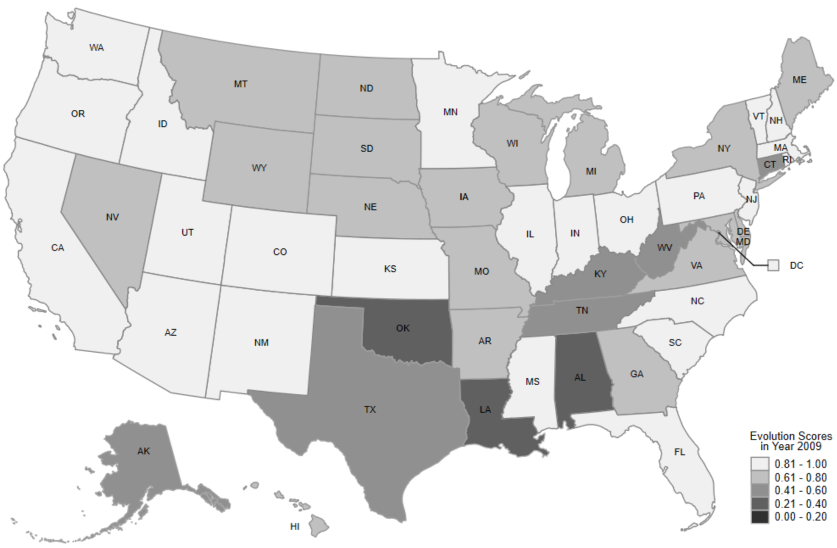
Note: Figure shows histogram of answer categories regarding agreement with the statement "When I do teach evolution, I focus heavily on what students need to know to meet state science standards". Data Source: National Survey for High School Biology Teachers, 2007.

Figure A.II
US Map of Evolution Scores

(A) Year 2000



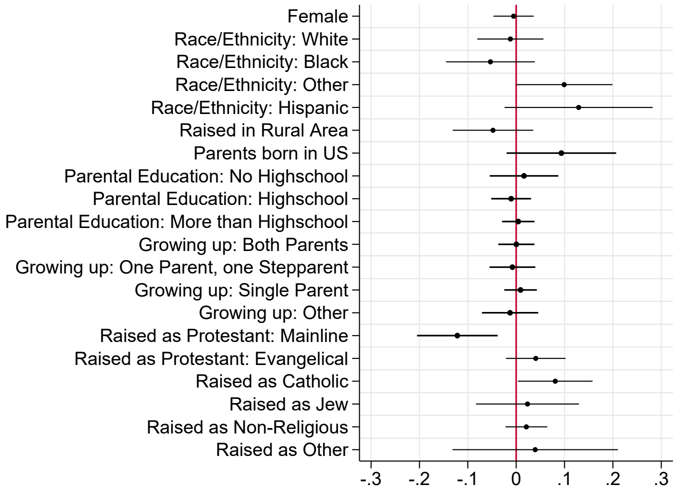
(B) Year 2009



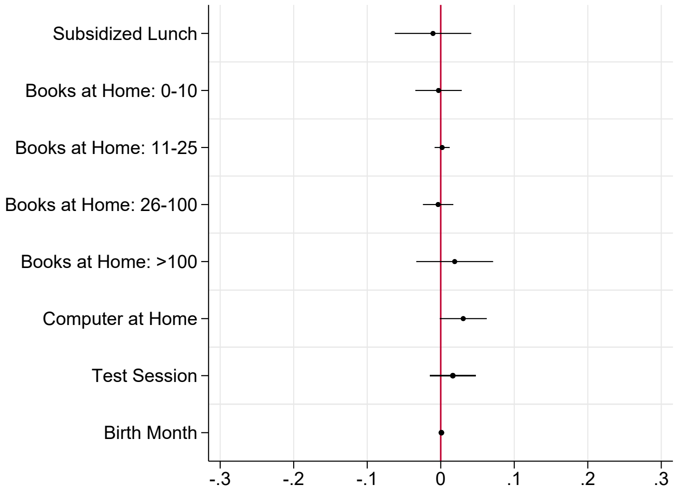
Note: Map depicts the evolution score of US States in 2000 and 2009, respectively. The evolution score measures the coverage of evolution in Science Standards, as reported in Lerner (2000) and Mead and Mates (2009). An evolution score of 0 indicates no or a creationist coverage of evolution, and a score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

Figure A.III
Correlates of Evolution Score with Covariates

(A) Panel A: Covariates from GSS

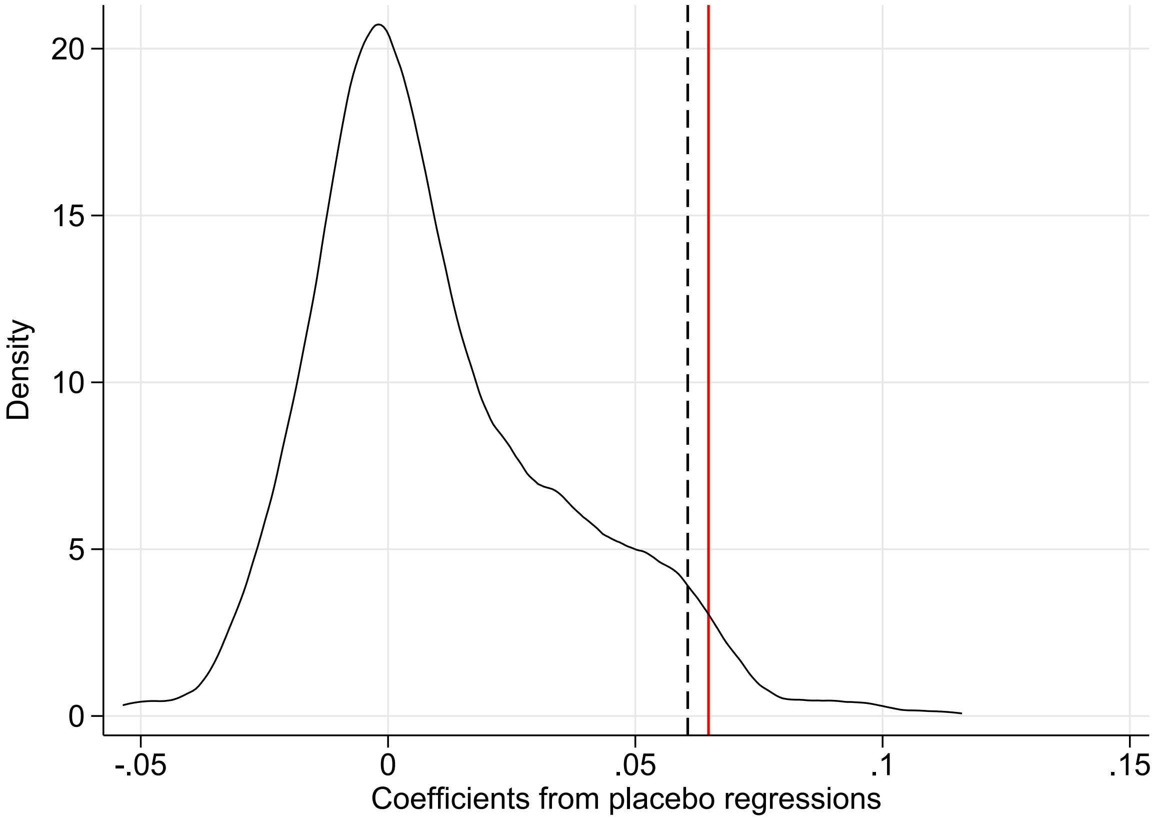


(B) Panel B: Covariates from NAEP



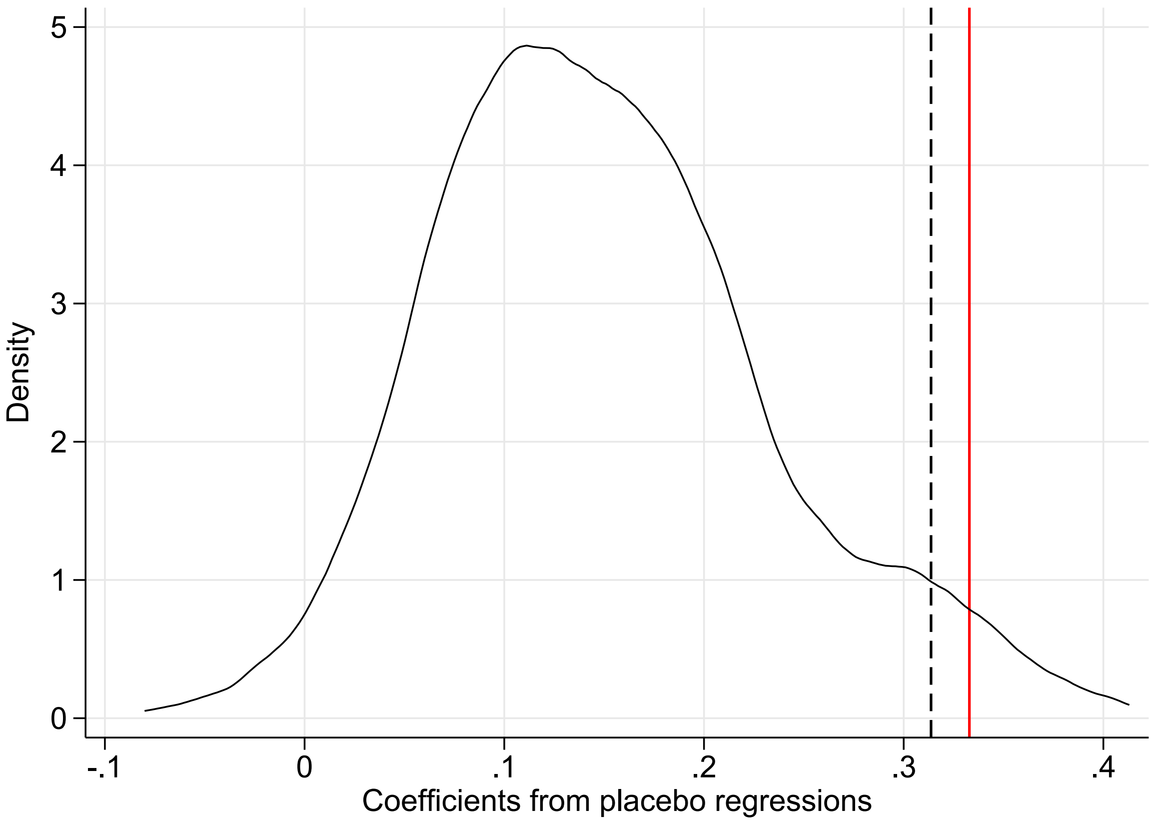
Note: Each row displays the estimate from a separate regression of the evolution score on the covariate defined along the vertical axis and survey year fixed effects. Panel A contains covariates from GSS as included in main regressions; Panel B contains additional covariates from NAEP as included in main regressions (other than those already shown in Panel A; note that ACS contains no covariates other than those already shown in Panel A). 95% confidence intervals displayed, calculated with clustering at the state level. Data sources: General Social Survey; 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.IV
Placebo Tests: Density Plot of Placebo Estimates of Effect of Evolution Coverage in Science Standards on Knowledge About Evolution, Based on 1000 Permutations



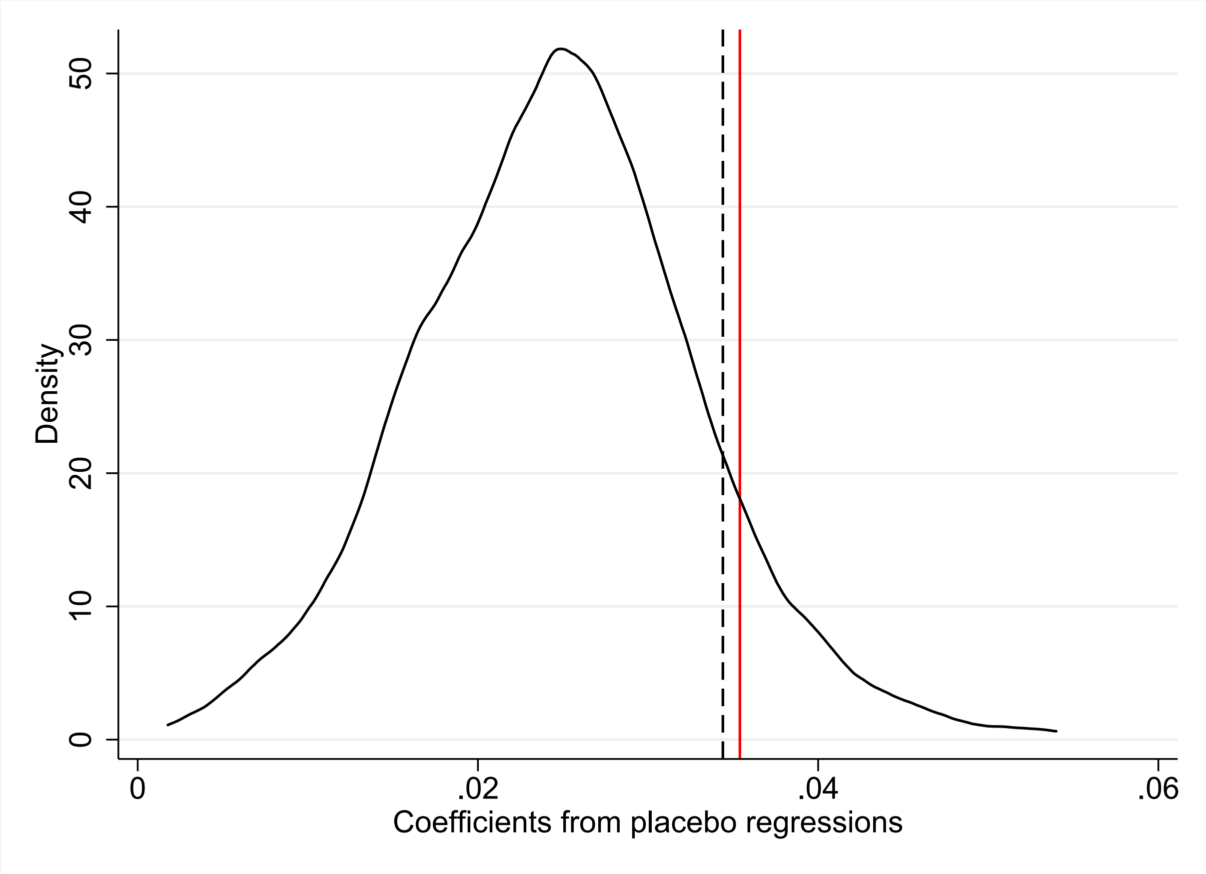
Note: Kernel density plot of TWFE OLS coefficients of effect of evolution coverage in Science Standards on knowledge about evolution from estimating equation 1, with randomly reshuffled reform years across reforming states (1000 permutations). Red solid vertical line indicates coefficient of reform effect from baseline model (0.065). Black dashed vertical line indicates 95^{th} percentile of the distribution of the 1000 placebo coefficients (0.061). Dependent variable: Share of questions about evolution answered correctly. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, home possessions (separate indicator variables for computer and books), birth month, test session, and fixed effects for state, cohort, 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.V
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 TWFE OLS coefficients of effect of evolution coverage in Science Standards on evolution belief from estimating equation 1, 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 (0.314). 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, cohort, and survey year. Standard errors clustered at the state level. Data source: General Social Survey

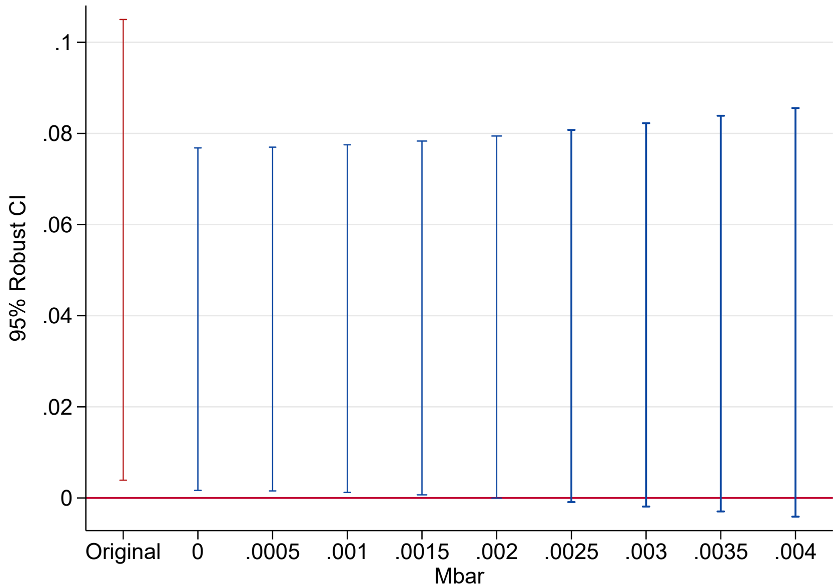
Figure A.VI
Placebo Tests: Density Plot of Placebo Estimates of Effect of Evolution Coverage in Science Standards on Probability of Working in Life Sciences, Based on 1000 Permutations



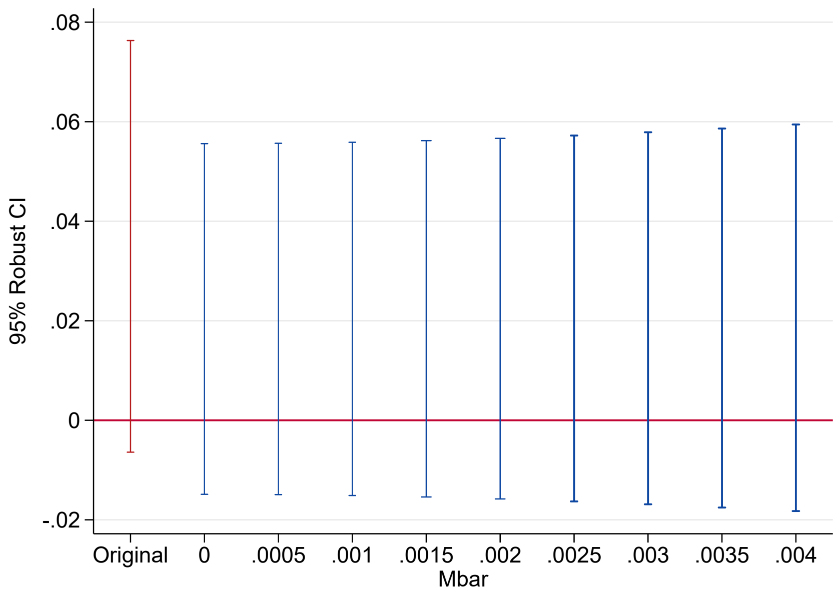
Note: Kernel density plot of TWFE OLS coefficients of effect of evolution coverage in Science Standards on probability of working in life sciences from estimating equation 1, with randomly reshuffled reform years across reforming states (1000 permutations). Red solid vertical line indicates coefficient 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, cohort, and survey year. Standard errors clustered at the state level. Data source: American Community Survey.

Figure A.VII
Sensitivity Analysis of Confidence Intervals for Non-linear Pre-trends: Effect of Evolution Coverage in Science Standards on Knowledge About Evolution

(A) States reducing evolution coverage



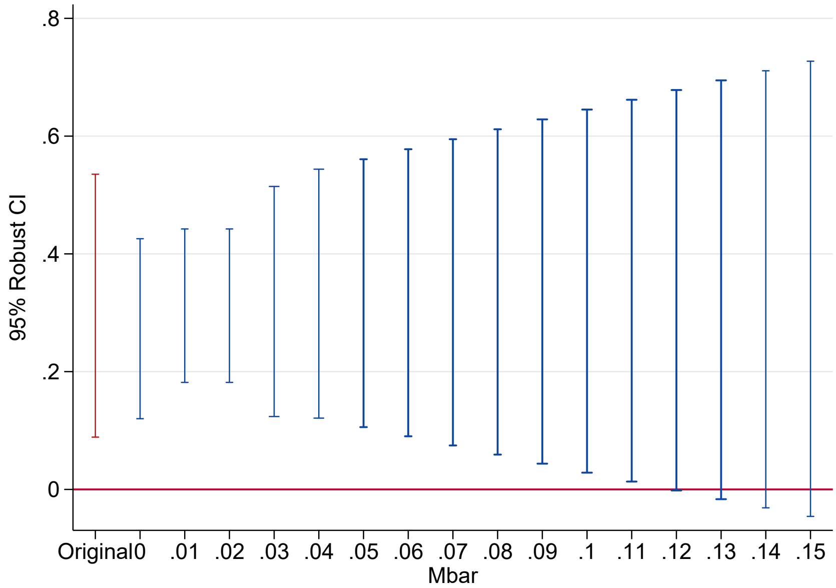
(B) States expanding evolution coverage



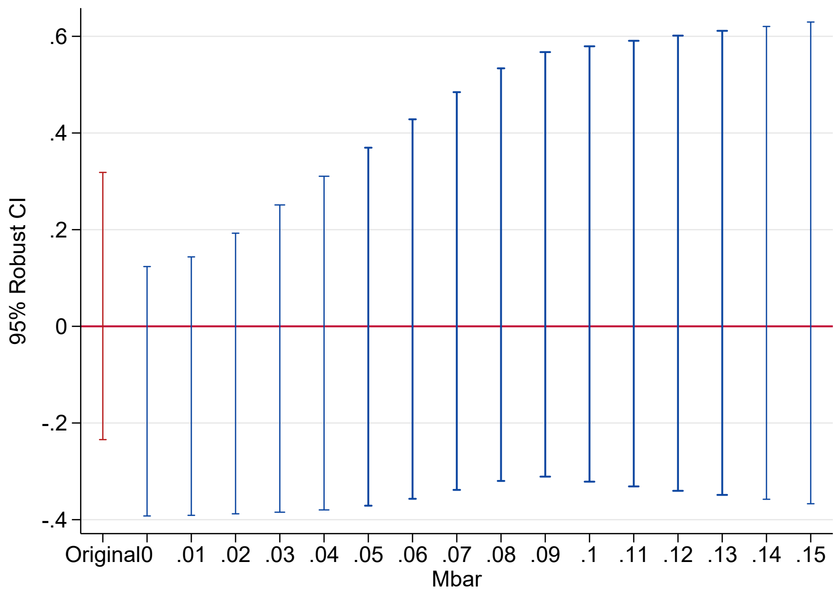
Note: Figure shows sensitivity analysis of estimated effect on evolution knowledge to potential violations of parallel trends assumption per Rambachan and Roth (2023); see also Ang (2020) and Biasi and Sarsons (2021). Red bar represents the 95% confidence interval of first post-reform TWFE event study estimate as displayed in Figure II. Blue bars represent corresponding 95% Fixed-Length Confidence Intervals (FLCI) when allowing for per-period violations of parallel trends of up to M. M = 0 corresponds to allowing a linear differential trend. M > 0 allows for increasingly more varied nonlinear trends. 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.VIII
Sensitivity Analysis of Confidence Intervals for Non-linear Pre-trends: Effect of Evolution Coverage in Science Standards on Evolution Belief in Adulthood

(A) States reducing evolution coverage



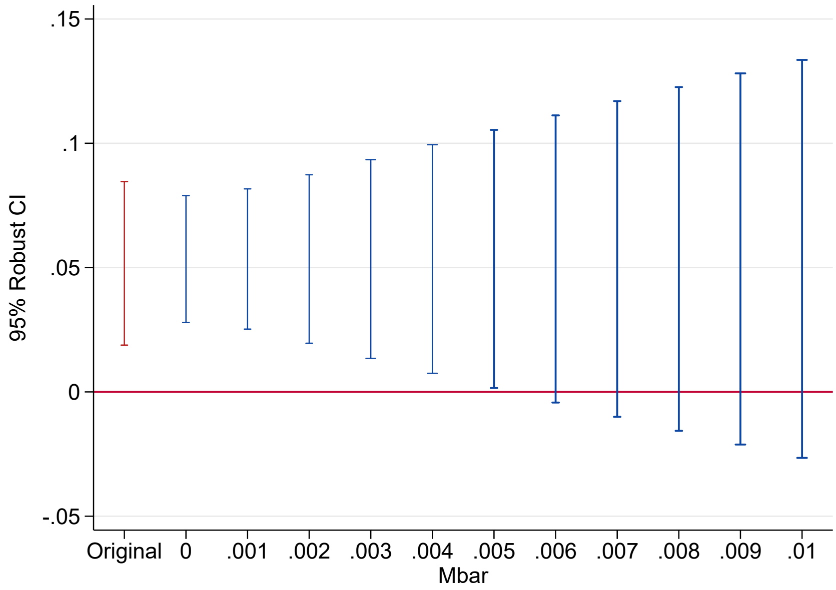
(B) States expanding evolution coverage



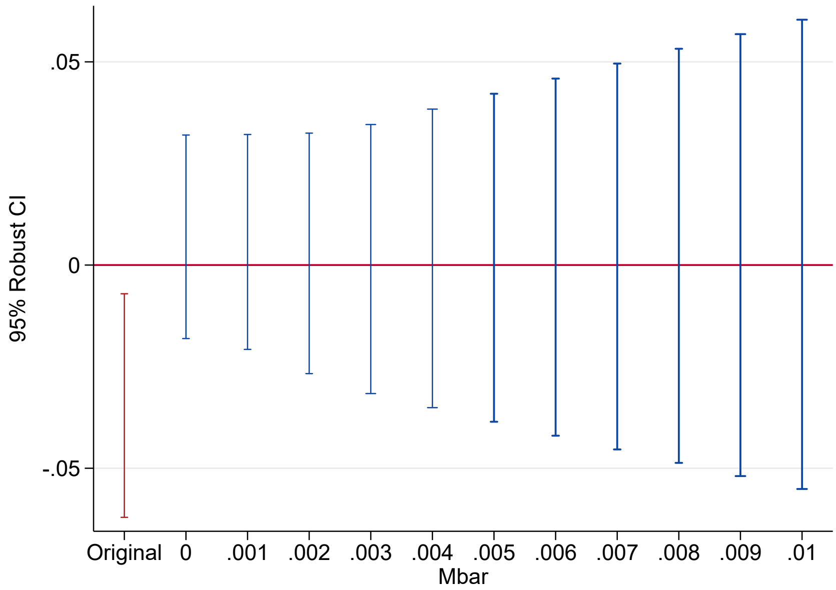
Note: Figure shows sensitivity analysis of estimated effect on evolution knowledge to potential violations of parallel trends assumption per Rambachan and Roth (2023); see also Ang (2020) and Biasi and Sarsons (2021). Red bar represents the 95% confidence interval of first post-reform TWFE event study estimate as displayed in Figure III. Blue bars represent corresponding 95% Fixed-Length Confidence Intervals (FLCI) when allowing for per-period violations of parallel trends of up to M. M = 0 corresponds to allowing a linear differential trend. M > 0 allows for increasingly more varied nonlinear trends. Data source: General Social Survey

Figure A.IX
Sensitivity Analysis of Confidence Intervals for Non-linear Pre-trends: Effect of Evolution Coverage in Science Standards on Probability of Working in Life Sciences

(A) States reducing evolution coverage



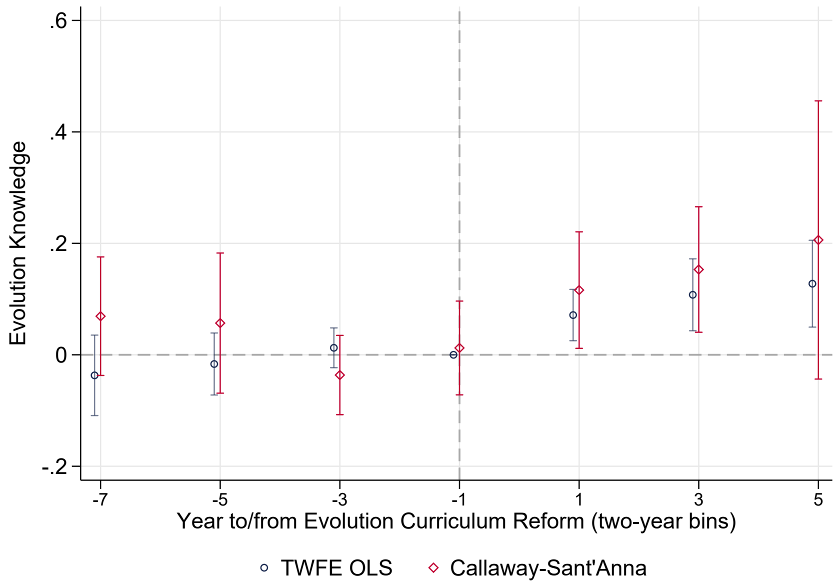
(B) States expanding evolution coverage



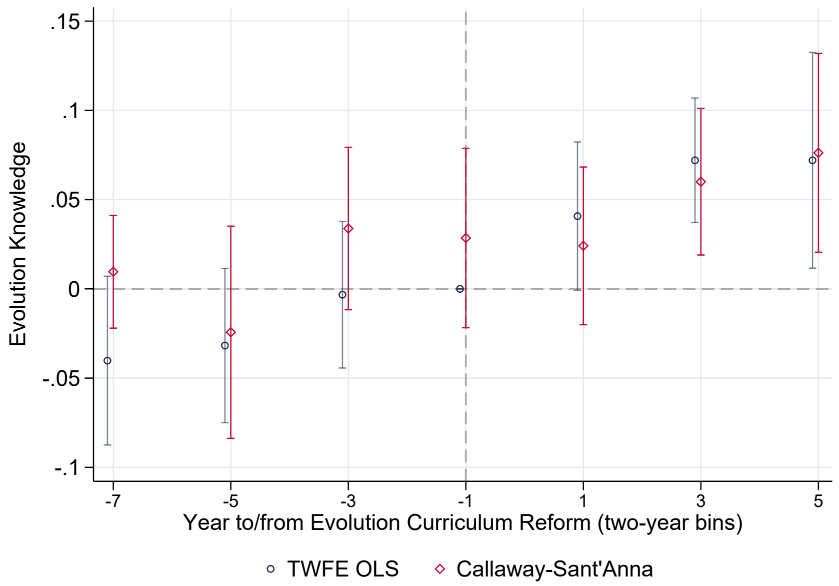
Note: Figure shows sensitivity analysis of estimated effect on evolution knowledge to potential violations of parallel trends assumption per Rambachan and Roth (2023); see also Ang (2020) and Biasi and Sarsons (2021). Red bar represents the 95% confidence interval of first post-reform TWFE event study estimate as displayed in Figure V. Blue bars represent corresponding 95% Fixed-Length Confidence Intervals (FLCI) when allowing for per-period violations of parallel trends of up to M. M = 0 corresponds to allowing a linear differential trend. M > 0 allows for increasingly more varied nonlinear trends. Data source: American Community Survey

Figure A.X
Robustness on Main Event Study Graphs (Evolution Knowledge): Adding Never-Treated States

(A) States reducing evolution coverage



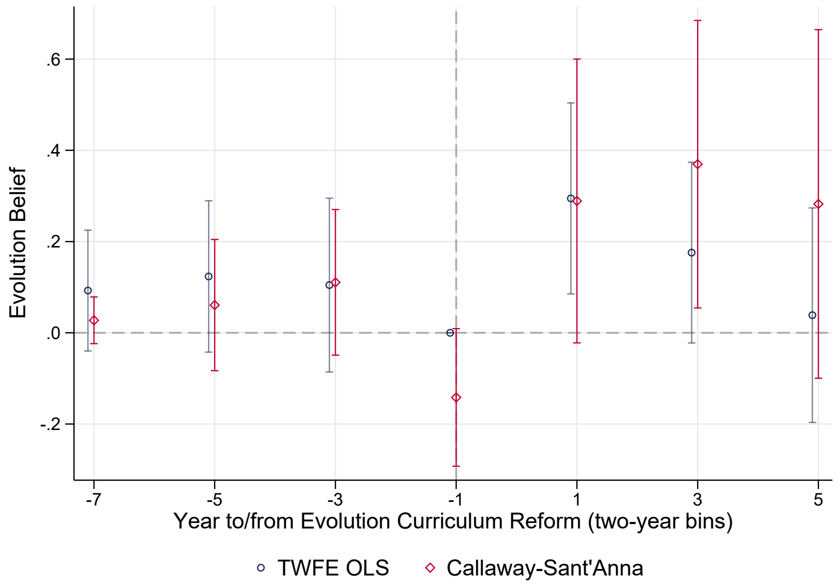
(B) States expanding evolution coverage



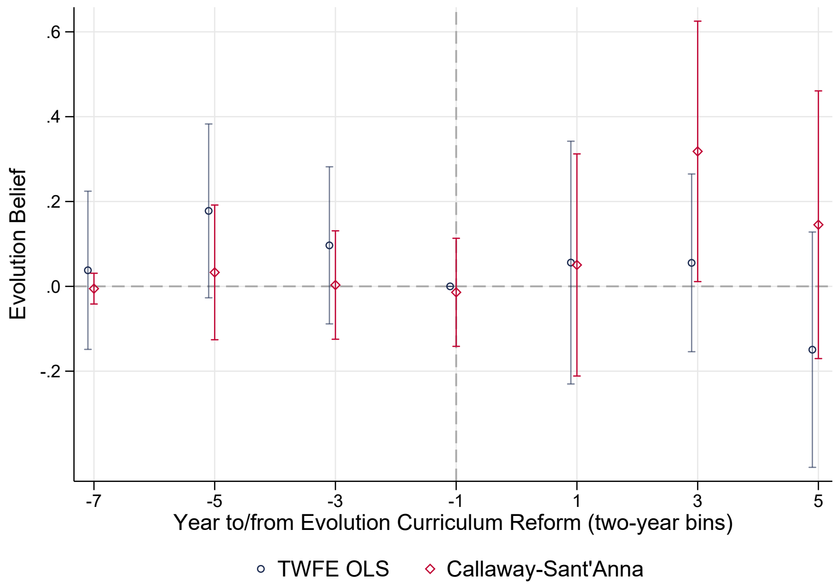
Note: Event study graphs with set of selected never-treated states (Nebraska, Oregon, Utah, Wisconsin) added to sample, otherwise as described in footnotes of Figure II.

Figure A.XI
Robustness on Main Event Study Graphs (Evolution Belief): Adding Never-Treated States

(A) States reducing evolution coverage



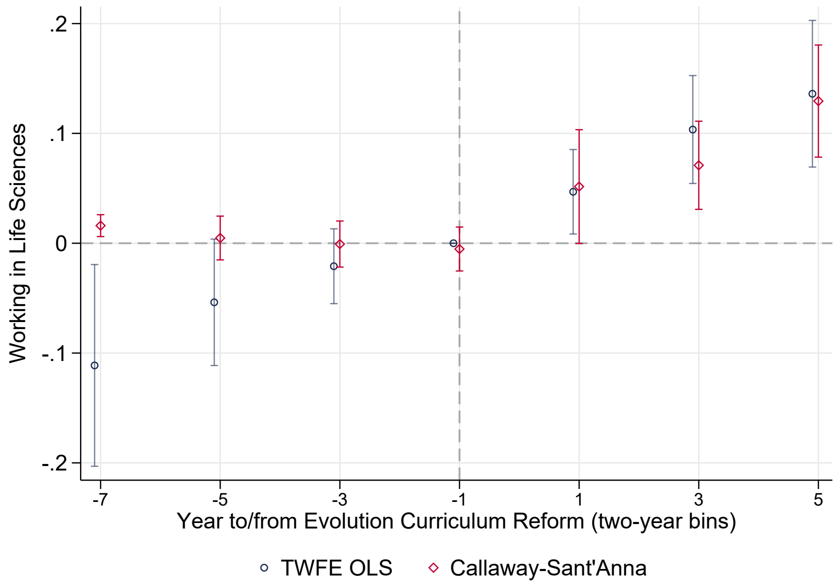
(B) States expanding evolution coverage



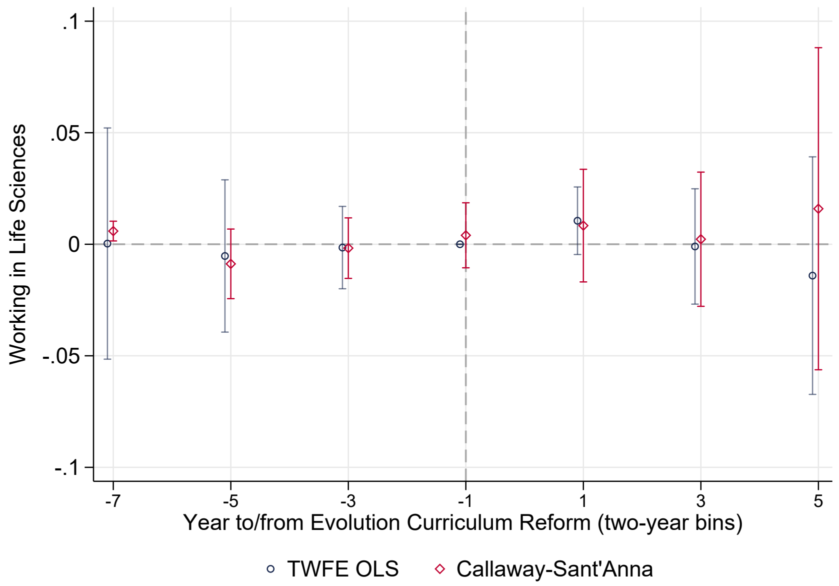
Note: Event study graphs with set of selected never-treated states (Nebraska, Oregon, Utah, Wisconsin) added to sample, otherwise as described in footnotes of Figure III.

Figure A.XII
Robustness on Main Event Study Graphs (Working in Life Sciences): Adding Never-Treated States

(A) States reducing evolution coverage



(B) States expanding evolution coverage



Note: Event study graphs with set of selected never-treated states (Nebraska, Oregon, Utah, Wisconsin) added to sample, otherwise as described in footnotes of Figure V.

Figure A.XIII
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.

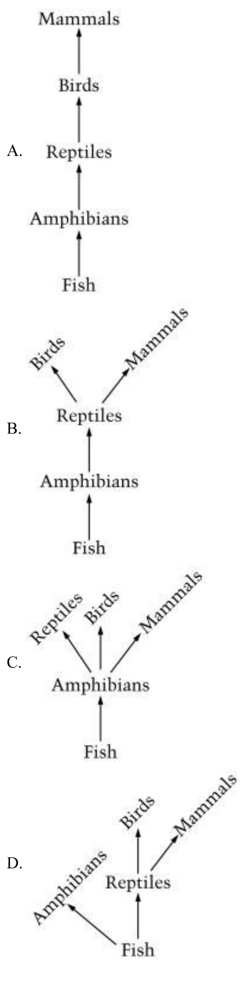
B. Some individuals possess features that enable them to survive better than individuals lacking those features.

C. More offspring are produced than can generally survive.

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

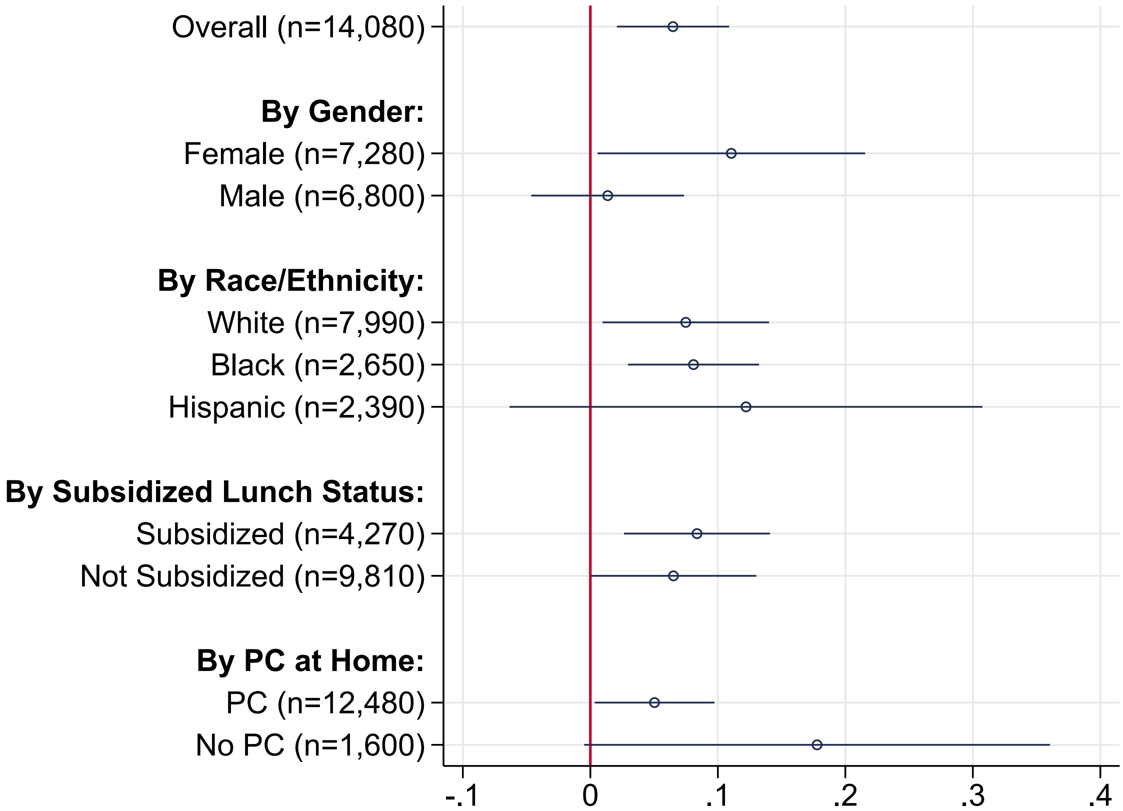
(B) Sample Question 2

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



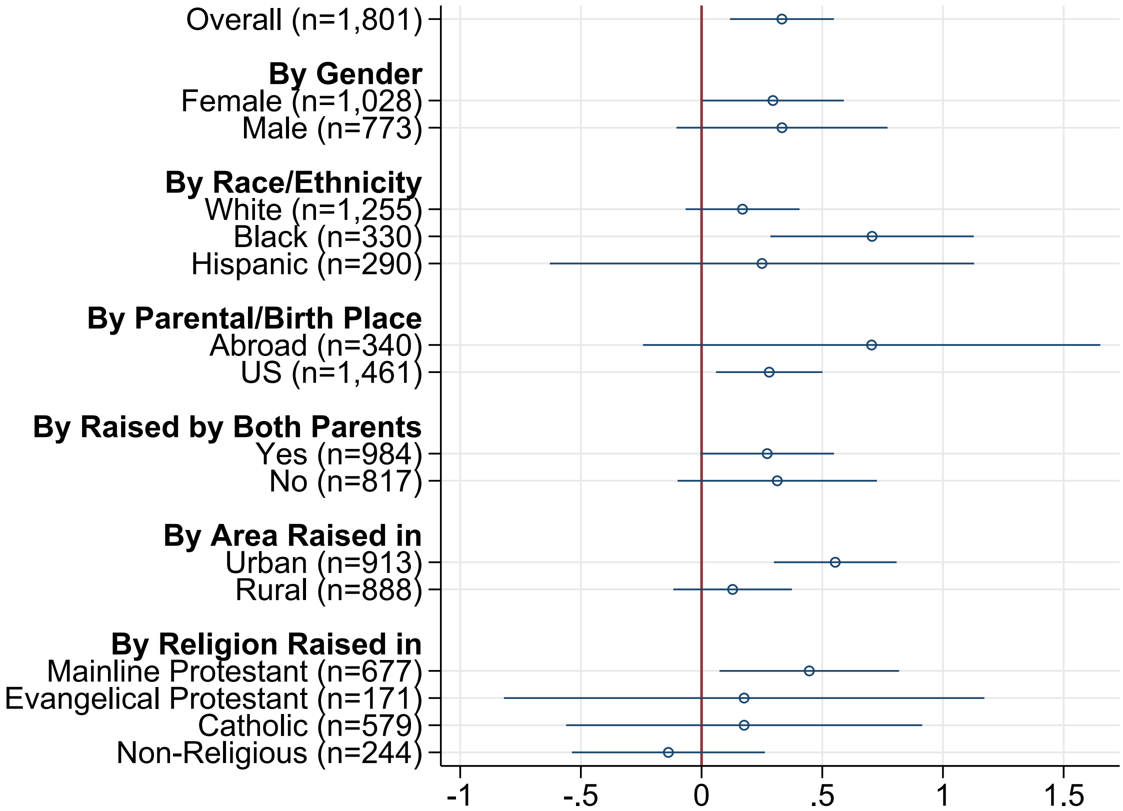
Note: Sample question on evolution knowledge from NAEP Science Test, Grade 12, Year 2000. Question also accessible online at NAEP question tool. Question 1: Answer D is correct. Question 2: Answer B is correct. Data source: U.S. Department of Education, National Center for Education Statistics, National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

Figure A.XIV
Effect of Evolution Coverage in Science Standards on Evolution Knowledge in School, by Subgroups



Note: Figure shows TWFE OLS coefficients and 95% confidence intervals of effect of evolution coverage in Science Standards on share of questions about evolution answered correctly from estimating equation 1, by individual subgroup as indicated along the vertical axis. Sample sizes of subgroups in parenthesis. Controls: Indicator variables for gender, races/ethnicities, subsidized lunch status, home possessions (separate indicator variables for computer and books), birth month, test session, and fixed effects for state, cohort, 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.XV
Effect of Evolution Coverage in Science Standards on Evolution Belief in Adulthood, by Subgroups



Note: Figure shows TWFE OLS coefficients and 95% confidence intervals of 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) from estimating equation 1, by individual subgroup as indicated along the vertical axis. 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, cohort, and survey year. Standard errors clustered at the state level. Data source: General Social Survey

References

Anderson, Michael L., “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 (2008), 1481–1495.

Andrilla, C. Holly A., Davis G. Patterson, Lisa A. Garberson, Cynthia Coulthard, and Eric H. Larson, “Geographic Variation in the Supply of Selected Behavioral Health Providers,” *American Journal of Preventive Medicine*, 54 (2018), 199–207.

Ang, Desmond, “The Effects of Police Violence on Inner-City Students,” *Quarterly Journal of Economics*, 136 (2020), 115–168.

Barro, Robert J., “Vaccination rates and COVID outcomes across U.S. states,” *Economics & Human Biology*, 47 (2022), 101–201.

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

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

Bertocchi, Graziella and Arcangelo Dimico, “The racial gap in education and the legacy of slavery,” *Journal of Comparative Economics*, 40 (2012), 581–595.

Biasi, Barbara and Heather Sarsons, “Flexible Wages, Bargaining, and the Gender Gap,” *Quarterly Journal of Economics*, 137 (2021), 215–266.

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

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

Chetty, Raj, John N Friedman, Michael Stepner, et al., “The economic impacts of COVID-19: Evidence from a new public database built using private sector data,” *Quarterly Journal of Economics,* (2023).

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter, "Race and economic opportunity in the United States: An intergenerational perspective," *Quarterly Journal of Economics*, 135 (2020), 711–783.

Cunningham, Brittany C, Kathleen Mulvaney Hoyer, and Dinah Sparks, "Gender Differences in Science, Technology, Engineering, and Mathematics (STEM) Interest, Credits Earned, and NAEP Performance in the 12th Grade. Stats in Brief. NCES 2015-075.," *National Center for Education Statistics* (2015).

Farsani, Danyal, Farzad Radmehr, Mohadaseh Alizadeh, and Yusuf Feyisara Zakariya, "Unpacking the black-box of students' visual attention in Mathematics and English classrooms: Empirical evidence using mini-video recording gadgets," *Journal of Computer Assisted Learning*, 37 (2021), 773–781.

Gershenson, Seth and Stephen B Holt, "Gender gaps in high school students' homework time," *Educational Researcher*, 44 (2015), 432–441.

Gibbons, Stephen and Olmo Silva, "Urban density and pupil attainment," *Journal of Urban Economics*, 63 (2008), 631–650.

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

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

Hoxby, Caroline M and Christopher Avery, "The missing one-offs: The hidden supply of high-achieving, low income students," *NBER Working Paper No. 18586* (2012).

Kahn, Shulamit and Donna Ginther, "Women and STEM," *NBER Working Paper No. 23525* (2017).

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

Litzler, Elizabeth, Cate C Samuelson, and Julie A Lorah, "Breaking it down: Engineering student STEM confidence at the intersection of race/ethnicity and gender," *Research in Higher Education*, 55 (2014), 810–832.

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

Pew Research Center, *About a quarter of rural Americans say access to high-speed internet is a major problem*, 2018.

Rambachan, Ashesh and Jonathan Roth, "A More Credible Approach to Parallel Trends," *Review of Economic Studies*, 90 (2023), 2555–2591.

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

Speer, Jamin D, "Pre-market skills, occupational choice, and career progression," *Journal of Human Resources*, 52 (2017), 187–246.

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

U.S. Department of Education, *Institute of Education Sciences, National Center for Education Statistics. National Assessment of Educational Progress (NAEP), 2009, 2011, 2015, and 2019 Science Assessments.* 2023, URL: https://nces.ed.gov/nationsreportcard/science/.