More than Sex:
Sex Education and Women's Labor Market Outcomes

Maddalena Grignani*

October 15, 2026

Abstract

The timing of fertility and gender norms are two important determinants of women's labor market outcomes. Sex education addresses both: it teaches students about contraception and conveys views about the roles of men and women. I construct a new panel of state sex education mandates in the United States from 1970 to 2022, coded from statutes and administrative rules, and link it to the 2000–2024 American Community Survey by birth state and cohort. Whether sex education is mandatory has no detectable effect on women's labor market outcomes, but its content does. Comprehensive mandates, which require teaching contraception rather than stressing abstinence on moral grounds, raise women's annual earnings by 4.2 log points and the share working full time by 1.8 percentage points, while men's earnings do not change. Traditional mandates, which stress abstinence on moral grounds instead, mostly until marriage, have no robust effect. Women exposed to a comprehensive mandate are more likely to complete college and less likely to marry by 23, and they hold more egalitarian views about women's market work. I also find suggestive evidence of lower teen fertility, more condom use among high-school students and later first births. The results point to a change in the life course young women plan for: schooling before marriage.

*Universitat Pompeu Fabra. maddalena.grignani@upf.edu

I am grateful to Libertad González for her guidance and support, and to Judit Vall-Castelló for her helpful comments. I thank Sasha Grenier (Oregon Department of Education), Laurie Dils (Washington OSPI), Susanne Schmal (North Carolina Department of Public Instruction), and Nawal Umar and Alison Macklin (SIECUS) for generously sharing their knowledge of the history of state sex education policy, and librarians at the Rutgers University Libraries for locating and digitizing archival material on New Jersey's family life education mandate. I also thank participants at the UPF Applied Lunch Seminar, the BSE PhD Jamboree, and the Thessaloniki 10th International PhD Meeting in Economics for helpful comments. All errors are my own.

1 Introduction

Women in the United States earn substantially less than men, and the gap widens over the course of their careers (Goldin, 2014). Two important contributors are motherhood (Kleven et al., 2019) and gender norms governing the division of paid work and domestic responsibilities (Fernández et al., 2004). Both are shaped partly during adolescence, when young people begin to form expectations about their future family and working lives. The career costs of motherhood are greater when childbearing occurs earlier (Miller, 2011; Adda et al., 2017). The timing of childbearing may depend both on adolescents' knowledge of contraception, which can affect the risk of unintended pregnancy, and on their expectations about adult life, including gender roles and the timing of parenthood (Goldin, 2006; Kearney and Levine, 2012; Billari et al., 2011). Schools can shape both: they provide information about sex and reproduction while also transmitting gender norms, explicitly or implicitly (Hara and Rodriguez-Planas, 2025). This paper studies the school subject where the two intersect: compulsory sex education.

About half of U.S. high school students in their final year report having had sexual intercourse (Centers for Disease Control and Prevention, 2024a), yet no federal mandate governs what they are taught about it. In some states sex education is still not mandatory, and where it is, its content varies widely. In 2022, 64 percent of schools nationally reported teaching students how to use a condom, but in some states fewer than 30 percent did (Centers for Disease Control and Prevention, 2023).[1] Many adolescents therefore become sexually active before receiving formal education how to prevent a pregnancy or an infection. In 2023, the birth rate among girls aged 15 to 19 was 13 per thousand, against 7 in the European Union and 5 in Canada (World Bank, 2024),[2] and nearly half of the 2.2 million cases of chlamydia, gonorrhea and syphilis reported in 2024 were among people aged 15 to 24 (Centers for Disease Control and Prevention, 2025). What sex education should contain remains contested. In 2025 alone, the Sexuality Information and Education Council of the United States (SIECUS) tracked more than 650 bills in state legislatures, nearly a quarter of which sought to limit or eliminate access to sex education and the rest to expand it (SIECUS, 2025).

In this paper, I study whether what high school sex education is required to teach—contraception, or sexual abstinence as a moral standard—affects women's earnings, employment and full-time work decades later, and I examine the timing of family formation, educational attainment and gender attitudes as mechanisms. To do so, I construct a new panel of binding state sex education mandates for all fifty states and the District of Columbia from 1970 to 2022, coded from primary statutes and administrative rules. I link it to the 2000–2024 American Community Survey, assigning each respondent the mandate in force in her birth state during her secondary school years, and follow the exposed cohorts into adulthood.

In the United States, education is a state responsibility: federal law bars federal officials from directing, supervising or controlling school curricula (20 U.S.C. 3403). Each state decides whether to mandate sex education and what it must contain, or leaves both choices to school districts. The panel records, for each state and year, whether sex education is mandatory, whether abstinence must be stressed on moral grounds, and whether contraception must be taught. From these provisions I define two types of mandate. A mandate is *traditional* if it presents abstinence as a moral standard—the expected or responsible behavior, in most states until marriage—and does not require contraception to be taught. It is *comprehensive* if it requires instruction on contraception and does not stress abstinence on moral grounds. Nine states adopted a comprehensive mandate between 1990 and 2022 and nine a traditional

[1] The national figure is likely an upper bound, since some conservative states that do not require instruction on contraception—Utah, Mississippi, Louisiana—are not in the School Health Profiles sample.

[2] By age 18, 3.1 percent of U.S. women have given birth (author's calculations from National Center for Health Statistics (2025)).

one between 1988 and 2013. The most recent comprehensive mandate, Washington's in 2022, leaves no exposed cohort old enough to be observed at working age, so the labor market estimates rest on eight comprehensive and nine traditional states.

The curricula taught on each side make the difference concrete. Abstinence programs often prescribe gender roles explicitly. *WAIT Training*, one of the most widely adopted, lists "Financial Support" among the "5 Major Needs of Women" and "Domestic Support" among the "5 Major Needs of Men." *Why kNOw* teaches that "women gauge their happiness and judge their success by their relationships," whereas "men's happiness and success hinge on their accomplishments."[3] Programs on the comprehensive side treat sex as part of growing up and teach how to prevent its risks: *Making Proud Choices!*, among the most widely used, teaches that "all people are sexual, and sexuality is a natural and normal part of being human," and devotes two modules to "Developing Condom Use and Negotiation Skills" (Jemmott et al., 2016). Most say nothing about gender, but a few address gender stereotypes directly: a lesson in *FLASH* teaches that gender roles "are learned" and "can be limiting," and asks pupils "Why might Mary think that only boys can be doctors? How could that stereotype limit her in life?" (Public Health – Seattle & King County, 2013).

I compare cohorts exposed to a mandate with cohorts born in the same state who finished secondary school shortly before it took effect, using the two-stage difference-in-differences estimator of Gardner (2022) with birth-state, birth-cohort and survey-year fixed effects. Labor market outcomes are measured between ages 25 and 59. Statutes set the curriculum, but what each cohort was actually taught is not observed, so the estimates are intention-to-treat. With eight or nine states per arm, conventional standard errors can be too small, so beneath every estimate I report a randomization-inference p-value, obtained by reassigning the states' policy histories at random, and I read significance from it.

The existing estimates for teen outcomes compare states with any mandate to states with none. Applied to the labor market, that comparison finds no detectable effect on any outcome. Separating the two curricula changes the picture. Exposure to a comprehensive mandate raises women's annual earnings by 4.2 log points, about \$1,300 on a pre-adoption mean near \$31,000, and the share working full time by 1.8 percentage points, from 59.5 percent. Employment rises by 0.9 points in the fully exposed cohorts, from 72.5 percent, and by about two points once birth-state trends absorb a downward drift in the adopting states before exposure. These effects are of the same order as those of policies that widened young women's options: early legal access to the pill raised women's hourly wages by 8 percent by their late forties (Bailey et al., 2012), and Title IX, by expanding girls' high-school sports, raised the share of women with at least four years of college by about two percentage points (Stevenson, 2010). Both changed what young women could do; a curriculum changes only what they are taught.

The results are robust. Randomization inference confirms the earnings and full-time estimates, as well as the employment estimates once trends are included, and the in-time placebos are close to zero. The earnings estimate also survives a wide range of further robustness checks (Section 6.2). Men's earnings and full-time work do not move and their employment falls slightly, as the model below predicts once the breadwinner prescription weakens, so the gap between men and women born in the same state and year narrows on all three margins. The traditional curriculum has no robust effect: earnings do not move, and full-time work falls in the baseline specification but not once birth-state trends are included.

To see why the content of a curriculum should affect adult earnings, I set out a model in which ado-

[3] Both passages are quoted in U.S. House of Representatives, Committee on Government Reform—Minority Staff, Special Investigations Division (2004). Federal grants paid for these programs in states with and without a traditional mandate: federally funded grantees used *WAIT Training* in 21 states between fiscal years 2003 and 2015, among them the traditional adopters Florida and Utah, and *Why kNOw* mainly in Tennessee (SIECUS, 2015); *Choosing the Best Life* contains similar material. Section 2.2 and Appendix A.3 document the curricula and where each operated.

lescent girls choose how much to invest in education and contraception, given their taste for a traditional life and the returns to working, which rise with college. Sex education works through two channels: what adolescents know about contraception, which sets the cost of avoiding an early birth, and the gender norms they hold, which shape the cost of pursuing a career, as in Akerlof and Kranton (2000). Either channel alone predicts that a comprehensive curriculum lowers early births and raises college going, employment and earnings, so the signs of these effects cannot tell the channels apart. Only the norm channel predicts two further results: a college response larger than the fall in early births can explain, and more egalitarian views about women's market work. A traditional curriculum has no effect if the prevailing norm is already traditional, and the opposite effects otherwise.

Exposed women finish college more often and marry later. College completion rises by 2.6 percentage points, from 33.9 percent, and the share married by 23—the age by which a four-year degree is normally complete—falls by 2.0 points, from 27.1 percent. Marriage is the outcome the statutes speak to most directly: most traditional mandates prescribe abstinence until marriage, and no comprehensive mandate ties sex to it. Yet only the comprehensive mandates move it; under the traditional ones, the share married by 23 does not change. College and marriage are both known to shape women's careers: women invest in college when they expect long careers (Goldin et al., 2006), later marriage and motherhood raise their wages (Loughran and Zissimopoulos, 2009; Miller, 2011), and early access to the pill raised women's wages through the same two margins (Goldin and Katz, 2002; Bailey et al., 2012). If exposed women enter the labor market with more schooling and start families later, the gains should show up in hours and pay and should persist with age. They do: the effects are concentrated on full-time work and earnings, and in the three earliest adopters, the only ones observed past 40, they show no sign of fading between the mid-thirties and the mid-forties, as expected when women build up experience before their first child (Troske and Voicu, 2013; Blundell et al., 2016). Natality data for the same three states point to later motherhood. First births at ages 25 to 29 fall by 7.9 log points, an estimate that survives birth-state trends and is more than twice its in-time placebo, and first births in the early thirties rise. The fall at 20–24 is smaller and less robust, so I read this as suggestive evidence that exposed women postpone motherhood.

Teen births and sexual behavior move in the same direction. Births to girls aged 15–19 fall by 2 to 4 per 1,000 girls a year, from about 41 before adoption, depending on the specification; the fall is driven largely by California, the largest adopter. In the Youth Risk Behavior Surveillance System, which covers four of the comprehensive adopters, the share of high-school students who have sex without a condom falls by 2.6 percentage points, about a sixth, both because fewer students are sexually active and because more of the sexually active use a condom. Averted teen births are far too few to account for the rise in college completion, which is what the norm channel predicts. The traditional curriculum has no robust effect on any of these margins.

Exposed women also hold more egalitarian views about women's market work. Asked in the American National Election Studies whether it is better for the man to work and the woman to take care of the home, they move 0.11 points up a 0–1 egalitarian scale, from a pre-adoption mean of 0.45, while men move by a smaller and imprecise 0.04. The women's estimate barely changes when I control for their work status, marital status and the number of children in the household, so the shift is not merely a by-product of their own employment and family. Nor is it a general turn away from traditional values: exposed women are no less likely to agree that the country would have fewer problems with more emphasis on traditional family ties. Because attitudes are measured decades after school and the in-time placebo reproduces about half of the women's estimate, I treat this evidence as suggestive.

The design compares cohorts within a state, so it is robust to anything that affected the whole

population at once. Changes that coincided with a mandate and reached the cohorts then in school could still bias the estimates. I consider three: school spending, state politics and access to fertility control. Per-pupil spending does not rise when a comprehensive mandate is adopted; the one exception, California, adopted in 2016 and is absent from the estimates for fully exposed cohorts. The partisan composition of state legislatures drifts toward the mandate's ideology before adoption but does not jump at it. Minors' legal access to contraception and abortion was settled by 1979, before the first mandate; enforcement of parental-involvement laws for abortion coincides with adoption in only three traditional states; and state mandates for insurance coverage of contraceptives reached the treated cohorts only as adults.

Contributions. The paper contributes first to the evaluation of sex education, where the evidence so far covers only short horizons and adolescent outcomes. Sex education, irrespective of its content, does not lower teen fertility (Oettinger, 1999; Sabia, 2006; Paton et al., 2020). Abstinence-based mandates do not either (Carr and Packham, 2017) and may even raise it (Santelli et al., 2017). By contrast, more comprehensive instruction, measured by program funding, lowers teenage pregnancies (Mark and Wu, 2022), and in Kenya information about differential HIV risk lowered them while an abstinence curriculum did not (Dupas, 2011; Duflo et al., 2015). Closest in design to this paper, Fung and Min (2025) find that the HIV/AIDS education mandates of 1987–88 that stressed abstinence raised teen births by about five per thousand, while those accompanied by sex education did not. Two gaps remain. First, the effects of state-mandated comprehensive and traditional curricula have never been estimated jointly: existing studies examine federal spending on programs that schools adopt voluntarily (Mark and Wu, 2022), abstinence mandates in isolation (Carr and Packham, 2017), or individual parts of curricula from the 2000s (Atkins and Bradford, 2021). Second, only one study follows exposed cohorts into adulthood. Lazuka and Elwert (2026) study the Swedish roll-out of sex education from 1942 to 1958 and find fewer unwanted pregnancies, postponed childbearing, higher female labor force participation at ages 30–35, and higher earnings at ages 48–50, with no significant difference between women and men. That reform, however, predates the pill and introduced a single curriculum, which promoted abstinence while also teaching family planning, so the effects of specific content cannot be separated. Whether sex education, and which kind, still matters once contraception is easy to obtain is an open question. The evidence that exists suggests that access to contraception alone does not lower births, and that without instruction it can even raise them. Three years of free modern contraception in rural Burkina Faso left birth rates unchanged, ruling out even modest effects (Dupas et al., 2025), and distributing condoms in U.S. schools raised teen fertility by about ten percent, an increase concentrated where distribution came without mandated counseling (Buckles and Hungerman, 2018). This paper estimates the effects of the two curricula jointly, in cohorts for whom contraception was legally available, and follows them into adulthood.

Second, the paper contributes to the literature on what shapes women's careers, in which two determinants stand out. One is the timing of family formation. Marrying and becoming a mother early lowers women's wages (Loughran and Zissimopoulos, 2009; Miller, 2011), and the gap between men's and women's earnings opens at the first birth (Kleven et al., 2019; Adda et al., 2017). Policies that reached adolescents—the pill, abortion access, Title IX—moved women's adult outcomes (Goldin and Katz, 2002; Bailey, 2006; Bailey et al., 2012; Stevenson, 2010), and girls' expectations about work and family drew them into college (Goldin, 2006; Goldin et al., 2006). The other is gender norms, which shape women's participation and earnings, are transmitted within the family and persist for decades (Fernández et al., 2004; Alesina et al., 2013; Jayachandran, 2020; Farré and Vella, 2013; Olivetti et al., 2020; Alesina and Fuchs-Schündeln, 2007). The existing evidence traces both to legal access to fertility control, to labor market prospects (Kearney and Levine, 2012; Jensen, 2012), to the family and, where schools teach gen-

der roles directly, to the classroom (Hara and Rodriguez-Planas, 2025). I show that a policy designed as fertility control, which regulates only what adolescents are taught about sex, also moves women's careers decades later, through margins that averted early births cannot account for on their own.

Finally, the paper contributes to the work on the long-run effects of schooling. Education is among the most consequential investments in young people: schooling raises earnings (Card, 1999) and brings non-pecuniary returns, including fewer teenage births (Oreopoulos and Salvanes, 2011; Black et al., 2008), and investments in its quality—preschool, classrooms and teachers, school spending—raise earnings and improve adult outcomes decades later (Heckman et al., 2010; Chetty et al., 2011, 2014; Jackson et al., 2016). These studies vary how much, and how well, students learn; what schools teach can also move expectations and norms. Prescribed curricular content shifts political attitudes (Cantoni et al., 2017), religious practice (Arold et al., 2025, 2026) and, closest to this paper, women's employment and earnings: the end of gender-segregated home economics in Japan raised both, with no change for men (Hara and Rodriguez-Planas, 2025). Targeted school programs on gender attitudes, harassment, dating violence or bullying shift adolescent beliefs (Dhar et al., 2022; Sharma, 2024; Foshee et al., 2004), and for the one program followed long enough, effects on wages are still visible at ages 27–29 (Braun et al., 2026). Sex education differs from these programs. It is designed, and has been evaluated, as fertility control, and its mandates regulate what adolescents are taught about sex rather than about the roles of men and women. Yet dropping a moral prescription on abstinence, and treating the timing of a birth as a girl's own decision, bears on norms as well as on knowledge. This paper shows that the content of sex education matters for women's careers: the comprehensive curriculum raises college completion, moves marriage past the college years, and raises earnings decades later, in cohorts for whom contraception was already legally available. The college response exceeds what averted teen births can deliver, and exposed women hold more egalitarian views about women's market work, which is consistent with the curriculum moving norms too. A mandate also reaches every school in the state, not only those that opt into a program, which makes it easy to scale.

The remainder of the paper is organized as follows. Section 2 describes sex education in the United States and the construction of the policy panel. Section 3 presents the model. Section 4 describes the identification strategy and Section 5 the outcome data. Section 6 reports the labor market results, Section 7 the mechanisms—the life course, including fertility and sexual behavior, and gender attitudes—and Section 8 the alternative explanations. Section 9 concludes.

2 Sex Education in the United States

2.1 Institutional setting

Sex education in the United States is governed by the states. Federal law prohibits the Department of Education from exercising direction, supervision or control over the curriculum of any school (20 U.S.C. 3403), so whether instruction is required, and what it must include, is decided state by state, through statute, binding administrative rules, or curriculum standards. Several states impose no requirement, others leave the decision to districts, and the remainder mandate instruction whose content differs sharply.

States that require instruction have generally adopted either a traditional or a comprehensive curriculum, and the choice between them has been contested since sex education entered American schools (Moran, 2000; Luker, 2006). Under the *traditional* model, statutes present abstinence as the behavior expected of adolescents, often until marriage, and do not require contraception to be taught; under the

comprehensive model, they require instruction on contraception and, when they cover abstinence, frame it in terms of efficacy rather than as a moral standard.

The statutes show the difference directly. Louisiana's statute (La. R.S. 17:281, 1991) states that "the major emphasis of any sex education instruction offered in the public schools of this state shall be to encourage sexual abstinence between unmarried persons," that instruction shall "emphasize abstinence from sexual activity outside of marriage as the expected standard for all school-age children," and that "no contraceptive or abortifacient drug, device, or other similar product shall be distributed at any public school." Oregon's statute (ORS 336.455) also promotes abstinence, but as "the most effective way to prevent pregnancy and the transmission of sexually transmitted diseases," a claim about efficacy, and it requires that "abstinence shall not be taught to the exclusion of other material and instruction on contraceptive and disease reduction measures"; its implementing standards add that programs "not, in any way, use shame or fear based tactics."[4] Abstinence appears in both statutes; what differs is its status—a moral standard tied to marriage in Louisiana, one option among several, taught alongside contraception and justified by its efficacy, in Oregon, which enters the comprehensive arm in 2010. That is the distinction the coding rules of Section 2.3 formalize, the reason it cannot be drawn from a keyword search for "abstinence," and the reason a design contrasting *any* mandate with none recovers little: the two arms impose opposite requirements on the same margin. Appendix A.2 quotes the operative language for every state in either arm.

2.2 What was actually taught

While the federal government cannot impose curricula, it has nevertheless shaped their content through school funds. For a single fiscal year, 1988, as states began to adopt HIV-education mandates, Congress required AIDS education paid for by the Departments of Health and Human Services and of Education to "emphasize abstinence from sexual activity outside a sexually monogamous marriage" (Pub. L. 100-202, §514); the clause lapsed after that year, but since 1994 federal education funds may not pay for sex or HIV education that omits "the health benefits of abstinence," nor for condom distribution in schools (20 U.S.C. 7906). Appendix A.4.1 traces these conditions year by year.

Federal support for abstinence-only instruction began in 1981, expanded when the 1996 welfare reform created the Title V abstinence education allotment, and again in 2001 with competitive grants awarded directly to organizations. Title V is allotted to states by formula, conditional on teaching "abstinence from sexual activity outside marriage as the expected standard for all school age children" (42 U.S.C. 710), and a governor can decline it. Cumulative federal spending on abstinence-only programs has exceeded $2 billion (Santelli et al., 2017), while federal funding for comprehensive instruction arrived only in 2010 and has remained small by comparison (Appendix A.4.2).

None of these conditions required a state to teach sex education, and none bound what states and districts taught with their own money: the 1994 law itself denies any federal official the power to "mandate, direct, review, or control" a school's curriculum. Nor did accepting the money mean spending it on school instruction: a state that took the Title V allotment chose who ran the funded programs and where, and many spent it on media or community campaigns that never reached a classroom (Section 2.3). Together, however, the conditions set the message of federally supported instruction—abstinence, framed in moral terms for one year in 1988 and again from 1996, with contraception at most described and, from 1994, not distributed in schools with federal funds. A state without a mandate left that message in place, and a traditional mandate restated it; a comprehensive mandate is the one policy in the data that departed

[4] Or. Admin. R. 581-022-2050 (Division 22 standards for public elementary and secondary schools).

from it.

Statutes set the boundary; curricula fill it. A 2004 congressional review of the curricula used by the hundred federally funded abstinence-only grantees found that eleven of the thirteen most widely adopted contained errors or distortions of public health information, and that those eleven were in use in 25 states by 69 grantees receiving over $32 million in fiscal year 2003 alone (U.S. House of Representatives, Committee on Government Reform—Minority Staff, Special Investigations Division, 2004). The passages on gender roles quoted in the introduction come from this group: federally funded grantees used *WAIT Training* in 21 states between fiscal years 2003 and 2015, and the *Choosing the Best* series, the most widely adopted of all, contains similar material. On the comprehensive side, grantees used *Making Proud Choices!* in 34 states between fiscal years 2009 and 2015 (SIECUS, 2015), and 47 Washington school districts used *FLASH* in 2009. Appendix A.3 reproduces the passages at length and documents where each program operated.

The geography of federally funded curricula does not coincide with the mandate arms: of the eighteen states in which a *Choosing the Best* grantee operated in fiscal year 2006, eleven never adopted a mandate of either kind, so pupils in control states also met abstinence-based curricula, though less often. The passages quoted in the introduction and in Appendix A.3 therefore show what abstinence curricula could tell adolescents, not what traditional-arm states taught relative to the control states. The curricula are evidence of what abstinence-stressed and comprehensive instruction says to adolescents, not a secondary treatment: schools developed their own curricula within the state mandate and could adopt existing ones such as these or not. Mandates bind: New Jersey rejected the Title V allotment in 2006 as inconsistent with its comprehensive health standards, and the programs it had funded were terminated.

2.3 Constructing the policy panel

Sources. I build the policy panel from primary legal sources, from the 1970s onward. For each state I trace the statute or administrative rule in force today back through its amendment history to the first version that contains each requirement, and where a statute or binding regulation makes a state board's curriculum or standards compulsory, I code content from those documents. I benchmark the result against the SIECUS State Profiles for 2003–2015 (SIECUS, 2015), the Guttmacher and KFF reviews of state requirements (Donovan, 1998; Gold and Nash, 2001; Kaiser Family Foundation, 2002), the coding in Fung and Min (2025) and the survey of state curriculum laws in Rosky (2017); where they disagree with the legal text, the text governs. Appendix A.1 describes the procedure in full.

Coding conventions. Four rules govern the classification. A mandate requires an *unconditional duty* to instruct students; permissive language, duties addressed only to state agencies or school staff, and requirements a district can waive do not qualify. Curriculum standards count only when they bind, through school approval, funds or accreditation. Abstinence is *stressed* when the text gives moral grounds for it—the expected or appropriate standard, responsible behavior, or a rule tied to marriage—and is otherwise *covered*.[5] Contraception is coded only when contraceptive methods are named as required content. Requirements are dated to the first school year in which they apply. A statute bounds rather than measures the curriculum a cohort met—public and religious private school textbooks overlap more than their

[5]The distinction between stressed and covered originates with Stanger-Hall and Hall (2011); drawing it on the grounds the text gives for abstaining is this paper's refinement, chosen because the mechanism studied here is the transmission of norms. Statements about efficacy ("the only certain way," "100 percent effective") do not qualify on their own; they count only when no contraception is required, so that abstinence is the only prevention message students receive, which is the case of Rhode Island and Kentucky.

stated purposes suggest (Adukia and Harrison, 2025), and comprehensive statutes are diluted between enactment and the classroom (Buzzell, 2024)—so the estimates are intention-to-treat. Appendix A.1.1 states the rules in full.

Policy dimensions. Two composite treatments carry the paper. **Mandatory comprehensive sex education**: sex education is compulsory *and* contraception is covered *and* abstinence is not stressed, with treatment beginning in the first year all three hold at once. **Mandatory traditional sex education**: sex education is compulsory *and* abstinence is stressed *and* contraception is not covered.[6] In practice the moral ground is mostly marriage. Seven of the nine traditional mandates invoke it: abstinence before or outside marriage, and fidelity within it, in Utah, Florida, Mississippi, North Carolina, Tennessee and Texas, and childbearing "best undertaken in marriage" in Georgia. The other two, Rhode Island and Kentucky, are traditional only because abstinence is the sole prevention message they require, with no contraception taught, and they are the weakest cases in the arm. No comprehensive mandate ties sex to marriage; California's, for instance, asks instruction to "teach respect for marriage and committed relationships".

A state can contribute to different groups for different cohorts whenever it switches mandate, but I use only the first switch.[7] States satisfying only some of the conditions—sex education mandatory, say, but abstinence not stressed and contraception not required—are neither treatment and are dropped, since putting them in the control pool would contaminate the counterfactual with partial versions of both curricula. Eight states satisfy only some of the conditions and are dropped on this rule.

The control group is the 23 states that never made sex education compulsory.[8] Never-mandate does not mean no instruction: districts in these states were free to teach sex education, and during the sample cohorts' school years fourteen of the states required HIV/AIDS instruction and seventeen required that abstinence be stressed whenever sex education was taught (Table A.1). What separates each arm from the controls is a compulsory curriculum of a given content, not the presence of instruction. Table C.2 compares the two arms with the control group before adoption.

[6] I use "traditional" rather than "abstinence-only" because what defines the arm is a moral prescription—abstinence as the expected or responsible behavior, often until marriage—rather than the exclusion of all other content, and because I also consider whether contraception is binding.

[7] Florida made traditional sex education mandatory in 1991 and removed the requirement in 2008; the switch occurred more than ten years after the first mandate and can therefore be ignored, since the event window ends nine years after adoption. Tennessee made traditional sex education compulsory in 1992 for counties with a teenage pregnancy rate above 19.5 per thousand; by 2002 over 88% of counties were still above that threshold, so the state is coded as treated. Rhode Island added a contraception requirement in 1999, eleven years after its traditional mandate; its cohorts from that year on are dropped. Wisconsin changed its mandate twice between 2007 and 2009 and is discarded.

[8] Alabama, Alaska, Arizona, Arkansas, Colorado, Connecticut, Idaho, Illinois, Indiana, Louisiana, Massachusetts, Michigan, Minnesota, Missouri, Montana, Nebraska, New York, North Dakota, Ohio, Oklahoma, Pennsylvania, South Dakota and West Virginia.

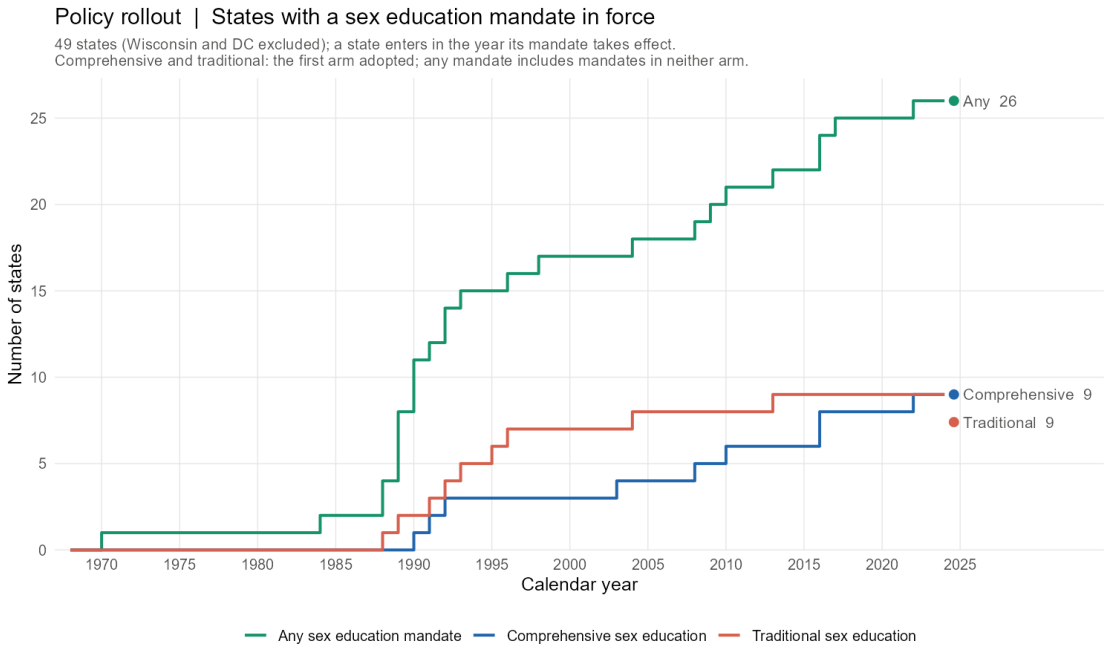


Figure 1: Sex education policy rollout: number of states with a sex education mandate in force, by calendar year.

Notes: A description of the treatment variation, not an estimate. Each line counts the states in which the mandate named in the legend is in force in that calendar year, from the policy panel of 49 states (Wisconsin and the District of Columbia are excluded); a state enters the year its mandate takes effect. *Any sex education mandate* counts every mandate, whatever its content: the two arms, the eight mandates that fit neither arm (dropped from the estimation), and the years in which a state later in an arm already had a mandate without that content (Maryland from 1970, New Jersey from 1984). Comprehensive and traditional sex education are the two arms used throughout the paper; as in the design, a state stays in the arm it adopted first and is not moved by a later repeal or switch. The state-by-state version, with the year each provision takes effect, is Table A.1. What identifies the effects is the spread of adoption dates within each arm, not the difference between the two groups of states.

Funding as a check on the coding. Federal money offers one check that the coding separates states by what they do and not only by the words of a statute, because a state can decline the Title V abstinence allotment. Across all state-years, a comprehensive mandate is associated with 0.25 (0.13) log points less Title V money and with no difference in federal comprehensive funding. Part of the gap is states declining the allotment outright: a state-year draws no Title V money 37 percent of the time under a comprehensive mandate and 1 percent under a traditional one, and 11 percent when the statute stresses abstinence against 33 percent otherwise. Within states, Title V falls when New Mexico and Maine adopt a comprehensive mandate, both times by refusal (Appendix A.4.3). The check speaks to how states handled federal money, not to what classrooms taught.

3 A Model of Sex Education and Women's Life-Course Choices

This section develops a two-period model of schooling, contraception and labor supply that organizes the empirical analysis. There are two life paths: a traditional one, in which a woman does not work and has children, and a career one, in which she works. A young woman chooses her education and her contraceptive effort before she knows how much she values the traditional path, as in two-stage models of reproductive choice (Ananat et al., 2009). An early birth closes the career path; a woman who avoids one learns her taste for the traditional path and chooses between the two. Sex education enters through two *costs*: a knowledge channel that shifts the marginal cost of effective contraception, in the tradition

of Goldin and Katz (2002), and a norm channel that shifts the cost of departing from the traditional female role, following the gender-identity literature (Akerlof and Kranton, 2000; Bertrand et al., 2015) and formalizing the "identity" element of the quiet revolution described by Goldin (2006). The model is interpretive rather than structural: it delivers signs rather than magnitudes, no parameter is estimated, and beyond signing the effects it says which findings separate the two channels. Derivations and proofs are in Appendix B.

3.1 Environment, timing and information

A woman i lives for two periods. Women differ in two independent respects: the cost of college, $k_i \sim G$, with G a strictly increasing distribution function, which a woman knows from the start; and the value she attaches to the traditional path, $\rho_i \sim U[\underline{\rho}, \bar{\rho}]$, with $\sigma \equiv \bar{\rho} - \underline{\rho}$, which she learns only as an adult. Because the two are independent, the known cost k_i carries no information about the unknown taste ρ_i. The timing is as follows; the first three steps are period 1 and the last is period 2.

1. In high school she is exposed to sex education of content $D \in [0, 1]$, a continuum from "traditional," normalized to $D = 0$, to "comprehensive," $D = 1$.
2. She chooses whether to invest in college, $s \in \{0, 1\}$, at the cost k_i, and her contraceptive effort, $e \in [0, 1]$.
3. With probability $f(e) = \bar{p}\,(1 - e)$, $\bar{p} \in (0, 1)$, she has an early birth (*derailment*), which truncates her education and leaves her on the traditional path with no further choice. The college cost is paid on every trajectory in which education was chosen, whether or not the degree was completed and whether or not it was used.
4. She learns ρ_i. If she was not derailed, she chooses between the traditional path, worth ρ_i, and the career.

The curriculum enters through two costs.

The knowledge channel. Contraceptive effort costs $c(e; D) = \gamma e^2/\big(2\kappa(D)\big)$, with $\gamma > 0$, $\kappa \in C^1$, $\kappa(0) = 1$ and $\kappa'(D) > 0$, so a more comprehensive curriculum (a higher D) flattens the marginal cost of effective contraception.[9]

The norm channel. The career is worth

$$a_s(D) \;\equiv\; w(s) - \tau(D), \tag{1}$$

where the wage $w(s)$, with $w(1) > w(0)$, depends on schooling alone and is known from period 1, and $\tau(D)$, with $\tau \in C^1$ and $\tau'(D) < 0$, is the cost of choosing the career: the prescription penalizes a woman who works rather than following the traditional path, and a more comprehensive curriculum weakens the prescription. The cost is in utility, not a tax on the wage: it changes whether she works, not what she earns when she does.

[9] The normalization $\kappa(0) = 1$ fixes units so that $D = 0$ reproduces the untreated cost $\gamma e^2/2$. Legal access (Goldin and Katz, 2002) and curricular information are, in this formulation, the same object operated by different levers.

Wages are fixed: the model is partial equilibrium and ignores any effect on $w(s)$ of the additional graduates and workers the curriculum produces. The only fertility event is an early birth in period 1; births in adulthood, and their timing, are outside the model. Section 3.2 discusses what the normalization $D = 0$ implies for the traditional arm.

The objective. Writing U_i for her period-2 payoff, the expected lifetime value at the start of period 1 is

$$V_i(e, s; D) \;=\; -\,k_i s \;-\; \frac{\gamma e^2}{2\kappa(D)} \;+\; f(e)\,\mathbb{E}[\rho_i] \;+\; \big[1 - f(e)\big]\,\mathbb{E}\big[\max\{a_s(D), \rho_i\}\big]. \tag{2}$$

Write

$$L(s; D) \;\equiv\; \max\big\{a_s(D) - \rho_i,\; 0\big\} \tag{3}$$

for the surplus of the career over the traditional path—what a derailment forfeits—and $\Lambda(s; D) \equiv \mathbb{E}[L(s; D)]$ for its expectation. Since $\max\{a_s(D), \rho_i\} = \rho_i + L(s; D)$, expected lifetime value can be rewritten as

$$V_i(e, s; D) \;=\; -\,k_i s \;-\; \frac{\gamma e^2}{2\kappa(D)} \;+\; \mathbb{E}[\rho_i] \;+\; \big[1 - f(e)\big]\,\Lambda(s; D), \tag{4}$$

and she chooses (e, s) to maximize it. The term $\mathbb{E}[\rho_i]$ depends on neither choice: what contraception protects, and what schooling raises, is the expected surplus $\Lambda(s; D)$.

Solution. Three results, derived by backward induction in Appendix B.1, carry every prediction. First, a woman who was not derailed takes the career if and only if $a_s(D) \geq \rho_i$, the college cost being sunk, so that before ρ_i is realized the probability of a career and the expected surplus are

$$P_s(D) \;=\; \frac{a_s(D) - \underline{\rho}}{\sigma}, \qquad \Lambda(s; D) \;=\; \frac{\sigma}{2}\,P_s(D)^2.$$

A graduate takes the career more often, $P_1(D) > P_0(D)$ (Figure B.1), so a weaker norm raises the stakes of avoiding derailment, $\Lambda(s; D)$, at both schooling levels and more for the graduate, who is more likely to collect it, and with them the premium to education, $\Delta\Lambda(D) \equiv \Lambda(1; D) - \Lambda(0; D) > 0$:

$$\frac{\partial \Lambda(s; D)}{\partial D} \;=\; \big(-\tau'(D)\big) P_s(D) \;>\; 0, \qquad \frac{d\,\Delta\Lambda(D)}{dD} \;=\; \big(-\tau'(D)\big)\big[P_1(D) - P_0(D)\big] \;>\; 0. \tag{5}$$

Second, optimal contraceptive effort equates the marginal cost $\gamma e/\kappa(D)$, which knowledge flattens, with the marginal benefit $\bar{p}\,\Lambda(s; D)$, which weaker norms raise:

$$e^*(s; D) \;=\; \frac{\kappa(D)\,\bar{p}\,\Lambda(s; D)}{\gamma},$$

and $f_s^*(D)$ is the resulting probability of an early birth. Third, at optimal effort college is worth $\max_e V_i(e, 1; D) - \max_e V_i(e, 0; D) = \Delta(D) - k_i$, so a woman goes to college with probability $\omega_1(D) \equiv G\big(\Delta(D)\big)$, with $\omega_0(D) \equiv 1 - \omega_1(D)$. The return to college is the premium discounted by the average equilibrium risk of derailment:

$$\Delta(D) \;=\; \big[1 - \bar{f}(D)\big]\,\Delta\Lambda(D) \;>\; 0, \qquad \bar{f}(D) \;\equiv\; \frac{f_1^*(D) + f_0^*(D)}{2}. \tag{6}$$

The probability that a woman works, $P(D)$, combines the three results (equation (19) in the appendix, where Assumption 1 keeps both choices interior). Each derivative below splits into a knowledge term, in

$\kappa'(D)$, and a norm term, in $\tau'(D)$.

3.2 Predictions

Each proposition is stated for a rise in D, the comprehensive curriculum; proofs are in Appendix B.2. The signs of Propositions 1–4 do not depend on the quadratic cost or the linear birth risk: they hold for any strictly convex effort cost and any decreasing, convex probability of an early birth (Appendix B.4).

What the model predicts for the traditional curriculum depends on the norm that prevails without a mandate. In the baseline reading that norm is itself traditional, a traditional curriculum restates it, and the traditional arm is the untreated state $D = 0$: the model predicts a null on every outcome. If the prevailing norm is instead neutral, a traditional curriculum is $D < 0$—it reinforces the traditional prescription and withholds contraceptive knowledge, $\tau(D) > \tau(0)$ and $\kappa(D) < 1$—and every sign below reverses, while the comprehensive curriculum keeps its signs with a smaller norm component, since there is less prescription to weaken (Appendix B.6). Table 1 reports both readings.

Proposition 1 (Contraception and early births)**.** *Comprehensive sex education raises contraceptive effort and lowers early births at either schooling level, through both channels:*

$$\frac{\partial e^*(s;D)}{\partial D} \;=\; \frac{\bar{p}}{\gamma}\Big[\underbrace{\kappa'(D)\,\Lambda(s;D)}_{\textit{knowledge}} \;+\; \underbrace{\kappa(D)\big(-\tau'(D)\big)P_s(D)}_{\textit{norms: the stakes rise}}\Big] \;>\; 0.$$

Contraception is worth most to a woman with a career to protect, and weaker norms raise that stake: this is the opportunity-cost view of teen childbearing (Kearney and Levine, 2012, 2014). The teen-birth response therefore reflects both channels and cannot on its own identify either.

Proposition 2 (The return to college)**.** *The return to college rises with comprehensive sex education, and so does the probability of completing college, $\omega_1(D)\big[1 - f_1^*(D)\big]$:*

$$\frac{d\Delta(D)}{dD} \;=\; \underbrace{\big[1-\bar{f}(D)\big]\,\frac{d\,\Delta\Lambda(D)}{dD}}_{\textit{norms: the premium rises}} \;+\; \underbrace{\Big(-\frac{d\bar{f}(D)}{dD}\Big)\,\Delta\Lambda(D)}_{\textit{both channels: survival rises}} \;>\; 0. \tag{7}$$

Only the second term is about pregnancy. The first is present only if norms move: the *value* of a degree rises because the path it opens has become cheaper to take.

Proposition 3 (Labor supply)**.** *Women's participation $P(D)$ rises with comprehensive sex education, and not only through schooling. Its derivative is the sum of three positive terms: a* composition *effect, as more women go to college and graduates are derailed less often and work more often ($P_1(D) > P_0(D)$); a* fewer-exits *effect, as early births fall at each schooling level; and a* direct norm *effect at given schooling, $\sum_s \omega_s(D)\big[1 - f_s^*(D)\big]\big(-\tau'(D)\big)/\sigma$.*

Proposition 4 (Family formation)**.** *Fewer women end on the traditional path under comprehensive sex education: at each schooling level the probability of a career, $\big[1 - f_s^*(D)\big]P_s(D)$, rises with D, and the probability of an early birth, $f_s^*(D)$, falls.*

Among the employed, the curriculum moves mean earnings only through composition: a working woman with schooling s earns $w(s)$ whatever D is, so mean earnings among the employed rise with D only if the share of graduates among them does, which requires college going to respond strongly enough to its return (Remark 1 in Appendix B.2).

Which findings separate the channels. Every sign in Propositions 1–4 holds with either channel alone, so the signs cannot tell knowledge from norms. Two predictions can, because they come from the norm term only. The first is the size of the college response. With knowledge alone the return to college rises only through survival, the second term of (7), and in proportion to it: completion rises because fewer of the women who invest are derailed, and investment rises by its elasticity with respect to the return times the proportional rise in survival. When early births are rare, survival rises little, so a knowledge-only account of a large college response requires schooling to be very elastic to its return. A college response large relative to the fall in early births is therefore evidence for the first term, the premium itself rising, unless that elasticity is implausibly large; Section 7.1 quantifies it. The second is attitudes. The knowledge channel leaves the prescription unchanged, so only the norm channel predicts that views about women's market work become more egalitarian (Section 7.2). A third norm term, the direct effect on participation at given schooling in Proposition 3, is not tested separately: the employment estimates combine it with composition and fewer exits.

Table 1 lists each prediction with the outcome that tests it and the section that reports the test.

		Trad., prevailing norm			
Prediction	Comp.	traditional	neutral	Outcome and data	Section
Prop. 1: contraceptive effort; early births	+; −	0	−; +	condom use (YRBS); teen births (NVSS)	7.1.4
Prop. 2: return to education	+	0	−	college degree, years of schooling (ACS)	7.1
Prop. 3: participation	+	0	−	earnings per woman, full-time, employed (ACS)	6
Prop. 4: traditional trajectory	−	0	+	married by 23 and by 30, age at first marriage (ACS)	7.1
Conditional earnings, composition only (Remark 1)	+	0	−	earnings among the employed (ACS)	6.1
Norm cost $\tau(D)$	−	0	+	agreement that a woman's place is the home (ANES)	7.2

Table 1: The model's predictions and where each is tested.

Notes: Each row pairs a comparative static of the model with the outcome that measures it in the paper; the signs are the predicted directions for the comprehensive arm, a rise in D, and for the traditional arm under the two readings of Section 3.2: a null (0) when the prevailing norm is traditional, and the reverse of the comprehensive sign when it is neutral and the traditional curriculum is $D < 0$. The conditional-earnings signs hold when college going responds strongly enough to its return (Remark 1 and Appendix B.3).

Two limits bound the interpretation. The norm cost $\tau(D)$ is not separately identified from beliefs: a formulation in which the curriculum corrects an adolescent's under-prediction of her own future attachment—Goldin's "horizon"—delivers identical comparative statics but a different welfare interpretation (Appendix B.8). And derailment is absorbing; costly re-entry would preserve every sign while compressing the magnitudes (Appendix B.7).

Men. The model is deliberately a model of women's decisions. For men the knowledge channel is absent, and the norm channel runs the other way: the male prescription is *breadwinner*, so a weaker prescription lowers the cost of *not* working, and male participation weakly falls in D (Appendix B.5). The model therefore predicts weak or ambiguous effects on men's labor supply, and makes no prediction about men's attitudes, which the same curriculum addresses directly.

4 Identification Strategy

I exploit the timing and content of the state mandates described in Section 2 to assign each birth cohort to the curriculum—comprehensive or traditional—that it was legally entitled to receive in secondary school. A two-stage difference-in-differences design compares cohorts exposed to a mandate with cohorts born in the same state that completed secondary school shortly before the mandate took effect. The coding of the two arms is described in Section 2.3.

4.1 Treatment Assignment

Treatment is defined at the birth-cohort and state level. A birth cohort c born in state s is assigned to comprehensive (traditional) sex education if the relevant mandate was in force in state s no later than the calendar year in which that cohort turned eighteen (the year it should have left secondary school):

$$\text{Compr. Sex Ed}_{cs} = \mathbf{1}[\text{Compr. Mandate Start}_s \leq c + 18] \tag{8}$$

where c is the year of birth, and Compr. Mandate Start$_s$ is the first year in which state s had all three conditions simultaneously active: mandatory sex education, contraception coverage, and no requirement to stress abstinence on moral grounds. An analogous indicator is constructed for traditional exposure. Cohorts born in states that never mandated sex education form the control group. In three comprehensive states—New Jersey, Maryland and Maine—the comprehensive onset upgrades an existing sex education mandate, while in the other five it introduces mandatory instruction. In one of these, California, the content requirement came first: from 2004 any instruction given had to cover contraception, and the onset, 2016, is the year instruction became compulsory (Table A.1). The comprehensive estimates therefore combine new content with new instruction, and with eight states the arm is too small to separate the two.

Another relevant dimension is exposure time,

$$\ell_{cs} \;=\; (c + 18) \;-\; \text{Compr. Mandate Start}_s, \tag{9}$$

the number of years between the mandate's arrival and the cohort's final year of school. A cohort at $\ell = 0$ turned eighteen in the year the mandate took effect and was therefore in its last year of secondary school when the curriculum changed; a cohort at $\ell = 4$ was fourteen at adoption and is the first to have spent all four years of high school under the new curriculum. Exposure phases in mechanically between $\ell = 0$ and $\ell = 4$ and is complete thereafter. Where a mandate covers only the lower grades, as North Carolina's runs through ninth grade, the cohorts at $\ell \geq 4$ were fourteen or younger at adoption and were taught under it (Appendix A.2).

4.2 Two-Stage Difference-in-Differences

The labor market analyses are conducted at the birth-state × birth-cohort × survey-year level. Absent a mandate, outcomes are additive in birth-state, birth-cohort and survey-year effects,

$$Y_{csy}(0) \;=\; \gamma_s + \delta_c + \eta_y + \varepsilon_{csy}, \tag{10}$$

and the estimand is the average effect of each curriculum on the cohorts exposed to it, relative to the never-mandated baseline. With staggered adoption and effects that vary across states and exposure times, a two-way fixed effects regression of Y_{csy} on the two treatment indicators estimates a weighted average of these effects that includes forbidden comparisons of later adopters with already-treated states (Callaway and Sant'Anna, 2021; Goodman-Bacon, 2021; Baker et al., 2022). I therefore use the two-stage estimator of Gardner (2022), which fits the fixed effects of (10) on untreated observations only, estimates the treatment effects on what remains, and accommodates the repeated observation of each birth-state × cohort cell across survey years.

Stage 1. Estimate the fixed effects on the untreated observations—the never-mandating states and the pre-mandate cohorts of adopting states, for which $\text{SexEd}_{cs} \equiv \max(\text{Compr. Sex Ed}_{cs}, \text{Trad. Sex Ed}_{cs}) = 0$:

$$\hat{\gamma}_s,\ \hat{\delta}_c,\ \hat{\eta}_y \ = \ \arg\min_{\gamma,\delta,\eta} \sum_{\{c,s,y:\,\text{SexEd}_{cs}=0\}} \left(Y_{csy} - \gamma_s - \delta_c - \eta_y\right)^2. \tag{11}$$

As a robustness check—and in particular when the pre-trend diagnostic fails—I re-estimate adding birth-state linear trends, one slope per state in the birth cohort, so that a state whose outcomes were already drifting across successive cohorts has that drift absorbed before the treatment coefficients are estimated. The trend runs in *cohort* time (the interaction of a birth-state dummy with year of birth). Appendix C.2 gives the specification.

Stage 2. Compute the residual $\tilde{Y}_{csy} = Y_{csy} - \hat{\gamma}_s - \hat{\delta}_c - \hat{\eta}_y$ for all observations, then estimate the event study:

$$\tilde{Y}_{csy} = \sum_{\ell \neq -1} \beta_\ell^C \, \mathbf{1}[\ell_{cs}^C = \ell] + \sum_{\ell \neq -1} \beta_\ell^T \, \mathbf{1}[\ell_{cs}^T = \ell] + \varepsilon_{csy},$$

where ℓ_{cs}^C and ℓ_{cs}^T are the exposure times of (9) for the comprehensive and the traditional mandate, and $\ell = -1$, the last unexposed cohort, is the omitted event time, plotted at zero. Because the first stage is fit on every untreated observation, each coefficient is the mean first-stage residual at that event time, so the pre-adoption coefficients measure deviations from each state's untreated history as a whole rather than from $\ell = -1$, as in the imputation estimator of Borusyak et al. (2024). Standard errors account for the two-stage procedure as in Gardner (2022). Both arms are estimated jointly, against a common comparison group: the never-mandate states and the cohorts of adopting states that left school before their mandate, which enter the first stage with roughly equal weight. With eight and nine identifying states, fitting the arms together disciplines the nuisance fixed effects with the whole panel rather than re-estimating them separately for each treated group.

The event study is estimated on $\ell \in [-7, 9]$ but reported on $[-6, 8]$, the end points being accumulation bins spanning over decades. The ATT windows are the equal-weighted mean of the reported dynamic coefficients, with delta-method standard errors:

$$\widehat{\text{ATT}}[\underline{\ell}, \bar{\ell}] \ = \ \frac{1}{\bar{\ell} - \underline{\ell} + 1} \sum_{\ell=\underline{\ell}}^{\bar{\ell}} \hat{\beta}_\ell. \tag{12}$$

Because exposure phases in over four years and implementation may itself have been gradual, I report the fully exposed window $[4, \bar{\ell}]$ throughout alongside $[0, \bar{\ell}]$. The states behind each event time differ: California and Hawaii, which adopt in 2016, enter only at $\ell = 0$ and 1, and the fully exposed window rests on the six states that adopted by 2010.

4.3 Contemporaneous Outcomes

Teen births, adult first births and per-pupil spending are recorded for a state either in a calendar year or for a cohort, so there is only one time dimension. For these I use the group-time estimator of Callaway and Sant'Anna (2021): each state adopting in year g is compared with the never-treated states, period by period, and the resulting $ATT(g, t)$ aggregated into an event study with a universal base period and then

into the windows of (12).[10] Observations are weighted by the relevant population—female population of the relevant ages for births, average daily attendance for spending—and standard errors are clustered on state.

4.4 Identifying Assumptions

Identification requires four assumptions. *Parallel trends*: absent treatment, outcomes evolve according to additive state, cohort and survey-year effects, which excludes unit-specific trends and time-varying confounders correlated with adoption timing; the trend specification relaxes this to allow a linear trend in cohort for each birth state. *No anticipation*: cohorts do not adjust before a mandate arrives, which here would require anticipatory relocation across states. *Rank*: enough untreated observations to estimate the first-stage fixed effects, satisfied by the never-treated states and pre-reform cohorts. *SUTVA*: treatment in state s does not affect cohorts born elsewhere, so that post-entry migration generates no spillovers. The first two are assessed by event-study pre-trends and by the in-time placebo of Section 4.5. Additive effects do not absorb shocks specific to a birth state at the time of observation or a nonlinear age profile. Adding age dummies to the first stage leaves every comprehensive estimate within a tenth of a point of the baseline; adding birth-state $\times$ survey-year effects as well keeps them positive but smaller: among the fully exposed, 4.3 (0.9) log points for earnings, 2.1 (0.7) points for full-time work and 0.6 (0.3) for employment, against 5.2, 2.5 and 0.9 at baseline (Appendix C.2).

Main confounding concern. The central threat is that mandates coincide with broader educational reforms affecting women. Fixed effects absorb common trends at the state, cohort, and survey year level; because the two arms are regionally concentrated, I also add census-region $\times$ birth-cohort effects (Section 6.2); and the comparison across genders separates the channels: human capital would move men and women alike, a gender-norm or fertility channel primarily women. Section 8 takes up the specific candidates—per-pupil spending, which does not break at adoption, and state legislative composition, which shifts gradually rather than discretely, and access to reproductive control tools. Section 8.3 also considers state policies that reached the same cohorts later, as young adults.

4.5 Inference and Sensitivity with Few Treated States

Nine states ever adopt a comprehensive mandate and nine a traditional one, and not all contribute identifying variation: a mandate arriving after 2017 has no cohort old enough to observe in the labor market, so the ACS design rests on eight comprehensive and nine traditional identifying states, the natality data on adult births on three comprehensive states, and the survey data on sexual behavior on four. Cluster-robust standard errors are badly sized in this regime (MacKinnon and Webb, 2018). The main tables report the analytic standard error in parentheses and the randomization-inference p-value in brackets beneath it; the stars refer to the analytic standard errors, and the text reads the bracketed p-values, treating an estimate as significant when its randomization p-value is below 0.10. Every estimate, in either arm, is read against the same three diagnostics: the randomization p-value, the average pre-adoption coefficient and the in-time placebo described below.

For each estimate I permute the entire policy *bundle* attached to a state—the year it first mandated

[10] Estimation is by regression adjustment rather than the doubly robust score; with treated groups this small and an unbalanced panel the doubly robust estimator is numerically unstable, and since the main specification carries no covariates the two coincide in any case.

sex education and the type—across states in 500 draws, refitting the specification each time. This holds the number of adopters of each type and the calendar timing of adoption fixed, so the permuted rollout retains the structure of the real one. The test is of the sharp null that the mandates had no effect in any state, under which every such reassignment of the observed bundles is equally likely. The p-value is the share of draws whose estimate is at least as large in absolute value as the real one. Because states differ in size and the regressions are weighted, the permutation test is not exact—a test built on the coefficient can over- or under-reject when treated clusters differ in size (Ferman and Pinto, 2019; MacKinnon and Webb, 2020)—but it does not rely on the number of treated clusters growing, as the analytic standard errors do.

Four further checks discipline the estimates. I summarize the pre-adoption coefficients by their average over $[-6, -2]$ to evaluate pre-trends, but the diagnostic is not used as a gate (Roth, 2022). Where it fails I add birth-state linear trends, reported beside the baseline rather than in place of it. For significant estimates I report the relative-magnitudes breakdown value $\bar{M}$ of Rambachan and Roth (2023): the largest post-adoption violation of parallel trends, as a multiple of the largest pre-adoption one, under which the estimate stays significant. It falls as the window lengthens, because the bound allows a separate violation at each post-adoption horizon. And every adoption date is moved seven years earlier and the design re-estimated: because treatment is anchored at eighteen, the cohorts this exercise labels treated were past school age when the real mandate arrived, so a placebo of comparable size indicates a trajectory rather than a break at exposure. The placebo ATT averages event times 0 to 6, the horizons that lie entirely before the true adoption, so that no genuinely treated cohort enters it. For the biennial YRBS waves and state legislative elections the shift is six years, three waves or three elections, because an odd shift would place the fake event times between waves or elections; the placebo window is then $[0, 5]$, the six years from the fake date to the true one, which hold the waves at event times 0, 2 and 4 and the three elections before the true adoption. The reference period is the last unexposed cohort, $\ell = -1$, everywhere except in the YRBS, where waves are two years apart and the last wave before adoption is $\ell = -2$, and in the ANES, whose event time is grouped into two-year bins so that the reference is the bin $[-2, -1]$.

A state-clustered bootstrap with 999 replications is reported alongside as a diagnostic rather than as corrected inference, considering the small number of clusters. Appendix C.2 sets out the construction of each check.

5 Data

The treatment for every result in the paper is the policy panel of Section 2.3. This section describes the outcome data. The principal source is the American Community Survey, which supplies the labor market, schooling and marriage outcomes. Four further sources—natality microdata, the Youth Risk Behavior Surveillance System, the American National Election Studies, and two administrative finance files—are each used for one purpose only and are introduced where they are needed.

5.1 American Community Survey

The outcome data come from the 2000–2024 American Community Survey (U.S. Census Bureau, 2026), using the public-use microdata distributed by IPUMS-USA (Ruggles et al., 2024). The sample covers individuals born between 1950 and 2004 and observed between the ages of 18 and 59, the upper bound removing the retirement margin.

The design operates on cells rather than individuals: I collapse the microdata to cells defined by birth state s, birth cohort c and survey year y, each holding the weighted mean of the outcome using ACS person weights. Cells with fewer than 20 respondents are dropped, which removes the District of Columbia entirely. Because every outcome is measured from age 23 at the earliest and most from 25, the estimated panel is substantially smaller than the collapsed one: 28,817 cells for women and 28,873 for men, across 41 states—nine comprehensive, nine traditional and the 23 never-mandating states as controls—50 cohorts and 25 survey years, resting on 8.5 million women and 8.5 million men, with a median of about 188 respondents per cell. Table C.1 in Appendix C.1 reports the sample at each stage and discusses the weights.

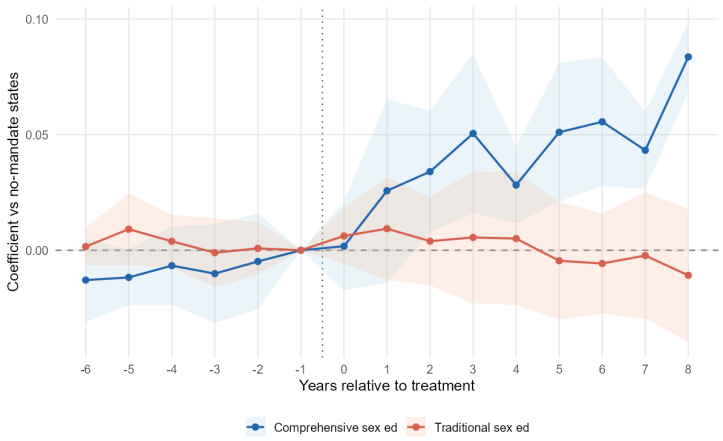
Outcomes. Three labor market outcomes carry the main results, all measured from age 25 so that education is complete or nearly so and current enrollment is not mistaken for non-employment. The main outcome is **log annual earnings**, the log of mean wage and salary income in the previous twelve months in the cell, with zeros for women who had none. Two further outcomes decompose it along the margins of labor supply: **full-time**, an indicator for having usually worked thirty-five hours a week or more in the previous twelve months, again with zeros for women who did not work; and **employed**, an employment-to-population ratio in which those out of the labor force count as zero.

Assigning zeros to women without earnings combines the extensive and intensive margins by construction, which is deliberate: mean earnings per woman is the share with earnings times mean earnings among them, so the effect on it splits into the two, and Section 6.1 performs that decomposition; because employment is measured in the survey week and earnings over the previous year, the split is approximate. The versions of full-time work and earnings conditional on working are reported alongside in the appendix. The same extract supplies the mechanism outcomes of Section 7.1—completed college by 25, years of schooling, and the marriage-timing measures, including marriage by 23—which are defined there.

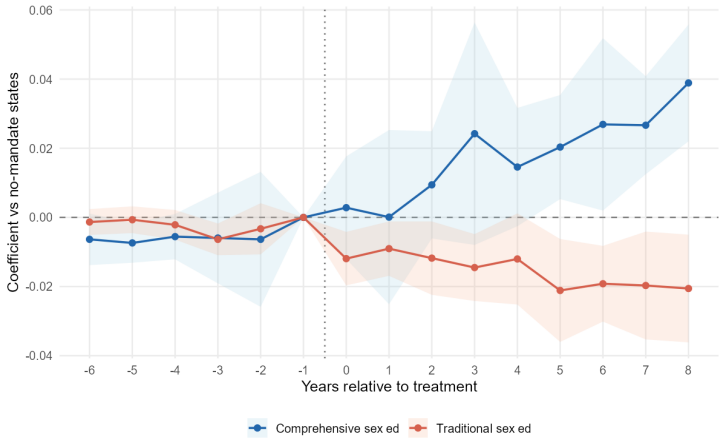
6 Results: Women's Labor Market Outcomes

Contrasting states with *any* sex education mandate against states with none, the comparison the literature on teen outcomes runs, produces nothing on any labor market outcome: the event studies are flat throughout (Figure C.1 in Appendix C.3). That comparison pools mandates whose content points in opposite directions; Section 6.3 returns to the mandates that require neither content.

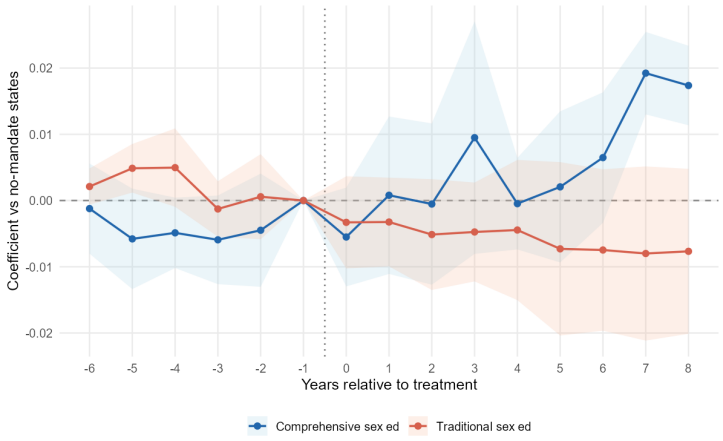
Separating the arms, women who reached the end of secondary school under a comprehensive mandate earn more and work full-time more often; women who reached it under a traditional mandate earn no differently, and the declines in their full-time work and employment do not survive the diagnostics below. Figure 2 shows the event studies and Table 2 the two ATT windows, each in the baseline specification and with birth-state linear trends.



(a) Log annual earnings (unconditional)



(b) Working full-time (unconditional)



(c) Employed

Figure 2: Women's labor market outcomes: event studies by sex education type.

Notes: Outcomes, top to bottom, are the log of mean annual wage and salary income per woman with zeros for those without earnings in the previous year, the share working thirty-five hours or more, and the share of women employed. The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the ACS 2000–2024 from age 25 and collapsed to birth-state × birth-cohort cells weighted by ACS person weights. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects and no state trends; the horizontal axis is the cohort's distance from the first cohort required to take the mandated curriculum, and the coefficient at −1 is omitted and plotted at zero (Section 4). Each arm—comprehensive and traditional sex education—is estimated against the same pool of never-mandate states. Shaded bands are 95 percent confidence intervals built from standard errors clustered on birth state. The same event studies with birth-state linear trends are in Figure C.2 and the in-time placebo, which moves every adoption date seven years earlier, in Figure C.3. Randomization-inference p-values for ATT $[0, 8]$ and $[4, 8]$, comprehensive then traditional: log earnings 0.014 and 0.010, 0.952 and 0.749; full-time 0.072 and 0.032, 0.038 and 0.036; employed 0.309 and 0.138, 0.176 and 0.136.

	Log earnings (all)		Full-time (all)		Employed	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
No state trends						
ATT $[0, 8]$	0.042^{***}	0.001	0.018^{**}	-0.016^{***}	0.005	-0.006
	(0.008)	(0.010)	(0.009)	(0.005)	(0.004)	(0.005)
	[0.014]	[0.952]	[0.072]	[0.038]	[0.309]	[0.176]
ATT $[4, 8]$	0.052^{***}	-0.004	0.025^{***}	-0.019^{***}	0.009^{**}	-0.007
	(0.008)	(0.012)	(0.009)	(0.007)	(0.004)	(0.006)
	[0.010]	[0.749]	[0.032]	[0.036]	[0.138]	[0.136]
With state trends						
ATT $[0, 8]$	0.047^{**}	0.002	0.030^{***}	-0.005	0.021^{**}	-0.010^{**}
	(0.019)	(0.007)	(0.008)	(0.003)	(0.009)	(0.004)
	[0.006]	[0.844]	[0.004]	[0.591]	[0.024]	[0.198]
ATT $[4, 8]$	0.056^{**}	-0.000	0.038^{***}	-0.006	0.026^{**}	-0.012^{**}
	(0.022)	(0.007)	(0.009)	(0.003)	(0.011)	(0.005)
	[0.006]	[0.982]	[0.006]	[0.605]	[0.024]	[0.224]
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states	8	9	8	9	8	9
Control states	23	23	23	23	23	23
Observations	28,817	28,817	28,817	28,817	28,817	28,817
Pre-adoption mean	10.349	10.159	0.595	0.623	0.725	0.706

Table 2: Women's labor market outcomes: average ATT, with and without birth-state linear trends.

Notes: Outcomes are the log of mean annual wage and salary income per woman with zeros for those without earnings in the previous year, the share working thirty-five hours or more, and the share of women employed. The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the ACS 2000–2024 from age 25 and collapsed to birth-state × birth-cohort cells weighted by ACS person weights. Estimated by two-stage difference-in-differences; each arm—comprehensive and traditional sex education—is estimated against the same pool of never-mandate states. ATT $[0, 8]$ averages the dynamic coefficients over the first nine post-adoption cohorts and ATT $[4, 8]$ over the cohorts that spent all four years of high school under the mandate. The upper block carries only the fixed effects listed at the foot; the lower block adds a linear trend in birth cohort for each birth state. Standard errors, clustered on birth state, in parentheses; randomization-inference p-values in brackets, from 500 draws that re-assign each state's policy bundle across states. Earnings and full-time work among the employed are in Table C.5, HonestDiD breakdown values and the in-time placebo in Table C.3, and alternative specifications in Table C.6. Stars refer to the analytic standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.1 Earnings, full-time work and employment

The main outcome is annual earnings per woman: the log of mean wage and salary income, with women who had no earnings in the previous year counted as zeros, so that both margins of labor supply enter. Exposure to a comprehensive mandate raises it by 4.2 (0.8) log points, about $1,300 on a pre-adoption mean of $31,000, and by 5.2 (0.8) among women who spent all of high school under the mandate. The pre-adoption coefficients are small, the in-time placebo is close to zero, and adding birth-state trends leaves the estimate at 4.7 (1.9). The traditional arm is a precise zero. For comparison, early legal access to the pill raised women's hourly wages by 8 percent by their late forties (Bailey et al., 2012).

Full-time work is where the two curricula diverge most. Comprehensive exposure raises the share of women working thirty-five hours or more by 1.8 (0.9) percentage points, from 59.5 percent, and by 2.5 (0.9) among the fully exposed. The adopters' pre-period drifts slightly downward, so birth-state trends raise the estimates to 3.0 and 3.8 points. In the traditional arm full-time work falls by 1.6 (0.5) points, but the decline disappears with trends; Section 6.5 returns to it.

Employment responds less. It rises by 0.5 (0.4) percentage points across post-adoption cohorts and by 0.9 (0.4) among the fully exposed, from 72.5 percent, estimates that randomization inference does not confirm. As with full-time work, the adopters were drifting downward before exposure. Once trends

absorb that drift, the estimates rise to 2.1 (0.9) and 2.6 (1.1) points, and randomization inference confirms them.

Mean earnings per woman are the share with earnings times mean earnings among them, so the earnings effect splits, approximately, into the two margins. With trends, the employment response of 2.1 points on a base of 72.5 percent is worth 2.9 log points, and earnings among the employed rise by 2.2 (1.0) log points (Table C.5). Roughly half of the gain therefore comes from women who would not otherwise have worked, and half from higher earnings among women who would have worked anyway, who also work full time more often. The model of Section 3 accounts for the first half, and for the second only through the schooling of the employed (Remark 1). The rest is what later marriage produces when women accumulate experience before forming a family (Section 7.1.5).

6.2 Robustness

The results hold up across the checks in Appendix C. Randomization inference, the cluster bootstrap and the HonestDiD sensitivity analysis support the earnings estimate throughout, and the full-time and employment estimates most clearly with birth-state trends (Table C.3).[11] Alternative control pools and a balanced event-time window leave all three estimates within a standard error of the baseline (Table C.6).

Because comprehensive adopters are concentrated in the Northeast and the West and traditional adopters in the South, I also compare each adopter only with states in its own census region, by adding census-region $\times$ birth-cohort effects. Earnings then rise by 3.1 (1.0) log points and full-time work by 1.1 (0.6) points, and employment is unchanged (Table C.6).

At baseline, the earnings estimate does not depend on any single state: dropping each treated state in turn keeps it between 3.5 and 4.6 log points (Table C.7). The full-time and employment estimates lean on New Jersey, the most populous early adopter. Without it, or with all states weighted equally, they lose significance at baseline, although the full-time estimate with trends survives.

The effects do not fade with age. On the three early adopters, which are observed at every age from 30 to 45, the point estimates for earnings and full-time work are positive at every age from 34 to 45, with no drift toward zero (Figure C.5); restricting the sample to ages 30 to 40 gives estimates close to the baseline (Table C.8). A flat profile is what one expects of an effect that works through full-time work, where experience accumulates and interruptions leave lasting gaps (Blundell et al., 2016; Bertrand et al., 2010).

6.3 Mandates that fit neither arm

Eight states made sex education compulsory without requiring either curriculum, and the main design drops them. Estimated as their own group against the same never-mandate states, with the comprehensive and traditional adopters removed, they show no effect on women's earnings, -0.0 (0.7) log points, full-time work, -0.5 (0.4) points, or employment, -0.1 (0.3) points, nor on teen births, -1.4 (1.3) log points.[12] The six whose mandate requires no content at all—Delaware, Iowa, Kansas, Nevada, New Hampshire and Wyoming—show nothing on any of these outcomes. Making sex education compulsory does not, by itself, move them.

[11] Across the eight main outcomes, every comprehensive estimate except employment remains significant at a ten percent false discovery rate (Table C.4).

[12] Labor market outcomes are estimated with the baseline two-stage design and teen births with the Callaway–Sant'Anna design of Section 4.3, over ATT $[0, 8]$; standard errors are clustered on state.

The other two, South Carolina and Virginia, have abstinence-plus mandates: they require contraception to be taught but stress abstinence until marriage on moral grounds, so they come closest to holding contraceptive instruction fixed while changing its framing. Women's earnings, full-time work and employment do not move there, 0.3 (0.6) log points, −0.4 (0.4) and 0.1 (0.2) points, but teen births fall by 2.8 (1.4) log points. With two states, both in the South and both adopting in 1990, this is descriptive, but it is the pattern the model predicts if contraceptive knowledge lowers early births while the labor market gains require a weaker norm against women's market work (Section 3).

6.4 Men and the gender gap

Men's labor market outcomes barely move. Their earnings and full-time work are unchanged, and their employment falls by 0.8 (0.3) points, just short of significance under randomization inference (Tables C.11 and C.12). A small decline is what the model predicts for men: a weaker breadwinner prescription lowers the cost of not working (Section 3.2). That women gain and men do not is what separates the norm channel from a general improvement in schooling in the adopting states, which would raise the outcomes of both.

Men's college completion does rise, though by less than women's (1.6 points), so men's outcomes alone cannot rule out a schooling change common to both sexes. The gap between men and women born in the same state and year addresses this, because it nets out anything that moved a birth state and cohort as a whole. In the comprehensive arm the gap narrows on all three margins: among the fully exposed, by 5.7 (2.4) log points in earnings, 2.9 (1.2) points in full-time work and 1.8 (0.7) points in employment, with a flat pre-period and a null placebo (Table C.9). Randomization inference confirms the narrowing of the full-time gap at baseline and of all three gaps with trends (Table C.10). In the traditional arm the gaps were already widening before adoption.

Together these deliver Proposition 3, and they account for the null the section opened with: pooling all mandates hides a positive comprehensive effect next to a traditional arm with no effect on earnings.

6.5 The traditional arm

The traditional arm has no effect on earnings, but its estimates for full-time work and employment are negative, and some are significant under randomization inference. At baseline, exposure lowers the share of women working full time by 1.6 (0.5) percentage points (Table 2). Among women aged 30 to 40 it also lowers employment and earnings (Table C.8), and the gender gap in full-time work widens (Table C.10). These are the signs the model predicts if the prevailing norm is neutral and a traditional curriculum reinforces the traditional prescription (Section 3.2).

The diagnostics, however, point to trajectories that predate the mandates. The full-time decline disappears with birth-state trends, and when each adopter is compared only with states in its census region, and its in-time placebo is two fifths of its size. In the 30–40 band the placebos are nearly as large as the estimates, and the gender gaps were already widening before adoption. Women in the traditional states, most of them in the South, were losing ground on full-time work before exposure, and the estimates after adoption continue that drift. I therefore do not read them as effects of the curriculum, although the design cannot rule out a small negative effect on full-time work.

7 Mechanisms

The model runs from contraception and the expected return to a career, through early births and schooling, to women's work. Section 7.1 shows that exposed women complete college more often and marry after the college years, sets out the weaker evidence on births and sexual behavior, and argues that this sequence is a plausible route to the labor market results. Section 7.2 reports gender attitudes.

7.1 A different life course: college, marriage and fertility

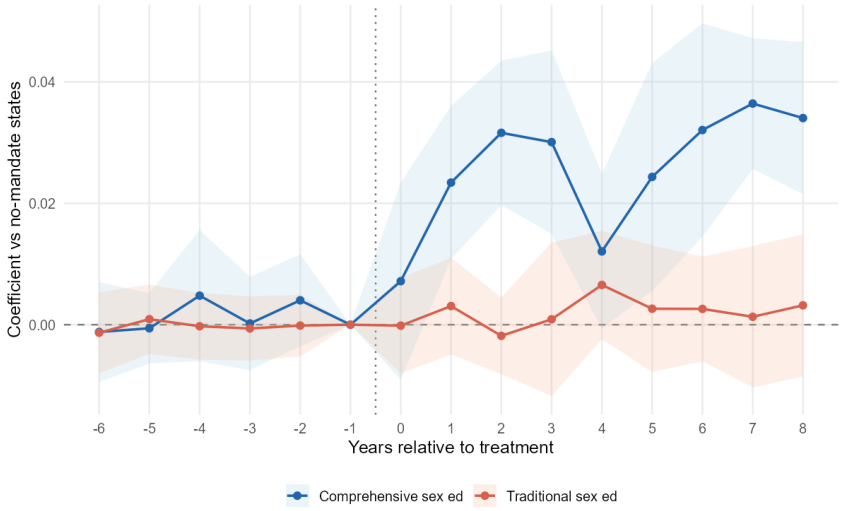
7.1.1 Data and specification

Proposition 2 predicts that education rises with comprehensive sex education, and Proposition 4 that fewer women end on the traditional path, which I measure by early marriage. Both are estimated on the ACS extract of Section 5.1 with the cohort design of Section 4.

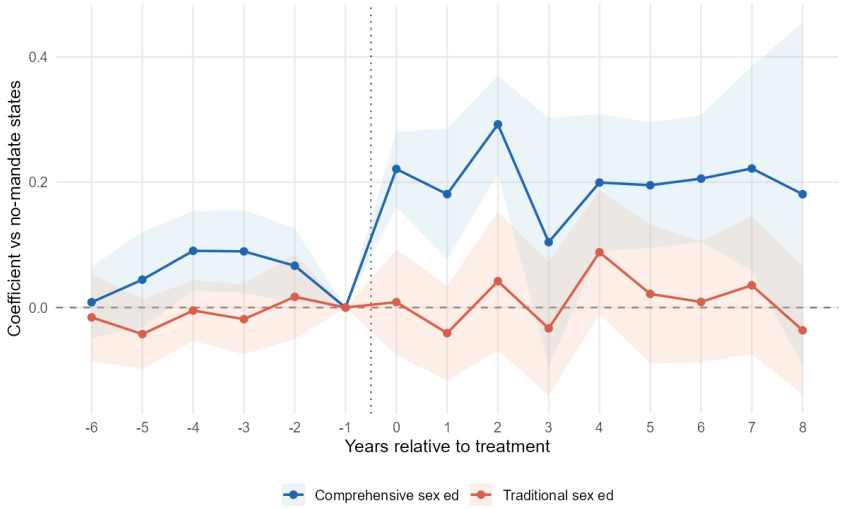
College degree is the share of women aged 25 or older holding at least a bachelor's degree; *years of education* is its continuous counterpart. For the timing of marriage I use the *restricted mean age at first marriage censored at 30*: the average number of years a woman spends unmarried between 15 and 30, so that a woman unmarried at 30 contributes 30 and no cohort is dropped for not yet having married. It comes from the ACS question on the year of marriage, asked from 2008.[13] *Married by 23* is the share of women married by 23, the age by which a four-year degree begun after high school is normally complete. Because it is observed at a younger age, it covers more cohorts and is identified by eight comprehensive states rather than six. The appendix also reports marriage by 30, the age at marriage among women married by 30, and the same measures censored at 35. The marriage measures are estimated in the baseline specification only, because they rest on a thin post-adoption panel (Appendix C.2).

[13] Age at first marriage is the year of marriage minus the year of birth for women married once. A woman married more than once reports only her most recent marriage; if it took place by the threshold, her first did too, and otherwise she is left out. Coding all of these women as married after the threshold, or all as married by it, leaves the comprehensive estimates of the same sign and significance (Tables D.5 and D.6).

7.1.2 Education



(a) College degree (ages 25 and over)



(b) Restricted mean age at first marriage, censored at 30

Figure 3: Education and the timing of marriage: event studies by sex education type.

Notes: Outcomes are the share of a birth-state × birth-cohort cell holding at least a bachelor's degree, measured from age 25 (panel a), and the restricted mean age at first marriage censored at 30—the average years a woman spends unmarried between 15 and 30, built from the year of first marriage (panel b). The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 and weighted by ACS person weights. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects and no state trends, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; the horizontal axis is the cohort's distance from the first cohort required to take the mandated curriculum. Shaded bands are 95 percent confidence intervals from standard errors clustered on birth state. The alternative measures—years of education, married by 30, age at marriage conditional on marrying by 30, and married by 23—are in Figure D.3. Randomization-inference p-values for ATT $[0, 8]$ and $[4, 8]$, comprehensive then traditional: college degree 0.008 and 0.016, 0.786 and 0.685; restricted mean age at first marriage 0.032 and 0.070, 0.876 and 0.731.

	College degree		Age at marriage (rest. mean, 30)	
	Comp.	Trad.	Comp.	Trad.
No state trends				
ATT $[0, 8]$	0.026***	0.002	0.200***	0.010
	(0.006)	(0.004)	(0.058)	(0.043)
	[0.008]	[0.786]	[0.032]	[0.876]
ATT $[4, 8]$	0.028***	0.003	0.201***	0.024
	(0.006)	(0.005)	(0.071)	(0.048)
	[0.016]	[0.685]	[0.070]	[0.731]
With state trends				
ATT $[0, 8]$	0.020**	0.000	—	—
	(0.009)	(0.005)		
	[0.010]	[1.000]		
ATT $[4, 8]$	0.020*	0.001	—	—
	(0.011)	(0.005)		
	[0.048]	[0.918]		
Fixed effects	Birth state, birth cohort, survey year			
Controls	None			
Weights	ACS person weights			
Treated states	8	9	6	8
Control states	23	23	23	23
Observations	28,817	28,817	18,179	18,179
Pre-adoption mean	0.339	0.279	26.318	25.425

Table 3: Education and family formation: average ATT for all cohorts and for fully exposed cohorts.

Notes: Outcomes are the share of a birth-state × birth-cohort cell holding at least a bachelor's degree, measured from age 25, and the restricted mean age at first marriage censored at 30. The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 and weighted by ACS person weights. Estimated by two-stage difference-in-differences, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. ATT $[0, 8]$ averages the dynamic coefficients over all post-adoption cohorts and ATT $[4, 8]$ over the cohorts that spent all four years of high school under the mandate. The upper block carries only the fixed effects listed at the foot; the lower block adds a linear trend in birth cohort for each birth state, and "—" marks a cell that is deliberately not estimated, since the marriage measures are reported in the baseline specification only (Section 7.1.1). Standard errors, clustered on birth state, in parentheses; randomization-inference p-values in brackets, from 500 draws that re-assign each state's policy bundle across states. The alternative measures of the same two margins are in Table D.5 and Figure D.3, the same measures censored at 35 and marriage by 23 in Table D.6, and HonestDiD and the in-time placebo in Table D.7. Stars refer to the analytic standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Comprehensive exposure raises the share of women with a bachelor's degree by 2.6 (0.6) percentage points, from 33.9 percent, a relative increase of about eight percent. The pre-period is flat, the in-time placebo is zero, birth-state trends change little, 2.0 (0.9), and years of education rise by 0.11 (0.04). No single state carries the estimate. Comparing each adopter only with states in its census region lowers it to 1.7 (1.0) points, and in that specification the traditional estimate turns positive too, so the college response is less precisely pinned down than the earnings result. Its magnitude is comparable to that of Title IX, which Stevenson (2010) estimates raised the share of women with at least four years of college by about two percentage points.

This is Proposition 2, and averted teen births cannot account for it on their own. At the pre-adoption rate of about 41 births per thousand girls a year, the estimate on data to 2019, 4.3 log points, removes at most nine births per thousand over the five teen years. The share of women who reach 20 without a birth therefore rises by at most 0.9 percentage points, about one percent. With contraceptive knowledge alone, completion rises by that one percent among the women who invest, and investment by its elasticity with respect to the return times the same one percent (Section 3.2); a rise of about eight percent would require an elasticity above five. The bound leaves out births in the early twenties, which the curriculum may also have averted, so the argument is indicative. It points to the first term of (7): the premium itself rising as the norm cost of a career falls. In the traditional arm the college estimate is a precise zero, 0.2

(0.4), with a flat pre-period. Men's college completion rises by 1.6 points, so part of the response may be a schooling change common to both sexes (Section 7.1.5).

7.1.3 Marriage

Marriage is the outcome the statutes speak to most directly. Seven of the nine traditional mandates prescribe abstinence until marriage, and no comprehensive mandate ties sex to marriage (Section 2.3). Marriage is not a choice in the model, but early marriage is the clearest marker in the data of the traditional path of Proposition 4, and its timing matters for women's careers through the timing of first births (Miller, 2011; Troske and Voicu, 2013).

Under comprehensive mandates marriage moves in the predicted direction, as Lazuka and Elwert (2026) find in Sweden. The restricted mean age at first marriage rises by 0.20 (0.06) years, about ten weeks, but its pre-adoption coefficients have the same sign, so part of the movement may continue an earlier one. Splitting the measure shows where it comes from. The share of women married by 30 falls by 2.5 (1.2) points, but this component fails the pre-trend test. The age at marriage among women who marry by 30 rises by 0.22 (0.04) years, with a flat pre-period and a null placebo. Exposure therefore delays marriage among women who marry young rather than preventing it.

The delay falls on the college years. Exposure lowers the share of women married by 23 by 2.0 (0.4) percentage points, from 27.1 percent, a fall of about seven percent (Table D.6). The pre-period drifts slightly downward, by about a quarter of the estimate, but randomization inference confirms the fall, and comparing each adopter only with states in its census region leaves 1.2 (0.5) points. A fall in marriage before 23 alongside a rise in college completion is what one would see if exposed women increasingly finished school before marrying.

The traditional mandates, the ones that speak of marriage, leave it unchanged. The share of women married by 23 moves by -0.2 (0.5) points, from 38.4 percent, and the restricted mean age at first marriage by 0.01 (0.04) years (Tables 3 and D.6); the other measures are small, and their in-time placebos are as large as the estimates or larger (Table D.7). Two features of the setting help explain the null. Women in the traditional states already married young before adoption—38 percent by 23, against 27 percent in the comprehensive states—so the mandates prescribed a path many of them already followed, the case in which the model predicts no effect. And many never-mandate states impose the same prescription whenever sex education is taught, so part of the comparison is between the same message delivered with a mandate and without one.

7.1.4 Fertility and sexual behavior: teen births, contraception and first births in adulthood

If exposed women finish school before they marry, their first births should arrive later. Teen births are the first link in the model's chain; first births in adulthood lie outside the model and are descriptive evidence on timing. Teen births come from NVSS natality files for 1980–2024 (National Center for Health Statistics, 2026), by the mother's state of residence, with Census population denominators (United States Census Bureau, 2021a,b). The outcome is the log of births per girl aged 15–19, with births dated by conception, so that a birth conceived before the mandate counts as pre-adoption. Adult first births are the log first-birth rate per woman by five-year age group, from NVSS microdata for 1980–2004 coded by the mother's state of birth (National Center for Health Statistics, 2024), and, for the later adopters that series cannot follow, from the 2000–2024 ACS by the own-children method (Appendix D.4.1). All are state-by-year panels estimated with the Callaway–Sant'Anna design of Section 4.3, weighted by the

female population at risk.

	log(teen birth share) Ages 15–19		log(teen birth share) Ages 15–17	
	Comp.	Trad.	Comp.	Trad.
Synthetic DiD				
ATT $[0, 8]$	-0.096^{***}	0.017	-0.123^{***}	0.006
	(0.024)	(0.017)	(0.028)	(0.023)
	[0.005]	[0.313]	[0.005]	[0.781]
ATT $[4, 8]$	-0.113^{***}	0.022	-0.160^{***}	0.012
	(0.032)	(0.024)	(0.037)	(0.031)
	[0.005]	[0.428]	[0.005]	[0.677]
Callaway–Sant'Anna				
ATT $[0, 8]$	-0.068^{*}	0.006	-0.098^{**}	0.005
	(0.035)	(0.013)	(0.041)	(0.013)
	[0.022]	[0.840]	[0.014]	[0.860]
ATT $[4, 8]$	-0.076^{*}	0.009	-0.127^{**}	0.009
	(0.040)	(0.019)	(0.050)	(0.019)
	[0.054]	[0.818]	[0.014]	[0.856]
Fixed effects	State and year			
Controls	None			
Weights	Teen female population			
Treated states	9	9	9	9
Control states	23	23	23	23
Observations	1,440	1,415	1,440	1,415
Pre-adoption mean	−3.192	−2.701	−3.787	−3.149

Table 4: Log teen birth share: average ATT by sex education type and age band.

Notes: The outcome is the log of births per girl in the age band named in the column head, dated to the year of conception, in a state-by-calendar-year panel, 1980–2024, of 49 states (Wisconsin and the District of Columbia excluded), from NVSS natality files by the mother's state of residence with Census population denominators. The upper block is the staggered synthetic difference-in-differences (Arkhangelsky et al., 2021), estimated separately for each adopting state at its own adoption year with the never-mandate states as donors and averaged by event year with teen-population weights; its standard errors are the standard deviation over 200 draws that give the adopters' adoption dates to randomly chosen donor states (Clarke et al., 2024). The lower block is the Callaway–Sant'Anna group-time estimator with never-treated controls, a universal base period and no covariates, weighted by the teen female population, with standard errors clustered on state. ATT $[0, 8]$ averages the first nine post-adoption years and ATT $[4, 8]$ the fully exposed ones. In brackets, the permutation p-value from the 200 placebo draws for the synthetic estimator and the randomization-inference p-value from 500 draws that re-assign adoption dates across states for Callaway–Sant'Anna. The in-time placebos and HonestDiD breakdown values are in Table D.1, the alternative specifications, including the data to 2019 and without California, in Table D.2, and the Callaway–Sant'Anna event study is Figure D.1. Stars refer to the standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

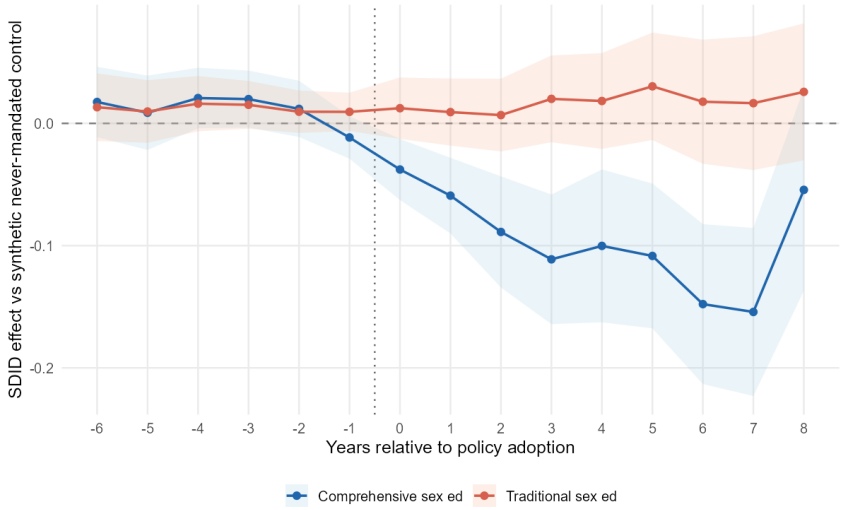


Figure 4: Log teen birth share, ages 15–19: staggered synthetic difference-in-differences by event year.

Notes: Each point is the teen-population-weighted average, across the adopting states observed at that event year, of the state's effect: its log teen birth share minus that of its synthetic control, net of their time-weighted pre-adoption difference. Births are dated to the year of conception; NVSS natality by state of residence, 1980–2024. Pre-adoption points are close to zero partly by construction, because the weights are chosen to match the pre-adoption path; the in-time placebo of Table D.1 is the test. California and Hawaii (2016) are observed up to seven years after adoption and Washington (2022) only in its first two. Shaded bands are 95 percent intervals from 200 placebo draws. Permutation p-values for ATT $[0, 8]$ and $[4, 8]$: 0.005 and 0.005 for the comprehensive arm, 0.313 and 0.428 for the traditional arm.

Teen births. Teen births in the comprehensive states had been converging toward those of the never-mandate states before adoption (Figure D.1), during the rise and long fall of U.S. teen births that began in the late 1980s (Kearney and Levine, 2012; Santelli et al., 2007). Synthetic difference-in-differences (Arkhangelsky et al., 2021) handles this by reweighting the never-mandate states, for each adopter, to reproduce its own pre-adoption path. It shows births to girls aged 15–19 falling by 9.6 (2.4) log points after adoption, about 3.8 fewer births per 1,000 girls a year at the pre-adoption rate of about 41, with an in-time placebo close to zero (Table 4, Figure 4).

Much of that fall rests on California, which adopts in 2016, carries the largest teen-population weight, and is observed through the 2020–21 fertility shock. The year effects absorb the national component of that shock, and ending the data in 2019 still gives a fall of 4.3 log points; weighting the adopting states equally gives 6.4, or 5.4 on data to 2019; all three are significant under the permutation test. Across these specifications, teen births fall by 2 to 4 per 1,000 girls a year. Only without California does the estimate, 3.2 log points, lose significance. The Callaway–Sant'Anna estimates point the same way but are smaller (Table D.2). The estimates are negative in every specification, the sign that Mark and Wu (2022) find for federal funding of comprehensive programs and that the experimental contrast in Dupas (2011) and Duflo et al. (2015) implies. The traditional arm shows nothing, in line with Carr and Packham (2017). Even the largest of these falls is too small to explain the rise in college completion, all the more so because a teen birth lowers schooling far less than raw comparisons suggest once selection into early motherhood is accounted for (Hotz et al., 2005; Ashcraft et al., 2013).

Sexual behavior. A fall in teen births can come from teenagers having less sex or from their protecting themselves better. The Youth Risk Behavior Surveillance System asks about both. I combine its state surveys from 1991 to 2023 (Centers for Disease Control and Prevention, 2024a,b) into a grade-level panel and estimate the two-stage design of Section 4 with state, survey-wave and grade fixed effects and contemporaneous exposure. Waves are biennial, so event time sits on a two-year grid and the omitted

reference is the last wave before adoption, $\ell = -2$.

Coverage is the binding constraint. Several adopters have no state survey, do not ask the sexual-behavior items, or are surveyed only on one side of adoption (Table 5). Four states identify the comprehensive arm—California, Hawaii, Maine and New Mexico—and three the traditional one—Mississippi, North Carolina and Texas. With so few treated states, I read these estimates as suggestive evidence on the behavioral channel.

	Sex without a condom (all)		Currently sexually active		Used a condom (sexually active)	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
ATT $[0, 6]$	-2.617^{***}	-1.020^{***}	-1.910^{***}	-1.556	3.375^{***}	0.398
	(0.434)	(0.381)	(0.643)	(1.403)	(0.556)	(1.327)
	[0.046]	[0.462]	[0.365]	[0.428]	[0.088]	[0.848]
ATT $[4, 6]$	-3.219^{***}	-1.221^{***}	-2.205^{**}	-2.128^{**}	4.460^{***}	0.506
	(0.516)	(0.359)	(0.972)	(1.009)	(0.796)	(1.265)
	[0.050]	[0.461]	[0.425]	[0.374]	[0.106]	[0.808]
Fixed effects	State, survey wave, grade					
Controls	None					
Weights	Respondents per state × wave × grade cell					
Treated states	4	3	4	3	4	3
Control states	22	22	22	22	22	22
Observations	1,232	1,232	1,459	1,459	1,232	1,232
Pre-adoption mean	15.691	18.726	28.483	44.152	50.653	58.776

Table 5: YRBS: sexual activity and contraceptive use, average ATT.

Notes: Outcomes, in percentage points, are the share of high-school students who are sexually active and did not use a condom at last intercourse, the share who report being currently sexually active, and the share of the sexually active reporting condom use at last intercourse; the share of all students using a condom is in Tables D.3 and D.4. The sample is the state Youth Risk Behavior Surveillance System surveys from 1991 to 2023, pooled from the DASH and SADC releases into a state × wave × grade panel, each cell weighted by its number of respondents; Wisconsin is excluded. Four states identify the comprehensive arm (California, Hawaii, Maine and New Mexico) and three the traditional one (Mississippi, North Carolina and Texas). Of the other adopters, Oregon and Washington have no state survey, Utah does not ask the sexual-behavior items, Vermont, New Jersey, Maryland, Rhode Island, Florida, Tennessee and Kentucky are first surveyed after adoption, and Georgia only before it; the estimates are suggestive. Estimated by two-stage difference-in-differences with state, survey-wave and grade fixed effects and contemporaneous rather than cohort exposure, each arm—comprehensive and traditional sex education—against the same pool of never-mandate states. Waves are biennial, so event time sits on a two-year grid with $\ell = -2$ as the omitted reference; ATT $[0, 6]$ averages the dynamic coefficients over the first seven post-adoption years and ATT $[4, 6]$ over the fully exposed ones. Standard errors, clustered on state, in parentheses; on three or four treated states they are unreliable and the randomization-inference p-values in brackets, from 500 draws that re-assign adoption dates across states, are the test to read. HonestDiD breakdown values and the in-time placebo are in Table D.3, event studies for condom use in Figure D.2 and alternative specifications in Table D.4. Stars refer to the analytic standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The share of all students who had sex without a condom, the outcome closest to the risk of an early birth, falls by 2.6 (0.4) percentage points from 15.7 percent, about a sixth, with a flat pre-period, and randomization inference confirms the fall (Table 5). Both margins contribute. The share of students who are sexually active falls by 1.9 (0.6) points, and among the sexually active, condom use at last intercourse rises by 3.4 (0.6) points, from 50.7 percent; randomization inference confirms the second but not the first.[14] The rise in protection is what reviews of comprehensive programs report (Santelli et al., 2017), and the opposite of what an abstinence rationale predicts.

The fall in sex without a condom is stable when any one treated state is left out and across specifications (Table D.4), but its in-time placebo is two thirds of the estimate. The placebo's own pre-adoption

[14] Conditioning on sexual activity selects on a margin the curriculum may move. If no student's condom use changed and the curriculum only kept 1.91 in every 100 students from being sexually active, selection alone would produce the estimated rise only if those students had used condoms at about 4 percent, against 51 percent among sexually active students before adoption, so most of the rise reflects more protection. The calculation is approximate: the two margins are estimated on slightly different samples of cells, and the fall in sexual activity is not confirmed by randomization inference.

coefficient is about as large, which points to a level difference at the reference wave rather than a break at the fake date (Appendix D.2). The rise in condom use among the sexually active passes the placebo. In the traditional arm, the estimates continue trajectories visible before adoption, and randomization inference finds nothing.

	Ages 20–24		Ages 25–29		Ages 30–34	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
ATT $[0, 6]$	-0.035^{**}	0.010	-0.079^{***}	0.038^{**}	0.094^{***}	-0.003
	(0.016)	(0.008)	(0.020)	(0.018)	(0.011)	(0.014)
	[0.060]	[0.515]	[0.082]	[0.248]	[0.056]	[0.944]
ATT $[4, 6]$	-0.040^{*}	0.015	-0.161^{***}	0.060^{**}	—	-0.086^{***}
	(0.021)	(0.013)	(0.029)	(0.030)		(0.015)
	[0.112]	[0.503]	[0.054]	[0.281]		[0.359]
Fixed effects	State and year (Callaway–Sant'Anna group-time ATTs)					
Controls	None					
Weights	Female population at risk					
Treated states	3	7	3	7	3	4
Control states	25	25	25	25	25	25
Observations	4,250	4,240	3,740	3,740	2,890	2,890
Pre-adoption mean	3.904	4.005	3.677	3.589	3.200	2.885

Table 6: Adult first births by age group, NVSS by state of birth: average ATT.

Notes: The outcome is the log first-birth rate per woman in the age group named in the column head, from NVSS natality microdata for 1980–2004 coded by the mother's state of birth, with exposure assigned at age 18. Estimated by the Callaway–Sant'Anna group-time estimator with never-treated controls, a universal base period and no covariates, weighted by the female population at risk; each arm is measured against the same pool of never-mandate states. Because the series ends in 2004, the comprehensive arm is identified by the three earliest adopters, Vermont, New Jersey and Maryland. ATT $[0, 6]$ averages the dynamic coefficients over the first seven exposed cohorts and ATT $[4, 6]$ over the fully exposed ones; — marks a window no treated state reaches at those ages. Standard errors, clustered on state, in parentheses; randomization-inference p-values in brackets, from 500 draws that re-assign each state's policy bundle across states. The in-time placebo, the bootstrap and the alternative specifications are in Table D.8, and the event studies in Figure D.4. Stars refer to the analytic standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

First births in adulthood. The natality data identify the three earliest comprehensive adopters, Vermont, New Jersey and Maryland. In them, first births fall by 3.5 (1.6) log points at 20–24 and 7.9 (2.0) at 25–29, and rise by 9.4 (1.1) at 30–34, all significant under randomization inference (Table 6). The diagnostics make this pattern suggestive rather than conclusive. With birth-state trends, the fall at 25–29 and the rise at 30–34 remain, but the fall at 20–24 disappears. In logs, the in-time placebos are well below the estimates at 25–29 and 30–34 and about three fifths of the estimate at 20–24, but in levels they reproduce most of the deficit in the twenties (Tables D.8 and D.9). The rise in the thirties passes the placebo but rests on the one exposed cohort observed at 30 before the series ends, Vermont's 1974 cohort. A synthetic difference-in-differences on the same three states gives the same pattern: a fall in the twenties about as large as its placebo, and a rise at 30–34 that comes almost entirely from Vermont.[15] The traditional arm shows nothing robust.

The natality series ends before the later adopters' exposed cohorts reach adulthood, so for them I reconstruct first births in the ACS from the eldest child living with the mother (Appendix D.4.1), on the six later comprehensive adopters: Maine, New Mexico, Oregon, California, Hawaii and Washington. The estimates suggest the same postponement, less precisely. First births fall at 20–24, by 8.8 (6.1) log points and by 16.2 (8.7) among the fully exposed, and rise at 25–29, by 4.1 (3.0) and 11.5 (2.2); under randomization inference the smallest p-value is 0.15 (Table D.10). Much of the fall at 20–24 was under

[15] First births fall by 3.2 (1.4) log points at 20–24 and 3.4 (2.4) at 25–29, against in-time placebos of -2.4 (3.2) and -5.5 (4.4), and rise by 15.3 (3.3) at 30–34, with a placebo of 0.0 (2.8).

way before adoption: its in-time placebo is −5.5 (1.0) (Table D.11). On the three early adopters, whose births in their thirties fall after 2000, the ACS also finds a rise at 30–34, of 2.6 (1.7) log points and 3.7 (1.8) among the fully exposed. In both sources first births move out of the early twenties. The traditional arm shows nothing in the ACS.

Across ages, the fertility evidence is suggestive of a shift in timing: fewer teen births, fewer first births in the twenties, and more later. A shift in the timing rather than the level of fertility is what Lazuka and Elwert (2026) find after the Swedish reform and Goldin and Katz (2002) after the pill, and Fung and Min (2025) find its mirror image for abstinence-based HIV-only mandates—first births brought forward, completed fertility unchanged.

7.1.5 From the life course to the labor market

Taken together, these margins describe one change, identified with different precision. Women exposed to a comprehensive mandate are more likely to finish college, less likely to marry before college would be over, and older at marriage when they marry young. College completion and marriage by 23 are the clearest results; the fertility evidence is consistent with them but weaker. The account that follows therefore rests on schooling and the timing of marriage: exposed women finish school before they marry, and marry later than they otherwise would have.

This sequence is the one through which the labor market literature expects women's careers to be shaped, and each of its links is established there rather than estimated here. Schooling raises earnings in its own right (Card, 1999), and women invest in college when they expect long careers: their college going overtook men's as young women came to anticipate working through adulthood and to marry later (Goldin et al., 2006). Marriage and a first birth each lower women's wages and wage growth (Loughran and Zissimopoulos, 2009), so delaying them pays: a year of delay in motherhood raises women's earnings by about nine percent (Miller, 2011), with the return concentrated among college graduates (Herr, 2016). The earnings gap between men and women opens at the first birth (Kleven et al., 2019), and the experience accumulated before it keeps women employed after it (Troske and Voicu, 2013). Early legal access to the pill moved women's careers through the same margins, raising their professional schooling, delaying their marriages and, decades later, raising their wages (Goldin and Katz, 2002; Bailey et al., 2012). A woman who reaches motherhood with a degree and several years of full-time work has more to return to and more to lose by leaving: experience accumulates mainly in full-time employment (Blundell et al., 2016), participation is state dependent (Hyslop, 1999), and even short interruptions open earnings gaps that do not close (Bertrand et al., 2010). The labor market results of Section 6 have the shape this account predicts. The effects are concentrated on full-time work and earnings rather than on participation alone; earnings rise among the employed as well as across all women (Table C.5); and the effects show no sign of fading with age—on the three early adopters, the point estimates for earnings and full-time work are positive at every age from 34 to 45 (Figure C.5). An effect that merely moved work from the late twenties, before motherhood, would have the opposite profile.

Men fit this reading. They too complete college more often, by 1.6 (0.6) points, and marry later, by 0.12 (0.04) years, as they must if their partners do; but their earnings and full-time work do not move, and their employment falls slightly. For men a later marriage and a later first child do not change attachment to the labor market, which childbearing does not interrupt (Kleven et al., 2019).

The model of Section 3 captures the first part of this account—more schooling, and fewer women derailed before it is complete—but not the second: with two periods it has no experience to accumulate

before a birth and no return to work after one. The life-course evidence therefore points beyond the model. What a comprehensive curriculum appears to change is not only whether a young woman avoids an early birth but the order in which she completes school, marries and has children, and, through that order, her attachment to the labor market. No single margin is identified as cleanly as the earnings result, so this is a reading of the pattern across margins rather than a decomposition, but it is the pattern the account implies.

7.2 Gender attitudes

The two curricula say different things about what women's lives are for. Abstinence-based programs do not merely omit contraception; they often convey traditional gender norms, teaching that men and women differ in temperament and desires.[16] Comprehensive curricula are not their mirror image: most say nothing about gender at all, and the few that take it up are the exception rather than the type.[17] The contrast the design estimates is therefore not instruction in gender equality against silence, but silence against an explicit traditional script, since many never-mandate states require that abstinence be stressed whenever a district teaches sex education (Table A.1). A comprehensive mandate removes the prescription, and the content it requires bears on norms in its own right: teaching that the timing of a birth is something a girl decides is a statement about whose trajectory is hers to plan. The attitude test below therefore does not presume that the curricula taught gender equality.

The model's second channel is a norm cost $\tau(D)$ attached to market work, and (5) makes its placement consequential: a cost attached to the career *state* cancels out of the education comparison, whereas a cost attached to the *behavior* raises both the stakes of avoiding derailment and the premium to education. The test is therefore whether views about women's market work become more egalitarian among the exposed cohorts.

I test it in the American National Election Studies (American National Election Studies, 2025), assigning respondents to the state where they grew up and to a birth cohort, with exposure at eighteen. The main item asks whether it is better for the family if the man works outside the home and the woman takes care of the home. Its wording changed in 2016, but both versions have a traditional, a middle and an egalitarian answer, so I pool them in an index from 0 (traditional) to 1 (egalitarian), with survey-year fixed effects absorbing the change. The second item asks whether *this country would have many fewer problems if there were more emphasis on traditional family ties*, on the same scale. Because cells are thin, event time is grouped into two-year bins, so the omitted reference is the bin $[-2, -1]$, and the trend specification uses census-division linear trends, since cells are too thin for state-specific slopes. Appendix D.5 gives the wording of every item and explains why two further items cannot be used.

[16] For instance, *WAIT Training* teaches that "a male is usually less discriminating about those to whom he is sexually attracted" and that "while a man needs little or no preparation for sex, a woman often needs hours of emotional and mental preparation"; *Why kNOw*, that "the father gives the bride to the groom because he is the one man who has had the responsibility of protecting her throughout her life" (U.S. House of Representatives, Committee on Government Reform—Minority Staff, Special Investigations Division, 2004).

[17] *Making Proud Choices!*, among the most widely used, mentions gender stereotypes only in passing; *FLASH* devotes a lesson to teaching that gender roles "are learned" and "can be limiting" (Appendix A.3.3); *SiHLE* names the Theory of Gender and Power as its theoretical basis; *¡Cuídate!* builds in *machismo* and Latino gender-role expectations, but in order to present abstinence and condom use as culturally acceptable rather than to question the expectations themselves. *Gender Matters*—EngenderHealth's Teen Pregnancy Prevention grantee in Austin, Texas—states as its objective influencing "traditional beliefs about masculinity and femininity" and advancing "more equitable attitudes about relationships and the balance of power within them" (SIECUS, 2015), but it is one federal grantee, not a curriculum any state mandated. The systematic reviews find that gender- and power-aware instruction is what distinguishes the comprehensive programs that work (Goldfarb and Lieberman, 2021; Goldfarb et al., 2026); that is a finding about effectiveness, not a description of what the mandated curricula contained.

	Man works, woman home (index, 0–1)		Traditional family ties (index, 0–1)	
	Comp.	Trad.	Comp.	Trad.
ATT $[0, 6]$	0.109**	−0.018	−0.001	0.004
	(0.047)	(0.034)	(0.025)	(0.021)
	[0.018]	[0.701]	[0.980]	[0.898]
ATT $[4, 6]$	0.108*	0.009	−0.015	0.022
	(0.059)	(0.045)	(0.047)	(0.032)
	[0.072]	[0.858]	[0.835]	[0.679]
Fixed effects	Grew-up state, birth cohort, survey year			
Controls	None			
Weights	ANES survey weights			
Treated states	9	9	9	9
Control states	24	24	24	24
Observations	5,256	5,256	5,649	5,649
Pre-adoption mean	0.454	0.467	0.414	0.286

Table 7: Gender-role attitudes (ANES), women: average ATT on the egalitarian index.

Notes: Outcomes are the egalitarian index of the item named in the column head, on a 0–1 scale: 0 for the traditional response, 0.5 for the middle category ("makes no difference", "neither") and 1 for the egalitarian response. A cell mean is therefore the item's whole three-category distribution in one number, and a coefficient is the movement of the average respondent along the scale, one step being 0.5. The man-works index pools the 2004/08 statement and the 2016–2024 question; the traditional-family-ties index is the same five-point statement in every wave, with agreement as the traditional end. The man-works index on the 2016–2024 waves alone and its indicator version are in Panel B of Table D.12. The sample is women in the American National Election Studies cumulative file, assigned to the state where they grew up and to a birth cohort, exposed at age 18 with single-year cohort resolution, and weighted by ANES survey weights. Estimated by two-stage difference-in-differences with grew-up-state, birth-cohort and survey-year fixed effects, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. Cells are thin, so event time is grouped into two-year bins with $[-2, -1]$ as the omitted reference and the window stops at ± 6; ATT $[0, 6]$ averages the dynamic coefficients over all post-adoption bins and ATT $[4, 6]$ over the fully exposed ones. Standard errors, clustered on the grew-up state, in parentheses; with nine and nine treated states, the randomization-inference p-values in brackets, from 500 draws that re-assign adoption dates across states, are the test to read. HonestDiD breakdown values and the in-time placebo are in Table D.12, the event study for *man works, woman home* in Figure D.5, alternative specifications in Table D.13, and the same table for men in Table D.14. Stars refer to the analytic standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The market-work item moves (Table 7). Women exposed to a comprehensive mandate move 0.109 (0.047) up the egalitarian scale, from a pre-adoption mean of 0.454: about a fifth of a step on the three-category item, with a flat pre-period. The estimate holds under the alternative specifications, when any one treated state is left out, and on the 2016–2024 question alone, 0.102 (0.048), which avoids the change in wording (Tables D.12 and D.13). Its in-time placebo is about half the estimate, but the placebo is not a like-for-like test: moving adoption seven years earlier more than doubles California's share of the treated pool, and California was already moving on these attitudes before its 2016 mandate. Without California the placebo falls to 0.032 (0.040) and the estimate rises to 0.144 (0.043). The estimate barely changes when the women's own work status, children and marital status are controlled for, so the shift is not merely a by-product of their current employment and family. Men's estimate is smaller, 0.042 (0.047), and imprecise (Table D.14).

The statement on traditional family ties does not move, −0.001 (0.025). It is less a statement about women's roles than a marker of broader moral and political commitments—its correlation with Republican identification is about twice that of the market-work index—and a curriculum that challenges gendered prescriptions about work gives no argument against family ties. The traditional arm moves neither item.

8 Ruling Out Alternative Explanations

8.1 School funding

The most immediate alternative is a resources story: a state that mandates comprehensive sex education may also spend more on its schools, so the treated cohorts would be better educated for reasons unrelated to what they were taught. That story predicts a break in school budgets at adoption. I test it with the same Callaway–Sant'Anna event study on log real current expenditure per pupil, from the NPEFS state-year finance panel for fiscal years 1987–2023 (National Center for Education Statistics, 2024), for the states that adopted through 2015 (Appendix E.1 explains the cap and reports the full sample).

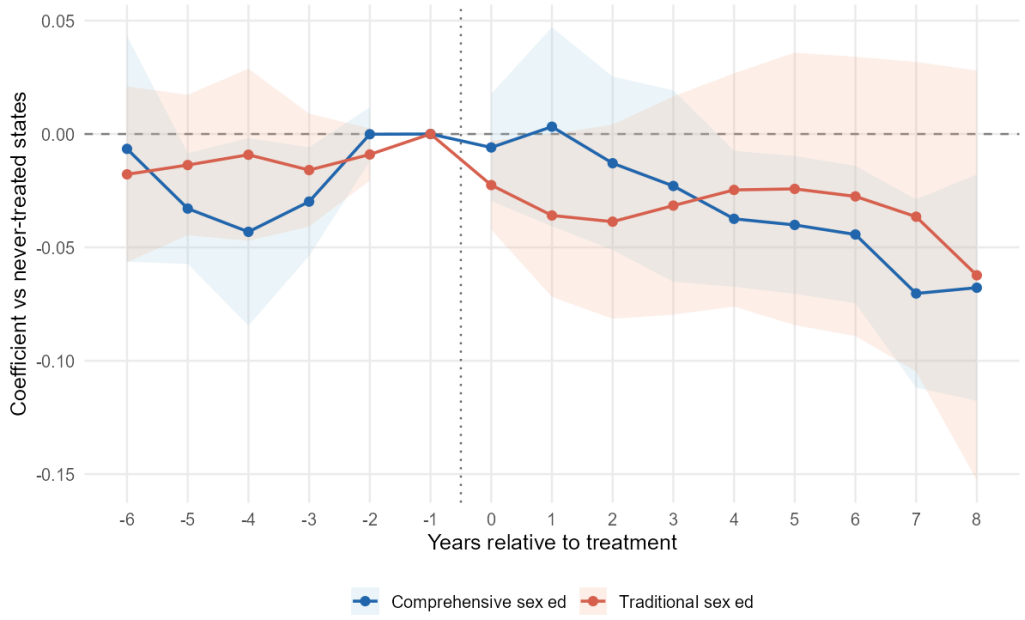


Figure 5: Fiscal placebo: current expenditure per pupil around sex education mandate adoption, adoptions through 2015.

Notes: The outcome is log real current expenditure per pupil in average daily attendance, from the NPEFS state-year school finance panel for fiscal years 1987–2023, on the 51-state panel with the adoption dates and control pool of every other battery. The sample caps adoptions at 2015, dropping California and Hawaii (2016) and Washington (2022); see Section 8.1. Washington's onset falls after the panel ends, so it adds no post-adoption year in either sample. Estimated by the Callaway–Sant'Anna group-time estimator with never-treated controls, a universal base period and no covariates, weighted by mean fall enrollment, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; the horizontal axis is years since the mandate took effect. Shaded bands are 95 percent confidence intervals from standard errors clustered on state. This is a placebo on the mechanism, not an outcome of interest: a resources story requires a break in school budgets at adoption, and the figure is what its absence looks like. The robustness tables are Tables E.1 and E.2, the same event study on the full sample including California Figure E.1b, and the state-by-state ATT Figure E.2.

There is no fiscal increase at adoption (Figure 5). If anything, comprehensive adopters' spending falls behind the control states' in the following years, from a pre-period that was already lower, and the traditional arm looks the same. Adding the 2016 adopters turns the comprehensive estimate positive, but the increase is California's, whose real per-pupil spending grew faster than any other state's after 2013 (Appendix E.1); California contributes several post-adoption years only to teen births, which Section 7.1.4 reports with and without it. Comprehensive adopters are therefore not states that raised per-pupil spending when they changed their curriculum.

One school finance reform does coincide with an onset: New Jersey's 1990 *Abbott v. Burke* decision

and Quality Education Act came the year before its comprehensive mandate, and higher school spending raises adult earnings (Jackson et al., 2016). New Jersey's per-pupil spending does not rise after adoption, by −1.1 (1.0) log points (Figure E.2), but the reform redirected money toward poor districts, which a state average can hide. The earnings result does not depend on New Jersey: without it the estimate is 3.5 log points (Table C.7). The full-time and employment estimates, which do depend on New Jersey, are less secure on this count.

8.2 Electoral cycles

A second concern is that a mandate is passed by a legislature that has just changed hands, and that the same turnover changes everything else about the state. What would make that story work is a *sudden* partisan shift at adoption. What I find is a slow one.

I rebuild the partisan composition of every state legislature from the candidate-level election returns of Klarner et al. (2013), spliced with the seat counts of Klarner (2013) where the returns do not reach, and measure it by the two-party seat share of the party whose ideology matches the mandate: the Democratic share in the comprehensive arm and the Republican share in the traditional arm.[18] I estimate the Callaway–Sant'Anna event study of Section 4.3 on this outcome, each arm against the never-treated states, with time measured in elections rather than years: $e = -1$ is the legislature elected in the two years before the mandate takes effect, the one that can have passed it, and every coefficient is measured against it (Figure 6).

[18] Construction and estimation are in Appendix E, Section E.2.

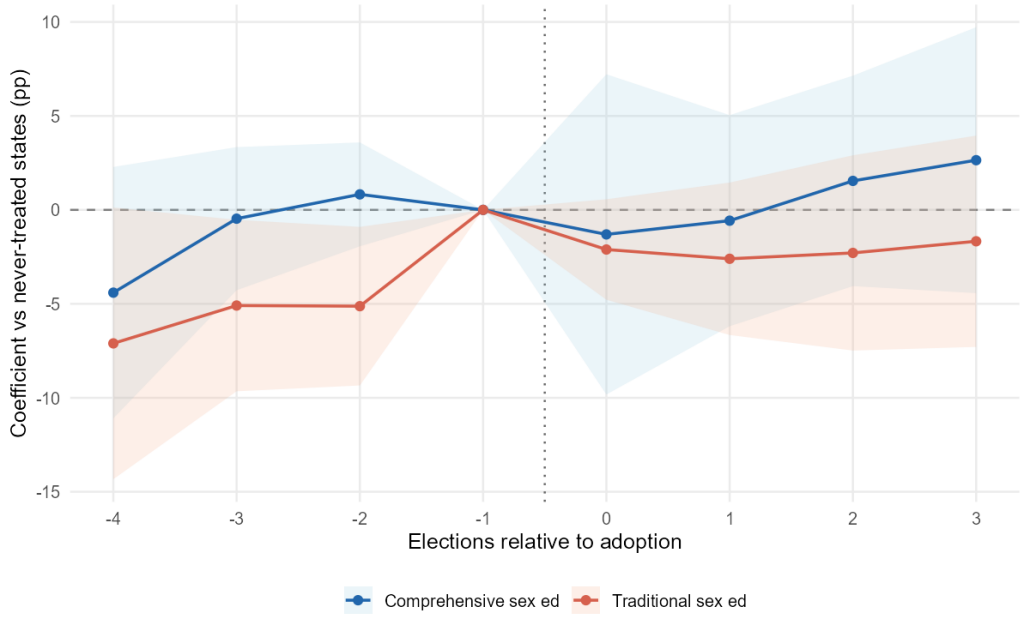


Figure 6: Legislative composition around sex education mandate adoption: two-party seat share of the party aligned with the mandate.

Notes: The outcome is the two-party seat share $D/(D+R)$ of the Democratic party in the comprehensive arm and of the Republican party in the traditional arm, in percentage points, averaged unweighted across the two chambers and rebuilt from Klarner et al. (2013) and Klarner (2013) for 1974–2016 on a two-year election grid. The horizontal axis counts elections relative to adoption: -1 is the legislature elected in the two years before the mandate took effect, the reference, and 0 the first one elected after it. Estimated by the Callaway–Sant'Anna group-time estimator with never-treated controls, a universal base period and no covariates, unweighted, each arm—the eight comprehensive and nine traditional adopters observed before 2016—against the same 30 states that never adopt either arm, which, unlike the control pool of the cohort designs, include the eight states whose mandate fits neither arm. Shaded bands are 95 percent confidence intervals from standard errors clustered on state. The estimates are in Table E.3 and the raw levels in Table E.4.

Composition moves toward the ideology of the mandate before adoption, not at it. In comprehensive states the Democratic share rises between four and three elections before the mandate and then stays flat; in traditional states the legislature that passed the mandate is the most Republican of the sequence. What does not happen is a break at adoption. The composition that passed a comprehensive mandate persists afterwards, and in traditional states the Republican share recedes part of the way (Table E.3).

A gradual realignment preceding adoption by several years would appear in the cohort event studies as a pre-trend, not as a change at the moment a cohort becomes exposed. And because it operates on the whole state at once, it would move the outcomes of cohorts already past school age, which is what the in-time placebo tests, and does not find.

8.3 Minors' access to contraception and abortion

The other policies aimed at adolescents' fertility are legal access to contraception and to abortion, and a mandate coinciding with a change in either would be attributed to the curriculum. Legal access to the pill is not at issue. The state-level expansion of minors' legal access to the pill that Goldin and Katz (2002) and Bailey (2006) exploit—through lower ages of majority, mature-minor doctrines and family-planning statutes—was complete in every state by the mid-1970s, so the cohorts it reached were born through the late 1950s, whereas the first cohort treated here is born in 1969. *Carey v. Population Services* (1977) and the 1978 amendment to Title X then extended confidential access to minors nationally, and the coding of Myers (2022) records no change in minors' access to contraception after 1979 in any comprehensive

adopter.

Later state policies on contraception reached the treated cohorts only as adults. The one with real overlap is the mandate that private insurers cover prescription contraceptives, adopted by Maryland in 1998, Vermont in 1999 and New Jersey in 2005–06, when their first treated cohorts were in their mid-twenties to early thirties. Exposure to these mandates changes gradually from one cohort to the next rather than at the cohort cutoff, and their estimated effect on births is insignificant (Mulligan, 2015). Paid family leave arrived in New Jersey in 2009, when its first treated cohort was 36.

The remaining reproductive-policy confounder is parental involvement in a minor's abortion. Legal access itself is not at issue—the coding underlying Myers (2017) has every state at its terminal regime by 1979, touching only cohorts born through 1965, whereas the first cohort treated here is born in 1969—but *enforcement* of parental consent and notification requirements continued for four decades, because many statutes sat enjoined until *Hodgson* (1990) and *Casey* (1992) settled the standard. Table 8 sets the enforcement dates of Myers (2022) against the mandate onsets.

State	Mandate year	First enforced PI law	Event time
Comprehensive arm			
Vermont	1990	never	—
New Jersey	1991	never	—
Maryland	1992	1992*	+0
Maine	2003	1989*	-14
New Mexico	2008	never	—
Oregon	2010	never	—
California	2016	never	—
Hawaii	2016	never	—
Washington	2022	never	—
Traditional arm			
Rhode Island	1988	1982	-6
Utah	1989	1980	-9
Florida	1991	2005	+14
Tennessee	1992	1992	+0
Kentucky	1993	1989†	-4
Georgia	1995	1991	-4
North Carolina	1996	1995	-1
Texas	2004	2000	-4
Mississippi	2013	1993	-20

Table 8: Sex education mandates and enforced parental-involvement laws.

Notes: A comparison of policy dates, not an estimate. "Mandate year" is the year each state's sex education mandate takes effect, and "first enforced PI law" the first year a parental consent or notification requirement for a minor's abortion was actually enforced, from Myers (2022); many statutes sat enjoined for years after passage, so enforcement rather than enactment is the relevant date. "Event time" is the second date minus the first, so 0 means the two policies change in the same year and "—" that the state never enforced such a law. States are grouped by the arm their mandate belongs to, comprehensive or traditional. The estimation window runs from event time −6 to +8, so only the rows inside that range represent a coincidence of timing that the design cannot separate from the curriculum. * Maryland's law is a notification requirement, and Maine's a consent requirement that a physician can waive for a mature minor. † Enforced in 1989 and again from 1994, the year after the mandate.

Two things follow. First, the two arms differ systematically in abortion policy. Seven of the nine comprehensive adopters never enforced a parental-involvement requirement between 1980 and 2020; all nine traditional adopters did. Birth-state fixed effects absorb this level difference, which reflects the different policy environments the two arms come from rather than a mandate standing in for abortion access.

Second, the timing coincides in a few cases, mostly on the traditional side. Tennessee, North Carolina and, through renewed enforcement, Kentucky changed both policies within a year of each other, so that arm's teen-birth estimate could not have been attributed to the curriculum had it been significant; it is not. On the comprehensive side, only Maryland has a change within the estimation window: a notification law in the year of its mandate. A restriction of that kind would raise early births rather than postpone them.

9 Conclusions

This paper asks whether what compulsory sex education is required to teach—contraception, or abstinence as a moral standard—affects women's labor market outcomes decades later. Using a new panel of binding state mandates for 1970–2022, linked to the 2000–2024 American Community Survey through the mandate in force in each woman's birth state during her school years, it finds that it does. Exposure to a comprehensive mandate raises women's annual earnings by 4.2 log points, relative to a pre-adoption mean near $31,000, and the share working full time by 1.8 percentage points, and employment rises by 2.1 to 2.6 points once the adopting states' prior downward drift is accounted for. The earnings estimate survives randomization inference on eight treated states, alternative control pools, comparisons within census regions and dropping any one state. Men's earnings and full-time work are unaffected, and a traditional mandate has no robust effect. Comparing states with any mandate to states with none yields a null, because that comparison mixes the comprehensive mandates with traditional ones and with mandates of neither kind, and dates exposure from each state's first mandate rather than from its content.

The evidence on mechanisms points to a change in the life course. Comprehensive exposure raises college completion by 2.6 points from a base of 33.9 percent, lowers the share of women married by 23—the age by which college would be over—by 2.0 points from 27.1 percent, and raises the age at marriage of women who marry young. Women who finish school before they marry start their careers with more schooling, so the gains should show up in hours and pay and persist with age. They do: the effects are concentrated on full-time work and earnings and show no sign of fading through the forties. Teen births fall by 2 to 4 per 1,000 girls a year, a fall driven largely by California, and the share of teenagers having sex without a condom falls by 2.6 percentage points in the four states the youth surveys can follow; natality data for the earliest adopters suggest that first births move from the late twenties to the early thirties. Women's attitudes also move toward the egalitarian answer on whether a man's place is at work and a woman's at home, although part of that movement predates exposure. Two features of this pattern govern its interpretation. The college response is too large to come from averted teen births alone, which is the model's signature that the value of education itself rose, and not only the probability of completing it. And the effects on earnings and full-time work are concentrated on women, which separates a channel operating on what girls expect of their own trajectories from a general improvement in schooling quality, although men's college completion rises too. The evidence supports the norm channel without isolating it: a beliefs formulation in which the curriculum corrects an adolescent's under-prediction of her own labor market attachment delivers identical comparative statics, and the data cannot separate a girl who faces a weaker prescription from one who has revised her forecast.

The traditional arm has no robust effect. Earnings, teen births and gender attitudes do not move, and full-time work falls in the baseline specification but not once state trends are included. The model admits two readings of that null. If the norm that prevails in the absence of a mandate is itself traditional, a curriculum that restates it is the untreated state and a null is what the model predicts; if the prevailing norm were neutral, a traditional curriculum would reinforce it and reverse every comprehensive effect.

The estimates are consistent with the first reading and too imprecise to rule out the second entirely. A third reading is that traditional mandates changed little of what was taught, since abstinence curricula also ran in states without a mandate; without a measure of instruction the data cannot tell it apart from the first.

Several limits temper the reading. Eight comprehensive and nine traditional states identify the labor market estimates, so inference rests on randomization rather than cluster asymptotics, and several results, full-time work and employment among them, depend on New Jersey. A statute bounds rather than measures the curriculum a cohort actually met, so the estimates are intention-to-treat. And the mechanisms are not identified equally well: college completion and the timing of marriage are the clearest, while the teen-birth estimate leans on California, the survey evidence on sexual behavior covers four states, and attitudes are measured decades after school.

Two implications follow for policy. First, the margin that matters is content, not presence. The question that has occupied state legislatures for four decades—whether to require sex education—matters less than what the requirement contains: whether it covers contraception and whether it presents abstinence as the standard adolescents are morally expected to meet. Second, the comprehensive effects arise in an environment in which contraception was already legal and available, so what the classroom transmits is plausibly not biology alone but what a girl is told to expect of her own trajectory. The evidence does not speak to family size: the natality data suggest that comprehensive instruction moves first births later, but completed fertility cannot yet be observed for most exposed cohorts.

Two extensions follow directly. The framework applies to intimate partner violence, where the same norm channel has an established link (González and Rodríguez-Planas, 2020) and where comprehensive curricula explicitly target consent and healthy relationships; that is the subject of a companion project. And the mechanism invites work that can separate identity from beliefs, plausibly through data on what adolescents expect of their own futures at the moment the schooling decision is made, rather than what they report about gender roles decades later.

References

Adda, J., C. Dustmann, and K. Stevens (2017). The career costs of children. *Journal of Political Economy 125*(2), 293–337.

Adukia, A. and E. Harrison (2025). Separation of church and state curricula? examining public and religious private school textbooks. Technical report, National Bureau of Economic Research.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *Quarterly Journal of Economics 115*(3), 715–753.

Alesina, A. and N. Fuchs-Schündeln (2007). Good-bye lenin (or not?): The effect of communism on people's preferences. *American Economic Review 97*(4), 1507–1528.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the origins of gender roles: Women and the plough. *The quarterly journal of economics 128*(2), 469–530.

American National Election Studies (2025). ANES Time Series Cumulative Data File (1948–2024) [dataset and documentation].

Ananat, E. O., J. Gruber, P. B. Levine, and D. Staiger (2009). Abortion and selection. *Review of Economics and Statistics 91*(1), 124–136.

Arkhangelsky, D., S. Athey, D. A. Hirshberg, G. W. Imbens, and S. Wager (2021). Synthetic difference-in-differences. *American Economic Review 111*(12), 4088–4118.

Arold, B., A. Alves Monteiro, and C. Yannelis (2026). The long term effects of school prayer.

Arold, B. W., L. Woessmann, and L. Zierow (2025). Can schools change religious attitudes?: Evidence from german state reforms of compulsory religious education. *Journal of Human Resources*.

Ashcraft, A., I. Fernández-Val, and K. Lang (2013). The consequences of teenage childbearing: Consistent estimates when abortion makes miscarriage non-random. *The Economic Journal 123*(571), 875–905.

Atkins, D. N. and W. D. Bradford (2021). The effect of state-level sex education policies on youth sexual behaviors. *Archives of Sexual Behavior 50*(6), 2321–2333.

Bailey, M. J. (2006). More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply. *Quarterly Journal of Economics 121*(1), 289–320.

Bailey, M. J., B. Hershbein, and A. R. Miller (2012). The opt-in revolution? contraception and the gender gap in wages. *American Economic Journal: Applied Economics 4*(3), 225–254.

Baker, A. C., D. F. Larcker, and C. C. Y. Wang (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics 144*(2), 370–395.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological) 57*(1), 289–300.

Benjamini, Y. and D. Yekutieli (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics 29*(4), 1165–1188.

Bertrand, M., C. Goldin, and L. F. Katz (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics 2*(3), 228–255.

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender identity and relative income within households. *Quarterly Journal of Economics 130*(2), 571–614.

Billari, F. C., A. Goisis, A. C. Liefbroer, R. A. Settersten, A. Aassve, G. Hagestad, and Z. Spéder (2011). Social age deadlines for the childbearing of women and men. *Human Reproduction 26*(3), 616–622.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2008). Staying in the classroom and out of the maternity ward? the effect of compulsory schooling laws on teenage births. *Economic Journal 118*(530), 1025–1054.

Blundell, R., M. Costa Dias, C. Meghir, and J. Shaw (2016). Female labor supply, human capital, and welfare reform. *Econometrica 84*(5), 1705–1753.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies 91*(6), 3253–3285.

Braun, T., A. Costa-Ramón, A. Rodríguez-González, U. Schaede, and C. Salmivalli (2026). Non-toxic classrooms: Long-run returns from an anti-bullying program. Working paper. AEA RCT Registry ID 0010879.

Buckles, K. S. and D. M. Hungerman (2018). The incidental fertility effects of school condom distribution programs. *Journal of Policy Analysis and Management 37*(3), 464–492.

Buzzell, P. P. (2024). *Reviewing Inclusive Sex Education Curriculum: A Policy Analysis of California's Healthy Youth Act (2015)*. Ph. D. thesis, City University of New York.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of econometrics 225*(2), 200–230.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics 90*(3), 414–427.

Cantoni, D., Y. Chen, D. Y. Yang, N. Yuchtman, and Y. J. Zhang (2017). Curriculum and ideology. *Journal of Political Economy 125*(2), 338–392.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3A, pp. 1801–1863. Elsevier.

Carr, J. B. and A. Packham (2017). The effects of state-mandated abstinence-based sex education on teen health outcomes. *Health Economics 26*(4), 403–420.

Centers for Disease Control (1988, December). HIV-related beliefs, knowledge, and behaviors among high school students. *Morbidity and Mortality Weekly Report 37*(47).

Centers for Disease Control and Prevention (2023). School health profiles 2022: Characteristics of health programs among secondary schools.

Centers for Disease Control and Prevention (2024a). Youth risk behavior surveillance system (yrbss): State and local datasets, 1991–2023.

Centers for Disease Control and Prevention (2024b). Youth risk behavior survey combined datasets (sadc), 1991–2023.

Centers for Disease Control and Prevention (2025). Sexually transmitted infections surveillance, 2024.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How does your kindergarten classroom affect your earnings? evidence from Project STAR. *Quarterly Journal of Economics 126*(4), 1593–1660.

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

Clarke, D., D. Pailañir, S. Athey, and G. Imbens (2024). On synthetic difference-in-differences and related estimation methods in Stata. *The Stata Journal 24*(4), 557–598.

Dhar, D., T. Jain, and S. Jayachandran (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American economic review 112*(3), 899–927.

Donovan, P. (1998). School-based sexuality education: The issues and challenges. *Family Planning Perspectives 30*(4), 188–193.

Duflo, E., P. Dupas, and M. Kremer (2015). Education, HIV, and early fertility: Experimental evidence from kenya. *American Economic Review 105*(9), 2757–2797.

Dupas, P. (2011). Do teenagers respond to HIV risk information? evidence from a field experiment in kenya. *American Economic Journal: Applied Economics 3*(1), 1–34.

Dupas, P., S. Jayachandran, A. Lleras-Muney, and P. Rossi (2025). The negligible effect of free contraception on fertility: Experimental evidence from Burkina Faso. *American Economic Review 115*(8), 2659–2688.

Farré, L. and F. Vella (2013). The intergenerational transmission of gender role attitudes and its implications for female labour force participation. *Economica 80*(318), 219–247.

Ferman, B. and C. Pinto (2019). Inference in differences-in-differences with few treated groups and heteroskedasticity. *Review of Economics and Statistics 101*(3), 452–467.

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review 103*(1), 472–500.

Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *Quarterly Journal of Economics 119*(4), 1249–1299.

Fogli, A. and L. Veldkamp (2011). Nature or nurture? learning and the geography of female labor force participation. *Econometrica 79*(4), 1103–1138.

Foshee, V. A., K. E. Bauman, S. T. Ennett, G. F. Linder, T. Benefield, and C. Suchindran (2004). Assessing the long-term effects of the safe dates program and a booster in preventing and reducing adolescent dating violence victimization and perpetration. *American journal of public health 94*(4), 619–624.

Fung, R. Y. L. and J. Min (2025). Does curriculum matter? how sex education impacts teen sexual activity and lifecycle fertility. Working paper, University of Missouri and Princeton University, 20 August 2025.

Gardner, J. (2022). Two-stage differences in differences. *arXiv preprint arXiv:2207.05943*.

Gold, R. B. and E. Nash (2001). State-level policies on sexuality, STD education. *The Guttmacher Report on Public Policy 4*(4), 4–7.

Goldfarb, E. S. and L. D. Lieberman (2021). Three decades of research: The case for comprehensive sex education. *Journal of Adolescent Health 68*(1), 13–27.

Goldfarb, E. S., L. D. Lieberman, and K. Conklin (2026). Promising approaches to comprehensive sex education: Evidence from three decades of research. *Journal of Adolescent Health 78*(1), 19–34.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review 96*(2), 1–21.

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.

Goldin, C. and L. F. Katz (2002). The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy 110*(4), 730–770.

Goldin, C., L. F. Katz, and I. Kuziemko (2006). The homecoming of American college women: The reversal of the college gender gap. *Journal of Economic Perspectives 20*(4), 133–156.

González, L. and N. Rodríguez-Planas (2020). Gender norms and intimate partner violence. *Journal of Economic Behavior & Organization 178*, 223–248.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of econometrics 225*(2), 254–277.

Hara, H. and N. Rodriguez-Planas (2025). Long-term consequences of teaching gender roles: Evidence from desegregating industrial arts and home economics in japan. *Journal of Labor Economics 43*(2), 349–389.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics 94*(1–2), 114–128.

Herr, J. L. (2016). Measuring the effect of the timing of first birth on wages. *Journal of Population Economics 29*(1), 39–72.

Hotz, V. J., S. W. McElroy, and S. G. Sanders (2005). Teenage childbearing and its life cycle consequences: Exploiting a natural experiment. *Journal of Human Resources 40*(3), 683–715.

Hyslop, D. R. (1999). State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica 67*(6), 1255–1294.

Jackson, C. K., R. C. Johnson, and C. Persico (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *Quarterly Journal of Economics 131*(1), 157–218.

Jayachandran, S. (2020). Social norms as a barrier to women's employment in developing countries. Technical report, National Bureau of Economic Research.

Jemmott, L. S., J. B. Jemmott, III, and K. A. McCaffree (2016). *Making Proud Choices! An Evidence-Based, Safer-Sex Approach to Teen Pregnancy and HIV/STD Prevention* (Fifth ed.). Scotts Valley, CA: ETR Associates.

Jensen, R. (2012). Do labor market opportunities affect young women's work and family decisions? experimental evidence from india. *The Quarterly Journal of Economics 127*(2), 753–792.

Kaiser Family Foundation (2002, October). Sex education in the U.S.: Policy and politics. Issue update, Henry J. Kaiser Family Foundation, Menlo Park, CA.

Kearney, M. S. and P. B. Levine (2012). Why is the teen birth rate in the united states so high and why does it matter? *Journal of Economic Perspectives 26*(2), 141–166.

Kearney, M. S. and P. B. Levine (2014). Income inequality and early nonmarital childbearing. *Journal of Human Resources 49*(1), 1–31.

Klarner, C. (2013). State partisan balance data, 1937–2011. *Harvard Dataverse 1*, 2013.

Klarner, C., W. Berry, T. Carsey, M. Jewell, R. Niemi, L. Powell, and J. Snyder (2013). State legislative election returns (1967-2010). *ICPSR Data Holdings*.

Kleven, H., C. Landais, and J. E. Søgaard (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Lazuka, V. and A. Elwert (2026). Life-cycle effects of comprehensive sex education. arXiv working paper 2310.11151, version 3. Earlier version: IZA Discussion Paper 16622 (2023).

Loughran, D. S. and J. M. Zissimopoulos (2009). Why wait? The effect of marriage and childbearing on the wages of men and women. *Journal of Human Resources 44*(2), 326–349.

Luker, K. (2006). *When Sex Goes to School: Warring Views on Sex—and Sex Education—Since the Sixties*. New York: W. W. Norton.

MacKinnon, J. G. and M. D. Webb (2018). The wild bootstrap for few (treated) clusters. *Econometrics Journal 21*(2), 114–135.

MacKinnon, J. G. and M. D. Webb (2020). Randomization inference for difference-in-differences with few treated clusters. *Journal of Econometrics 218*(2), 435–450.

Mark, N. D. and L. L. Wu (2022). More comprehensive sex education reduced teen births: Quasi-experimental evidence. *Proceedings of the National Academy of Sciences 119*(8), e2113144119.

Miller, A. R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics 24*(3), 1071–1100.

Moran, J. P. (2000). *Teaching Sex: The Shaping of Adolescence in the 20th Century*. Cambridge, MA: Harvard University Press.

Mulligan, K. (2015). Contraception use, abortions, and births: The effect of insurance mandates. *Demography 52*(4), 1195–1217.

Myers, C. K. (2017). The power of abortion policy: Reexamining the effects of young women's access to reproductive control. *Journal of Political Economy 125*(6), 2178–2224.

Myers, C. K. (2022). Confidential and legal access to abortion and contraception in the usa, 1960–2020. *Journal of Population Economics 35*(4), 1385–1441.

National Center for Education Statistics (2024). National public education financial survey (npefs), fiscal years 1987–2023.

National Center for Health Statistics (2024). Natality detail files, 1980–2004 [restricted-use microdata, by mother's state of birth].

National Center for Health Statistics (2025). CDC WONDER Online Database: Natality.

National Center for Health Statistics (2026). National vital statistics system: Natality data, 1975–2026. Accessed May 18, 2026.

Oettinger, G. S. (1999). The effects of sex education on teen sexual activity and teen pregnancy. *Journal of Political Economy 107*(3), 606–644.

Olivetti, C., E. Patacchini, and Y. Zenou (2020). Mothers, peers, and gender-role identity. *Journal of the European Economic Association 18*(1), 266–301.

Oreopoulos, P. and K. G. Salvanes (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives 25*(1), 159–184.

Paton, D., S. Bullivant, and J. Soto (2020). The impact of sex education mandates on teenage pregnancy: International evidence. *Health Economics 29*(7), 790–807.

Public Health – Seattle & King County (2013). *FLASH: Family Life and Sexual Health, Grades 4, 5 and 6*. Seattle, WA: Public Health – Seattle & King County, Family Planning Program. First published 1986; revised 2013.

Rambachan, A. and J. Roth (2023). A more credible approach to parallel trends. *Review of Economic Studies 90*(5), 2555–2591.

Rosky, C. (2017). Anti-gay curriculum laws. *Columbia Law Review 117*(6), 1461–1542.

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights 4*(3), 305–322.

Ruggles, S., S. Flood, M. Sobek, D. Backman, A. Chen, G. Cooper, S. Richards, R. Rodgers, and M. Schouweiler (2024). IPUMS USA: Version 15.0 [dataset].

Sabia, J. J. (2006). Does sex education affect adolescent sexual behaviors and health? *Journal of Policy Analysis and Management 25*(4), 783–802.

Santelli, J. S., L. M. Kantor, S. A. Grilo, I. S. Speizer, L. D. Lindberg, J. Heitel, A. T. Schalet, M. E. Lyon, A. J. Mason-Jones, T. McGovern, C. J. Heck, J. Rogers, and M. A. Ott (2017). Abstinence-only-until-marriage: An updated review of u.s. policies and programs and their impact. *Journal of Adolescent Health 61*(3), 273–280.

Santelli, J. S., L. D. Lindberg, L. B. Finer, and S. Singh (2007). Explaining recent declines in adolescent pregnancy in the United States: The contribution of abstinence and improved contraceptive use. *American Journal of Public Health 97*(1), 150–156.

Sharma, K. (2024). Tackling sexual harassment: short and long-run experimental evidence from india. *Available at SSRN 6460764*.

SIECUS (2015). State profiles 2003-2015.

SIECUS (2025). The path forward: 2025 SIECUS legislative mid-year report. Released 27 August 2025.

Stanger-Hall, K. F. and D. W. Hall (2011). Abstinence-only education and teen pregnancy rates: Why we need comprehensive sex education in the U.S. *PLoS ONE 6*(10), e24658.

Stevenson, B. (2010). Beyond the classroom: Using title ix to measure the return to high school sports. *The Review of Economics and Statistics 92*(2), 284–301.

Troske, K. R. and A. Voicu (2013). The effect of the timing and spacing of births on the level of labor market involvement of married women. *Empirical Economics 45*(1), 483–521.

United States Census Bureau (2021a). Decennial census of population and housing, 1970–2020. Accessed May 18, 2026.

United States Census Bureau (2021b). Intercensal population estimates. Accessed May 18, 2026.

U.S. Census Bureau (2026). American Community Survey 2000-2024. Data Tables, Table B01003. Accessed: May 1, 2026.

U.S. Department of Health and Human Services (2024). Tracking accountability in government grants system (taggs).

U.S. House of Representatives, Committee on Government Reform—Minority Staff, Special Investigations Division (2004, December). The content of federally funded abstinence-only education programs. Technical report, Prepared for Rep. Henry A. Waxman, Washington, DC.

World Bank (2024). Adolescent fertility rate (births per 1,000 women ages 15–19).

A Appendix: Sex Education Policies, Curricula and Funding

This appendix documents the policy variation the paper is built on. Section A.1 describes how the policy panel is coded, states the coding rules and reports the coded provisions; Section A.2 documents, for every state in either treatment arm, the instruments used and the operative language; Section A.3 documents what the curricula taught under the two kinds of mandate and where they operated; and Section A.4 describes the federal funding streams for sex education and their association with state policy.

A.1 Coding state sex education policies

The policy data record, for each state and the District of Columbia, the first school year in which a given requirement applied to public schools: a sex education mandate, an HIV/AIDS education mandate, the treatment of abstinence (stressed or covered), coverage of contraception, content on gender and sexual orientation, consent, and dating violence. Each cell is built from primary legal sources. I start from the statute or administrative rule in force today and trace it back through its amendment history (session laws, history notes in the annotated codes, archived versions of state codes and administrative registers, and state digital libraries) until I reach the first version that contains the requirement.

For many states the first version of a law is not available in electronic form, but a later amending act or an annotated code is. Because amending acts mark every change explicitly (inserted text underlined or capitalized, deleted text struck through or bracketed), and annotated codes list every amending act in the section's history note, I reconstruct the earlier text from the later document: language that appears unmarked in an amending act was already part of the law before the amendment. When the history note leaves it unclear which of two acts introduced a clause, I code the earliest year consistent with the evidence, so that no possibly treated cohort is classified as untreated.

Section A.1.1 states the classification rules. Under them, seven of the nine traditional states (Florida, Georgia, Mississippi, North Carolina, Tennessee, Texas and Utah) stress abstinence on moral grounds. Rhode Island and Kentucky enter on the secondary criterion: their binding text ranks abstinence above other methods—"the preferred method" in Rhode Island's statute, "the only sure means" in Kentucky's Program of Studies, whose curriculum material also frames sexual activity as a risk to students' "physical and emotional health"—and requires no instruction on contraception, so abstinence is the only prevention message students receive. Where the operative requirement is set by the state board of education and a statute or binding regulation makes it compulsory, I code content from the board's documents themselves, as for Georgia's Quality Core Curriculum, North Carolina's Standard Course of Study and Virginia's Family Life Education Standards of Learning.

I benchmark the resulting coding against the existing compilations of state policy: the state profiles published by the Sexuality Information and Education Council of the United States (SIECUS; SIECUS, 2015), the Guttmacher Institute's review of state policies (Gold and Nash, 2001) and its earlier list of state requirements (Donovan, 1998), the coding in Fung and Min (2025), and the survey of state curriculum laws in Rosky (2017), which provides session-law citations for the laws adopted between 1987 and 1996. Where my coding and these sources disagree, I return to the legal text, and the primary source governs; the compilations are used as search indices and as corroboration, not as sources of treatment status. The KFF/Guttmacher *Sex Education in the U.S.* snapshot (Kaiser Family Foundation, 2002) provides a further cross-check for the years before 2003, in which several states adopt. For several states SIECUS changed how it described abstinence—from "covered" to "stressed," for instance—with no accompanying change in the law, which suggests a change in its own definitions rather than in policy, as a member of

the SIECUS team confirmed in a phone call; in those cases I apply one definition throughout.

When the legal text itself cannot be located, I rely on secondary sources to document the instrument and its date: contemporaneous publications of state education agencies and professional associations (many of them archived in ERIC), academic papers, and newspaper coverage of state board and legislative decisions. Cells that rest on secondary sources only are flagged in Section A.2.

Finally, I asked the department of education of every state for historical curricula, board regulations and policy documents, and the material received is incorporated in the coding. Every coded cell is linked to its source in the replication archive, together with a log of all changes made relative to the compilations used as benchmarks.

A.1.1 Coding rules

The following rules are applied uniformly across states.

Mandates. A sex education (or HIV education) mandate requires an unconditional duty to instruct students. Permissive provisions ("may"), duties addressed only to state agencies or school staff, and requirements a district can waive do not count. Board standards or curricula count only where a statute or binding regulation makes them compulsory. A binding duty adopted from 1985 onwards to teach the prevention of communicable disease in a context that explicitly addresses sexual transmission counts as HIV education even if HIV is not named.

Content rules. Content columns record requirements that apply to any instruction a school gives, whether or not instruction is mandatory. Statutory "should include" language is treated as a prescription. Contraception is coded only when contraceptive methods are named as required content; generic references to "methods of prevention" or "ways to reduce the risks" do not count, and a requirement limited to failure rates subordinated to abstinence is a restriction, not a contraception requirement. Where a state adopts binding sexuality standards that specify contraceptive content and do not rank abstinence, abstinence and contraception are coded from those standards even if an earlier HIV-instruction rule that stresses abstinence remains in force; the later, specific standards govern sexuality instruction and the HIV rule governs the HIV unit.

Abstinence. The primary criterion is moral framing. Abstinence is *stressed* when the instrument (i) promotes it on moral grounds: as the expected or appropriate standard, as responsible behavior, or tied to marriage ("until marriage", "premarital", "outside of marriage", "unmarried"); (ii) presents adolescent sexual activity as harmful to students' mental or emotional health; or (iii), as a secondary criterion, ranks abstinence above other methods while no contraception is required (this applies only to Rhode Island and Kentucky among the traditional states). Statements about efficacy ("only certain way", "most effective", "100 percent effective", "best method"), generic rankings ("the preferred method") and bare directives to "stress" abstinence count only under (iii), that is, when contraception is not required. A statement that abstinence is the only certain way, combined with a requirement to teach contraceptive methods as prevention, is coded as *covered*. The definition follows the mechanism studied in the paper, the transmission of norms: abstinence-based programs present abstinence as a value or as the only acceptable option, while comprehensive programs present it as one effective method among others. The state's own curriculum materials, even when not binding, are used to corroborate the classification but do not by themselves create a stress requirement.

Timing. Requirements are dated by their effective date, not by later compliance deadlines, and coded

to the first school year in which they apply (effective year +1; a requirement effective on 1 January, with most of the school year remaining, is coded to that year).

A.1.2 The coded provisions

Table A.1 reports, for every state and the District of Columbia, the year in which each provision first takes effect, and Section A.2 quotes, for each of the nineteen mandates, the clause that creates the duty to instruct and the clause that places the mandate in its arm. Section 2.3 defines the two arms.

Table A.1: State sex-education policies: year each provision takes effect.

Notes: The policy record every treatment variable in the paper is built from; no estimation is involved. Each cell is the calendar year in which that provision first takes effect in that state, and "—" means the state has no such provision. Provisions are coded from state statutes and binding administrative rules, cross-checked against the SIECUS State Profiles; Section 2.3 gives the coding rules and defines the two arms, and the eight states whose record does not resolve into either are excluded from the estimation sample. Cohorts are treated from the first cohort that would have met the mandated curriculum in high school, at age 18.

State	Sex ed	HIV ed	Abst. stressed	Abst. covered	Compreh. focus	Contracep.	Contracep. restriction
Alabama	—	1988	1993	—	—	1993	—
Alaska	—	—	—	—	—	—	—
Arizona	—	—	1992	—	—	—	—
Arkansas	—	1987	1992	—	—	—	Local option only
California	2016	1993	1990	2004	2004	2004	—
Colorado	—	—	1991	2008	2020	2008	—
Connecticut	—	1989	—	—	—	—	—
Delaware	1992	1989	—	1992	—	—	—
Florida	1991	1989	1989	—	—	—	—
Georgia	1989	1989	1995	1989	—	—	Local option only
Hawaii	2016	1997	1996	2016	2016	1996	—
Idaho	—	—	—	—	—	—	—
Illinois	—	1988	1988	2014	2014	2014	—
Indiana	—	1988	1988	—	—	—	—
Iowa	1989	1989	—	2008	—	—	—
Kansas	1989	1989	—	—	—	—	—
Kentucky	1993	1993	1993	—	—	—	—
Louisiana	—	2006	1993	—	—	—	—
Maine	1998	2003	—	2003	2003	2003	—
Maryland	1970	1988	—	—	1992	1992	—
Massachusetts	—	—	—	1990	—	—	—
Michigan	—	1988	1994	—	—	—	—
Minnesota	—	1989	1999	—	—	—	—
Mississippi	2013	—	1999	—	—	—	Failure rates required
Missouri	—	—	2000	—	—	2000	—
Montana	—	2000	—	—	—	—	—
Nebraska	—	—	1987	—	—	—	—
Nevada	1988	1988	—	—	—	—	—
New Hampshire	2009	1988	—	2009	—	—	—
New Jersey	1984	1991	—	2004	1991	1991	—
New Mexico	2008	1989	1989	2008	2008	2008	—
New York	—	1988	1988	—	—	—	—
North Carolina	1996	1988	1996	—	—	2011	Failure rates required; Local consent for referrals
North Dakota	—	—	2013	—	—	—	—
Ohio	—	2002	2000	—	—	—	—
Oklahoma	—	1988	1988	—	—	—	—
Oregon	2010	1994	1994	2010	1994	1994	—

continued on next page

Table A.1, continued

State	Sex ed	HIV ed	Abst. stressed	Abst. covered	Compreh. focus	Contracep.	Contracep. restriction
Pennsylvania	—	2000	2000	—	—	—	—
Rhode Island	1988	1988	1988	1999	—	1999	—
South Carolina	1990	1990	1990	—	—	1990	Family-planning context only
South Dakota	—	—	1998	—	—	—	—
Tennessee	1992	1992	1990	—	—	—	—
Texas	2004	—	1992	—	—	2022	Failure rates required
Utah	1989	1989	1989	—	—	—	Advocacy prohibited
Vermont	1990	1990	—	1990	1990	1990	—
Virginia	1990	1990	1990	—	—	1990	—
Washington	2022	1989	1989	2008	2008	2008	—
Washington DC	1979	1989	—	—	1979	1979	—
West Virginia	—	1990	—	—	—	1990	—
Wisconsin	—	1990	2013	—	2011	—	—
Wyoming	2017	2017	—	—	—	—	—

A.2 State-by-state documentation of the treatment arms

This section documents, for every state in the traditional and in the comprehensive arm, the instruments used and the operative language. The traditional arm comprises states with a sex education mandate and stressed abstinence before any contraception requirement; the comprehensive arm comprises states with a sex education mandate and a contraception requirement while abstinence is not stressed. The arm onsets are stored directly in the data (columns `first_trad_sex_ed`, `first_comp_sex_ed` and their end years), alongside the component variables from which they are computed. Years are coded as the first academic year in which the requirement applied. Quotations are verbatim; ellipses mark omissions. Where a primary text could not be retrieved, the secondary source used is cited.

A.2.1 Traditional arm

Florida (sex education 1991; abstinence stressed 1989). The mandate comes from ch. 90-358 (1990), which amended Fla. Stat. §233.067(4)(c): the comprehensive health education program "shall include the following in all public and laboratory schools: ... Instruction in reproductive health, interpersonal skills, and parenting to reduce teenage pregnancy ... for all students in kindergarten through grade 12, beginning with the 1991-1992 school year", and "Instruction in the benefits of sexual abstinence and consequences of teenage pregnancy ... in life management skills and other appropriate high school courses". The same act requires every district to schedule comprehensive health education in the middle grades from 1990–91, with instruction that "Abstinence from sexual activity outside of marriage is the expected standard for all school age children" (§230.2319). The later clause of §233.061 that lists comprehensive health education "including an awareness of the benefits of sexual abstinence as the expected standard" among the subjects every teacher must teach is not part of the 1990 amendment to that section (s. 60, ch. 90-288, which adds the Declaration of Independence, the Federalist Papers and the Constitution) and came with a later amendment. Abstinence stress comes from §233.0672(2)(a), traced through its history note (ch. 87-329, ch. 88-380, ch. 89-350) to ch. 88-380 (1988): a school whose instruction contains human sexuality shall "Teach abstinence from sexual activity outside of marriage as the expected standard for all school-age children while teaching the benefits of monogamous heterosexual marriage." No contraception requirement applies during the window. The 1990 act's middle-grades instruction includes "human sexuality and pregnancy prevention"; because contraceptive methods are not named, this is not coded as a contraception requirement, a borderline case under the named-contraception rule. The abstinence clauses tie abstinence to marriage ("outside of marriage ... the expected standard"), so Florida is traditional.

Georgia (sex education 1989; abstinence covered 1989, stressed 1995). O.C.G.A. §20-2-143 (Ga. L. 1988, p. 868) requires that "Each local board of education shall prescribe a course of study in sex education and AIDS prevention instruction," to be implemented "not later than July 1, 1989"; its content includes "abstinence from sexual activity as an effective method of prevention", which lists abstinence without ranking it (cover). The statewide Quality Core Curriculum is binding under §20-2-140 (1985): "Each local unit of administration shall include this uniformly sequenced core curriculum as the basis for its own curriculum." Its health objectives, verified in the 1994 printing (local implementation Fall 1994), read: "Identifies methods of preventing pregnancy and sexually transmitted diseases and their degrees of effectiveness or lack thereof (Abstinence is the only sure way to prevent pregnancy and sexually transmitted diseases)." This ranks abstinence (stress from 1995); because contraceptive methods appear only through their failure rates, the objective is not coded as a contraception requirement. The same

objectives state that the student "Recognizes that having children is best undertaken in marriage", which ties the abstinence message to marriage, so abstinence is also stressed on moral grounds from 1995.

Kentucky (sex education and abstinence-based content 1993). The Program of Studies (704 KAR 3:303, authorized by KRS 156.160, first effective 1 August 1992) is binding on all schools. Its 1998 version (archived on the Department of Education's website in 1999) requires a one-half credit health course for graduation, covering "human growth and development" and HIV/AIDS and STD prevention, and requires students to "identify abstinence as the only sure means of preventing pregnancy and STDs"; the 1996 Core Content for assessment adds "strategies for delaying sexual activity" and STD prevention "(with an emphasis on abstinence)". The 1992 text has not been retrieved, so 1993 is a lower bound (the 2004 SIECUS profile reports the same wording, SIECUS, 2015). These texts rank abstinence by efficacy rather than on moral grounds; an explicitly moral statute follows only later, KRS 158.1415 (2018 Ky. Acts ch. 156): "Abstinence from sexual activity is the desirable goal for all school-age children" (2019). Kentucky is coded as abstinence-based, and therefore traditional, from 1993 because the binding Program of Studies stresses abstinence ("the only sure means") and requires no contraception. The state's own curriculum material points the same way: the 1996 Core Content frames sexual activity as a risk to be delayed for "physical and emotional health". Consistent with this classification, Kentucky accepted Title V abstinence-only-until-marriage funds from 1998.

Mississippi (sex education 2013; abstinence stressed 1999). Miss. Code §37-13-171 as amended by H.B. 999 (Laws 2011, ch. 430) provides: "The local school board of every public school district shall adopt a policy to implement abstinence-only or abstinence-plus education into its curriculum by June 30, 2012, which instruction in those subjects shall be implemented not later than the start of the 2012-2013 school year"; "Abstinence-only education shall remain the state standard for any sex-related education taught in the public schools", and such education "Teaches that abstinence from sexual activity before marriage, and fidelity within marriage, is the only certain way to avoid out-of-wedlock pregnancy, sexually transmitted diseases and related health problems". Contraceptives may be discussed "only if that discussion includes a factual presentation of the risks and failure rates" (restriction, not a contraception requirement). The unmarked text in the 2011 act shows that the abstinence standard was already in the 1998 statute. The original 1998 version (H.B. 1304) allowed a local board to authorize sex education without abstinence emphasis by majority vote, so it does not constitute a mandate; its abstinence clause is used only for the content variable (1999).

North Carolina (sex education and abstinence stressed 1996). Session Law 1995-534, adding N.C.G.S. §115C-81(e1), requires a comprehensive school health education program "to be developed and taught to pupils of the public schools of this State from kindergarten through ninth grade", including "Abstinence until marriage education", and provides that "Each local school administrative unit shall provide a comprehensive school health education program that meets all the requirements of this subsection"; instruction "shall teach that abstinence from sexual activity until marriage is the only certain means of avoiding out-of-wedlock pregnancy, sexually transmitted diseases, and other associated health and emotional problems". Local boards had to implement it "by the beginning of the 1996-97 school year". The requirement runs only through ninth grade, so the cohorts already past it in 1996 were not taught under it; the fully exposed cohorts of ATT $[4, 8]$ were fourteen or younger at adoption and were. The 1985 Standard Course of Study listed health content only as "the suggested cognitive scope and sequence", and the Teacher Handbook "provides recommended goals and objectives"; neither is binding

sexuality content. Contraception is coded only from S.L. 2009-213 (2011): earlier objectives present contraceptive methods through their "effectiveness and failure rates" within the abstinence-until-marriage regime.

Rhode Island (sex education, HIV education and abstinence stress 1988; contraception 1999). R.I. Gen. Laws §16-22-18 (P.L. 1987, ch. 575): "Every secondary school teaching courses in family life or sex education within this state shall include, as part of the course instruction, abstinence from sexual activity and refraining from sexual intercourse as the preferred method of pregnancy prevention and the prevention of sexually transmitted diseases." The wording "the preferred method" appears unmarked in the 1996 amending act (P.L. 1996, ch. 400), so it predates 1996; the section's history lists only P.L. 1987, ch. 575 and P.L. 1988, ch. 84 before 1996. Because it cannot be established which of the two introduced the clause, the earliest consistent year (1988) is coded. The 1996 act changed "a" to "the preferred means" only in the AIDS section (§16-22-17). Contraception enters in 1999, outside the window, through the Comprehensive Health Instructional Outcomes (rev. January 1998), made mandatory by the Rules and Regulations for School Health Programs (R16-21-SCHO, §5.1: the curriculum "shall be based on" the state framework "and consistent with the mandated health instructional outcomes"); the 1998 content is inferred from the 2015 edition ("contraception as a means to prevent pregnancy; methods; advantages and disadvantages") and corroborated by Gold and Nash (2001). "The preferred method" ranks abstinence among prevention methods but carries no normative content (no reference to marriage, responsibility or an expected standard), so abstinence is not stressed; but abstinence is ranked first and no contraception is required until 1999, so Rhode Island's content is abstinence-based and it is traditional from 1988 to 1998, and comprehensive from 1999.

Tennessee (sex education 1992; abstinence stressed 1990). The Family Life Curriculum Act (Tenn. Code Ann. §49-6-1301 et seq., Acts 1989, ch. 565) provides: "Beginning with the 1991-1992 school year, if the most recent, annual data maintained by the department of health ... indicate that pregnancy rates in any county exceeded nineteen and five tenths (19.5) pregnancies per one thousand (1,000) females aged fifteen (15) through seventeen (17), then every LEA within the county shall locally devise, adopt and implement a program of family life instruction"; the program "shall emphasize abstinence from sexual relations outside of marriage" and "shall also include a component that specifically addresses the nature and prevention of AIDS and other sexually transmitted diseases". Abstinence stress comes from §49-6-1008 (Acts 1989, ch. 215): all AIDS and STD prevention material directed at students in grades K–12 "shall place primary emphasis on abstinence from premarital intimacy." Because the mandate is triggered by county pregnancy rates, Tennessee is excluded in a robustness check.

Texas (sex education 2004; abstinence stressed 1992; contraception 2022). Texas has no statute requiring sex education; the mandate rests on the state curriculum standards. S.B. 815 (78th Legislature, effective 1 September 2003) made the essential knowledge and skills of the enrichment curriculum, which includes health, a condition of accreditation (19 TAC §74.1(b), as amended effective 3 October 2004: "A school district must provide instruction in the essential knowledge and skills ... in the foundation and enrichment curriculum"); before 2003 these standards "serve[d] as guidelines to school districts" (22 TexReg 7740, 1997). Districts offering grades 6–8 "must provide instruction in the required curriculum" (19 TAC §74.3(a)(1)), and the 1998 middle-school health standards cover "information related to understanding puberty and the reproductive process", pregnancy and abstinence; a one-half

credit of Health 1 was also required for graduation until 2010–11.[19] Abstinence is stressed under Health and Safety Code §163.002(1) (Acts 1991, ch. 14): course materials and instruction "should include ... an emphasis on sexual abstinence as the only completely reliable method", and on moral grounds under Education Code §28.004 (Acts 1995, ch. 260): instruction must "present abstinence from sexual activity as the preferred choice of behavior ... for unmarried persons of school age". Contraception appears only through "effectiveness and ineffectiveness ... keeping in mind the effectiveness of remaining abstinent until marriage" (Health 1, 1998) until the 2020 standards, so Texas is traditional from 2004 to 2021. The mandate is contested: the Texas Education Agency's general counsel described it in 2010 as "an open question" whether districts must teach the human-sexuality parts of the curriculum, because Education Code §28.004 leaves the adoption of human-sexuality instruction to local boards, and no compiler records a Texas mandate. Results are reported with and without Texas.

Utah (sex education, HIV education and abstinence stressed 1989). Laws of Utah 1988, ch. 80 (Public School Curriculum Amendments), as codified at Utah Code §53A-13-101: the State Board shall establish curriculum requirements that include instruction in "community and personal health, physiology, personal hygiene, and prevention of communicable disease"; "That instruction shall stress the importance of abstinence from all sexual activity before marriage and fidelity after marriage as methods for preventing certain communicable diseases." The 1988 session-law text is not available online; the year is dated through the session-law citation in Rosky (2017) and corroborated by Gold and Nash (2001). Utah requires parental consent before students receive this instruction. Later versions of the section also prohibit materials involving "the advocacy or encouragement of the use of contraceptive methods or devices" or "the advocacy of homosexuality" (Utah Code 53A-13-101, 2004).

States with a mandate but in neither arm. South Carolina (1990) and Virginia (1990) combine stressed abstinence with a contraception requirement from the start (for example, South Carolina's Comprehensive Health Education Act requires instruction to "stress the importance of abstaining from sexual activity until marriage" while explaining "methods of contraception and the risks and benefits of each method"). Delaware (1992) has a mandate whose abstinence wording, "an HIV prevention program that promotes abstinence" (1991) and "stresses the benefits of abstinence from high-risk behaviors" (2000 onward), neither ranks abstinence nor frames it morally, and no instrument names contraception; it is coded as covering abstinence without contraception.

A.2.2 Comprehensive arm

District of Columbia (sex education and contraception 1979). The coding rests on the 1979 Competency-Based Curriculum of the D.C. Public Schools. No primary text has been located; the onset is supported by secondary sources that list the District among the jurisdictions requiring sex education by the end of the 1970s (Donovan, 1998).

Maryland (sex education 1970; HIV education 1988; contraception and comprehensive 1992). COMAR 13A.04.18 (Program in Comprehensive Health Education, adopted 18 July 1990 and

[19] A health course was required in high school (from at least the 1998 graduation programs) and in grades 6–8 before 2003, but its content was binding only once S.B. 815 made the health standards a condition of accreditation; both levels are therefore coded from 2004. The State Board removed the high school health credit from 2010–11, so from then on the requirement operates only through middle-school health. This weakens treatment for cohorts reaching high school from 2011 but does not end the mandate.

effective 15 October 1990; 17:8 Md. R. 991–993 and 17:20 Md. R. 2430) requires annual comprehensive health education for all students in grades K–8, under a curriculum that "shall include" the subgoal "Develop and use skills for making responsible decisions about family planning and preventing pregnancy" (§.01B(6)(i), later §.02F(9)). Because "family planning and preventing pregnancy" does not name contraceptive methods (as with Florida's "pregnancy prevention"), contraception is dated from the 1991 revision (effective 13 May 1991). It requires every local school system to offer a course on advanced human sexual behavior (Focus Area Three), "offered several times each school year to make it available to as many students as possible", whose content "shall be included in each program of the local school system": "... (iv) Contraception; ... (vii) Family planning". The course is elective and requires prior parental consent, but every school system had to provide it, so the regulation made contraception part of the sex education content of every district rather than a local option. The compilations read Maryland's requirement the same way: Gold and Nash (2001) record contraception as covered in both its sexuality and its STD education, and the SIECUS State Profiles list "contraception" and "family planning" among the topics of the family-life goal of the required health program (SIECUS, 2015). The regulation contains no abstinence wording. HIV education comes from the AIDS Prevention Education bylaw (COMAR 13A.04.15; emergency rule effective 6 October 1987, permanent 2 May 1988): "local school systems shall provide annual instruction in acquired immune deficiency syndrome to all students at least once in grades 3 to 6, in grades 6 to 9, and in grades 9 to 12". The State Department of Education's 1988 guidelines, which state that abstinence "prior to marriage is physically, mentally, morally, and ethically the appropriate decision", are non-binding ("curriculum development and implementation ... is left solely to each local school system") and are not coded. The 1970 mandate rests on State Board Bylaw 720:3, quoted in *Cornwell v. State Board of Education*, 428 F.2d 471 (4th Cir. 1970): "It is the responsibility of the local school system to provide a comprehensive program of family life and sex education in every elementary and secondary school for all students as an integral part of the curriculum including a planned and sequential program of health education." Parents challenged the bylaw on First and Fourteenth Amendment grounds; the district court dismissed the action and the Court of Appeals affirmed on 9 June 1970. The bylaw, later codified as the Family Life and Human Development bylaw (COMAR 13A.04.01), was therefore in force before June 1970; the quoted text does not name contraception. The 2019 replacement regulation requires that teaching "emphasize that refraining from sexual activity is the best method to avoid sexually transmitted infections, including HIV, and unintended pregnancy" alongside "medically accurate information about contraception and condoms"; this is an efficacy statement (not moral) and falls outside the window.

Vermont (sex education, HIV education, contraception and comprehensive 1990). Public Acts 1988, No. 270 repealed 16 V.S.A. §134(a) ("Nothing in this subchapter shall constitute or authorize the imposition of mandatory requirements on any public or private school") and added comprehensive health education to the minimum course of study in 16 V.S.A. §906, effective 1 September 1989. The definition in §131 includes HIV infection and other STDs, "decision making about sexual activity including abstinence", and "information regarding the possible outcomes of premature sexual activity, contraceptives, adolescent pregnancy, childbirth, adoption, and abortion." The comprehensive health education subchapter itself dates from 1978 (Act No. 131 of the 1977 adjourned session), but §134(a) of that act provided that nothing in the subchapter "shall constitute or authorize the imposition of mandatory requirements on any public or private school", so no onset is coded before the 1988 act.

New Jersey (sex education 1984; HIV education, contraception and comprehensive 1991). N.J.A.C. 6:29-7.1 (1980) requires every district to provide a family life education program, defined as

instruction whose "program of studies shall include the physiological and psychological basis of human development, sexuality, and reproduction"; full implementation was required by September 1983. Contraception and the comprehensive classification are dated from the 1990 rule on comprehensive health education (N.J.A.C. 6:29-4, adopted 5 March 1990, 22 N.J.R. 793), which introduced HIV education. The rule leaves the curriculum to districts, but it twice obliges the State Department of Education to provide them with "instructional guides"—for HIV prevention education (6:29-4.1(e)3) and for family life education (6:29-4.2(i))—a phrase the Board substituted for "guidelines" at adoption "to clarify the fact that actual publications are being used". The Departments of Health and Education's *AIDS: Instructional Guide for Teachers, Grades 9 through 12* (1987) teaches contraception throughout: students are to "explain how condoms can help prevent infection with HIV if used during anal or vaginal intercourse", discuss "what keeps them from using condoms", and rehearse "asking partner to wear a condom"; abstinence is one prevention method among several, never a moral standard. The coding is corroborated by Gold and Nash (2001), who record contraception as covered in both New Jersey's sexuality and STD education. The AIDS Prevention Act (P.L. 2001, c. 303) later required that instruction "stress that abstinence from sexual activity is the only completely reliable means" of preventing pregnancy and STDs, an efficacy statement that does not change the classification.

Maine (sex education 1998; contraception and comprehensive 2003). The mandate comes from the education title: 20-A M.R.S. §4723 (P.L. 1983, c. 859) requires that "The secondary course of study shall include instruction in health, safety and physical education, as prescribed by the commissioner", and the commissioner's prescription is the system of Learning Results (20-A M.R.S. §6209, P.L. 1995, c. 649), made binding by Department of Education rule ch. 127. The July 1997 Learning Results contain only minimal sexuality content: students in grades 5–8 "Analyze the effects that risky behaviors have on personal health (e.g., tobacco, drugs, poor nutrition, sexual activity . . .)" and "Identify the characteristics and stages of human growth and development"; they do not mention HIV, abstinence or contraception. Content comes from 22 M.R.S. §1902(1-A) (P.L. 2001, c. 654, 2002), which defines comprehensive family life education as "education in kindergarten to grade 12 regarding human development and sexuality, including education on family planning and sexually transmitted diseases, that is medically accurate and age appropriate; . . . that promotes responsible sexual behavior with an emphasis on abstinence; that addresses the use of contraception; . . . ". Abstinence is emphasized but neither ranked above contraception nor framed morally, and is coded as covered. The 2002 act does not itself oblige schools to teach: §1910 directs the health commissioner to "undertake initiatives to implement effective, comprehensive family life education services" by contracting with family planning programs for training, staff, forums and grants, and §1911 provides an opt-out "To the extent that comprehensive family life education takes place in a school". Following the 2004 SIECUS profile (SIECUS, 2015) and later secondary sources, the definition is read as governing the content of school instruction; unlike SIECUS, the mandate is not attributed to the 2002 act. This is consistent with Gold and Nash (2001), who record a Maine sex education requirement in August 2001, before the 2002 act existed. The comprehensive classification therefore rests on reading the Title 22 definition as governing the content of school instruction, a reading shared by all the secondary sources but not stated in the education title itself. It differs from the agency-only provisions not counted elsewhere (Missouri's state-prepared HIV programs, Texas Health and Safety Code §85.007, Louisiana Act 534 of 2014) because in Maine the duty to teach comes from the education title, and the 2002 definition only specifies what that instruction contains. Dropping Maine from the comprehensive arm is reported as a robustness check.

New Mexico (sex education, contraception and comprehensive 2008). The 14 February 2006 amendment of the health performance standards (6.30.2.19 NMAC) requires, in grades 9–12, that students "understand human reproduction and how pregnancy can be prevented through the use of various methods of contraception including barrier and hormonal methods" and "identify alternatives to health risk behaviors ... (i.e. abstinence, condom use, other pregnancy prevention methods ...)". The earlier HIV rule (6.12.2.10 NMAC, 1988), requiring instruction on "ways to reduce the risks of getting HIV/AIDS, stressing abstinence", remains in force but governs the HIV unit; under the later, specific standards abstinence is not ranked. No benchmark in the 2006 health standards ranks abstinence or frames it morally; abstinence always appears alongside condom use and other methods. Gold and Nash (2001) and later Guttmacher compilations code New Mexico's abstinence as stressed, presumably on the basis of the HIV rule; dropping New Mexico from the comprehensive arm is reported as a robustness check.

Oregon (sex education 2010; contraception 1994; abstinence stressed 1994–2009, covered from 2010). ORS 336.455 (1993 c. 775, §1) sets criteria for "all human sexuality education courses that discuss human sexuality in public elementary and secondary schools": instruction shall "Promote abstinence for school-age youth ... as the safest and most responsible sexual behavior. However, abstinence shall not be taught to the exclusion of other material and instruction on contraceptive and disease reduction measures," and "Pupils shall be provided with statistics based on the latest medical information regarding both the possible side effects and health benefits of all forms of contraceptives." Subsection (2) adds that such courses "shall emphasize that abstinence from sexual contact is the only method that is 100 percent effective ... Abstinence is to be stressed, but not to the exclusion of other material and instruction on contraceptive and disease reduction measures." Because contraceptive methods must be taught as prevention alongside the abstinence message, abstinence is coded as covered (the same rule is applied to California's S.B. 71); contraception and cover start in 1994. The mandate comes from House Bill 2509 (2009), which inserted into ORS 336.455: "Each school district shall provide age-appropriate human sexuality education courses in all public elementary and secondary schools as an integral part of the health education curriculum"; the amendments "first apply to the 2009-2010 school year" (sex education 2010). The 1993 wording "the safest and most responsible sexual behavior" is a value judgment, so abstinence is stressed from 1994; House Bill 2509 (2009) replaced it with "the most effective way to prevent pregnancy and the transmission of sexually transmitted diseases", an efficacy statement, ending the moral regime in 2010, the year the mandate begins. Oregon enters the comprehensive arm in 2010.

California (sex education 2016; abstinence stressed 1990–2003; contraception and cover 2004). Senate Bill 71 (Stats. 2003, ch. 650, effective 1 January 2004) repealed Education Code §51553, whose 1988 version required that classes "emphasize that abstinence from sexual intercourse is the only protection that is 100 percent effective" (stress, 1990–2003). Under S.B. 71, instruction must "teach that abstinence from sexual intercourse is the only certain way to prevent unintended pregnancy ... and provide information about the value of abstinence while also providing medically accurate information on other methods" and cover all FDA-approved contraceptive methods; abstinence is therefore coded as covered. The mandate comes from the California Healthy Youth Act (A.B. 329, 2015, effective 1 January 2016). Its Legislative Counsel's Digest records that existing law "authorizes school districts to provide comprehensive sexual health education ... and requires school districts to ensure that all pupils in grades 7 to 12, inclusive, receive HIV/AIDS prevention education", and that the bill "would instead require school districts to ensure that all pupils in grades 7 to 12, inclusive, receive comprehensive sexual health education and HIV prevention education". Education Code §51934(a) now reads: "Each school district

shall ensure that all pupils in grades 7 to 12, inclusive, receive comprehensive sexual health education and HIV prevention education from instructors trained in the appropriate courses. Each pupil shall receive this instruction at least once in junior high or middle school and at least once in high school." HIV education alone had been required since 1992–93 (Educ. Code §51201.5, Stats. 1991, ch. 818: "Commencing in the 1992-93 school year, school districts shall ensure that all pupils in grades 7 to 12, inclusive ... receive AIDS prevention instruction"). The 1988 criteria ("abstain from sexual intercourse until they are ready for marriage"; "honor and respect for monogamous heterosexual marriage") stress abstinence from 1990 until S.B. 71 took effect in 2004.

Hawaii (sex education 2016; contraception 1996; abstinence stressed 1996–2015; comprehensive 2016). Board of Education Policy 2110 (1995): "the Department of Education shall instruct students that abstention from sexual intercourse is the surest and most responsible way to prevent unintended pregnancies, sexually transmitted diseases such as HIV/AIDS ... The abstinence-based education program shall: ... provide youth with information on and skill development in the use of protective devices and methods for the purpose of preventing sexually transmitted diseases and pregnancy." Policy 2110 is read as a content rule rather than a mandate. "Most responsible" is a moral judgment, so abstinence is stressed from 1996 to 2015. The mandate comes from Policy 103-5 (approved 16 June 2015): "the Department of Education shall provide sexual health education to include age appropriate, medically accurate, health education that: (1) Includes education on abstinence, contraception, and methods of infection prevention to prevent unintended pregnancy and sexually transmitted infection, including human immunodeficiency virus; ... Instruction will emphasize that abstention from sexual intercourse is the surest way to prevent unintended pregnancies, sexually transmitted infections such as HIV/AIDS, and consequent emotional distress." The 2015 wording keeps only the efficacy claim ("surest way") together with required contraceptive content, so abstinence is covered from 2016.

Washington (sex education 2022; contraception, cover and comprehensive 2008). The Healthy Youth Act (Laws 2007, ch. 265; RCW 28A.300.475): every school that offers sexual health education must ensure that it "includes information about abstinence and other methods of preventing unintended pregnancy and sexually transmitted diseases" and "Abstinence may not be taught to the exclusion of other materials and instruction on contraceptives and disease prevention." The 1988 AIDS Act (Laws 1988, ch. 206, §402) still requires the AIDS unit to "stress that abstinence from sexual activity is the only certain means" of preventing sexual transmission; it governs the HIV unit only. The mandate comes from S.B. 5395 (Laws 2020, ch. 188, approved by voters as Referendum 90), which amended RCW 28A.300.475: "every public school shall provide comprehensive sexual health education to each student by the 2022-23 school year", and "Beginning in the 2021-22 school year, comprehensive sexual health education must be provided to all public school students in grades six through twelve" (coded 2022).

A.3 Curricula and programs

Statutes set the boundary; curricula fill it. The most systematic evidence on what abstinence-only programs contained comes from a 2004 review prepared for Rep. Henry A. Waxman by the Special Investigations Division of the House Committee on Government Reform (U.S. House of Representatives, Committee on Government Reform—Minority Staff, Special Investigations Division, 2004). The review obtained program summaries from all 100 organizations receiving SPRANS community-based abstinence funding in fiscal year 2003 and examined every curriculum named by at least five grantees—thirteen in

all. Eleven of the thirteen were found to contain errors or distortions of public health information.

The scale is what makes these curricula relevant rather than anecdotal. Those eleven curricula were used *in 25 states by 69 grantees*, among them state health departments, school districts and hospitals as well as religious and pro-life organizations. The 69 grantees received over $32 million in SPRANS funding in fiscal 2003 alone and over $90 million since fiscal 2001. The most widely adopted were *Choosing the Best Life* (32 grantees), *Choosing the Best Path* (28), *A.C. Green's Game Plan* (23) and *WAIT Training* (19). Nor did they disappear with the program that funded them: SIECUS's catalogue of federally funded programs still lists the *Choosing the Best* series—six grade-specific volumes running from sixth grade through high school—in fiscal year 2016.

A.3.1 Abstinence-only curricula

The gender content of these curricula is the reason they bear on this paper's norm channel. *WAIT Training* lists "Financial Support" among the "5 Major Needs of Women" and "Domestic Support" among the "5 Major Needs of Men," and puts "sexual fulfillment" and "physical attractiveness" among the men's needs. It explains that "just as a woman needs to feel a man's devotion to her, a man has a primary need to feel a woman's admiration," and teaches that "a male is usually less discriminating about those to whom he is sexually attracted," while "women usually have greater intuitive awareness of how to develop a loving relationship" and "while a man needs little or no preparation for sex, a woman often needs hours of emotional and mental preparation." *Why kNOw* instructs that "women gauge their happiness and judge their success by their relationships. Men's happiness and success hinge on their accomplishments," and explains that when women entered the labor force "companies noticed that women were not as concerned about preparing for retirement," which "stems from the priority men and women place on the past, present, and future." Discussing wedding traditions, the same curriculum teaches that "the father gives the bride to the groom because he is the one man who has had the responsibility of protecting her throughout her life," and is "now giving his daughter to the only other man who will take over this protective role."[20] *Choosing the Best Life* tells instructors that "generally, guys are able to focus better on one activity at a time and may not connect feelings with actions," whereas "girls access both sides of the brain at once, so they often experience feelings and emotions as part of every situation."

SIECUS's contemporaneous state reviews reproduce the same material and add to it. Of *Choosing the Best LIFE*, the reviews find that it "suggests that sexually active teens will never have a happy future, and implies that only teens with low self-esteem and poor judgment become sexually active," quoting: "relationships often lower the self-respect of both partners—one feeling used, the other feeling like the user. Emotional pain can cause a downward spiral leading to intense feelings of lack of worthlessness."[21] Of *WAIT Training*, that it "explains, 'When it comes to sex, men are like microwaves and women are like crockpots. . .[M]en respond sexually by what they see and women respond sexually by what they hear and how they feel about it,"' alongside "messages promoting biased views of gender, sexual orientation, and family structure."[22]

These are not incidental passages. They state, as fact and to entire cohorts of adolescents, that women's returns come from relationships and men's from accomplishments—which is the prescription

[20] U.S. House of Representatives, Committee on Government Reform—Minority Staff, Special Investigations Division (2004), citing *WAIT Training* pp. 196 and 199 and *Why kNOw* pp. 61 and 122.

[21] SIECUS (2015), FY2004–FY2006 profiles; the passage is reproduced verbatim in each state where a *Choosing the Best* grantee operated.

[22] SIECUS (2015), Colorado profile, FY2011.

the norm cost $\tau(D)$ in Section 3 is meant to capture, delivered at exactly the age at which the model places the schooling decision.

A.3.2 Where they were taught

The state-level reviews make the geography explicit, because SIECUS documents each federally funded grantee in the state where it operates. The two curricula are of different vintages, and the earlier one is the one that overlaps the cohorts this paper identifies from. *Choosing the Best* peaks in the mid-2000s: its grantees appear in nine state profiles in FY2004—Alabama, Arizona, Colorado, Florida, Georgia, Iowa, Missouri, Oklahoma and Texas—and in eighteen by FY2006, adding Arkansas, Indiana, Kansas, Kentucky, Michigan, Mississippi, New Jersey, Ohio and Tennessee. *WAIT Training* is the later of the two, appearing in FY2011 in Arizona, Colorado, Florida, Georgia, Maryland, Michigan, Nebraska, New Hampshire, Pennsylvania and Utah. *Why kNOw*, produced in Chattanooga, is concentrated in Tennessee throughout.[23]

This map does *not* line up with the mandate arms. Of the eighteen states in which a *Choosing the Best* grantee operated in FY2006, eleven never adopt a mandate of either kind over the sample period (nine adopt none, and Iowa and Kansas adopt one that fits neither arm); six are traditional adopters (Florida, Georgia, Kentucky, Tennessee, Texas and, later, Mississippi); and one—New Jersey—is a comprehensive adopter. Federal abstinence dollars were allocated by application and by formula, not by statutory content, so the curricula they funded travelled across the arms. Section A.4.3 shows this directly: mandate status explains little of the funding variation, and what it does explain runs through the mandate label rather than through the content indicators on their own: comprehensive states draw less Title V money and refuse it more often, while traditional states, whose curriculum the federal programs mirrored, take it almost every year.

The consequence for identification is that these curricula are evidence about *content*—what abstinence-stressed instruction says to adolescents about gender—and not a second measure of treatment. Reading them as treatment would be a mistake in both directions: it would attribute to mandates a curriculum that never-adopting states also used, and it would miss that the mandate is what governs when the two conflict.

New Jersey is the instructive case, and it cuts the way the design requires. It appears in the FY2004–FY2006 profiles as a Title V state whose sub-grantees used *Choosing the Best* among seven other curricula—not as state policy, but as one option among many at the sub-grantee level. In October 2006 the state rejected the Title V allotment outright, on the stated ground that the federal eight-point definition of abstinence education "is not consistent with the Comprehensive Health and Physical Education standards and New Jersey's AIDS Prevention Act of 1999," and that accepting the money might cost the state more because "students may require additional sexuality education to clarify the partial information and correct the misinformation" taught under it.[24] The curriculum is gone from the New Jersey profile thereafter. This is the mechanism behind the funding result in Section 2.3—that comprehensive states are distinguished less by attracting comprehensive grants than by declining abstinence ones—observed in a single state, and it is evidence that where the mandate and the federal curriculum conflict, the mandate binds.

[23] Author's tabulation of the state profiles in SIECUS (2015). Counts are of states in which a grantee using the curriculum is named, not of students reached, and understate coverage where districts adopted a curriculum without federal funding.

[24] SIECUS (2015), New Jersey profiles FY2006 and FY2008, quoting the letter from the Commissioners of Health and Senior Services and of Education to HHS. By the FY2008 profile *Choosing the Best* no longer appears in New Jersey.

A.3.3 Comprehensive curricula

Programs on the comprehensive side are built around skills and negotiation rather than prescription. *Making Proud Choices!*, an eight-session curriculum for adolescents aged 11–13 delivered in both school and community settings, consists of activities designed to "increase comfort with practicing condom use, address concerns about negative effects of practicing safer sex, and build skills in condom use and negotiation."[25] The manual itself states the goal as to "increase knowledge and perception of personal vulnerability, develop positive attitudes toward safer sex, and build the skills and confidence needed to abstain successfully or to use safer sex practices willingly and effectively if they choose to be sexually active." In the fifth edition two of its fourteen modules are devoted to "Developing Condom Use and Negotiation Skills," and a myths-and-facts exercise tells students that sexuality "is a part of who we are as human beings" and that "all people are sexual, and sexuality is a natural and normal part of being human." Gender enters only in passing: a module goal is to "confront stereotypes about who becomes infected with HIV," and the same exercise notes that in pornography "you often see women in stereotypical roles, as objects for men's pleasure, which encourages violence against women" (Jemmott et al., 2016). Federally funded grantees used it in 34 states between fiscal years 2009 and 2015.

FLASH (Family Life and Sexual Health), written by the Family Planning Program of Public Health – Seattle & King County and used by 47 of the Washington districts that taught sex education in 2009, is the clearest case of a comprehensive curriculum that takes on gender roles as a subject. Lesson 4 of the fourth- to sixth-grade volume, "Gender Roles," defines them as "the rules people think they should follow because of their sex... ways they act because 'boys and men are supposed to' or 'girls and women ought to,'" and uses a case study "to teach the concepts that (a) gender roles are learned and (b) roles can be limiting." In the case study a girl tells a boy that "only boys are doctors," and the class is asked: "Why might Mary think that only boys can be doctors? How could that stereotype limit her in life?" and "If Mary keeps believing that girls cannot be doctors, do you think she will do well in math and science classes in high school?" The lesson closes by having pupils look for women and men in each profession and discuss "how a lack of role models in a field, whether the field is parenting or politics, can limit a person's imagination about what he or she can become" (Public Health – Seattle & King County, 2013). The lesson carries a 1986 copyright, so unlike the other comprehensive examples it predates the comprehensive mandates in the sample, although the text available is the 2013 revision.

Gender Matters, delivered to 14–15 year-olds in Austin, Texas under a Tier 2 teen pregnancy prevention grant, is described as an "innovative, science-based intervention designed to address... the impact of gender norms on the sexual and reproductive health behaviors of youth," aiming to "[influence] traditional beliefs about masculinity and femininity, [advance] more equitable attitudes about relationships and the balance of power within them, and [influence] other gender specific views, such as responsibility for pregnancy prevention."[26] That a comprehensive program names the gender-norm channel as its explicit target, while the abstinence-only curricula transmit the norm as content, is the sharpest available statement of what separates the two arms. The difference is not primarily about contraception: it is whether the curriculum tells a girl what her role is or equips her to negotiate within it.

The documentary record is asymmetric. Federal money in this period was earmarked for abstinence-only instruction—SPRANS, Title V and CBAE—so the curricula on the traditional side were catalogued grantee-by-grantee and state-by-state, while comprehensive programming was locally funded and largely undocumented until the Teen Pregnancy Prevention Initiative created a federal record in 2010. The comprehensive examples above, *FLASH* aside, are therefore later than the mandates they illustrate and

[25] SIECUS (2015), FY2009 profiles (e.g. Delaware, Massachusetts, Ohio).
[26] SIECUS (2015), New York profile, FY2011, describing EngenderHealth's grant.

national rather than state programs: evidence of what comprehensive instruction consists of, not of its geographic reach.

A.4 Federal sex education funding

A.4.1 Conditions attached to federal funds

Before federal money paid for abstinence education as such, Congress attached content conditions to the money that paid for HIV education. For fiscal year 1988, Section 514 of the appropriations act for the Departments of Labor, Health and Human Services and Education (Pub. L. 100-202), an amendment by Senator Jesse Helms, required that the AIDS education it funded "emphasize—(1) abstinence from sexual activity outside a sexually monogamous marriage," on pain of repayment; it reached schools through CDC, which funded the education departments of fifteen states and twelve large cities from the fall of 1987 and of every state from the fall of 1988 (Centers for Disease Control, 1988). The acts for fiscal years 1989–1991 replaced it with a clause without the marriage formula or the enforcement, barring CDC-funded curricula "designed to promote or encourage ... sexual activity" and requiring information on "the health risks of promiscuous sexual activity"; the AIDS Amendments of 1988 (Pub. L. 100-607) made a similar clause permanent for Public Health Service AIDS programs (42 U.S.C. 300ee). From 1994, the Improving America's Schools Act (Pub. L. 103-382, §14511) bars federal school funds from sex or HIV education that is not "age appropriate" and omits "the health benefits of abstinence," and from condom distribution in schools, while denying federal officials the power to "mandate, direct, review, or control" a school's curriculum; the provision is now 20 U.S.C. 7906. From 1996, Title V (42 U.S.C. 710; Section A.4.2) brought back the moral formula of 1987, "abstinence from sexual activity outside marriage as the expected standard for all school age children," attached to a stream a state can decline. None of these provisions enters the estimation: each applies to every state at once, so the birth-cohort fixed effects absorb it; what they document is the message federally supported instruction carried, against which a state mandate either restated the abstinence prescription or departed from it.

A.4.2 The funding streams

Federal money for sex education arrives through programs that differ in kind, and the difference matters: a correlation between state policy and any one of them means something different depending on who applies for the money and who can turn it down. Figure A.1 and the tables below report three streams.

Title V is a **state formula allotment** for abstinence education (listing number 93.235, FY1998–2020). Each state is allotted a share in proportion to its share of the nation's low-income children, a state that accepts it must put up three non-federal dollars, in cash or in kind, for every four federal dollars, and a governor can simply decline it, so a state's total measures political willingness as much as need. Neither the competitive abstinence grants nor the later comprehensive programs require a match. Cumulative federal spending on abstinence-only programs has exceeded $2 billion (Santelli et al., 2017), and the figure roughly doubles once state and private expenditures are counted. Accepting the allotment does not determine how it is spent: the state chooses who runs the funded programs and in which settings. The listing was renamed "Title V State Sexual Risk Avoidance Education" in FY2018, so its awards carry SRAE wording from then on; it remains the formula allotment and is not the competitive program of that name.

CBAE and its successor *SRAE* are **competitive abstinence grants**, and they are reported as one

stream. They are the same money under two names—Congress terminated CBAE and the competitive abstinence appropriation reappeared as SRAE in FY2012—and both are awarded to organizations rather than to states, so a state's total depends on who happened to apply there and reading either as a policy choice by the state would be a mistake. CBAE ran FY2001–09 and SRAE FY2012–20; FY2010–11 carry no competitive abstinence money at all, so the pooled series is zero in those two years by construction rather than by measurement. In its first four years, FY2001–04, CBAE was a set-aside inside listing number 93.110; it received its own number, 93.010, in FY2005, and the stream includes both. SRAE is 93.060.

PREP (93.092) and *TPP* (93.297) are the two **comprehensive** programs, FY2010–20, and they are likewise reported as one stream. They are not alike—PREP is a population formula and TPP a grant competition, so pooling them averages two allocation rules—but 93.092 is not a pure formula grant either, and splitting them leaves too few states in each to support a separate reading.

Two aggregation points. "All abstinence funds" is Title V plus the competitive grants plus the adolescent family life program (93.995), which mixes a formula a state can refuse with competitions it does not control; it is the right aggregate for asking how much abstinence money a state ends up with, and the wrong one for asking what the state chose. The family life program (Title XX) bundles abstinence prevention with *care* for teens who are already pregnant or parenting and is 4.9 percent of all abstinence funding over 1998–2020. For FY2003–15 the obligations are spliced with the organization-level awards listed in the SIECUS State Profiles, which reproduce SIECUS's own state totals in 79 percent of state-years for the competitive abstinence grants and in 97 percent for TPP and PREP; the FY2008 profiles list only "some examples" of grantees and are the main source of the gap.

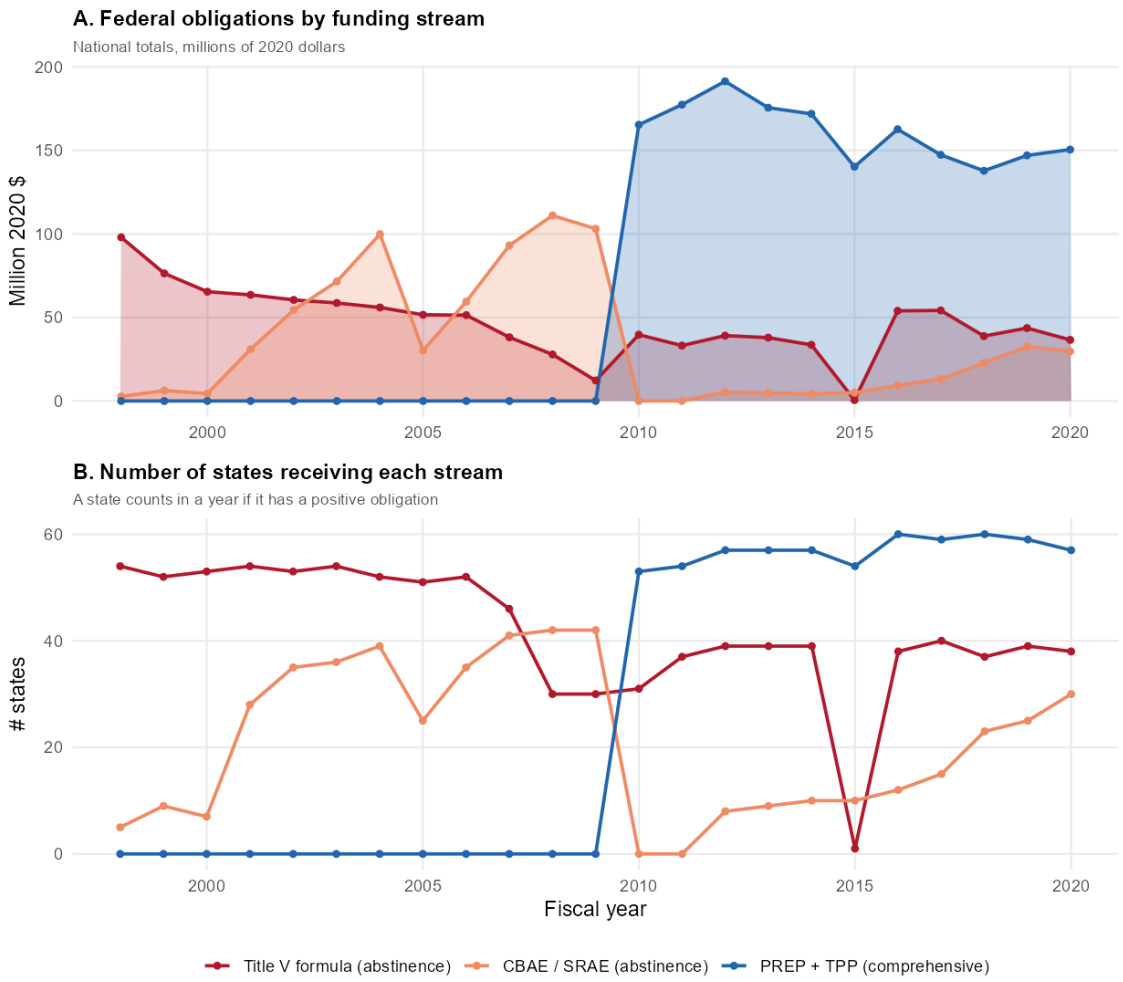


Figure A.1: Federal sex education funding by stream and fiscal year: national obligations in real 2020 dollars (panel A) and the number of recipient states (panel B).

Notes: This is a description of the funding data, not an estimate. Panel A plots total federal obligations in each stream, deflated to 2020 dollars; panel B counts a state in a year if it has a positive obligation in that stream, and the count can exceed 51 because territorial and tribal grantees appear in TAGGS. The three streams are the Title V formula allotment, the competitive abstinence grants (CBAE and its successor SRAE pooled) and the comprehensive programs (PREP and TPP pooled); the gap in the competitive abstinence series in FY2010–11 is real, as neither program was funded in those years. Obligations come from the TAGGS grant database and the Title V allotment, spliced with the SIECUS annual tables for FY2003–15. Section A.4.2 describes each program.

A.4.3 Funding and state policy

Table A.2 regresses each funding stream, in state-year obligations per youth aged 15–19 assembled from the TAGGS grant database (U.S. Department of Health and Human Services, 2024) and the Title V allotment, on the policy indicators. The coefficients are descriptive, not causal: there is no source of identifying variation, and they are read as a correlation between the statutory coding and a fiscal choice the state makes, not as an effect of a mandate on funding. A comprehensive mandate is associated with 0.25 (0.13) log points less Title V money, the one stream a state itself accepts or declines, with 0.25 (0.18) log points less abstinence-only funding overall, and with no difference in comprehensive funding, −0.04 (0.12). On curriculum content rather than the composite labels, stressing abstinence carries 0.20 (0.19) log points more Title V money and 0.29 (0.20) more abstinence funding overall, neither significant, while requiring contraception carries 0.41 (0.24) log points less Title V money and nothing on comprehensive

funding.

Declining the Title V allotment outright is a real margin: part of the association between a comprehensive mandate and Title V dollars is a state refusing the allotment rather than drawing a smaller one. Declining is not rare, and it follows the mandate label more closely than the content coding. Over the fiscal years the panel covers, 1998–2017, a state-year draws no Title V money 37 percent of the time under a comprehensive mandate (six of the nine states decline at some point), 19 percent with no mandate (25 of 39 states) and 1 percent under a traditional mandate; the only traditional decline is Tennessee in 2008–09. Sorted by content, states whose statute stresses abstinence decline 11 percent of the time and the rest 33 percent. In a linear probability model with the specification of Table A.2, stressing abstinence lowers the probability of declining by 0.03 (0.15) and a traditional mandate by 0.27 (0.12), while the comprehensive-mandate coefficient, 0.04 (0.11), is not significant once year effects are in: declining is mostly a post-2007 phenomenon, after the federal evaluation of the program and the lapse of its authorization in 2009, and about a third of no-mandate states declined too. What sets comprehensive states apart is timing—Maine declined from 2005, and three of the five states holding a comprehensive mandate in 2007 had declined by then, against 16 percent of no-mandate states—and the gap narrows by 2010.

Mandate switches. The associations above compare states. Within states the funding record allows few comparisons, because it begins in FY1998 and most mandates predate it: six states adopt a comprehensive mandate inside the Title V window—Rhode Island (1999, switching from its traditional mandate), Maine (2003), New Mexico (2008), Oregon (2010), California and Hawaii (2016)—and Texas (2004) and Mississippi (2013) adopt a traditional one before the series ends in FY2017. Title V falls in two of the six comprehensive switchers, both by refusal. New Mexico's governor withdrew from the program in December 2007, as its mandate took effect, and the state took no Title V money until 2012, when it resumed it for community-based programs only; Maine took the allotment for two years, for a media campaign, and declined it from 2005. Each fall, measured as the change in log funding per youth relative to the never-mandate states between the three years before adoption and the four years from adoption, is larger than the change in any of the 31 never-mandate states over the same years. Oregon kept its allotment, and kept it in classrooms, but from 2010 opened it to its own curriculum; Hawaii's allotment did not change; California had taken none since fiscal year 1999. Pooling the six in a Callaway–Sant'Anna event study against the never-mandate states gives -0.38 (0.11) log points of Title V money over the seven years from adoption, with an average pre-adoption coefficient of 0.05 ($p = 0.54$); the two-way fixed-effects analogue is -0.19 (0.15), and the joint pre-trend test, computed from six treated states, rejects.

Comprehensive funding cannot be studied the same way. It begins in 2010, so only California and Hawaii switch inside its window, each with two post-adoption years. California's rises slightly after 2016, within the range of the never-mandate states; Hawaii's fell by nine tenths in FY2015, a year before its mandate, when its teen pregnancy prevention grants ended and none of the 2015 round went to the state. Mississippi's and Texas's traditional mandates left their Title V allotments on the path of the never-mandate states, and its comprehensive funding rose in FY2015 with the same round of teen pregnancy prevention grants. A mandate switch, in short, predicts a refusal of abstinence money where the federal conditions conflicted with a comprehensive state law, and little else; the cross-state differences in Table A.2 reflect persistent differences between states more than responses to the mandates.

	Title V allotment	All abstinence funds	All comprehensive funds
Mandate type			
Comprehensive sex-ed mandate	-0.249*	-0.249	-0.042
	(0.130)	(0.176)	(0.121)
Traditional sex-ed mandate	0.419***	0.399**	0.038
	(0.160)	(0.159)	(0.105)
Observations	980	980	392
Curriculum content (separate specification)			
Abstinence stressed	0.203	0.292	0.104
	(0.191)	(0.202)	(0.136)
Contraception covered	-0.412*	-0.372	-0.039
	(0.240)	(0.246)	(0.107)
Observations	980	980	392

Table A.2: Federal sex-education funding and state sex-education policy.

Notes: Outcome is $\log(1 + \text{real 2020 \$ per youth aged 15–19})$ of the funding stream or total named in the column head; the sample is a state-fiscal-year panel of 49 states over the years in which the programs in that column operated. Each column is a separate OLS regression on the policy indicators shown, with year fixed effects, weighted by the state population aged 15–19; standard errors clustered by state in parentheses. The mandate panel and the content panel are separate specifications, and the two mandate coefficients are each relative to states with no mandate. Policy indicators are time-varying, so a state contributes to each regime in the years it was in it. Title V is the state formula allotment for abstinence education (FY1998–2017 in this panel): each state is allotted a share in proportion to its share of the nation's low-income children, a state that accepts it must put up three non-federal dollars, in cash or in kind, for every four federal dollars, and a governor can decline the allotment outright, so it is the one stream whose amount reflects a choice by the state rather than by grantees. Abstinence funds are Title V plus the competitive abstinence grants plus the adolescent family life program, so the first column is contained in the second; comprehensive funds are PREP and TPP pooled. Appendix A.4.2 describes the streams. These are descriptive associations, not causal estimates: only six comprehensive states and two traditional states switch inside the funding window. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Model: Derivations, Proofs and Extensions

Section B.1 solves the model of Section 3: three lemmas derive the period-2 choice, contraceptive effort and the return to college, the objects every prediction uses. Section B.2 proves the propositions from them, together with Remark 1 on earnings among the employed, Section B.3 gives a sufficient condition for that remark, and Section B.4 shows that the signs of Propositions 1–4 do not depend on the functional forms. The remaining sections treat men, the traditional arm, the model's limits and an equivalent formulation in which the curriculum acts on beliefs.

B.1 Solving the model

I solve the model of Section 3 by backward induction. The curriculum enters only through $\kappa(D)$ and $\tau(D)$: every prediction is the sign of the derivative of one of the objects below with respect to D, and each derivative splits into a knowledge term, in $\kappa'(D)$, and a norm term, in $\tau'(D)$. Two restrictions on parameters keep both choices interior.

Assumption 1 (Interior choices)**.** *(a) $\underline{\rho} < a_0(D) < a_1(D) < \bar{\rho}$ for all D, so that at either schooling level some women choose each path; (b) $\kappa(D)\,\bar{p}\,\Lambda(1;D) < \gamma$, so that contraceptive effort is interior.*

Notation. In this appendix the argument D is dropped from objects evaluated at the current curriculum: I write $V_i(e,s) \equiv V_i(e,s;D)$, $\Lambda_s \equiv \Lambda(s;D)$, $e^*_s \equiv e^*(s;D)$ and $v_s \equiv v(s;D)$, and a_s, P_s, f^*_s, ω_s, P, Δ, $\Delta\Lambda$, $\bar{f}$, κ, κ' and τ' for the values of $a_s(D)$, $P_s(D)$, and so on. The argument is kept only where D takes a particular value, as in $\tau(0)$, or where a function is first defined. Derivatives d/dD hold s fixed.

Lemma 1 (Period 2)**.** *A woman who was not derailed chooses the career if and only if it is worth at least the traditional path,*

$$\textit{career} \iff a_s \geq \rho_i. \tag{13}$$

Before ρ_i is realized, the probability of a career and the expected surplus are

$$P_s(D) \;\equiv\; \Pr(\textit{career} \mid s) \;=\; \frac{a_s(D) - \underline{\rho}}{\sigma}, \tag{14}$$

$$\Lambda(s;D) \;=\; \frac{\big(a_s(D) - \underline{\rho}\big)^2}{2\sigma} \;=\; \frac{\sigma}{2}\,P_s(D)^2. \tag{15}$$

The surplus rises in D, $\partial\Lambda_s/\partial D = (-\tau')P_s > 0$, and the college premium in expected surplus is positive and rises in D, as in (5)*:*

$$\Delta\Lambda \;=\; \big[w(1) - w(0)\big]\,\frac{P_1 + P_0}{2} \;>\; 0, \qquad \frac{d\,\Delta\Lambda}{dD} \;=\; \big(-\tau'\big)\big(P_1 - P_0\big) \;>\; 0.$$

Proof. A woman who was not derailed compares $a_s - k_i s$ with $\rho_i - k_i s$. The college cost enters both and cancels, so the rule is $a_s \geq \rho_i$. Under $\rho_i \sim U[\underline{\rho}, \bar{\rho}]$, $\Pr(\rho_i \leq a_s) = (a_s - \underline{\rho})/\sigma$, which lies in $(0,1)$ by Assumption 1(a). The surplus (3) equals $a_s - \rho_i$ for $\rho_i \leq a_s$ and zero above, so

$$\Lambda_s \;=\; \frac{1}{\sigma}\int_{\underline{\rho}}^{a_s} \big(a_s - \rho\big)\,d\rho \;=\; \frac{\big(a_s - \underline{\rho}\big)^2}{2\sigma} \;=\; \frac{\sigma}{2}\,P_s^2,$$

using $a_s - \underline{\rho} = \sigma P_s$. Differentiating, $\partial\Lambda_s/\partial a_s = P_s$, and $da_s/dD = -\tau'$. Finally, since $a_1 - a_0 =$

$w(1) - w(0)$,

$$\Delta\Lambda \;=\; \frac{\big(a_1 - \underline{\rho}\big)^2 - \big(a_0 - \underline{\rho}\big)^2}{2\sigma} \;=\; \big(a_1 - a_0\big)\,\frac{\big(a_1 - \underline{\rho}\big) + \big(a_0 - \underline{\rho}\big)}{2\sigma} \;=\; \big[w(1) - w(0)\big]\,\frac{P_1 + P_0}{2},$$

and $d\,\Delta\Lambda/dD = \partial\Lambda_1/\partial D - \partial\Lambda_0/\partial D = (-\tau')(P_1 - P_0)$. □

Figure B.1 draws the lemma. In panel (a), a woman with schooling s takes the career when $\rho_i \le a_s$, with probability P_s; education moves the career payoff from a_0 to a_1, so $P_1 > P_0$. The wage and the norm cost enter the choice only through a_s, so a unit fall in τ raises the probability of a career exactly as much as a unit rise in the wage, by $1/\sigma$, and by the same amount at both schooling levels (the two strips); unlike the wage, τ does not enter what a woman earns once she works. In panel (b), the surplus falls one-for-one in ρ_i until it reaches zero at a_s; realizations above a_s are traditional-bound and lose nothing from a derailment. Its expectation is the area of the triangle, scaled by the density $1/\sigma$, and the hatched band is $\Delta\Lambda$. A weaker norm slides both triangles to the right, and because a triangle grows faster the larger it is, the graduate's grows more: the marginal value of a higher career payoff is the probability of taking it, $\partial\Lambda_s/\partial a_s = P_s$, and $P_1 > P_0$, which is $d\,\Delta\Lambda/dD > 0$. The two panels therefore rest on the same gap, $P_1 > P_0$. A woman may invest and then choose the traditional path: she invests before ρ_i is known, so this is ex-post regret that was priced ex ante, not an inconsistency.

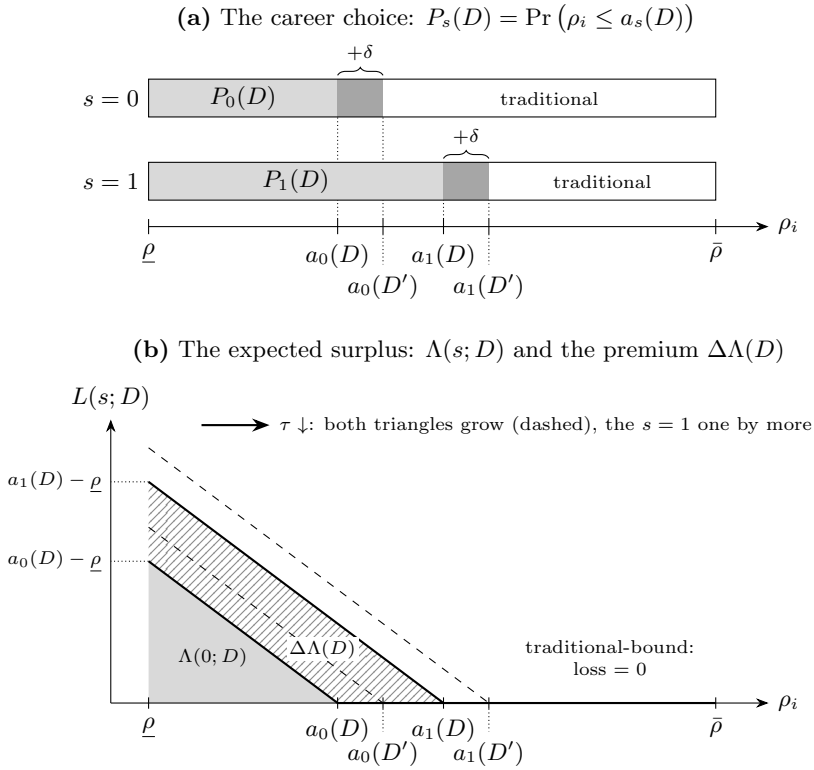


Figure B.1: The career choice and the expected surplus from the career.

Notes: A schematic of the model of Section 3 in the notation of the text, at a curriculum D and, marked by primes, at a more comprehensive curriculum D', with τ falling by 0.6 on a support of length $\sigma = 7.5$; no data or estimates enter it. Panel (a): each bar is the uniform distribution of the value of the traditional path, ρ_i, for a woman with schooling s, who takes the career if $\rho_i \le a_s$; the light gray share of the bar is the career probability $P_s = (a_s - \underline{\rho})/\sigma$, larger at $s = 1$, and the dark strips of mass δ are its common response to a weaker norm, $dP_s/dD = -\tau'/\sigma$ at both schooling levels—a smaller proportional rise for graduates, who already work more often. Panel (b): the surplus of the career over the traditional path falls one-for-one in ρ_i and is zero above a_s; its expectation, $\Lambda_s = (a_s - \underline{\rho})^2/(2\sigma)$, is the area of the triangle scaled by the density $1/\sigma$, and the hatched band is the premium $\Delta\Lambda$. When τ falls both triangles grow (dashed), and the graduate's grows more: $\partial\Lambda_s/\partial D = (-\tau')P_s$ and $d\,\Delta\Lambda/dD = (-\tau')(P_1 - P_0) > 0$.

At each schooling level she chooses effort optimally. The college cost and $\mathbb{E}[\rho_i]$ in (4) do not depend on effort and pass through the maximum:

$$\begin{aligned} \max_e V_i(e,s;D) \;&=\; -k_i s \;+\; \mathbb{E}[\rho_i] \;+\; v(s;D), \\ v(s;D) \;&\equiv\; \max_e \Big\{ -\frac{\gamma e^2}{2\kappa(D)} + \big[1-f(e)\big]\Lambda(s;D)\Big\}, \end{aligned} \tag{16}$$

where $v(s;D)$ is the part common to all women.

Lemma 2 (Contraceptive effort)**.** *For each s the effort problem in v_s is strictly concave, with the unique interior solution*

$$e^*(s;D) \;=\; \frac{\kappa(D)\,\bar{p}\,\Lambda(s;D)}{\gamma}, \qquad f_s^*(D) \;\equiv\; f\big(e^*(s;D)\big) \;=\; \bar{p} - \frac{\bar{p}^2}{\gamma}\,\kappa(D)\,\Lambda(s;D), \tag{17}$$

where $f_s^*(D)$ is the equilibrium probability of an early birth. A woman who invests in college has more to protect and uses more contraception, $e_1^* > e_0^*$ and $f_1^* < f_0^*$, and effort rises in D:

$$\frac{\partial e_s^*}{\partial D} \;=\; \frac{\bar{p}}{\gamma}\Big[\kappa'\Lambda_s + \kappa\big(-\tau'\big)P_s\Big] \;>\; 0.$$

At the optimum, $v_s = (1-\bar{p})\Lambda_s + \kappa\bar{p}^2\Lambda_s^2/(2\gamma)$.

Proof. With $f(e) = \bar{p}(1-e)$, $[1-f(e)]\Lambda_s = (1-\bar{p})\Lambda_s + \bar{p}\,e\,\Lambda_s$, so the objective $-\gamma e^2/(2\kappa) + [1-f(e)]\Lambda_s$ has derivative $-\gamma e/\kappa + \bar{p}\Lambda_s$ and second derivative $-\gamma/\kappa < 0$. The first-order condition gives $e_s^* = \kappa\bar{p}\Lambda_s/\gamma$, which lies in $(0,1)$ because $\Lambda_s > 0$ and, by Assumption 1(b), $\kappa\bar{p}\Lambda_s \le \kappa\bar{p}\Lambda_1 < \gamma$. Substituting into $f_s^* = \bar{p}(1-e_s^*)$ gives early births. The derivative of e_s^* follows from Lemma 1, and $e_1^* > e_0^*$ because $\Lambda_1 > \Lambda_0$. At the optimum the effort cost is $\kappa\bar{p}^2\Lambda_s^2/(2\gamma)$ and the gain from effort, $\bar{p}\,e_s^*\Lambda_s$, is $\kappa\bar{p}^2\Lambda_s^2/\gamma$; adding $(1-\bar{p})\Lambda_s$ gives v_s. □

Comparing the two schooling levels, $\mathbb{E}[\rho_i]$ cancels, since it is collected whatever she chooses, and the college cost remains:

$$\begin{aligned} \max_e V_i(e,1;D) - \max_e V_i(e,0;D) \;&=\; \big[-k_i + \mathbb{E}[\rho_i] + v(1;D)\big] - \big[\mathbb{E}[\rho_i] + v(0;D)\big] \\ &=\; \Delta(D) - k_i, \end{aligned} \tag{18}$$

where $\Delta(D) \equiv v(1;D) - v(0;D)$ is the *return to college*. Because k_i is independent of ρ_i, the return is common to all women, who differ only in the cost they weigh against it, so the probability that a woman goes to college is $\omega_1(D) \equiv G\big(\Delta(D)\big)$, with $\omega_0(D) \equiv 1 - \omega_1(D)$.

Lemma 3 (The return to college)**.** *A woman invests in college if and only if $\Delta \ge k_i$, where the return to college is $\Delta = \big(1-\bar{f}\big)\Delta\Lambda > 0$, as in* (6)*. Its derivative is* (7)*, with both terms positive. When only the knowledge channel operates ($\tau' = 0$), the first term vanishes and $\big(d\Delta/dD\big)/\Delta = \big(-d\bar{f}/dD\big)/\big(1-\bar{f}\big)$.*

Proof. By (16) and Lemma 2,

$$\max_e V_i(e,s) \;=\; -k_i s \;+\; \mathbb{E}[\rho_i] \;+\; (1-\bar{p})\Lambda_s \;+\; \frac{\kappa\bar{p}^2}{2\gamma}\Lambda_s^2,$$

so that, by (18),

$$\max_e V_i(e,1) - \max_e V_i(e,0) \;=\; (1-\bar{p})\,\Delta\Lambda + \frac{\kappa\bar{p}^2}{2\gamma}\big(\Lambda_1^2 - \Lambda_0^2\big) - k_i \;=\; \Delta\Lambda\left[(1-\bar{p}) + \frac{\kappa\bar{p}^2}{2\gamma}\big(\Lambda_1 + \Lambda_0\big)\right] - k_i,$$

factoring $\Lambda_1^2 - \Lambda_0^2 = \Delta\Lambda(\Lambda_1 + \Lambda_0)$. She therefore invests if and only if the first term, the return Δ, is at least k_i. Since $1 - f_s^* = 1 - \bar{p} + (\bar{p}^2/\gamma)\kappa\Lambda_s$, the bracket is the average of $1 - f_1^*$ and $1 - f_0^*$, that is, $1 - \bar{f}$. The average is exact, not an approximation: v_s is quadratic in Λ_s, so its difference across schooling levels factors through $\Lambda_1 + \Lambda_0$. $\Delta > 0$ because $\bar{f} < \bar{p} < 1$ and $\Delta\Lambda > 0$ (Lemma 1). Equation (7) is the product rule; its first term is positive by Lemma 1 and its second because f_s^* falls in D at each s (Lemma 2). With $\tau' = 0$, $\Delta\Lambda$ does not depend on D, only the second term remains, and dividing by Δ gives the last statement. □

A degree is worth its full premium only to a woman certain of avoiding an early birth, so fertility and schooling are not separate outcomes: the same $\bar{f}$ prices the investment and is itself chosen. The two terms of (7) are the two ways in which the curriculum raises the return to college. The second holds the premium fixed and asks how often it is collected: fewer of the women who invest are derailed. With the knowledge channel alone it is the only term, and by Lemma 3 the proportional rise in college going is then the elasticity of college going with respect to its return times the proportional rise in survival, $-d\bar{f}/(1-\bar{f})$; completion rises by that plus the mechanical gain from fewer derailments among those who invest. Both are tied to the fall in early births. The first term says the value of the degree itself rises: a weaker norm lowers the cost of the career, which is worth more to the woman whose career pays a_1 than to the one whose career pays a_0 (Figure B.1).

Only women who were not derailed and chose the career work, so the probability that a woman works is

$$P(D) \;=\; \sum_{s\in\{0,1\}} \omega_s(D)\left[1 - f_s^*(D)\right] P_s(D), \tag{19}$$

and a working woman with schooling s earns $w(s)$ whatever the curriculum.

B.2 Proofs of the propositions

Every derivative below is signed for all D on which Assumption 1 holds, not only for $D \in [0,1]$; the results therefore apply unchanged, with the sign of dD reversed, to the traditional curriculum modeled as $D < 0$ in Section B.6.

Proof of Proposition 1. The derivative of e_s^* is Lemma 2. Since $f_s^* = f(e_s^*)$ with $f' = -\bar{p} < 0$, early births fall in D at each schooling level. □

Proof of Proposition 2. $d\Delta/dD > 0$ is Lemma 3. Completion, $\omega_1\big(1 - f_1^*\big)$, is the product of two positive, strictly increasing functions of D: the first because $\omega_1 = G(\Delta)$ with G strictly increasing, the second by Lemma 2. □

Proof of Proposition 3. Differentiating (19), with $d\omega_1/dD = -d\omega_0/dD = G'(\Delta)\, d\Delta/dD$ and $dP_s/dD = -\tau'/\sigma$ from (14),

$$\frac{dP}{dD} \;=\; \underbrace{G'(\Delta)\,\frac{d\Delta}{dD}\Big[\big(1 - f_1^*\big)P_1 - \big(1 - f_0^*\big)P_0\Big]}_{\text{composition}} \;+\; \underbrace{\sum_s \omega_s\Big(-\frac{df_s^*}{dD}\Big)P_s}_{\text{fewer exits}} \;+\; \underbrace{\sum_s \omega_s\big(1 - f_s^*\big)\frac{-\tau'}{\sigma}}_{\text{direct norm effect}}.$$

The composition term is positive because $d\Delta/dD > 0$ (Lemma 3) and both factors in the bracket are larger at $s = 1$: $f_1^* < f_0^*$ (Lemma 2) and $P_1 > P_0$. The second term is positive by Lemma 2 and the third because $\tau' < 0$. □

Proof of Proposition 4. $\frac{d}{dD}\big[(1-f_s^*)P_s\big] = \big(-df_s^*/dD\big)P_s + (1-f_s^*)(-\tau')/\sigma > 0$ and $df_s^*/dD < 0$, both by Lemma 2 and $\tau' < 0$. The share of women on the traditional path is $1-P$, which falls by Proposition 3. □

Remark 1 (The intensive margin)**.** *A working woman with schooling s earns $w(s)$ whatever D is, so mean earnings among the employed move with D only through the share of graduates among them. They rise with comprehensive sex education if college going responds strongly enough to its return, and always when only the knowledge channel operates.*

Proof of Remark 1. Conditional on schooling s and on working, a woman earns $w(s)$, since τ enters the choice (13) but not the wage. Let $A_s \equiv \big(1-f_s^*\big)P_s$ be the probability that a woman with schooling s works, and $\theta \equiv \omega_1 A_1/\big(\omega_1 A_1 + \omega_0 A_0\big)$ the share of graduates among working women. Mean earnings among the employed are $w(0) + \theta\,\big[w(1)-w(0)\big]$, which depend on D only through θ and rise with it. In log-odds,

$$\frac{d}{dD}\ln\frac{\theta}{1-\theta} \;=\; \underbrace{\frac{d}{dD}\ln\frac{\omega_1}{\omega_0}}_{\text{schooling}} \;+\; \underbrace{\frac{d}{dD}\ln\frac{1-f_1^*}{1-f_0^*}}_{\text{survival}} \;+\; \underbrace{\frac{-\tau'}{\sigma}\Big(\frac{1}{P_1}-\frac{1}{P_0}\Big)}_{\text{direct norm effect}}.$$

The schooling term equals $G'(\Delta)\,(d\Delta/dD)/(\omega_1\omega_0) > 0$ by Lemma 3. The direct norm term is negative: a weaker norm raises P_1 and P_0 by the same amount, $-\tau'/\sigma$, and so raises the probability of working proportionally more among non-graduates. For the survival term, $1-f_s^* = 1-\bar p + (\bar p^2/\gamma)\kappa\Lambda_s$, so that

$$\frac{d}{dD}\ln\big(1-f_s^*\big) \;=\; \frac{\bar p^2}{\gamma}\,\frac{\kappa'\Lambda_s + \kappa(-\tau')P_s}{1-\bar p + (\bar p^2/\gamma)\kappa\Lambda_s}.$$

The knowledge part, $\kappa'\Lambda_s/\big(1-\bar p + (\bar p^2/\gamma)\kappa\Lambda_s\big)$, is increasing in Λ_s and therefore larger at $s=1$. The norm part, $\kappa(-\tau')P_s/\big(1-\bar p + (\bar p^2/\gamma)\kappa\sigma P_s^2/2\big)$, is increasing in P_s whenever $1-\bar p > \bar p\, e_s^*$, which holds for $\bar p \le 1/2$ since $e_s^* < 1$. The survival term is therefore positive through the knowledge channel, and through the norm channel as well when $\bar p \le 1/2$; the direct norm effect is then the only force that lowers θ. The schooling term is proportional to $G'(\Delta)$, while the other two terms do not depend on G, so for a strong enough response of college going to its return the schooling term outweighs them and θ rises; Section B.3 gives a sufficient condition. With $\tau' = 0$ the direct norm term and the norm part of the survival term vanish, and θ rises. □

B.3 The intensive margin: a sufficient condition

The condition in Remark 1 can be stated in terms of the elasticity of college going with respect to its return, $\varepsilon \equiv \Delta\, G'(\Delta)/\omega_1$.

Lemma 4 (A bound for the intensive margin)**.** *If $\bar p \le 1/2$, mean earnings among the employed rise with D whenever*

$$\varepsilon \;\ge\; \frac{\omega_0}{2}\left(\frac{P_1}{P_0}-\frac{P_0}{P_1}\right). \tag{20}$$

Proof. With $\bar p \le 1/2$ the survival term in the proof of Remark 1 is positive, so it suffices that the schooling term outweigh the direct norm term. Both are proportional to $-\tau'$. By Lemma 3, $d\Delta/dD \ge \big(1-\bar f\big)\,d\,\Delta\Lambda/dD = \big(1-\bar f\big)(-\tau')(P_1-P_0)$, the survival part of (7) being positive, so that

$$\underbrace{\frac{G'(\Delta)}{\omega_1\omega_0}\,\frac{d\Delta}{dD}}_{\text{schooling}} \;+\; \underbrace{\frac{-\tau'}{\sigma}\Big(\frac{1}{P_1}-\frac{1}{P_0}\Big)}_{\text{direct norm effect}} \;\ge\; (-\tau')\big(P_1-P_0\big)\left[\frac{G'(\Delta)\big(1-\bar f\big)}{\omega_1\omega_0} - \frac{1}{\sigma P_0 P_1}\right].$$

Write $G'(\Delta) = \varepsilon\,\omega_1/\Delta$ and use $\Delta/(1-\bar{f}) = \Delta\Lambda = [w(1)-w(0)](P_1+P_0)/2$ (Lemma 1) together with $w(1)-w(0) = a_1 - a_0 = \sigma(P_1 - P_0)$ from (14). The first fraction in the bracket is then $\varepsilon/(\omega_0\,\Delta\Lambda) = 2\varepsilon/\big[\omega_0\,\sigma(P_1-P_0)(P_1+P_0)\big]$, and the bracket is nonnegative if and only if

$$\varepsilon \;\geq\; \omega_0\,\frac{P_1^2 - P_0^2}{2P_0P_1} \;=\; \frac{\omega_0}{2}\left(\frac{P_1}{P_0} - \frac{P_0}{P_1}\right).$$

Under this bound every term of the log-odds derivative that is not positive is outweighed, so θ and mean earnings among the employed rise with D. □

The bound is sufficient rather than necessary, since it discards the survival terms. It depends only on the share of women without a degree and on how much more often graduates work, and it is small when the two groups work at similar rates, because a weaker norm then draws in workers whose schooling mix is close to that of women already working, and there is little to dilute. Two qualifications apply. The restriction $\bar{p} \leq 1/2$—an early birth is less likely than not without any contraception—is used only here: it ensures that fewer early births raise graduates' probability of working proportionally more than non-graduates', so that the fall in early births does not itself dilute the graduate share. And because Δ is a return measured in utility, ε is an elasticity with respect to a utility return; it coincides with the elasticity with respect to a monetary return only if utility is proportional to money. Given functional forms that are only indicative, (20) is a qualitative condition—college going must respond strongly enough to its return—rather than a number to take to the data.

B.4 General effort costs and birth risks

The closed forms (17) and (6) use the quadratic cost and the linear birth risk; the signs of Propositions 1–4 do not. Let the effort cost be $C(e)/\kappa(D)$, with C increasing and strictly convex, and the probability of an early birth $f(e)$, decreasing and convex, and suppose effort is interior. The objective $-C(e)/\kappa + [1-f(e)]\Lambda_s$ is then strictly concave in e; its cross-partial in (e, D), $C'(e)\kappa'/\kappa^2 + \big(-f'(e)\big)(-\tau')P_s$, is positive, and so is its cross-partial in (e, Λ_s), $-f'(e)$. Effort therefore rises in D and is higher at $s = 1$, and early births fall in D and are lower at $s = 1$, which is Proposition 1. By the envelope theorem,

$$\frac{d\Delta}{dD} \;=\; \frac{\kappa'}{\kappa^2}\Big[C\big(e_1^*\big) - C\big(e_0^*\big)\Big] \;+\; \big(-\tau'\big)\Big[\big(1-f_1^*\big)P_1 - \big(1-f_0^*\big)P_0\Big] \;>\; 0,$$

both brackets being positive because $e_1^* > e_0^*$, $f_1^* < f_0^*$ and $P_1 > P_0$; and $\Delta > 0$, because a woman who invests could choose e_0^* and would then gain $\big(1-f_0^*\big)\Delta\Lambda > 0$ over not investing. This is Proposition 2. The proofs of Propositions 3 and 4 use only these signs and (14), and carry over unchanged. Only the exact factorization (6), and with it the decomposition (7) into a premium term and a survival term, is specific to the quadratic cost; the envelope expression above splits instead into a knowledge term, in κ', and a norm term, in τ'.

B.5 Men

For men the knowledge channel is absent—they bear no biological fertility cost and their education is rarely interrupted by parenthood—and the norm channel is ambiguous in sign, because the male prescription is *breadwinner*, so weaker prescriptions lower the cost of *not* working. Formally, if the male deviation cost attaches to non-participation, the male career threshold becomes $w^m(s) + \tau^m(D) \geq \rho$ with $\tau^{m\prime}(D) < 0$, so male participation and the male analogue of $\Delta\Lambda$ *weakly fall* in D. The model therefore

predicts strong effects for women and weak or ambiguous effects for men's labor supply, and does not claim to explain male outcomes; the estimated male college response, positive but fragile to state trends, is noted rather than targeted.

B.6 The traditional arm and the prevailing norm

The normalization $D = 0$ in Section 3.1 is a statement about the norm that prevails in the absence of any mandate. In the baseline reading that norm is traditional: $\tau(0) = \bar{\tau}$ is the cost of departing from a prescription the environment already imposes, and a traditional curriculum, which restates it and adds no contraceptive knowledge, is the untreated state itself. A traditional mandate then moves nothing—$dD = 0$—and the model predicts a null on every outcome for that arm relative to the never-mandate states, which is the pattern Sections 6 and 7 report.

Suppose instead that the prevailing norm is neutral, $\tau(0) = \tau_N < \bar{\tau}$, and extend the index to $D \in [-1, 1]$ with the same primitives, $\kappa' > 0$ and $\tau' < 0$ throughout, so that $D < 0$ carries $\kappa(D) < 1$ and $\tau(D) > \tau_N$: a traditional curriculum now withholds contraceptive knowledge and reinforces the traditional prescription. Because every comparative static in Sections B.1 and B.2 is signed on the whole domain, a fall in D reverses each of Propositions 1–4: effort falls and early births rise, the return to education and college completion fall, participation falls, and more women end on a traditional trajectory. The comprehensive curriculum keeps its signs but, on the norm channel, a smaller magnitude: with $\tau \geq 0$ the reduction available to it is bounded by $\tau(0) - \tau(1) \leq \tau_N < \bar{\tau}$, so there is less prescription to weaken. The two readings differ only in what they predict for the traditional arm—a null, or the mirror image of the comprehensive effects—and the empirical results for that arm are read against both.

B.7 Limitations and an extension

The norm cost $\tau(D)$ is not separately identified from beliefs (Section B.8), and the two readings differ in welfare: a utility cost lowered, or an investment better informed. The intensive-margin result of Remark 1 follows from the two-period structure: with one working period, the timing of motherhood cannot reach the wage of a woman who works. The strongest assumption is that derailment is absorbing. Allowing re-entry at a cost would preserve the sign of every result while compressing the magnitudes, predict the largest effects where re-entry is hardest—at the bottom of the income distribution—and let a later first birth raise the earnings of the employed through experience accumulated before the interruption, the half of the earnings effect that Section 6.1 finds and that the present version cannot account for.

B.8 An equivalent beliefs formulation

The norm channel can equally be placed in beliefs. Let the second branch point be a chance node at which a traditional trajectory realizes with believed probability $\tilde{q}(D)$, $\tilde{q}'(D) < 0$, and true probability q: the adolescent under-predicts her own future attachment—Goldin's "horizon"—and the curriculum moves $\tilde{q}$ toward the truth, the learning mechanism of Fernández (2013) and Fogli and Veldkamp (2011) driven here by a curriculum rather than by intergenerational observation. With W_0 the value of the traditional path after a derailment and $W_T(s) < W_C(s)$ the values of the traditional and career paths with schooling s, the derailment loss becomes $\widetilde{\Lambda}(s; D) = \tilde{q}(D)\big[W_T(s) - W_0\big] + \big(1 - \tilde{q}(D)\big)\big[W_C(s) - W_0\big]$, and the effort problem is that of Lemma 2 with $\widetilde{\Lambda}$ in place of Λ, so $e_s^* = \kappa\bar{p}\widetilde{\Lambda}_s/\gamma$. Because $\partial\widetilde{\Lambda}_s/\partial D = -\tilde{q}'\big[W_C(s) - W_T(s)\big] > 0$ and, when schooling raises career values more than traditional ones, $d\,\Delta\widetilde{\Lambda}/dD = -\tilde{q}'\big[\Delta W_C - \Delta W_T\big] > 0$,

effort, completion and participation rise in D exactly as in Propositions 1–4. The formulations differ only in welfare: this one requires that priors were mistaken, the historical claim of Goldin (2006), and its gain is better-informed investment.

C Appendix: Main Results

C.1 Sample construction

	Women			Men		
	Panel (age $\geq$ 18)	Age $\geq$ 25	Estimation sample	Panel (age $\geq$ 18)	Age $\geq$ 25	Estimation sample
Cells	48,167	39,764	28,817	48,335	39,930	28,873
States	49	49	41	49	49	41
Birth cohorts	1950–2004	1950–1999	1950–1999	1950–2004	1950–1999	1950–1999
Respondents	13,007,301	10,505,753	8,546,940	13,041,592	10,443,721	8,483,204
Survey years	2000–2024					
Minimum cell size	20 respondents					

Table C.1: ACS sample, before and after the age restriction.

Notes: Counts of the American Community Survey 2000–2024 extract that every labor market result is estimated on; no estimation is involved. A cell is a birth state $\times$ birth cohort $\times$ survey year, and both sexes are counted from age 18. "Age $\geq$ 25" applies the age restriction used for every labor market outcome; "estimation sample" additionally drops the eight states whose curriculum cannot be classified into one of the two arms and trims the never-mandate control pool to the event-time window spanned by the observed adoption dates. The marriage-timing outcomes are observed for fewer cohorts and rest on a smaller sample still, reported at the foot of each table that uses them.

	Comprehensive			Traditional		
	Arm	Never-mandate	Difference	Arm	Never-mandate	Difference
ACS, women aged 25 and over						
Log annual earnings (all)	10.373	10.364	0.009	10.204	10.347	−0.143
			(0.038)			(0.039)
Full-time (all)	0.598	0.627	−0.028	0.630	0.618	0.012
			(0.006)			(0.005)
Employed	0.728	0.749	−0.021	0.716	0.740	−0.024
			(0.007)			(0.006)
College degree by 25	0.353	0.393	−0.040	0.299	0.374	−0.076
			(0.017)			(0.014)
Years of education	13.88	14.08	−0.20	13.55	14.00	−0.46
			(0.09)			(0.08)
Married by 23	0.256	0.264	−0.008	0.363	0.304	0.060
			(0.019)			(0.020)
Married by 30	0.603	0.629	−0.027	0.666	0.653	0.014
			(0.018)			(0.017)
Age at first marriage (restricted to 30)	26.49	26.34	0.15	25.64	26.10	−0.46
			(0.18)			(0.19)
States	9	23		9	23	
NVSS teen births, 1980–2024						
Births per 1,000 girls aged 15–19	28.4	30.8	−2.5	65.0	47.1	17.9
			(2.4)			(3.1)
Log share of girls giving birth	−3.644	−3.593	−0.052	−2.741	−3.098	0.356
			(0.078)			(0.067)
States	9	23		9	23	

Table C.2: Pre-adoption levels by sex education type.

Notes: Each arm's pre-adoption cells against never-mandate cells of the same birth cohort and survey year (ACS) or of the same year (teen births). ACS: women aged 25 and over in the estimation sample, every pre-adoption cohort, cells weighted by ACS person weights. Teen births: NVSS natality files for 1980–2024 by the mother's state of residence, births dated by conception, the seven years before adoption, weighted by the female population aged 15–19. Never-mandate cells are reweighted to the arm's weight in each cohort × survey-year (or year) cell and only cells observed in both groups are kept, so both means share the arm's composition; the arm means therefore differ slightly from the pre-adoption means of the main tables, which average every pre-adoption cell. The difference is the coefficient on an arm indicator in a regression with cohort × survey-year (or year) fixed effects and these weights, which equals the difference between the two means. Standard errors, clustered on state, in parentheses; with few treated states they are indicative only. The state fixed effects of every specification absorb these levels.

Weights. Estimation is weighted by $\sum_{i\in(c,s,y)} \texttt{perwt}_i$, which is proportional to the ACS-estimated population of the cell. ACS weights are calibrated to state-of-*residence* totals, whereas treatment is assigned by state of *birth*, so weight and treatment-relevant population do not correspond exactly for women who have moved. This is standard in birth-state designs (e.g., Goldin and Katz, 2002), and for share outcomes the resulting bias is mechanically bounded and small.

C.2 Inference and sensitivity: construction

This appendix gives the construction of the checks summarized in Section 4.5.

Randomization inference. The permutation reassigns each state's entire policy bundle to a randomly drawn state; permuting the comprehensive and traditional onsets independently would test a null nobody holds. The reported p-value is two-sided, computed on the ATT of (12), with the usual finite-sample correction,

$$p \;=\; \frac{1 + \#\{\,|\hat{b}_{\text{perm}}| \geq |\hat{b}_{\text{obs}}|\,\}}{1 + n_{\text{valid}}},$$

so with 500 draws the smallest attainable value is 1/501.

Pre-trends. The average pre-adoption coefficient over $[-6, -2]$ is a one-degree-of-freedom test. The joint Wald statistic over the same coefficients rejects more often: on the women's outcomes it rejects in eight of thirteen cases where the average test rejects in one; with few treated states the joint test relies on clustered standard errors that over-reject in this regime. Neither is used as a gate—pre-tests are underpowered against exactly the violations that matter, and conditioning on having passed one distorts the estimate that follows (Roth, 2022).

Birth-state linear trends. Each birth state receives its own linear slope in the birth cohort, estimated on the clean sample alongside the other first-stage parameters of (11):

$$\hat{\gamma}_s,\ \hat{\delta}_c,\ \hat{\eta}_y,\ \hat{\psi}_s \ = \ \arg\min_{\gamma,\delta,\eta,\psi} \sum_{\{c,s,y:\,\mathrm{SexEd}_{cs}=0\}} \big(Y_{csy} - \gamma_s - \delta_c - \eta_y - \psi_s\, c\big)^2.$$

This adds one parameter per state, not the $|S| \times |Y|$ of a saturated state-by-year specification, which would leave no clean variation to identify β_ℓ.[27] Trends are double-edged: with short pre-periods they can extrapolate into the post-period and manufacture or destroy an effect. The marriage-timing outcomes sit on a thin post-adoption panel and their traditional arm sign-flips with stars on both sides once trends are added, and are for that reason reported in the baseline specification only.

Relative-magnitudes sensitivity. $\bar{M}$ is the largest value at which the robust confidence set of Rambachan and Roth (2023) still excludes zero, where $\bar{M}$ bounds the post-treatment violation of parallel trends as a multiple of the largest violation observed before adoption. With around fifty state clusters, effects of a few percentage points and a nine-year post window, breakdown values in this design come out uniformly low, and they fall as the window lengthens because the bound permits a separate violation at each post-period. For reference, college completion breaks down at $\bar{M} = 0.25$ on a three-horizon window.

Bootstrapped variance. Each battery re-computes its variance by bootstrap with 999 replications clustered on state—a multiplier bootstrap where the Callaway–Sant'Anna machinery supports it, a cluster resample refitting both stages for the two-stage estimator. The coefficient is identical to the analytic fit by construction, so what the bootstrap adds is a standard error and an equal-tailed percentile p-value on the recentred draws,

$$p \ = \ \frac{1 + \#\{\,|\hat{b}^* - \hat{b}| \geq |\hat{b}|\,\}}{1 + B},$$

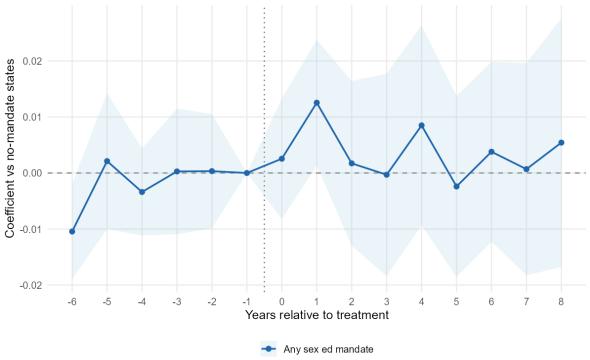
the recentring converting draws distributed around $\hat{b}$ into a null distribution. Both specifications are bootstrapped wherever the design has two: bootstrapping only the alternative-specification battery would leave the baseline column unpriced, which for employment and full-time work is precisely the column whose pre-period fails.

[27] In the estimation the slope is written in $c + 18$, the calendar year in which cohort c turned eighteen, which differs from c by a state-invariant constant and is therefore the same specification once γ_s is included. Age $= y - c$ is collinear with the cohort and survey-year effects only in its linear part, so the first stage absorbs neither a nonlinear age profile nor shocks specific to a birth state at the time of observation. Adding age dummies leaves every comprehensive estimate within a tenth of a point of the baseline; adding birth-state × survey-year effects as well gives, across all post-adoption cohorts, 3.3 (0.9) log points for earnings, 1.4 (0.7) points for full-time work, 0.3 (0.3) for employment, 2.3 (0.5) for college and −1.9 (0.4) for marriage by 23, and among the fully exposed 4.3 (0.9), 2.1 (0.7), 0.6 (0.3), 2.4 (0.5) and −1.9 (0.4). In the ANES, where cells are thin and some treated states have as few as two pre-treatment cohorts, fifty state slopes are overfit; that battery puts the trend on census division instead.

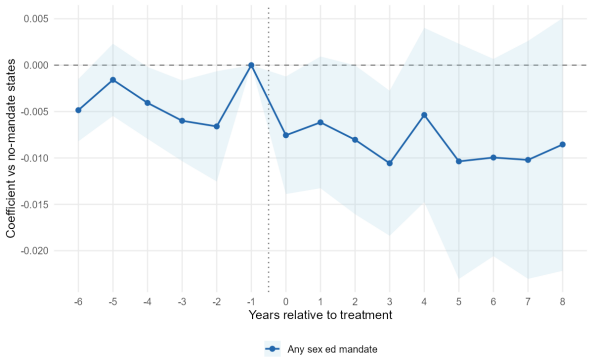
Why the bootstrap is a diagnostic. What governs cluster-robust inference is the number of *treated* clusters—eight or nine per arm in the ACS, three in the YRBS. MacKinnon and Webb (2018) show the cluster-robust t over-rejects in this regime and that the pairs cluster bootstrap does not repair it, since resampling clusters with replacement varies the number of treated clusters across draws. The failure is visible here: the bootstrap distribution of the average ATT has excess kurtosis near 3.7, about one draw in two hundred loses an event-time coefficient entirely, and in the YRBS more than half of all resamples fail that way, so no bootstrap is computed and the analytic standard error is reported in its place. A bootstrapped standard error much larger than the analytic one—the median ratio is about 1.2 across the ACS and ANES batteries, reaching 2.1—is informative in the direction of caution. But the bootstrapped p-value is not corrected inference: the pairs bootstrap does not impose the null, and a wild cluster bootstrap-t imposing H_0 (Cameron et al., 2008) is unavailable for the two-stage estimator as implemented. In practice the bootstrap and randomization inference agree, and the estimates the bootstrap declines to confirm are almost exactly those randomization inference had already declined to confirm.

C.3 Any mandate versus no mandate

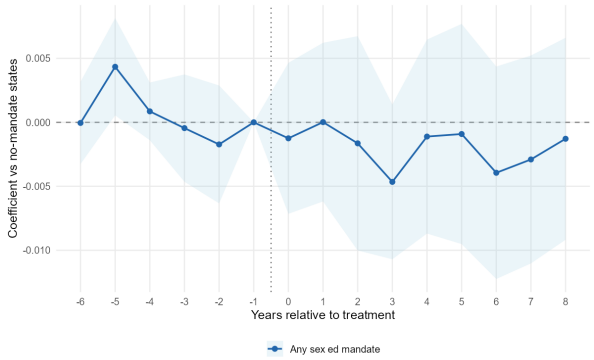
This is the comparison the literature usually runs: states with *any* sex education mandate against states with none, whatever the mandate requires, each dated from its first sex education mandate. It is reported first, so the presence margin can be judged on its own.



(a) Log annual earnings (all women)



(b) Working full-time (all women)

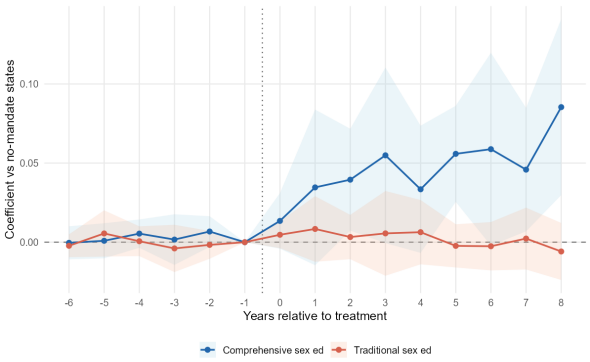


(c) Employed

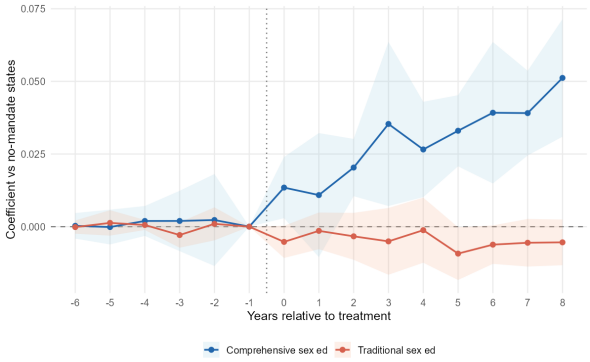
Figure C.1: Women's labor market outcomes: any sex education mandate versus no mandate.

Notes: Outcomes are the log of mean annual wage and salary income per woman with zeros for those without earnings in the previous year, the share working thirty-five hours or more, and the share of women employed. The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 from age 25, collapsed to birth-state × birth-cohort cells and weighted by ACS person weights. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects and no state trends, exactly as in Figure 2; the only change is the treatment definition, which here makes every state with a mandate of either kind one treated group against the states that never adopt one. The pooled group has 25 treated states, against the eight comprehensive and nine traditional adopters of Table 2: it adds states whose mandate requires neither content, and dates every state from its first sex education mandate rather than from the content it later required. Shaded bands are 95 percent confidence intervals from standard errors clustered on birth state.

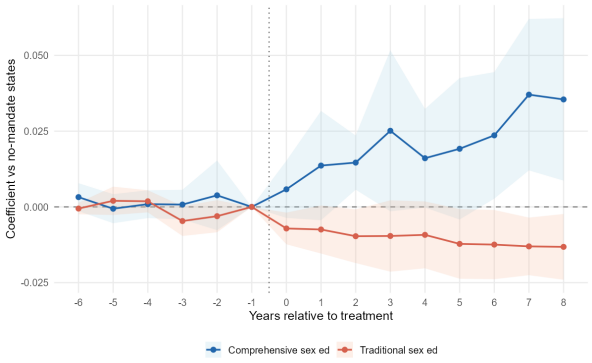
C.4 Women's labor market outcomes



(a) Log annual earnings (all)



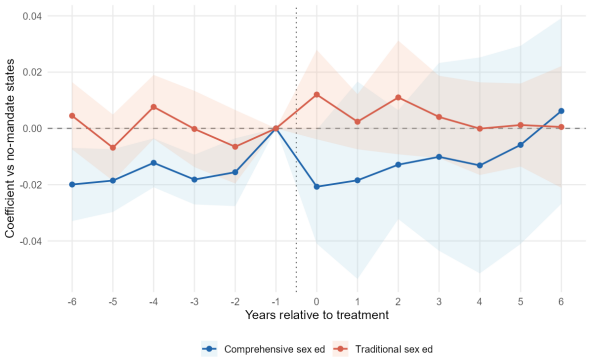
(b) Full-time (all)



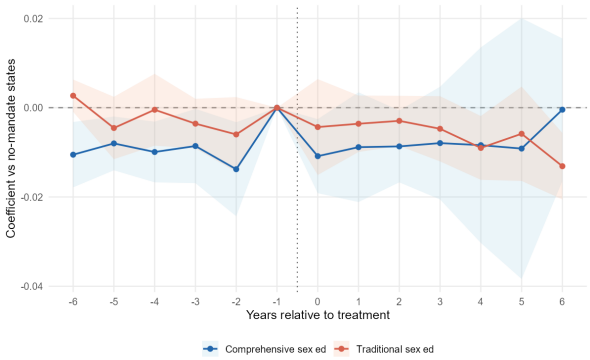
(c) Employed

Figure C.2: Women's labor market outcomes, with birth-state linear trends.

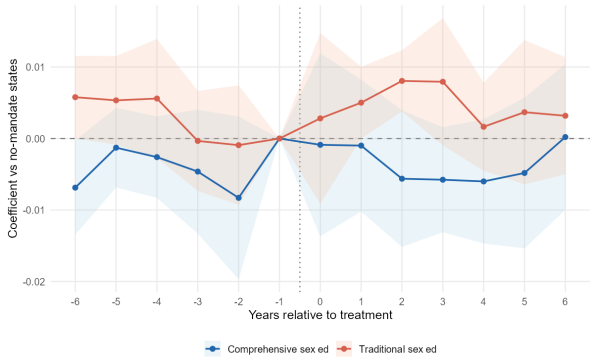
Notes: Figure 2 re-estimated with a linear trend in birth cohort for each birth state added to the fixed effects. Outcomes are the log of mean annual wage and salary income per woman with zeros for the non-employed, the share working thirty-five hours or more, and the share of women employed; the sample is women born 1950–1999 in the 41 classified states, observed in the American Community Survey 2000–2024 from age 25, collapsed to birth-state × birth-cohort cells and weighted by ACS person weights. Estimated by two-stage difference-in-differences, with each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states, and event time counted from the first cohort required to take the mandated curriculum. Shaded bands are 95 percent confidence intervals from standard errors clustered on birth state. Randomization-inference p-values for ATT $[0, 8]$ and $[4, 8]$ with trends, comprehensive then traditional: log earnings 0.006 and 0.006, 0.844 and 0.982; full-time 0.004 and 0.006, 0.591 and 0.605; employed 0.024 and 0.024, 0.198 and 0.224.



(a) Log annual earnings (all)



(b) Full-time (all)



(c) Employed

Figure C.3: Women's labor market outcomes, in-time placebo: every adoption date moved seven years earlier, no state trends.

Notes: Figure 2 re-estimated after moving every adoption date seven years earlier and dropping the genuinely treated cohorts, so the estimate is identified on pre-adoption data alone and should be zero. Outcomes, sample, estimator and treatment arms are those of Figure 2: women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 from age 25, two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects and no state trends, comprehensive and traditional sex education each against the never-mandate pool. Shaded bands are 95 percent confidence intervals from standard errors clustered on birth state.

	Log earnings (all)		Full-time (all)		Employed	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
Panel A: No state trends						
ATT $[0, 8]$	0.042^{***}	0.001	0.018^{**}	-0.016^{***}	0.005	-0.006
analytic SE	(0.008)	(0.010)	(0.009)	(0.005)	(0.004)	(0.005)
RI p-value	[0.014]	[0.952]	[0.072]	[0.038]	[0.309]	[0.176]
HonestDiD $\bar{M}$ (10%)	0.40	n/a	0.00	0.20	n/a	n/a
HonestDiD $\bar{M}$ (5%)	0.30	n/a	0.00	0.10	n/a	n/a
ATT $[4, 8]$	0.052^{***}	-0.004	0.025^{***}	-0.019^{***}	0.009^{**}	-0.007
analytic SE	(0.008)	(0.012)	(0.009)	(0.007)	(0.004)	(0.006)
RI p-value	[0.010]	[0.749]	[0.032]	[0.036]	[0.138]	[0.136]
HonestDiD $\bar{M}$ (10%)	0.30	n/a	0.10	0.20	0.00	n/a
HonestDiD $\bar{M}$ (5%)	0.30	n/a	0.00	0.10	0.00	n/a
Panel B: With birth-state linear trends						
ATT $[0, 8]$	0.047^{**}	0.002	0.030^{***}	-0.005	0.021^{**}	-0.010^{**}
analytic SE	(0.019)	(0.007)	(0.008)	(0.003)	(0.009)	(0.004)
RI p-value	[0.006]	[0.844]	[0.004]	[0.591]	[0.024]	[0.198]
HonestDiD $\bar{M}$ (10%)	0.20	n/a	0.40	n/a	0.20	0.10
HonestDiD $\bar{M}$ (5%)	0.10	n/a	0.30	n/a	0.10	0.10
ATT $[4, 8]$	0.056^{**}	-0.000	0.038^{***}	-0.006	0.026^{**}	-0.012^{**}
analytic SE	(0.022)	(0.007)	(0.009)	(0.003)	(0.011)	(0.005)
RI p-value	[0.006]	[0.982]	[0.006]	[0.605]	[0.024]	[0.224]
HonestDiD $\bar{M}$ (10%)	0.20	n/a	0.40	n/a	0.20	0.10
HonestDiD $\bar{M}$ (5%)	0.10	n/a	0.30	n/a	0.10	0.00
Placebo: adoption moved 7 years earlier; baseline specification						
ATT $[0, 6]$	-0.011	0.004	-0.008	-0.006^{*}	-0.003	0.005
	(0.015)	(0.005)	(0.007)	(0.003)	(0.004)	(0.003)
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states	8	9	8	9	8	9
Control states	23	23	23	23	23	23
Observations	28,817	28,817	28,817	28,817	28,817	28,817
Pre-adoption mean	10.349	10.159	0.595	0.623	0.725	0.706

Table C.3: Women's labor market outcomes: ATT, randomization-inference p-values, HonestDiD breakdown values and the in-time placebo.

Notes: Outcomes, sample, estimator and treatment arms are those of Table 2: log mean annual earnings, the share working thirty-five hours or more, and the share employed per woman, for women born 1950–1999 in the 41 states of the estimation sample and observed in the American Community Survey 2000–2024 from age 25, estimated by two-stage difference-in-differences with comprehensive and traditional sex education each measured against the never-mandate pool. Panel A carries no state trends and Panel B adds a linear trend in birth cohort for each birth state. Analytic standard errors, clustered on birth state, in parentheses; randomization-inference p-values, in brackets, from 500 draws that re-assign adoption dates across states holding the rollout structure fixed. The HonestDiD row is the breakdown $\bar{M}$: the largest post-treatment violation of parallel trends, as a multiple of the largest pre-treatment violation, at which the effect stays significant. "n/a" means the estimate is not significant even at $\bar{M} = 0$, and "—" that the leg was not run for that specification. $\bar{M}$ falls mechanically with the length of the post-treatment window, because the relative-magnitudes bound permits a separate violation at each post-period, so the $[0, 8]$ and $[4, 8]$ values rest on nine and five post-periods and are not comparable to each other. The placebo is estimated on the baseline specification only and is therefore reported once, at the foot of the table. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Comprehensive sex education				Traditional sex education			
	ATT $[0, 8]$		ATT $[4, 8]$		ATT $[0, 8]$		ATT $[4, 8]$	
	RI p	BH q	RI p	BH q	RI p	BH q	RI p	BH q
Log annual earnings, all	0.014	**0.037**	0.010	0.064	0.952	0.952	0.749	0.749
Share working full-time, all	0.072	0.082	0.032	0.080	0.038	0.303	0.036	0.287
Share employed	0.309	0.309	0.138	0.138	0.176	0.703	0.136	0.543
College degree	0.008	**0.037**	0.016	0.064	0.786	0.952	0.685	0.749
Married by 23	0.020	**0.040**	0.046	0.080	0.830	0.952	0.587	0.749
Married by 30	0.056	0.075	0.070	0.080	0.429	0.952	0.505	0.749
Age at first marriage, restricted mean to 30	0.032	0.051	0.070	0.080	0.876	0.952	0.731	0.749
Age at first marriage, conditional on marrying by 30	0.014	**0.037**	0.056	0.080	0.591	0.952	0.505	0.749
Outcomes significant at 5 percent	5	4	4	0	1	0	1	0
Outcomes tested	8	8	8	8	8	8	8	8

Table C.4: Women's outcomes: randomization-inference p-values adjusted for multiple testing.

Notes: Each cell pairs the randomization-inference p-value of the baseline estimate (RI p, from 500 draws that re-assign adoption dates across states) with its Benjamini–Hochberg adjusted value (BH q), computed across the eight main outcomes of the women's battery within each arm and window, so that a family is one results column. The eight outcomes are the labor market outcomes of Table 2 and, from Section 7.1, college completion by 25, marriage by 23 and by 30, and the age at first marriage to 30, as a restricted mean and among women married by 30, all on the baseline specification without state trends. Their robustness versions (earnings and full-time work among the employed, years of education, and the age at first marriage to 35) and the state-trend and secondary-estimator legs are not counted as separate hypotheses. The Benjamini–Hochberg procedure controls the false discovery rate and is valid when the tests are positively dependent (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001), which is the case here because the outcomes are built from the same birth-state by cohort cells; a family-wise correction that admits arbitrary dependence would over-penalise such a battery and is not reported. Bold marks $q < 0.05$. The two ATT windows are not independent tests, since $[4, 8]$ averages horizons inside $[0, 8]$ on overlapping data, and are adjusted separately.

	Log earnings (employed)		Full-time (employed)	
	Comp.	Trad.	Comp.	Trad.
Panel A: Women				
No state trends				
ATT $[0, 8]$	0.037^{***}	0.008	0.018^{**}	-0.015^{***}
	(0.006)	(0.008)	(0.009)	(0.003)
ATT $[4, 8]$	0.044^{***}	0.005	0.024^{***}	-0.018^{***}
	(0.007)	(0.009)	(0.009)	(0.004)
With state trends				
ATT $[0, 8]$	0.022^{**}	0.017^{***}	0.020^{***}	0.004^{**}
	(0.010)	(0.006)	(0.002)	(0.002)
ATT $[4, 8]$	0.025^{**}	0.017^{***}	0.025^{***}	0.004^{**}
	(0.011)	(0.006)	(0.002)	(0.002)
Panel B: Men				
No state trends				
ATT $[0, 8]$	0.011	0.028^{**}	0.004^{***}	0.001
	(0.017)	(0.011)	(0.001)	(0.002)
ATT $[4, 8]$	0.009	0.024^{**}	0.006^{***}	0.000
	(0.021)	(0.012)	(0.002)	(0.002)
With state trends				
ATT $[0, 8]$	0.002	0.008	0.006^{***}	-0.001
	(0.008)	(0.011)	(0.001)	(0.002)
ATT $[4, 8]$	-0.003	0.003	0.009^{***}	-0.002
	(0.011)	(0.015)	(0.002)	(0.003)
Panel C: Gender gap				
No state trends				
ATT $[0, 8]$	-0.027^{*}	0.019^{**}	-0.015^{*}	0.016^{***}
	(0.015)	(0.009)	(0.009)	(0.004)
ATT $[4, 8]$	-0.036^{**}	0.018	-0.018^{*}	0.018^{***}
	(0.017)	(0.012)	(0.009)	(0.004)
With state trends				
ATT $[0, 8]$	-0.021^{**}	-0.009	-0.014^{***}	-0.005^{**}
	(0.008)	(0.009)	(0.003)	(0.002)
ATT $[4, 8]$	-0.028^{***}	-0.015	-0.017^{***}	-0.006^{*}
	(0.009)	(0.013)	(0.003)	(0.003)
Fixed effects	Birth state, birth cohort, survey year			
Controls	None			
Weights	ACS person weights			
Treated states	8	9	8	9
Control states	23	23	23	23
Observations	28,817	28,817	28,817	28,817
Pre-adoption mean	10.633	10.467	0.769	0.822

Table C.5: Labor market outcomes conditional on working: average ATT.

Notes: Outcomes are the share working thirty-five hours or more and log mean annual wage and salary income, both computed over employed people only, for women, for men and as the men-minus-women gap. The sample is people born 1950–1999 in the 41 classified states, observed in the American Community Survey 2000–2024 from age 25 and weighted by ACS person weights; estimation is by two-stage difference-in-differences, with comprehensive and traditional sex education each measured against the never-mandate pool, in the upper block without and in the lower block with birth-state linear trends. Standard errors, clustered on birth state, in parentheses. Conditioning on employment selects on a margin the mandate itself moves, so these coefficients mix a compositional shift with the effect on hours and pay; they are reported for completeness and the paper's headline numbers are the unconditional versions in Tables 2 and C.9. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Log earnings (all)		Full-time (all)		Employed	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
Main (baseline 2SDID)	0.042***	0.001	0.018**	−0.016***	0.005	−0.006
	(0.008)	(0.010)	(0.009)	(0.005)	(0.004)	(0.005)
Bootstrap SE (999 reps)	0.042**	0.001	0.018	−0.016**	0.005	−0.006
	(0.018)	(0.012)	(0.016)	(0.006)	(0.009)	(0.005)
	[0.049]	[0.946]	[0.194]	[0.017]	[0.267]	[0.220]
No weights	0.035**	−0.007	0.012	−0.007	−0.002	−0.006
	(0.015)	(0.016)	(0.009)	(0.009)	(0.008)	(0.009)
Balanced [-5,5]	0.037***	−0.006	0.019**	−0.017***	0.005	−0.007
	(0.008)	(0.009)	(0.009)	(0.006)	(0.004)	(0.005)
Drop HIV-mandate control states	0.043***	0.005	0.021**	−0.011**	0.005	−0.004
	(0.009)	(0.012)	(0.010)	(0.006)	(0.004)	(0.005)
HIV-mandate control states only	0.044***	0.003	0.019**	−0.015***	0.006	−0.005
	(0.008)	(0.010)	(0.009)	(0.005)	(0.004)	(0.005)
Controls truncated at own content rule	0.040***	0.004	0.018*	−0.013**	0.005	−0.005
	(0.008)	(0.012)	(0.010)	(0.005)	(0.004)	(0.005)
Census region × birth cohort	0.031***	0.004	0.011*	−0.003	0.006	−0.002
	(0.010)	(0.010)	(0.006)	(0.005)	(0.004)	(0.005)
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states	8	9	8	9	8	9
Control states	23	23	23	23	23	23
Observations	28,817	28,817	28,817	28,817	28,817	28,817
Pre-adoption mean	10.349	10.159	0.595	0.623	0.725	0.706

Table C.6: Women: ATT $[0, 8]$ under alternative specifications.

Notes: Each row re-estimates the ATT $[0, 8]$ of Table 2 under one departure from the main specification: with the variance from a 999-draw pairs cluster bootstrap, unweighted, on a balanced event-time window $[-5, 5]$, dropping states with an active HIV-education mandate from the control pool, keeping only such states in the control pool (which holds HIV education fixed and isolates the contraception content, since the early comprehensive adopters enacted their contraception requirements in the wake of their HIV-education mandates), truncating each control state at the first year it carried any binding content rule—abstinence stressed, contraception required or a comprehensive-focus standard—so that it contributes only cohorts schooled before it, and adding census-region × birth-cohort effects to the first stage, so that each adopter is compared only with states in its own census region (Northeast, Midwest, South, West). The in-time placebo is in Table C.3. Outcomes, sample, estimator and treatment arms are otherwise those of Table 2, estimated without state trends. Standard errors, clustered on birth state, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

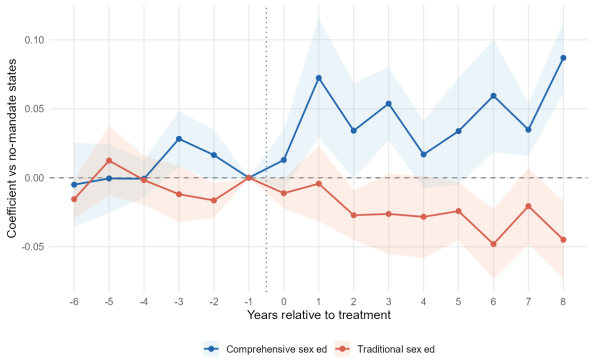
	Log earnings (all)	Full-time (all)	Employed	College degree
Panel A: No state trends				
ATT $[0, 8]$, all treated states	0.042***	0.018**	0.005	0.026***
	(0.008)	(0.009)	(0.004)	(0.006)
range across single-state drops	[0.035, 0.046]	[0.001, 0.021]	[0.000, 0.007]	[0.023, 0.028]
most influential state	New Jersey	New Jersey	New Jersey	California
ATT without it	0.035 [$p = 0.02$]	0.001 [$p = 0.89$]	0.000 [$p = 0.97$]	0.023 [$p = 0.00$]
drops losing significance	none	Maine, Maryland, New Jersey, Texas	none	none
ATT $[4, 8]$, all treated states	0.052***	0.025***	0.009**	0.028***
	(0.008)	(0.009)	(0.004)	(0.006)
range across single-state drops	[0.047, 0.055]	[0.009, 0.029]	[0.004, 0.010]	[0.024, 0.031]
most influential state	California	New Jersey	New Jersey	California
ATT without it	0.047 [$p = 0.00$]	0.009 [$p = 0.24$]	0.004 [$p = 0.60$]	0.024 [$p = 0.00$]
drops losing significance	none	New Jersey	Maryland, New Jersey	none
Panel B: With birth-state linear trends				
ATT $[0, 8]$, all treated states	0.047**	0.030***	0.021**	0.020**
	(0.019)	(0.008)	(0.009)	(0.009)
range across single-state drops	[0.018, 0.064]	[0.017, 0.036]	[0.007, 0.028]	[0.007, 0.028]
most influential state	New Jersey	New Jersey	New Jersey	New Jersey
ATT without it	0.018 [$p = 0.32$]	0.017 [$p = 0.00$]	0.007 [$p = 0.21$]	0.007 [$p = 0.33$]
drops losing significance	Maine, New Jersey	none	New Jersey	New Jersey
ATT $[4, 8]$, all treated states	0.056**	0.038***	0.026**	0.020*
	(0.022)	(0.009)	(0.011)	(0.011)
range across single-state drops	[0.024, 0.077]	[0.025, 0.045]	[0.010, 0.036]	[0.003, 0.031]
most influential state	New Jersey	New Jersey	New Jersey	New Jersey
ATT without it	0.024 [$p = 0.26$]	0.025 [$p = 0.00$]	0.010 [$p = 0.21$]	0.003 [$p = 0.69$]
drops losing significance	New Jersey	none	New Jersey	none
Fixed effects		Birth state, birth cohort, survey year		
Controls		None		
Weights		ACS person weights		
Treated states dropped in turn	18	18	18	18

Table C.7: Women's labor market outcomes and college, comprehensive arm: leaving one treated state out at a time.

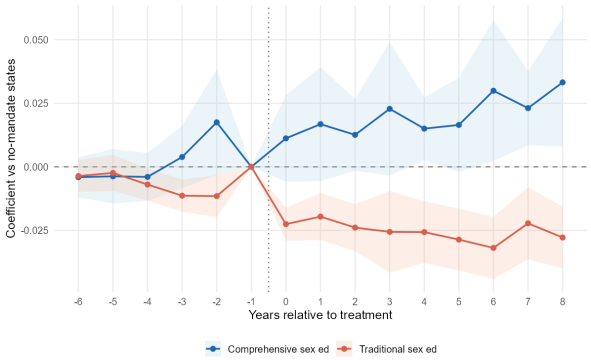
Notes: A leverage diagnostic, not a test. For each outcome and window the first row repeats the ATT of Table 2 (Panel A, no state trends) or its birth-state linear trend counterpart (Panel B), with the standard error clustered on birth state in parentheses; the rows beneath report the range of the ATT when each treated state, of either arm, is dropped in turn and the design re-estimated, the single state whose removal moves the estimate most and the ATT and analytic p-value without it, and the drops, if any, that take the analytic p-value from below to above 0.05. The sample, estimator and treatment arms are those of Table 2: women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 from age 25, two-stage difference-in-differences with comprehensive and traditional sex education each against the never-mandate pool. With eight identifying comprehensive states the jackknife has few points that matter and its spread is not a sampling distribution; the question it answers is whether any one state flips the sign, removes the significance or moves the estimate by more than a standard error. A drop that removes significance without moving the point estimate much is a loss of power: New Jersey and Maryland are the most populous early adopters, so dropping either removes a large share of the treated weight and most of the cohorts observed at long horizons. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.5 Labor market outcomes at ages 30–40

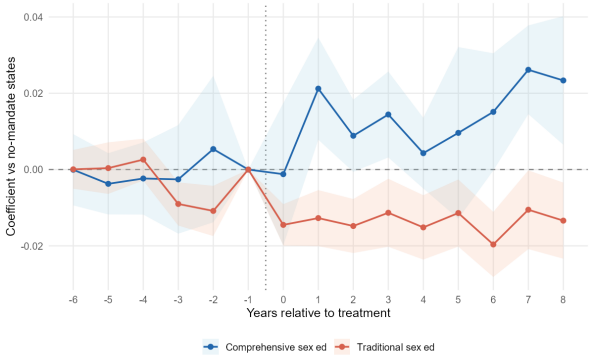
The headline outcomes are measured over everyone aged 25 and above, a band that mixes the early-career transition with settled mid-career work. Re-measuring them over ages 30–40 asks whether the estimated effect is a timing effect that fades once the late twenties are removed, or a level effect on mid-career labor supply and pay. Two cautions attach to the comparison. The 30–40 band is not the 25-and-over sample re-weighted: it is a smaller panel (11,030 cells against 28,817) resting on six comprehensive and eight traditional treated states rather than eight and nine, because at $k = 18$ the two most recent comprehensive adopters and the most recent traditional one have no treated cohort yet observed at age 30, so part of any difference between the bands is cohort composition. And the band is estimated on the baseline fixed effects only: layering the state-trend extrapolation on top of a sample restriction would confound two sources of fragility, neither of which could then be read off the result. Table C.8 reports both bands side by side and Figure C.4 the event studies.



(a) Log annual earnings (all)



(b) Full-time (all)



(c) Employed

Figure C.4: Women's labor market outcomes measured at ages 30–40: event studies by sex education type.

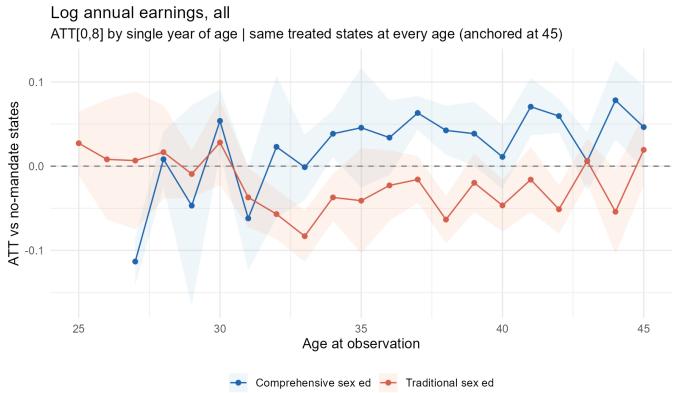
Notes: Outcomes are the log of mean annual wage and salary income per woman with zeros for those without earnings in the previous year, the share working thirty-five hours or more, and the share of women employed, each computed over women observed at ages 30–40 rather than from age 25. The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024, collapsed to birth-state × birth-cohort cells and weighted by ACS person weights; the age restriction leaves 11,030 cells and six comprehensive and eight traditional treated states. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects and no state trends, with each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states and event time counted from the first cohort required to take the mandated curriculum. Shaded bands are 95 percent confidence intervals from standard errors clustered on birth state. Randomization-inference p-values for ATT $[0, 8]$ and $[4, 8]$, comprehensive then traditional: log earnings 0.026 and 0.066, 0.104 and 0.070; full-time 0.158 and 0.188, 0.010 and 0.024; employed 0.112 and 0.114, 0.030 and 0.044.

	Log earnings (all)		Full-time (all)		Employed	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
Ages 25 and over						
ATT $[0, 8]$	0.042^{***}	0.001	0.018^{**}	-0.016^{***}	0.005	-0.006
	(0.008)	(0.010)	(0.009)	(0.005)	(0.004)	(0.005)
ATT $[4, 8]$	0.052^{***}	-0.004	0.025^{***}	-0.019^{***}	0.009^{**}	-0.007
	(0.008)	(0.012)	(0.009)	(0.007)	(0.004)	(0.006)
Pre-adoption average	-0.009	0.003	-0.006	-0.003	-0.004	0.002
	(0.008)	(0.003)	(0.005)	(0.002)	(0.003)	(0.001)
Ages 30–40						
ATT $[0, 8]$	0.045^{***}	-0.026^{***}	0.020^{**}	-0.025^{***}	0.014^{***}	-0.014^{***}
	(0.010)	(0.009)	(0.009)	(0.005)	(0.005)	(0.003)
RI p-value	[0.026]	[0.104]	[0.158]	[0.010]	[0.112]	[0.030]
ATT $[4, 8]$	0.046^{***}	-0.033^{***}	0.024^{**}	-0.027^{***}	0.016^{**}	-0.014^{***}
	(0.010)	(0.011)	(0.009)	(0.006)	(0.006)	(0.004)
RI p-value	[0.066]	[0.070]	[0.188]	[0.024]	[0.114]	[0.044]
Pre-adoption average	0.008	-0.007	0.002	-0.007	-0.001	-0.003
	(0.006)	(0.004)	(0.005)	(0.003)	(0.004)	(0.002)
Placebo, -7 yrs, ATT $[0, 6]$	0.025^{**}	-0.020^{**}	0.006	-0.022^{***}	0.004	-0.011^{***}
	(0.011)	(0.008)	(0.009)	(0.005)	(0.006)	(0.004)
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states, 25+	8	9	8	9	8	9
Treated states, 30–40	6	8	6	8	6	8
Cells, 25+	28,817	28,817	28,817	28,817	28,817	28,817
Cells, 30–40	11,030	11,030	11,030	11,030	11,030	11,030

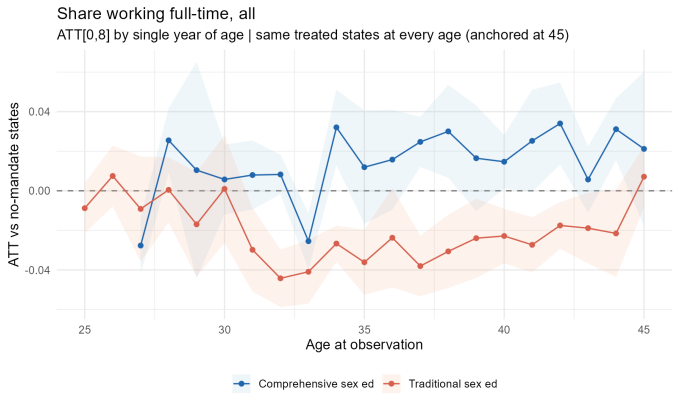
Table C.8: Labor market outcomes measured from age 25 and over ages 30–40: average ATT, women.

Notes: Outcomes are the log of mean annual wage and salary income per woman with zeros for those without earnings in the previous year, the share working thirty-five hours or more, and the share employed, computed first over everyone aged 25 and above and then over ages 30–40 only. The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024, collapsed to birth-state × birth-cohort cells and weighted by ACS person weights. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects and no state trends in either band, with each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. ATT $[0, 8]$ averages the dynamic coefficients over the first nine post-adoption cohorts and ATT $[4, 8]$ over the cohorts that spent all four years of high school under the mandate; the pre-adoption average is the mean dynamic coefficient over event times -6 to -2, and the placebo row moves every adoption date seven years earlier so that it is identified on pre-adoption data alone. Standard errors, clustered on birth state, in parentheses; randomization-inference p-values, in brackets, from 500 draws that re-assign adoption dates across states holding the rollout structure fixed, and they, not the analytic p-values, are what the text reads, because the 30–40 band rests on six comprehensive and eight traditional treated states. The two bands are not the same sample: the restriction costs two comprehensive and one traditional treated state and roughly two-thirds of the cells, as the foot of the table records. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

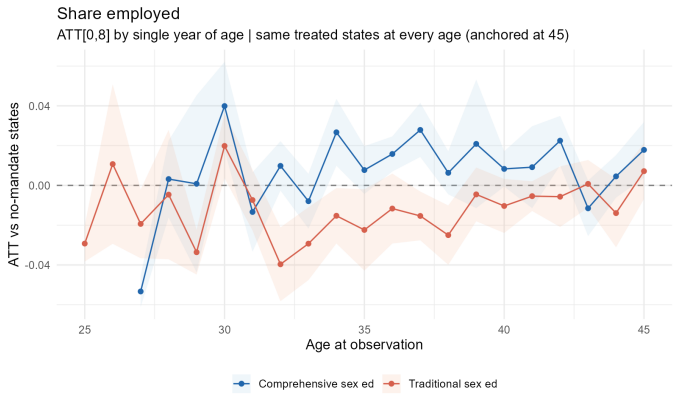
C.6 The age profile of the effect



(a) Log annual earnings (all)



(b) Full-time (all)



(c) Employed

Figure C.5: Women's labor market outcomes: ATT $[0, 8]$ by single year of age.

Notes: Each point is a separate two-stage difference-in-differences estimate of ATT $[0, 8]$ on the birth-state $\times$ birth-cohort slice of the ACS 2000–2024 observed at exactly that age, with birth-state and birth-cohort fixed effects (the survey-year effect is the cohort effect on a single-age slice), the same event-time clock as Figure 2, and each arm—comprehensive and traditional sex education—against the never-mandate pool. Outcomes are the log of mean annual wage and salary income per woman with zeros for those without earnings in the previous year, the share working thirty-five hours or more, and the share employed; weights are ACS person weights. The treated set is held fixed at the states with a treated cohort observed at age 45—Maryland, New Jersey and Vermont in the comprehensive arm; Florida, Georgia, Kentucky, North Carolina, Rhode Island, Tennessee and Utah in the traditional—so the same states contribute at every age from 30; their treated cohorts enter the ACS in their late twenties, and at 28 and 29 only one or two of the three comprehensive states have both an untreated and a treated cohort at that age. Shaded bands are 95 percent confidence intervals from standard errors clustered on birth state.

C.7 Gender gaps

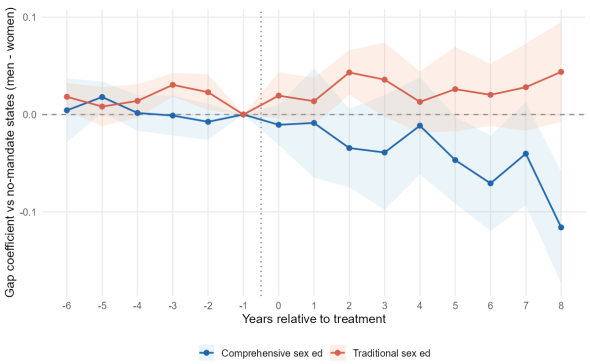
The outcome in this section is the men-minus-women gap within a birth-state × birth-cohort cell, so a negative coefficient is a narrowing gap; Section 6.1 summarizes the results. In the comprehensive arm the earnings gap narrows by 4.2 (2.2) log points across post-adoption cohorts and by 5.7 (2.4) among the fully exposed, against a pre-adoption gap of 46; the full-time gap by 2.3 (1.1) and 2.9 (1.2) percentage points, against 19.3; and the employment gap by 1.4 (0.6) and 1.8 (0.7) points, against 9.0 (Table C.9). With birth-state trends the estimates tighten, and the employment gap grows: −4.9 (1.8) and −6.4 (2.1) for earnings, −2.8 (0.6) and −3.3 (0.8) for full-time work, −2.5 (0.7) and −3.1 (0.9) for employment. The comprehensive pre-adoption coefficients are flat on every outcome without trends—none exceeds 0.003 in absolute value—and the in-time placebo is a precise zero on all three margins (Figure C.6 and Table C.10).

At baseline, resampling confirms little: randomization p-values are 0.148 and 0.100 for the earnings gap, 0.098 and 0.080 for full-time work and 0.108 and 0.072 for employment (Table C.10). With trends they fall to 0.010 and 0.010, 0.006 and 0.006, and 0.006 and 0.008. The narrowing of all three gaps therefore survives randomization inference once the state trends are absorbed. In the traditional arm the earnings and full-time gaps were already widening before adoption—average pre-adoption coefficients of 0.019 ($p = 0.001$) and 0.009 ($p = 0.001$)—and the in-time placebo reproduces the drift, 0.032 (0.010) and 0.017 (0.004), so the traditional estimates are not interpreted.

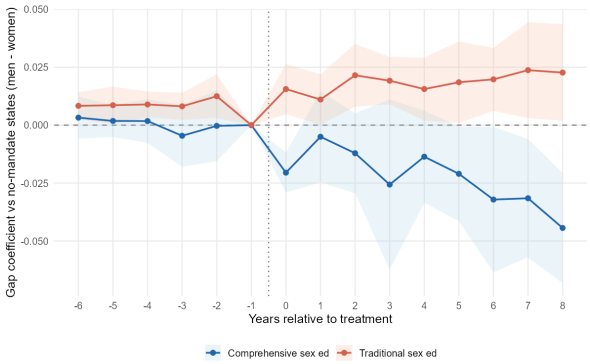
	Log earnings gap (all)		Full-time gap (all)		Employment gap	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
No state trends						
ATT [0, 8]	−0.042*	0.027*	−0.023**	0.019***	−0.014**	0.006
	(0.022)	(0.015)	(0.011)	(0.006)	(0.006)	(0.006)
ATT [4, 8]	−0.057**	0.026	−0.029**	0.020**	−0.018***	0.007
	(0.024)	(0.019)	(0.012)	(0.008)	(0.007)	(0.008)
With state trends						
ATT [0, 8]	−0.049***	−0.006	−0.028***	−0.003	−0.025***	0.000
	(0.018)	(0.011)	(0.006)	(0.004)	(0.007)	(0.004)
ATT [4, 8]	−0.064***	−0.012	−0.033***	−0.005	−0.031***	−0.001
	(0.021)	(0.015)	(0.008)	(0.006)	(0.009)	(0.005)
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states	8	9	8	9	8	9
Control states	23	23	23	23	23	23
Observations	28,760	28,760	28,760	28,760	28,760	28,760
Pre-adoption mean	0.457	0.521	0.193	0.180	0.090	0.099

Table C.9: Gender gaps in labor market outcomes: average ATT, with and without birth-state linear trends.

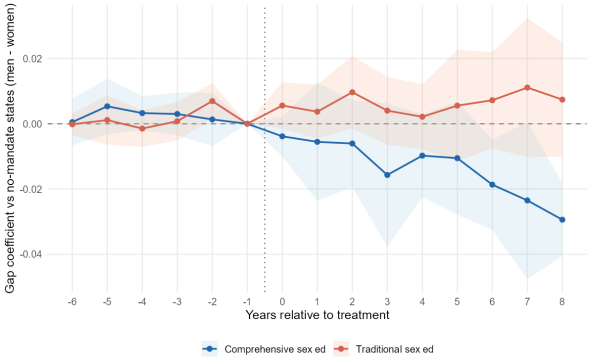
Notes: The outcome is the men-minus-women gap within a birth-state × birth-cohort cell in, respectively, log mean annual wage and salary income, the share working thirty-five hours or more, and the share employed; a negative coefficient is a narrowing gap. The sample is men and women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 from age 25 and weighted by ACS person weights. Estimated by two-stage difference-in-differences; each arm—comprehensive and traditional sex education—is estimated against the same pool of never-mandate states. ATT [0, 8] averages the dynamic coefficients over the first nine post-adoption cohorts and ATT [4, 8] over the cohorts that spent all four years of high school under the mandate. The upper block carries only the fixed effects listed at the foot; the lower block adds a linear trend in birth cohort for each birth state. Standard errors, clustered on birth state, in parentheses. The event studies are in Figure C.6, the conditional-on-working gaps in Table C.5, and randomization inference, HonestDiD and the in-time placebo in Table C.10; men's own outcomes, which the gap and the women's results jointly imply, are in Section C.8. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



(a) Log annual earnings gap (all)



(b) Full-time gap (all)



(c) Employment gap

Figure C.6: Gender gaps in labor market outcomes: event studies by sex education type.

Notes: The outcome is the men-minus-women gap within a birth-state × birth-cohort cell in, respectively, log mean annual wage and salary income, the share working thirty-five hours or more, and the share employed; a negative coefficient is a narrowing gap. The sample is men and women born 1950–1999 in the 41 classified states, observed in the American Community Survey 2000–2024 from age 25 and weighted by ACS person weights. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects and no state trends, with each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states and event time counted from the first cohort required to take the mandated curriculum. Shaded bands are 95 percent confidence intervals from standard errors clustered on birth state. The average ATTs are in Table C.9. Randomization-inference p-values for ATT $[0, 8]$ and $[4, 8]$, comprehensive then traditional: earnings gap 0.148 and 0.100, 0.240 and 0.343; full-time gap 0.098 and 0.080, 0.068 and 0.092; employment gap 0.108 and 0.072, 0.295 and 0.353.

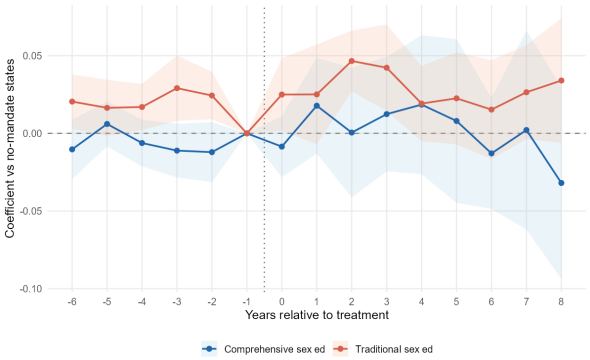
	Log earnings gap (all)		Full-time gap (all)		Employment gap	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
Panel A: No state trends						
ATT $[0, 8]$	-0.042^{*}	0.027^{*}	-0.023^{**}	0.019^{***}	-0.014^{**}	0.006
analytic SE	(0.022)	(0.015)	(0.011)	(0.006)	(0.006)	(0.006)
RI p-value	[0.148]	[0.240]	[0.098]	[0.068]	[0.108]	[0.295]
HonestDiD $\bar{M}$ (10%)	0.00	0.00	0.10	0.20	0.10	n/a
HonestDiD $\bar{M}$ (5%)	0.00	n/a	0.00	0.10	0.00	n/a
ATT $[4, 8]$	-0.057^{**}	0.026	-0.029^{**}	0.020^{**}	-0.018^{***}	0.007
analytic SE	(0.024)	(0.019)	(0.012)	(0.008)	(0.007)	(0.008)
RI p-value	[0.100]	[0.343]	[0.080]	[0.092]	[0.072]	[0.353]
HonestDiD $\bar{M}$ (10%)	0.10	n/a	0.10	0.10	0.10	n/a
HonestDiD $\bar{M}$ (5%)	0.00	n/a	0.00	0.10	0.10	n/a
Panel B: With birth-state linear trends						
ATT $[0, 8]$	-0.049^{***}	-0.006	-0.028^{***}	-0.003	-0.025^{***}	0.000
analytic SE	(0.018)	(0.011)	(0.006)	(0.004)	(0.007)	(0.004)
RI p-value	[0.010]	[0.705]	[0.006]	[0.689]	[0.006]	[0.982]
HonestDiD $\bar{M}$ (10%)	0.20	n/a	0.40	n/a	0.40	n/a
HonestDiD $\bar{M}$ (5%)	0.10	n/a	0.30	n/a	0.40	n/a
ATT $[4, 8]$	-0.064^{***}	-0.012	-0.033^{***}	-0.005	-0.031^{***}	-0.001
analytic SE	(0.021)	(0.015)	(0.008)	(0.006)	(0.009)	(0.005)
RI p-value	[0.010]	[0.555]	[0.006]	[0.583]	[0.008]	[0.924]
HonestDiD $\bar{M}$ (10%)	0.20	n/a	0.40	n/a	0.40	n/a
HonestDiD $\bar{M}$ (5%)	0.10	n/a	0.30	n/a	0.30	n/a
Placebo: adoption moved 7 years earlier; baseline specification						
ATT $[0, 6]$	0.003	0.032^{***}	-0.000	0.017^{***}	0.002	0.003
	(0.012)	(0.010)	(0.007)	(0.004)	(0.004)	(0.002)
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states	8	9	8	9	8	9
Control states	23	23	23	23	23	23
Observations	28,760	28,760	28,760	28,760	28,760	28,760
Pre-adoption mean	0.457	0.521	0.193	0.180	0.090	0.099

Table C.10: Gender gaps: ATT, randomization inference, HonestDiD and placebo.

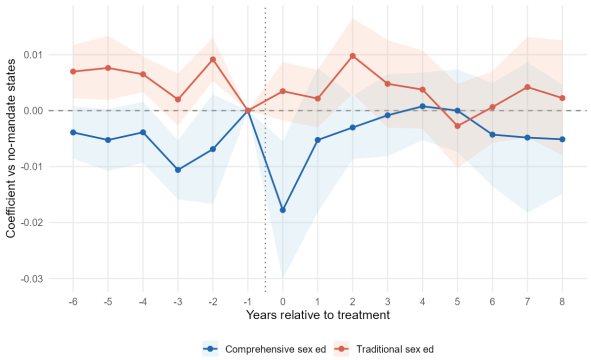
Notes: Outcomes, sample, estimator and treatment arms are those of Table C.9: the men-minus-women gap in log mean annual earnings, the share working thirty-five hours or more, and the share employed, for people born 1950–1999 in the 41 states of the estimation sample and observed in the American Community Survey 2000–2024 from age 25, estimated by two-stage difference-in-differences with comprehensive and traditional sex education each measured against the never-mandate pool. Panel A carries no state trends and Panel B adds a linear trend in birth cohort for each birth state. Analytic standard errors, clustered on birth state, in parentheses; randomization-inference p-values, in brackets, from 500 draws that re-assign adoption dates across states. The HonestDiD rows would report the breakdown $\bar{M}$ described in the notes to Table C.3; the breakdown search is indeterminate for every gap outcome in both specifications, so they show n/a. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.8 Men's labor market outcomes

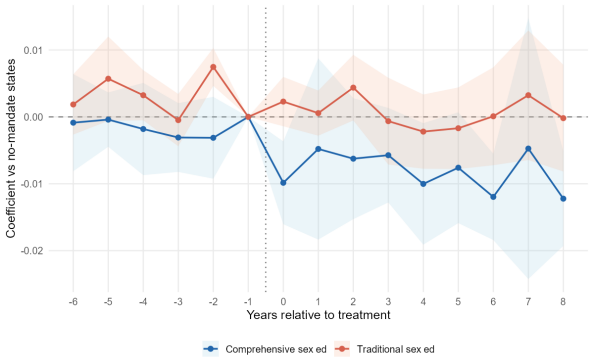
Men's results are reported here rather than in the main text, since they add nothing beyond what the women's and gap results already show.



(a) Log annual earnings (all)



(b) Full-time (all)



(c) Employed

Figure C.7: Men's labor market outcomes: event studies by sex education type.

Notes: Outcomes are the share of men employed, the share working thirty-five hours or more, and the log of mean annual wage and salary income per man with zeros for those without earnings in the previous year. The sample is men born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 from age 25, collapsed to birth-state × birth-cohort cells and weighted by ACS person weights. Estimated by two-stage difference-in-differences, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. Fixed effects are birth state, birth cohort and survey year, with no state trends; the horizontal axis is the cohort's distance from the first cohort required to take the mandated curriculum. Shaded bands are 95 percent confidence intervals from standard errors clustered on birth state. Men are the control condition for the norm channel: a general improvement in schooling quality would move them too, and a curriculum acting on what girls expect of their own trajectories would not. Randomization-inference p-values for ATT $[0, 8]$ and $[4, 8]$, comprehensive then traditional: log earnings 0.990 and 0.932, 0.184 and 0.389; full-time 0.485 and 0.749, 0.469 and 0.733; employed 0.106 and 0.128, 0.874 and 0.966.

	Log earnings (all)		Full-time (all)		Employed	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
No state trends						
ATT $[0, 8]$	0.001	0.028**	−0.004*	0.003	−0.008**	0.001
	(0.020)	(0.013)	(0.003)	(0.003)	(0.003)	(0.002)
ATT $[4, 8]$	−0.003	0.023*	−0.003	0.002	−0.009**	−0.000
	(0.026)	(0.014)	(0.004)	(0.003)	(0.004)	(0.002)
With state trends						
ATT $[0, 8]$	−0.001	−0.003	0.003	−0.008***	−0.004	−0.010**
	(0.007)	(0.010)	(0.003)	(0.003)	(0.004)	(0.004)
ATT $[4, 8]$	−0.006	−0.012	0.005*	−0.010***	−0.005	−0.012**
	(0.011)	(0.014)	(0.003)	(0.004)	(0.005)	(0.005)
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states	8	9	8	9	8	9
Control states	23	23	23	23	23	23
Observations	28,873	28,873	28,873	28,873	28,873	28,873
Pre-adoption mean	10.805	10.684	0.787	0.802	0.815	0.806

Table C.11: Men's labor market outcomes: average ATT, with and without birth-state linear trends.

Notes: Table 2 estimated on men. Outcomes are the log of mean annual wage and salary income per man with zeros for those without earnings in the previous year, the share working thirty-five hours or more, and the share employed; the sample is men born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 from age 25 and weighted by ACS person weights. Estimated by two-stage difference-in-differences, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. ATT $[0, 8]$ averages the dynamic coefficients over the first nine post-adoption cohorts and ATT $[4, 8]$ over the cohorts that spent all four years of high school under the mandate. The upper block carries only the fixed effects listed at the foot; the lower block adds a linear trend in birth cohort for each birth state. Standard errors, clustered on birth state, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

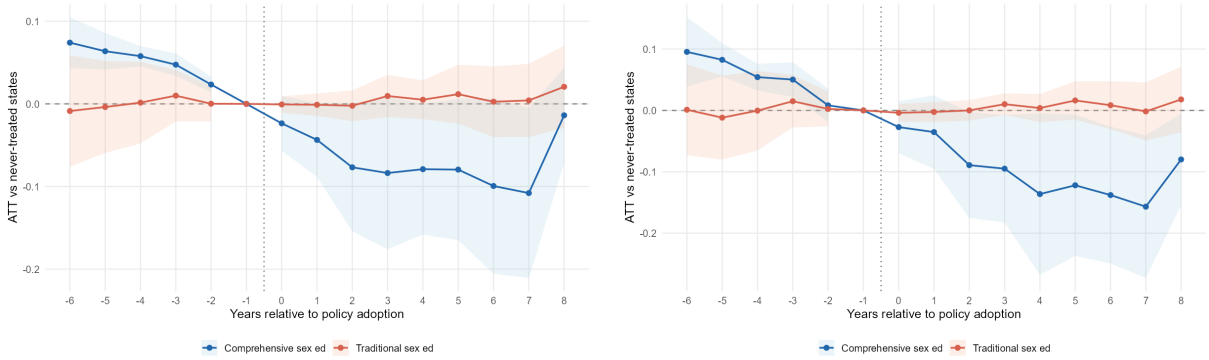
	Log earnings (all)		Full-time (all)		Employed	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
ATT $[0, 8]$	0.001	0.028**	−0.004*	0.003	−0.008**	0.001
analytic SE	(0.020)	(0.013)	(0.003)	(0.003)	(0.003)	(0.002)
RI p-value	[0.990]	[0.184]	[0.485]	[0.469]	[0.106]	[0.874]
HonestDiD $\bar{M}$ (10%)	n/a	0.10	0.00	n/a	0.10	n/a
HonestDiD $\bar{M}$ (5%)	n/a	0.00	n/a	n/a	0.00	n/a
ATT $[4, 8]$	−0.003	0.023*	−0.003	0.002	−0.009**	−0.000
analytic SE	(0.026)	(0.014)	(0.004)	(0.003)	(0.004)	(0.002)
RI p-value	[0.932]	[0.389]	[0.749]	[0.733]	[0.128]	[0.966]
HonestDiD $\bar{M}$ (10%)	n/a	0.00	n/a	n/a	0.00	n/a
HonestDiD $\bar{M}$ (5%)	n/a	n/a	n/a	n/a	0.00	n/a
Placebo: adoption moved 7 years earlier						
ATT $[0, 6]$	−0.009	0.036***	−0.008*	0.011***	−0.002	0.008***
	(0.015)	(0.011)	(0.005)	(0.003)	(0.004)	(0.003)
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states	8	9	8	9	8	9
Control states	23	23	23	23	23	23
Observations	28,873	28,873	28,873	28,873	28,873	28,873
Pre-adoption mean	10.805	10.684	0.787	0.802	0.815	0.806

Table C.12: Men's labor market outcomes: ATT, randomization inference, HonestDiD and placebo.

Notes: Table C.3 estimated on men: the same three outcomes, for men born 1950–1999 in the 41 states of the estimation sample and observed in the American Community Survey 2000–2024 from age 25, by two-stage difference-in-differences with comprehensive and traditional sex education each measured against the never-mandate pool, without state trends. Analytic standard errors, clustered on birth state, in parentheses; bootstrap standard errors and percentile p-values from 999 pairs cluster resamples; randomization-inference p-values, in brackets, from 500 draws that re-assign adoption dates across states. The HonestDiD row is the breakdown $\bar{M}$ defined in the notes to Table C.3, and the placebo at the foot moves every adoption date seven years earlier. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Appendix: Mechanisms

D.1 Teenage pregnancies



(a) Ages 15–19 **(b)** Ages 15–17

Figure D.1: Log teen birth share, ages 15–19 and 15–17: Callaway–Sant'Anna event studies by sex education type.

Notes: The outcome is the log of births per girl in the age band, dated to the year of conception; ages 15–17 isolate the ages at which a birth is most likely to truncate schooling. The sample is a state-by-calendar-year panel, 1980–2024, of 49 states (Wisconsin and the District of Columbia excluded), built from NVSS natality files by the mother's state of residence with Census population denominators. Estimated by the Callaway–Sant'Anna group-time estimator with never-treated controls, a universal base period and no covariates, weighted by the teen female population; each arm—comprehensive and traditional sex education—is measured against the same pool of never-mandate states, and the horizontal axis is years since the mandate took effect. Shaded bands are 95 percent confidence intervals from standard errors clustered on state. The synthetic difference-in-differences event study is Figure 4. Randomization-inference p-values for ATT $[0, 8]$ and $[4, 8]$, comprehensive then traditional: ages 15–19 0.022 and 0.054, 0.840 and 0.818; ages 15–17 0.014 and 0.014, 0.860 and 0.856.

	log(teen birth share) Ages 15–19		log(teen birth share) Ages 15–17	
	Comp.	Trad.	Comp.	Trad.
Panel A: Synthetic DiD				
ATT $[0, 8]$	-0.096^{***}	0.017	-0.123^{***}	0.006
placebo-in-space SE	(0.024)	(0.017)	(0.028)	(0.023)
permutation p-value	[0.005]	[0.313]	[0.005]	[0.781]
ATT $[4, 8]$	-0.113^{***}	0.022	-0.160^{***}	0.012
placebo-in-space SE	(0.032)	(0.024)	(0.037)	(0.031)
permutation p-value	[0.005]	[0.428]	[0.005]	[0.677]
Panel B: Callaway–Sant'Anna				
ATT $[0, 8]$	-0.068^{*}	0.006	-0.098^{**}	0.005
analytic SE	(0.035)	(0.013)	(0.041)	(0.013)
RI p-value	[0.022]	[0.840]	[0.014]	[0.860]
HonestDiD $\bar{M}$ (10%)	0.00	n/a	0.10	n/a
HonestDiD $\bar{M}$ (5%)	n/a	n/a	0.00	n/a
ATT $[4, 8]$	-0.076^{*}	0.009	-0.127^{**}	0.009
analytic SE	(0.040)	(0.019)	(0.050)	(0.019)
RI p-value	[0.054]	[0.818]	[0.014]	[0.856]
HonestDiD $\bar{M}$ (10%)	0.00	n/a	0.10	n/a
HonestDiD $\bar{M}$ (5%)	n/a	n/a	0.10	n/a
Placebo: adoption moved 7 years earlier				
Synthetic DiD, ATT $[0, 6]$	0.006	0.010	−0.028	0.008
	(0.026)	(0.021)	(0.028)	(0.025)
	[0.821]	[0.756]	[0.323]	[0.841]
Callaway–Sant'Anna, ATT $[0, 6]$	−0.024	−0.001	−0.047	−0.006
	(0.023)	(0.027)	(0.030)	(0.028)
Fixed effects	State and year			
Controls	None			
Weights	Teen female population			
Treated states	9	9	9	9
Control states	23	23	23	23
Observations	1,440	1,415	1,440	1,415
Pre-adoption mean	−3.192	−2.701	−3.787	−3.149

Table D.1: Teen births: synthetic difference-in-differences and Callaway–Sant'Anna, with permutation and randomization inference, HonestDiD and in-time placebos, by age band and sex education type.

Notes: The outcome is the log of births per girl in the age band named in the column head, dated to the year of conception. The sample is a state-by-calendar-year panel, 1980–2024, of 49 states (Wisconsin and the District of Columbia excluded), built from NVSS natality files by the mother's state of residence with Census population denominators, and weighted by the teen female population. Panel A is the staggered synthetic difference-in-differences of Table 4: one synthetic control per adopting state at its own adoption year, aggregated by event year with teen-population weights; its standard errors and permutation p-values come from 200 placebo draws that give the adopters' dates to randomly chosen never-mandate states, and its stars from the ratio of the estimate to that standard error. Panel B is the Callaway–Sant'Anna group-time estimator with never-treated controls and a universal base period; analytic standard errors, clustered on state, in parentheses, and randomization-inference p-values, in brackets, from 500 draws that re-assign adoption dates across states holding the rollout structure fixed. The HonestDiD row is the breakdown $\bar{M}$ defined in the notes to Table C.3; it has no counterpart for the synthetic estimator, whose pre-adoption path is matched by construction. The placebos at the foot move every adoption date seven years earlier, so that they are identified on pre-adoption data alone. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Synthetic DiD		Callaway–Sant'Anna	
	Comp.	Trad.	Comp.	Trad.
Main specification	−0.096***	0.017	−0.068*	0.006
	(0.024)	(0.017)	(0.035)	(0.013)
	[0.005]	[0.313]		
Data to 2019	−0.043*	0.016	−0.015	0.004
	(0.025)	(0.017)	(0.013)	(0.012)
	[0.065]	[0.348]		
Without California	−0.032	0.017	−0.001	0.006
	(0.027)	(0.017)	(0.011)	(0.013)
	[0.224]	[0.313]		
No population weights	−0.064***	0.015	−0.029*	0.013
	(0.024)	(0.017)	(0.015)	(0.016)
	[0.010]	[0.378]		
No population weights, data to 2019	−0.054**	0.010	−0.016	0.008
	(0.025)	(0.018)	(0.012)	(0.015)
	[0.025]	[0.557]		
+ state unemployment rate	−0.102***	0.020	−0.072**	−0.016
	(0.025)	(0.021)	(0.035)	(0.035)
	[0.005]	[0.328]		
Drop HIV-mandate control states	−0.080**	0.023	−0.058	0.002
	(0.035)	(0.029)	(0.036)	(0.013)
	[0.030]	[0.458]		
Balanced $[-5, 5]$	−0.098***	0.017	−0.069**	0.006
	(0.025)	(0.017)	(0.035)	(0.013)
	[0.005]	[0.313]		
Not-yet-treated controls	—	—	−0.060*	−0.003
			(0.033)	(0.012)
Multiplier bootstrap SE (999 reps)	—	—	−0.068	0.006
			(0.047)	(0.014)
			[0.148]	[0.696]
Fixed effects	State and year			
Weights	Teen female population, except where the row says otherwise			
Treated states	9	9	9	9
Donor or control states	23	23	23	23
Pre-adoption mean	−3.192	−2.701	−3.192	−2.701

Table D.2: Teen births 15–19: ATT $[0, 8]$ under alternative specifications, synthetic difference-in-differences and Callaway–Sant'Anna.

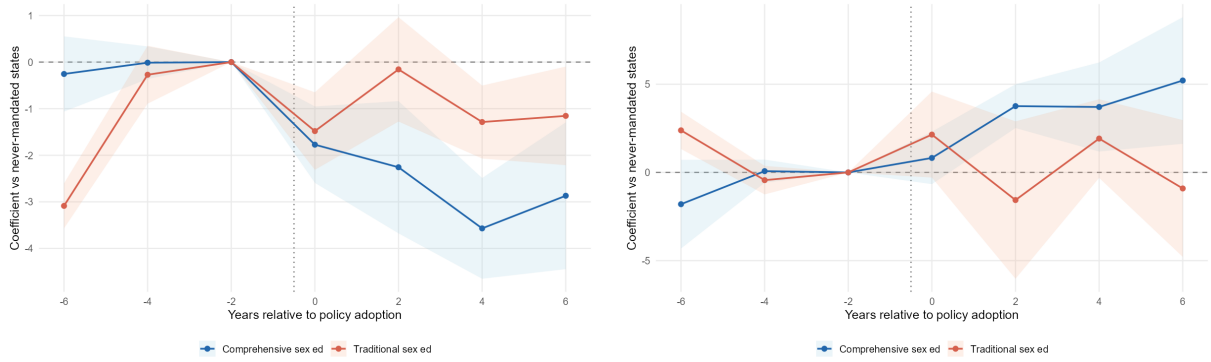
Notes: Each row re-estimates the ATT $[0, 8]$ of Table 4 for ages 15–19 under one departure from the main specification. *Data to 2019* ends the series before the 2020–21 fertility shock, so the 2016 adopters (California, Hawaii) are observed for at most three years and Washington (2022) not at all. *Without California* drops the 2016 adopter that carries the largest teen-population weight among the comprehensive adopters, and whose school spending rose faster than any other state's after 2013 (Section 8.1). *No population weights* weights the adopting states equally. The unemployment row conditions on the state unemployment rate, through regression adjustment in Callaway–Sant'Anna and, in the synthetic estimator, by residualizing the outcome on it with a coefficient estimated on the untreated state-years. Dropping the HIV-mandate states leaves nine donors, as many as the adopters, so the synthetic estimator's placebo draws give each adopter's date to a donor drawn separately. *Balanced* keeps the states observed at every event time from −5 to 5. The not-yet-treated control group and the multiplier bootstrap apply to Callaway–Sant'Anna only. The outcome is the log of births per girl aged 15–19, dated to the year of conception. The sample is a state-by-calendar-year panel, 1980–2024, of 49 states (Wisconsin and the District of Columbia excluded), built from NVSS natality files by the mother's state of residence with Census population denominators, and weighted by the teen female population. Synthetic difference-in-differences as in Table D.1, with permutation p-values in brackets; Callaway–Sant'Anna with standard errors clustered on state, and the bootstrap's percentile p-value in brackets. Every other specification in the paper is estimated without covariates: unemployment enters here and nowhere else, because on the ACS outcomes employment is itself an outcome and conditioning on it would be a bad control. Counts at the foot refer to the main specification. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.2 Sexual behavior (YRBS)

Condom use among all students. The share of all students who used a condom at last intercourse is the product of the share who are sexually active and the share of them who used a condom, so it nets the two margins of Section 7.1.4 against each other. It is small and insignificant, +0.4 (0.8) points on a base of 15.1 percent (Table D.3), which is what the two margins imply: at the pre-adoption means, 28.48 and

50.65 percent, the product is 14.43 percent, and with the estimated changes, −1.91 and +3.38 points, it becomes $0.2657 \times 54.03 = 14.36$ percent, an implied change of −0.07 points (+0.06 for the fully exposed, against an estimate of +0.6 (1.1)). In logs the decomposition is exact on the cells where both items are observed: the share falls by 0.058 (0.034) log points, the sum of −0.116 (0.029) for sexual activity and +0.058 (0.011) for condom use among the sexually active. Because it is the small difference of two offsetting effects, the estimate is imprecise and depends on which treated state is included, between −0.4 and +1.2 points when one is left out, and it cannot separate more protection from less sex; the share having sex without a condom, which adds the two margins, is the outcome reported in the main text.

The in-time placebo. For sex without a condom the placebo, which moves adoption six years earlier, gives −1.7 (0.7), two thirds of the estimate. Its own pre-adoption coefficient is also −1.7 (0.4), so relative to the single reference wave the placebo states sit at about the same level before and after the fake date: the placebo picks up a difference at the reference wave rather than a break, although it cannot rule out that part of the estimate reflects the same pattern. Leaving one treated state out moves the estimate between −2.33 (without Maine) and −3.16 (without New Mexico). For condom use among the sexually active, the pre-adoption coefficient, −0.86 (0.68), is insignificant and opposite in sign, and the placebo is +0.66 (1.24). In the traditional arm the pre-adoption coefficients and the placebos are significant and as large as the estimates, for example −1.7 (0.2) and −3.7 (0.3) for sex without a condom. With three or four treated states the cluster bootstrap cannot be computed: more than half the resamples lose an event-time coefficient.



(a) Sex without a condom, all respondents **(b)** Used a condom, sexually active respondents

Figure D.2: Sex without a condom and condom use at last intercourse.

Notes: Outcomes are the share of all high-school students who are sexually active and did not use a condom at last intercourse (panel a) and the share of the sexually active reporting condom use at last intercourse (panel b). The sample is the state Youth Risk Behavior Surveillance System surveys from 1991 to 2023, pooled from the DASH and SADC releases into a state × wave × grade panel, each cell weighted by its number of respondents; four states identify the comprehensive arm (California, Hawaii, Maine and New Mexico) and three the traditional one (Mississippi, North Carolina and Texas), the other adopters having no survey, no sexual-behavior items or no wave on one side of adoption (Section 7.1.4). Estimated by two-stage difference-in-differences with state, survey-wave and grade fixed effects and contemporaneous exposure, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; waves are biennial, so event time sits on a two-year grid with $\ell = -2$ as the omitted reference. Shaded bands are 95 percent confidence intervals from standard errors clustered on state, which on three or four treated states are unreliable—the randomization-inference p-values in Table D.3 are the test to read. Randomization-inference p-values for ATT $[0, 6]$ and $[4, 6]$, comprehensive then traditional: sex without a condom 0.046 and 0.050, 0.462 and 0.461; condom use among the sexually active 0.088 and 0.106, 0.848 and 0.808.

	Sex without a condom (all)		Currently sexually active		Used a condom (sexually active)		Used a condom (all)	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
ATT $[0, 6]$	-2.617^{***}	-1.020^{***}	-1.910^{***}	-1.556	3.375^{***}	0.398	0.428	-0.619
analytic SE	(0.434)	(0.381)	(0.643)	(1.403)	(0.556)	(1.327)	(0.798)	(1.221)
RI p-value	[0.046]	[0.462]	[0.365]	[0.428]	[0.088]	[0.848]		
HonestDiD $\bar{M}$ (10%)	1.00	0.00	0.10	n/a	0.30	n/a	n/a	n/a
HonestDiD $\bar{M}$ (5%)	0.80	0.00	0.10	n/a	0.20	n/a	n/a	n/a
ATT $[4, 6]$	-3.219^{***}	-1.221^{***}	-2.205^{**}	-2.128^{**}	4.460^{***}	0.506	0.558	-1.089
analytic SE	(0.516)	(0.359)	(0.972)	(1.009)	(0.796)	(1.265)	(1.074)	(0.897)
RI p-value	[0.050]	[0.461]	[0.425]	[0.374]	[0.106]	[0.808]		
HonestDiD $\bar{M}$ (10%)	1.00	0.00	0.10	0.00	0.20	n/a	n/a	n/a
HonestDiD $\bar{M}$ (5%)	0.80	0.00	0.00	0.00	0.20	n/a	n/a	n/a
Placebo: adoption moved 6 years earlier								
ATT $[0, 5]$	-1.704^{**}	-3.710^{***}	0.242	-5.389^{***}	0.656	2.029^{**}	0.945	-1.813^{***}
	(0.749)	(0.270)	(0.806)	(0.520)	(1.244)	(0.985)	(0.798)	(0.524)
Fixed effects	State, survey wave, grade							
Controls	None							
Weights	Respondents per state × wave × grade cell							
Treated states	4	3	4	3	4	3	4	3
Control states	22	22	22	22	22	22	22	22
Observations	1,232	1,232	1,459	1,459	1,232	1,232	1,232	1,232
Pre-adoption mean	15.691	18.726	28.483	44.152	50.653	58.776	15.121	25.426

Table D.3: YRBS: ATT, randomization inference, HonestDiD and placebo.

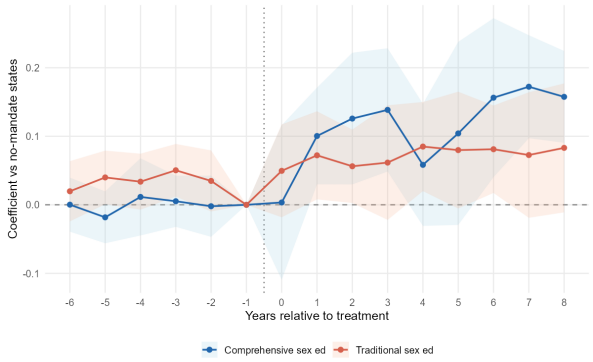
Notes: Outcomes, sample, estimator and treatment arms are those of Table 5: sex without a condom, sexual activity and condom use, among the sexually active and among all students, in percentage points, on the state Youth Risk Behavior Surveillance System surveys 1991–2023 pooled into a state × wave × grade panel, estimated by two-stage difference-in-differences with state, wave and grade fixed effects and contemporaneous exposure, comprehensive and traditional sex education each against the never-mandate pool. Analytic standard errors, clustered on state, in parentheses; randomization-inference p-values, in brackets, from 500 draws that re-assign adoption dates across states, and with three or four treated states per arm these are the test to read. The HonestDiD row is the breakdown $\bar{M}$ defined in the notes to Table C.3. The placebo moves adoption six years earlier rather than seven, because waves are biennial and an odd shift would put the fake event times between waves, and is reported at ATT $[0, 5]$, the waves at event times 0, 2 and 4, all before the true adoption. The cluster bootstrap is not reported: with three or four treated states among 29 clusters, fewer than half the resamples yield a full coefficient vector. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Sex without a condom (all)		Currently sexually active		Used a condom (sexually active)		Used a condom (all)	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
Main (2SDID)	−2.617***	−1.020***	−1.910***	−1.556	3.375***	0.398	0.428	−0.619
	(0.434)	(0.381)	(0.643)	(1.403)	(0.556)	(1.327)	(0.798)	(1.221)
Bootstrap SE (999 reps)	−2.617	−1.020	−1.910	−1.556	3.375	0.398	0.428	−0.619
	(0.434)	(0.381)	(0.643)	(1.403)	(0.556)	(1.327)	(0.798)	(1.221)
No weights	−2.750***	−1.179***	−1.720***	−2.267	3.579***	0.438	0.473	−1.028
	(0.402)	(0.389)	(0.658)	(1.786)	(0.646)	(1.572)	(0.996)	(1.690)
Balanced [-4,4]	−2.596***	−1.029***	−1.803***	−1.582	3.449***	0.382	0.647	−0.640
	(0.458)	(0.382)	(0.637)	(1.405)	(0.552)	(1.327)	(0.741)	(1.221)
Drop HIV-mandate control states	−2.528***	−0.443	−2.167***	−0.898	3.064***	0.044	0.046	−0.488
	(0.838)	(0.416)	(0.577)	(1.632)	(1.134)	(1.369)	(0.942)	(1.422)
Fixed effects	State, survey wave, grade							
Controls	None							
Weights	Respondents per state × wave × grade cell							
Treated states	4	3	4	3	4	3	4	3
Control states	22	22	22	22	22	22	22	22
Observations	1,232	1,232	1,459	1,459	1,232	1,232	1,232	1,232
Pre-adoption mean	15.691	18.726	28.483	44.152	50.653	58.776	15.121	25.426

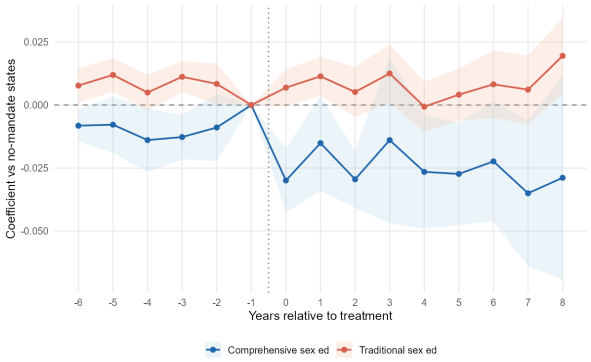
Table D.4: YRBS: ATT $[0, 6]$ under alternative specifications.

Notes: Each row re-estimates the ATT $[0, 6]$ of Table 5 under one departure from the main specification: unweighted, on a balanced event-time window, dropping states with an active HIV-education mandate from the control pool, and with adoption moved earlier in time. Outcomes are sex without a condom, sexual activity and condom use, among the sexually active and among all students, in percentage points, on the state Youth Risk Behavior Surveillance System surveys 1991–2023 pooled into a state × wave × grade panel, each cell weighted by its number of respondents; estimation is by two-stage difference-in-differences with state, wave and grade fixed effects, comprehensive and traditional sex education each against the never-mandate pool. The balanced window is $[-4, 4]$ and the placebo shift six years, both narrower than the ACS equivalents because the YRBS panel is shorter. Standard errors, clustered on state, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

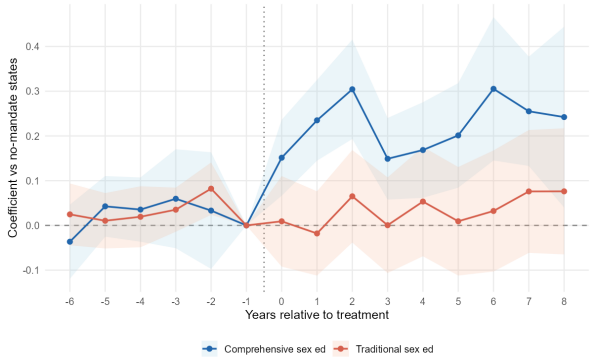
D.3 College and family formation



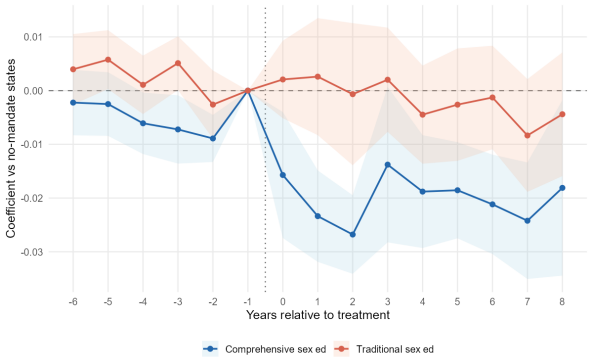
(a) Years of education



(b) Married by 30



(c) Age at first marriage, conditional on marrying by 30



(d) Married by 23

Figure D.3: Alternative measures of education and family formation.

Notes: Outcomes are the mean years of education in a birth-state × birth-cohort cell, the share of women married by 30, the mean age at first marriage among women who marry by 30, and the share married by 23, the age by which a four-year degree begun after high school would be complete; eight comprehensive and nine traditional states identify it. The marriage measures are built from the year of first marriage, asked from 2008. The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 (2008–2024 for the marriage measures) and weighted by ACS person weights. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. None carries birth-state linear trends, since all are age-based outcomes. The horizontal axis is the cohort's distance from the first cohort required to take the mandated curriculum. Shaded bands are 95 percent confidence intervals from standard errors clustered on birth state. The average ATTs are in Tables D.5 and D.6, and the two headline measures these decompose are in Figure 3. Randomization-inference p-values for ATT $[0, 8]$ and $[4, 8]$, comprehensive then traditional: years of education 0.040 and 0.044, 0.122 and 0.132; married by 30 0.056 and 0.070, 0.429 and 0.505; age at marriage conditional on marrying by 30 0.014 and 0.056, 0.591 and 0.505; married by 23 0.020 and 0.046, 0.830 and 0.587.

	Years of education		Married by 30		Age at marriage (cond., 30)	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
No state trends						
ATT $[0, 8]$	0.113^{***}	0.071^{**}	-0.025^{**}	0.008^{*}	0.224^{***}	0.034
	(0.042)	(0.034)	(0.012)	(0.005)	(0.044)	(0.050)
ATT $[4, 8]$	0.130^{***}	0.080^{**}	-0.028^{**}	0.007	0.234^{***}	0.049
	(0.044)	(0.038)	(0.014)	(0.006)	(0.056)	(0.057)
With state trends						
ATT $[0, 8]$	0.096^{**}	-0.024	—	—	—	—
	(0.045)	(0.036)				
ATT $[4, 8]$	0.098^{*}	-0.028	—	—	—	—
	(0.059)	(0.037)				
Lower bound, ATT $[0, 8]$	—	—	-0.035^{**}	0.017^{***}	—	—
			(0.015)	(0.006)		
Upper bound, ATT $[0, 8]$	—	—	-0.024^{**}	0.006	—	—
			(0.010)	(0.006)		
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states	8	9	6	8	6	8
Control states	23	23	23	23	23	23
Observations	28,817	28,817	18,179	18,179	18,179	18,179
Pre-adoption mean	13.817	13.426	0.619	0.679	23.811	22.864

Table D.5: Alternative measures of education and family formation: average ATT.

Notes: The three measures behind the two headline outcomes of Table 3: mean years of education alongside completed college, and the two components of the restricted mean age at marriage—whether a woman marries by 30 at all, and how old she is when she does. The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 and weighted by ACS person weights. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. ATT $[0, 8]$ averages the dynamic coefficients over all post-adoption cohorts and ATT $[4, 8]$ over the cohorts that spent all four years of high school under the mandate. The age-based columns carry no state-trends block, since those outcomes are not estimated with birth-state linear trends. Standard errors, clustered on birth state, in parentheses. The corresponding event studies are Figure D.3. The two bottom rows re-estimate the married-by column on all women, coding those whose first-marriage age is unknown—married more than once, with the most recent marriage after the threshold—as all married after it (lower bound) or all married by it (upper bound); the codings bracket each woman's outcome, not the effect. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Age at marriage (rest. mean, 35)		Age at marriage (cond., 35)		Married by 23	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
ATT $[0, 8]$	0.296^{***}	-0.033	0.250^{***}	-0.047	-0.020^{***}	-0.002
	(0.102)	(0.062)	(0.063)	(0.050)	(0.004)	(0.005)
ATT $[4, 8]$	0.282^{**}	-0.012	0.227^{***}	-0.040	-0.020^{***}	-0.004
	(0.120)	(0.078)	(0.079)	(0.060)	(0.004)	(0.005)
Lower bound, ATT $[0, 8]$	—	—	—	—	-0.028^{***}	0.011^{***}
					(0.008)	(0.004)
Upper bound, ATT $[0, 8]$	—	—	—	—	-0.022^{***}	-0.004
					(0.003)	(0.007)
Fixed effects	Birth state, birth cohort, survey year					
Controls	None					
Weights	ACS person weights					
Treated states	4	8	4	8	8	9
Control states	23	23	23	23	23	23
Observations	14,755	14,755	14,755	14,755	22,936	22,936
Pre-adoption mean	27.875	26.969	25.078	23.992	0.271	0.384

Table D.6: Family formation with the censoring age set to 35, and marriage by 23.

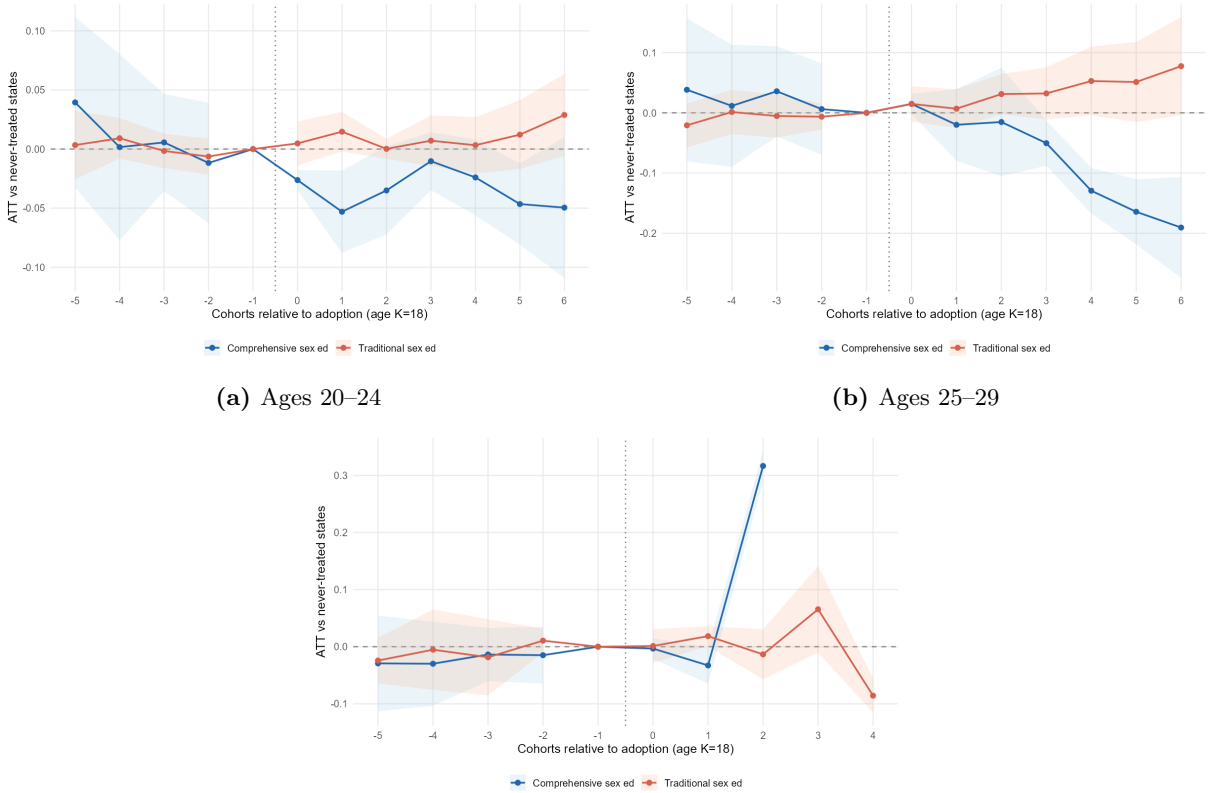
Notes: The marriage measures of Table 3 censored at 35 rather than 30—the restricted mean age at first marriage to 35 and the age at first marriage conditional on marrying by 35—and the share married by 23, the age by which college would be complete (event study in Figure D.3d). The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 and weighted by ACS person weights. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. No state-trends block is reported, because these are all age-based outcomes and are not run with birth-state linear trends. ATT $[0, 8]$ averages the dynamic coefficients over all post-adoption cohorts and ATT $[4, 8]$ over the fully exposed ones. Standard errors, clustered on birth state, in parentheses. The two bottom rows re-estimate the married-by column on all women, coding those whose first-marriage age is unknown—married more than once, with the most recent marriage after the threshold—as all married after it (lower bound) or all married by it (upper bound); the codings bracket each woman's outcome, not the effect. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	College degree		Age at marriage (rest. mean, 30)		Years of education		Married by 30		Age at marriage (cond., 30)	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
ATT $[0, 8]$	0.026^{***}	0.002	0.200^{***}	0.010	0.113^{***}	0.071^{**}	-0.025^{**}	0.008^{*}	0.224^{***}	0.034
analytic SE	(0.006)	(0.004)	(0.058)	(0.043)	(0.042)	(0.034)	(0.012)	(0.005)	(0.044)	(0.050)
RI p-value	[0.008]	[0.786]	[0.032]	[0.876]	[0.040]	[0.122]	[0.056]	[0.429]	[0.014]	[0.591]
HonestDiD $\bar{M}$ (10%)	0.40	n/a	0.30	n/a	0.30	0.10	0.10	0.00	0.20	n/a
HonestDiD $\bar{M}$ (5%)	0.30	n/a	0.20	n/a	0.20	0.00	0.00	n/a	0.20	n/a
ATT $[4, 8]$	0.028^{***}	0.003	0.201^{***}	0.024	0.130^{***}	0.080^{**}	-0.028^{**}	0.007	0.234^{***}	0.049
analytic SE	(0.006)	(0.005)	(0.071)	(0.048)	(0.044)	(0.038)	(0.014)	(0.006)	(0.056)	(0.057)
RI p-value	[0.016]	[0.685]	[0.070]	[0.731]	[0.044]	[0.132]	[0.070]	[0.505]	[0.056]	[0.505]
HonestDiD $\bar{M}$ (10%)	0.20	n/a	0.20	n/a	0.20	0.10	0.10	n/a	0.10	n/a
HonestDiD $\bar{M}$ (5%)	0.20	n/a	0.10	n/a	0.20	0.00	0.00	n/a	0.10	n/a
Placebo: adoption moved 7 years earlier										
ATT $[0, 6]$	-0.001	-0.002	0.058	-0.070^{*}	-0.007	0.054^{*}	-0.015	0.020^{***}	0.026	0.037
	(0.006)	(0.004)	(0.043)	(0.040)	(0.031)	(0.029)	(0.009)	(0.004)	(0.032)	(0.041)
Fixed effects	Birth state, birth cohort, survey year									
Controls	None									
Weights	ACS person weights									
Treated states	8	9	6	8	8	9	6	8	6	8
Control states	23	23	23	23	23	23	23	23	23	23
Observations	28,817	28,817	18,179	18,179	28,817	28,817	18,179	18,179	18,179	18,179
Pre-adoption mean	0.339	0.279	26.318	25.425	13.817	13.426	0.619	0.679	23.811	22.864

Table D.7: Education and family formation: ATT, randomization inference, HonestDiD and placebo, for the main and the alternative measures.

Notes: Outcomes are completed college degree, mean years of education, and the three marriage-timing measures. The sample is women born 1950–1999 in the 41 states of the estimation sample, observed in the American Community Survey 2000–2024 and weighted by ACS person weights. Estimated by two-stage difference-in-differences with birth-state, birth-cohort and survey-year fixed effects, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. No state trends are used, since all but the education outcomes are age-based. Analytic standard errors, clustered on birth state, in parentheses; randomization-inference p-values, in brackets, from 500 draws that re-assign adoption dates across states; the HonestDiD row is the breakdown $\bar{M}$ defined in the notes to Table C.3; and the placebo at the foot moves every adoption date seven years earlier so that it is identified on pre-adoption data alone. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.4 Adult first births



(a) Ages 20–24

(b) Ages 25–29

(c) Ages 30–34

Figure D.4: NVSS first births by state of birth: log first-birth rate, event studies by age group.

Notes: The outcome is the log first-birth rate per woman in the five-year age group named in each panel. The sample is a state-of-birth by year panel built from NVSS natality microdata for 1980–2004, weighted by the female population at risk. Estimated by the Callaway–Sant'Anna group-time estimator with never-treated controls and a universal base period, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. The horizontal axis is years since the mandate took effect. Shaded bands are 95 percent confidence intervals from standard errors clustered on state of birth. The average ATTs are in Table 6 and their in-time placebos in Table D.8. The comprehensive arm is identified by Vermont, New Jersey and Maryland, and at ages 30–34 only event times up to 2 are observed before the series ends, event time 2 for Vermont alone; Section 7.1.4 discusses the placebo and the pre-periods. Randomization-inference p-values for ATT $[0, 6]$ and $[4, 6]$, comprehensive then traditional: ages 20–24 0.060 and 0.112, 0.515 and 0.503; ages 25–29 0.082 and 0.054, 0.248 and 0.281; ages 30–34 0.056 (comprehensive, $[0, 6]$ only), 0.944 and 0.359.

	Ages 20–24		Ages 25–29		Ages 30–34	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
Main (CS)	−0.035**	0.010	−0.079***	0.038**	0.094***	−0.003
	(0.016)	(0.008)	(0.020)	(0.018)	(0.011)	(0.014)
Not-yet-treated controls	−0.035**	0.010	−0.079***	0.038**	0.093***	−0.003
	(0.016)	(0.008)	(0.020)	(0.018)	(0.011)	(0.014)
Multiplier bootstrap SE (999 reps)	−0.035**	0.010	−0.079***	0.038*	0.094***	−0.003
	(0.017)	(0.008)	(0.021)	(0.019)	(0.012)	(0.015)
	[0.038]	[0.205]	[0.000]	[0.050]	[0.000]	[0.855]
No weights	−0.029*	0.006	−0.059	−0.007	0.089***	−0.040***
	(0.017)	(0.014)	(0.058)	(0.020)	(0.014)	(0.015)
Balanced [-5,5]	−0.037**	0.013	−0.081***	0.035**	—	—
	(0.016)	(0.008)	(0.021)	(0.016)		
Drop HIV-mandate control states	−0.032**	0.005	−0.086***	0.009	0.110***	−0.013
	(0.016)	(0.008)	(0.024)	(0.025)	(0.013)	(0.013)
Outcome log(share), zero cells dropped	−0.036**	0.010	−0.082***	0.039**	0.095***	−0.002
	(0.016)	(0.008)	(0.020)	(0.019)	(0.011)	(0.016)
2SDID	−0.074***	0.048***	−0.172***	0.101*	0.134***	0.027*
	(0.015)	(0.018)	(0.060)	(0.055)	(0.029)	(0.015)
2SDID with state trends	0.014	−0.022*	−0.085***	−0.037	0.086***	−0.014
	(0.011)	(0.012)	(0.013)	(0.023)	(0.010)	(0.013)
Placebo shift -7	−0.022	−0.004	−0.032	0.032*	0.010	−0.011
	(0.023)	(0.019)	(0.020)	(0.017)	(0.012)	(0.015)
Fixed effects	State and year (Callaway–Sant'Anna group-time ATTs)					
Controls	None					
Weights	Female population at risk					
Treated states	3	7	3	7	3	4
Control states	25	25	25	25	25	25
Observations	4,250	4,240	3,740	3,740	2,890	2,890
Pre-adoption mean	3.904	4.005	3.677	3.589	3.200	2.885

Table D.8: NVSS first births by state of birth: ATT $[0, 6]$ under alternative specifications.

Notes: Each row re-estimates the ATT $[0, 6]$ of Table 6 under one departure from the main specification: a not-yet-treated rather than never-treated control group, the multiplier-bootstrap variance, dropping the population weights, a balanced event-time window, dropping states with an active HIV-education mandate from the control pool, the log first-birth share with zero cells dropped, and two-stage difference-in-differences in place of Callaway–Sant'Anna, without and with birth-state linear trends. The last row is the in-time placebo, which moves every adoption date seven years earlier and re-estimates on pre-adoption data alone. The outcome is the log first-birth rate per woman in the age group named in the column head; the sample is a state-of-birth by year panel built from NVSS natality microdata for 1980–2004, weighted by the female population at risk. The bootstrap row repeats the point estimate by construction, since the multiplier bootstrap re-computes only the variance; its percentile p-values are in brackets. — marks a window no treated state reaches. Standard errors, clustered on state of birth, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Comprehensive sex ed		Traditional sex ed	
	Estimate	Placebo (−7)	Estimate	Placebo (−7)
Panel A: First births per 1,000 women				
Ages 20–24	−9.0***	−6.0**	9.5***	7.2***
	(1.8)	(2.4)	(2.9)	(1.9)
Ages 25–29	−13.4***	−14.1***	12.3***	12.1***
	(3.4)	(3.8)	(3.5)	(2.3)
Ages 30–33	7.2***	3.1***	1.7	−2.3**
	(0.5)	(0.9)	(1.3)	(1.1)
By 34 (ages 20–33)	−15.2***	−16.9***	23.4***	17.0***
	(3.9)	(4.6)	(4.7)	(3.2)
Pre-adoption base, by 34	545		529	
Panel B: All births per 1,000 women				
Ages 20–24	−12.1***	−3.0	13.9**	12.2***
	(2.6)	(3.0)	(6.2)	(3.7)
Ages 25–29	−25.9***	−14.0***	19.9***	25.9***
	(5.4)	(5.0)	(4.6)	(3.7)
Ages 30–33	8.5***	2.7	5.9***	5.7**
	(1.9)	(2.6)	(1.6)	(2.3)
By 34 (ages 20–33)	−29.5***	−14.3**	39.7***	43.8***
	(6.3)	(6.4)	(7.9)	(5.7)
Pre-adoption base, by 34	1394		1447	
Fixed effects	State and year (two-stage DiD, one single age at a time)			
Controls	None			
Weights	Female population at risk			

Table D.9: NVSS first births and all births by state of birth: change in births per 1,000 women by age band, from the single-age design, with the in-time placebo.

Notes: Each entry is the sum over the ages in the band of the single-age ATT $[0, 6]$ on the level birth rate per 1,000 women, estimated by two-stage difference-in-differences on one single-age subpanel at a time, with state and year fixed effects and no state trends, weighted by cohort size, each arm against the never-mandate pool, so that the 20–33 row is the change in births by 34. An age identified by fewer than three treated states is dropped before summing: ages 31–34 in the comprehensive arm, whose 30–33 band is therefore age 30 alone and whose 20–33 row is births by 31, and ages 32–34 in the traditional arm. The comprehensive arm is the three states whose cohorts reach 30 before the series ends in 2004—Vermont, New Jersey and Maryland. The placebo moves every adoption date seven years earlier and is summed the same way; identified on earlier cohorts, it covers ages 20–33 in both arms. The standard error of a sum treats the per-age fits as independent; they share the same treated states, so it understates the uncertainty. The pre-adoption base is the cumulative rate over the band in the treated states' cohorts before the first treated one. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.4.1 ACS own-children data

The ACS asks women only about births in the twelve months before the survey, so a history of first births has to be reconstructed. I use the own-children method: a woman's age at her first birth is her age at the survey minus the age of the eldest of her own children living in the household. Pooling the 2000–2024 surveys and assigning each woman to her state of birth and birth cohort, as in the main ACS analysis, gives the number of first births at each single year of age per woman, weighted by ACS person weights; the outcome is the log of one plus this rate per 1,000, as in the natality series.

Observing only the children who live at home has two consequences. A first child who has left home, died or lives with the other parent is not seen: if a younger sibling is still at home, the first birth is dated to the sibling's birth and so looks later than it was; if no child is at home, the birth is missed. Both errors grow with the age of the children, and both are more common among women whose children leave home early or live apart, who are not a random group. To bound them, a first birth at age a is counted only for women surveyed while a child born then would be at most 17, and the women at risk at age a are restricted in the same way. Which surveys can see a birth at a given age depends only on the mother's cohort and age, so the restriction is the same for exposed and comparison states, and the errors bias the

estimates only to the extent that co-residence evolves differently in the two groups. The natality data have neither problem, which is why they are the main series, at the cost of ending in 2004.

The ACS is therefore used where the natality data cannot go. A birth from 2000 on is seen while the child is an infant; an earlier birth only through a child still at home when the surveys begin, and the early adopters' exposed cohorts had their births in their twenties largely before 2000. The tables below estimate the comprehensive arm on the six later adopters—Maine, New Mexico, Oregon, California, Hawaii and Washington—and drop Vermont, New Jersey and Maryland, the states the natality data identify, from the sample; the traditional arm is unchanged.

	Ages 20–24		Ages 25–29		Ages 30–34	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
ATT $[0, 6]$	−0.088	−0.005	0.041	0.002	−0.043	−0.017
	(0.061)	(0.013)	(0.030)	(0.009)	(0.038)	(0.021)
	[0.204]	[0.824]	[0.423]	[0.924]	[0.587]	[0.417]
ATT $[4, 6]$	−0.162*	−0.008	0.115***	0.013	0.143**	−0.013
	(0.087)	(0.015)	(0.022)	(0.016)	(0.060)	(0.023)
	[0.172]	[0.798]	[0.154]	[0.637]	[0.319]	[0.655]
Fixed effects	State and year (Callaway–Sant'Anna group-time ATTs)					
Controls	None					
Weights	Female population at risk					
Treated states	6	9	5	9	3	8
Control states	23	23	23	23	24	24
Observations	6,090	6,610	6,090	6,635	6,300	6,870
Pre-adoption mean	3.680	4.008	3.672	3.807	3.503	3.366

Table D.10: Adult first births (ACS own-children), later comprehensive adopters: average ATT on the log first-birth rate, by age group.

Notes: The comprehensive arm is the six later adopters (Maine, New Mexico, Oregon, California, Hawaii, Washington); Vermont, New Jersey and Maryland, identified in Table 6, are dropped from the sample. The outcome is the log first-birth rate per woman in the five-year age group named in the column head, with births inferred by the own-children method from the children living in a woman's household at the survey date. The sample is a state-of-birth by year panel built from the American Community Survey, weighted by the female population at risk. Estimated by the Callaway–Sant'Anna group-time estimator with never-treated controls and a universal base period, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states. ATT $[0, 6]$ averages the group-time effects over the first seven post-adoption years and ATT $[4, 6]$ over the fully exposed ones, the same windows as the NVSS series. Standard errors, clustered on state of birth, in parentheses; randomization-inference p-values in brackets, from 500 draws that re-assign each state's policy bundle across states. The in-time placebo, the bootstrap and the alternative specifications are in Table D.11. Appendix D.4.1 describes the construction and why observing only children at home can date first births too late. Stars refer to the analytic standard errors: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Ages 20–24		Ages 25–29		Ages 30–34	
	Comp.	Trad.	Comp.	Trad.	Comp.	Trad.
Main (CS)	−0.088	−0.005	0.041	0.002	−0.043	−0.017
	(0.061)	(0.013)	(0.030)	(0.009)	(0.038)	(0.021)
Not-yet-treated controls	−0.082	−0.007	0.041	0.002	−0.043	−0.015
	(0.061)	(0.013)	(0.032)	(0.008)	(0.038)	(0.020)
Multiplier bootstrap SE (999 reps)	−0.088	−0.005	0.041	0.002	−0.043	−0.017
	(0.065)	(0.013)	(0.032)	(0.009)	(0.043)	(0.023)
	[0.178]	[0.724]	[0.194]	[0.835]	[0.320]	[0.453]
No weights	−0.128	0.010	0.031	−0.006	−0.100*	0.007
	(0.180)	(0.016)	(0.052)	(0.019)	(0.055)	(0.023)
Balanced [-5,5]	−0.108**	−0.005	0.032	−0.002	—	−0.018
	(0.050)	(0.013)	(0.038)	(0.008)		(0.020)
Drop HIV-mandate control states	−0.044	−0.001	0.080	0.011	−0.017	−0.007
	(0.068)	(0.014)	(0.054)	(0.022)	(0.046)	(0.021)
Outcome log(share), zero cells dropped	−0.105**	−0.005	0.044	0.002	−0.022	−0.020
	(0.047)	(0.014)	(0.031)	(0.009)	(0.044)	(0.022)
2SDID	−0.102**	0.016	0.046*	0.035**	0.123*	−0.011
	(0.050)	(0.015)	(0.026)	(0.017)	(0.069)	(0.014)
2SDID with state trends	−0.035	−0.007	0.024	−0.013	−0.045**	−0.022
	(0.027)	(0.011)	(0.028)	(0.010)	(0.019)	(0.018)
Placebo shift -7	−0.055***	−0.010	−0.031	0.029*	0.023	−0.000
	(0.010)	(0.010)	(0.021)	(0.015)	(0.015)	(0.012)
Fixed effects	State and year (Callaway–Sant'Anna group-time ATTs)					
Controls	None					
Weights	Female population at risk					
Treated states	6	9	5	9	3	8
Control states	23	23	23	23	24	24
Observations	6,090	6,610	6,090	6,635	6,300	6,870
Pre-adoption mean	3.680	4.008	3.672	3.807	3.503	3.366

Table D.11: ACS own-children first births, later comprehensive adopters: ATT $[0, 6]$ under alternative specifications.

Notes: Each row re-estimates the ATT $[0, 6]$ of Table D.10, on the same sample, under one departure from the main specification, as in Table D.8: a not-yet-treated control group, the multiplier-bootstrap variance, no population weights, a balanced event-time window, no HIV-mandate states in the control pool, the log first-birth share with zero cells dropped, and two-stage difference-in-differences without and with birth-state linear trends; the last row is the in-time placebo, which moves every adoption date seven years earlier. Births are inferred by the own-children method from co-resident children, on a state-of-birth by year panel from the American Community Survey weighted by the female population at risk. The bootstrap row repeats the point estimate by construction; its percentile p-values are in brackets. Standard errors, clustered on state of birth, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.5 Gender attitudes

Question wording. In 2004 and 2008 respondents were asked whether they agree that "it is much better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family"; from 2016 they were asked whether it is "better, worse, or makes no difference for the family as a whole if the man works outside the home and the woman takes care of the home and family." The index scores the traditional answer 0, the middle category 0.5 and the egalitarian answer 1, so a cell mean is the whole distribution of responses in one number. The statement on traditional family ties is a five-point agree/disagree item with the same wording in every wave, scored on the same scale with agreement as the traditional end. Two further items cannot be used. One asks whether a working mother can establish as warm a relationship with her children as a mother who does not work, and its wording changed substantially after 2008; the other places the respondent on a seven-point scale between "women should have an equal role with men in running business, industry, and government" and "a woman's place is in the home," and was last asked in 2008. Neither reaches the treated cohorts: the comprehensive arm has six post-adoption women in three states on the first and ten in four states on

the second.

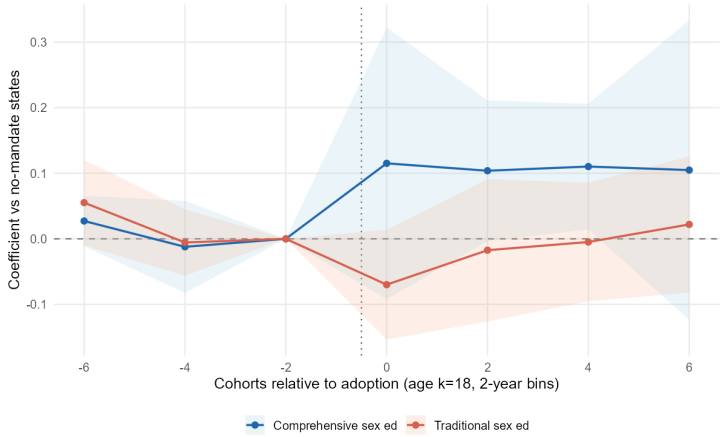


Figure D.5: "A man's place is working and a woman's place is the home": the egalitarian index, women.

Notes: The outcome is the egalitarian index of the item (0 traditional, 0.5 no difference, 1 egalitarian), pooling the 2004/08 statement and the 2016–2024 question. The sample is the American National Election Studies cumulative file, respondents assigned to the state where they grew up and to a birth cohort, exposed at age 18, and weighted by ANES survey weights. Estimated by two-stage difference-in-differences with grew-up-state, birth-cohort and survey-year fixed effects, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; cells are thin, so event time is grouped into two-year bins with $[-2, -1]$ as the omitted reference and the window stops at ± 6. The horizontal axis is the two-year bin of the cohort's distance from the first cohort required to take the mandated curriculum. Shaded bands are 95 percent confidence intervals from standard errors clustered on the grew-up state; with nine treated states in each arm these are unreliable and the randomization-inference p-values in Table D.12 are the test to read. Randomization-inference p-values for ATT $[0, 6]$ and $[4, 6]$: 0.018 and 0.072 for the comprehensive arm, 0.701 and 0.858 for the traditional arm.

Panel A: All waves

	Man works, woman home (index, 0–1)		Traditional family ties (index, 0–1)	
	Comp.	Trad.	Comp.	Trad.
ATT $[0, 6]$	0.109**	−0.018	−0.001	0.004
analytic SE	(0.047)	(0.034)	(0.025)	(0.021)
RI p-value	[0.018]	[0.701]	[0.980]	[0.898]
HonestDiD $\bar{M}$ (10%)	0.30	n/a	n/a	n/a
HonestDiD $\bar{M}$ (5%)	0.20	n/a	n/a	n/a
ATT $[4, 6]$	0.108*	0.009	−0.015	0.022
analytic SE	(0.059)	(0.045)	(0.047)	(0.032)
RI p-value	[0.072]	[0.858]	[0.835]	[0.679]
HonestDiD $\bar{M}$ (10%)	0.10	n/a	n/a	n/a
HonestDiD $\bar{M}$ (5%)	n/a	n/a	n/a	n/a
Placebo: adoption moved 7 years earlier				
ATT $[0, 6]$	0.055**	0.028	−0.001	0.014
	(0.024)	(0.036)	(0.034)	(0.030)
Treated states	9	9	9	9
Control states	24	24	24	24
Observations	5,256	5,256	5,649	5,649
Pre-adoption mean	0.454	0.467	0.414	0.286

Panel B: Man works, woman home: the 2016–2024 question alone

	Man works, woman home 2016–24 (index, 0–1)		Man works, woman home 2016–24: worse or no difference	
	Comp.	Trad.	Comp.	Trad.
ATT $[0, 6]$	0.102**	−0.019	0.156***	−0.044
analytic SE	(0.048)	(0.040)	(0.045)	(0.062)
RI p-value	[0.052]	[0.681]	[0.070]	[0.535]
HonestDiD $\bar{M}$ (10%)	0.20	n/a	0.30	n/a
HonestDiD $\bar{M}$ (5%)	0.10	n/a	0.20	n/a
ATT $[4, 6]$	0.098	0.000	0.147*	−0.018
analytic SE	(0.060)	(0.059)	(0.079)	(0.082)
RI p-value	[0.156]	[0.996]	[0.158]	[0.846]
HonestDiD $\bar{M}$ (10%)	0.00	n/a	0.10	n/a
HonestDiD $\bar{M}$ (5%)	n/a	n/a	n/a	n/a
Placebo: adoption moved 7 years earlier				
ATT $[0, 6]$	0.060***	−0.004	0.103***	−0.006
	(0.022)	(0.033)	(0.035)	(0.057)
Treated states	9	9	9	9
Control states	24	24	24	24
Observations	4,354	4,354	4,354	4,354
Pre-adoption mean	0.395	0.388	0.717	0.702
Fixed effects	Grew-up state, birth cohort, survey year			
Controls	None			
Weights	ANES survey weights			

Table D.12: ANES, women: ATT, randomization inference, HonestDiD and placebo.

Notes: Panel A: the two egalitarian indices of Table 7, *man works, woman home*, pooling the 2004/08 statement and the 2016–2024 question, and *traditional family ties*, asked with the same wording in every wave. Panel B: the man-works item on the 2016–2024 question alone, as the index and as an indicator equal to one if the respondent answers that the arrangement is worse for the family or makes no difference and zero if she answers that it is better. The sample is the American National Election Studies cumulative file, respondents assigned to the state where they grew up and to a birth cohort, exposed at age 18, and weighted by ANES survey weights. Estimated by two-stage difference-in-differences with grew-up-state, birth-cohort and survey-year fixed effects, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; cells are thin, so event time is grouped into two-year bins with $[-2, -1]$ as the omitted reference and the window stops at ± 6. Analytic standard errors, clustered on the grew-up state, in parentheses; randomization-inference p-values, in brackets, from 500 draws that re-assign adoption dates across states, and with nine treated states per arm, these are the test to read. The HonestDiD row is the breakdown $\bar{M}$ defined in the notes to Table C.3, and the placebo at the foot moves every adoption date seven years earlier. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Man works, woman home (index, 0–1)	
	Comp.	Trad.
Main (2SDID)	0.109**	−0.018
	(0.047)	(0.034)
Bootstrap SE (999 reps)	0.109*	−0.018
	(0.054)	(0.037)
	[0.051]	[0.625]
No weights	0.094***	−0.032
	(0.024)	(0.021)
Balanced [-4,4]	0.109**	−0.027
	(0.049)	(0.034)
Drop HIV-mandate control states	0.149***	−0.009
	(0.049)	(0.040)
Drop California	0.144***	−0.015
	(0.043)	(0.035)
Controls for own work status	0.094**	−0.017
	(0.047)	(0.031)
Controls for children in household	0.113**	−0.018
	(0.046)	(0.034)
Controls for own work status and children in household	0.099**	−0.017
	(0.046)	(0.031)
Controls for marital status	0.105**	−0.014
	(0.047)	(0.035)
Controls for work status, children and marital status	0.096**	−0.014
	(0.046)	(0.032)
Placebo shift -7	0.055**	0.028
	(0.024)	(0.036)
Placebo shift -7, dropping California	0.032	0.032
	(0.040)	(0.036)
Fixed effects	Grew-up state, birth cohort, survey year	
Controls	None, except where the row says otherwise	
Weights	ANES survey weights	
Treated states	9	9
Control states	24	24
Observations	5,256	5,256
Pre-adoption mean	0.454	0.467

Table D.13: ANES: ATT $[0, 6]$ under alternative specifications.

Notes: Each row re-estimates the ATT $[0, 6]$ of Table 7 under one departure from the main specification: unweighted, on a balanced event-time window $[-4, 4]$, dropping states with an active HIV-education mandate from the control pool, with the variance from a 999-draw pairs cluster bootstrap, with the respondent's own work status (working, unemployed, retired, disabled, homemaker, student or missing) added to the first stage, with the number of children aged 0–17 in the household (0, 1, 2, 3 or more, or missing), with both, with marital status (married, widowed, divorced or separated, never married, partnered, or missing), with all three, and with adoption moved earlier in time. The work-status, children and marital-status rows condition on outcomes the mandate itself can move, so they answer whether the attitude shifts over and above the change in who works, who raises children and who is married, not whether the estimate is cleaner. All three covariates are recorded in every wave in which the index is asked. The run covers the *man works, woman home* index. The sample is the American National Election Studies cumulative file, respondents assigned to the state where they grew up and to a birth cohort, exposed at age 18, and weighted by ANES survey weights. Estimated by two-stage difference-in-differences with grew-up-state, birth-cohort and survey-year fixed effects, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; cells are thin, so event time is grouped into two-year bins with $[-2, -1]$ as the omitted reference and the window stops at ± 6. Standard errors, clustered on the grew-up state, in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Man works, woman home (index, 0–1)		Traditional family ties (index, 0–1)	
	Comp.	Trad.	Comp.	Trad.
ATT $[0, 6]$	0.042	−0.009	0.034	−0.033
	(0.047)	(0.016)	(0.037)	(0.023)
ATT $[4, 6]$	0.023	−0.025	0.058	−0.055
	(0.066)	(0.030)	(0.056)	(0.036)
Fixed effects	Grew-up state, birth cohort, survey year			
Controls	None			
Weights	ANES survey weights			
Treated states	9	8	9	8
Control states	24	24	24	24
Observations	4,474	4,474	4,769	4,769
Pre-adoption mean	0.439	0.438	0.362	0.262

Table D.14: Gender-role attitudes (ANES), men: average ATT.

Notes: Table 7 estimated on men: the egalitarian index of each item on the 0–1 scale (0 traditional, 0.5 middle, 1 egalitarian). The sample is the American National Election Studies cumulative file, respondents assigned to the state where they grew up and to a birth cohort, exposed at age 18, and weighted by ANES survey weights. Estimated by two-stage difference-in-differences with grew-up-state, birth-cohort and survey-year fixed effects, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; cells are thin, so event time is grouped into two-year bins with $[-2, -1]$ as the omitted reference and the window stops at ± 6. ATT $[0, 6]$ averages the dynamic coefficients over all post-adoption bins and ATT $[4, 6]$ over the fully exposed ones. Standard errors, clustered on the grew-up state, in parentheses; the randomization-inference p-values in Table D.12 are the test to read. Reported for completeness: the argument in the text is about women's attitudes, and the men's estimates are the condition under which a movement among women can be read as specific to them. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Appendix: Ruling Out Alternative Explanations

E.1 School funding

Construction of the spending panel

The fiscal series is built from the National Public Education Financial Survey (NPEFS), the NCES collection of state-level public K–12 finance aggregates, for fiscal years 1987 through 2023. The outcome is current expenditure per pupil in average daily attendance (`TE5/ADA`), deflated to constant dollars: the definition the NCES *Digest of Education Statistics* publishes, and the only construction that reproduces its series, which it matches state by state in 1979–80, 1989–90 and 1998–99 (*Digest* table 194). Record layouts change three times over the period and are parsed from each year's layout files, and negative codes for missing values are set to missing. The one break in the series, around 1989–90, is a reporting change common to every state, which the year effects absorb. Four earlier years (1979–80, 1980–81, 1983–84 and 1985–86) are spliced on from *Digest* 1995 table 164 and extend the pre-period of the earliest adopters.

The estimator is the group-time design of Section 4.3 applied to the log of real per-pupil current expenditure, weighted by each state's mean fall enrollment, with never-treated controls, a universal base period, no covariates, and the adoption dates and control pool of the teen-birth and adult-birth batteries. The windows, the in-time placebo, the pre-trend diagnostic and the HonestDiD breakdown are those of every other outcome, and a balanced $[-5, 5]$ window is reported as a robustness column.

Results

Sample. The main sample caps adoptions at 2015, dropping the later adopters: California and Hawaii (2016) and Washington (2022). The cap matters for the first two. Between 2013 and 2019 California had the fastest growth in real per-pupil spending of any state, 37 percent from a level below most states, and Hawaii the third fastest, 26 percent, against a median of 8 (Figure E.1a); with enrollment weights California alone carries 57 percent of the comprehensive arm. The cap costs the rest of the paper little: California's first treated cohort, born in 1998, reaches 25 only in 2023, so it contributes labor-market cells at event times 0 and 1 alone, and the same holds for Hawaii.

Adopters through 2015. In the comprehensive arm the coefficients for the first three years are within two log points of zero; spending then falls behind the control states, to about seven log points below them by the seventh year, so the averages are negative: -0.033 (0.015) across the first nine post-adoption years and -0.052 (0.016) among the fully exposed (Table E.1). The decline is concentrated in Maryland and New Mexico, whose state-specific estimates are -0.071 and -0.066 (Figure E.2). The pre-period is not flat either—spending in these states was lower four to six years before adoption, an average pre-adoption coefficient of -0.022 (0.012)—so the negative averages should not be read as an effect of the mandate on budgets. The traditional arm gives -0.034 (0.026) and -0.035 (0.033), on a flat pre-period, -0.013 (0.012).

All adopters. Adding the later adopters turns the comprehensive estimate positive, $+0.064$ (0.039) across all post-adoption years and among the fully exposed alike, on a steeper pre-adoption trajectory, -0.049 (0.010) (Figure E.1b). The change comes from California, observed for eight years after adoption,

with the largest state-specific estimate, +0.143; without it, the enrollment-weighted mean of the state-specific estimates is −0.014.

In-time placebo. The placebo is not informative here: on the capped comprehensive sample it is −0.032 (0.049), on the three states with seven years of earlier data, and on the traditional sample +0.019 (0.015) (Table E.2). Spending grows steadily and differently across states, so a design that moves adoption seven years earlier picks up drift; what the capped estimate rules out is a jump at adoption, not a trend.

		Adopters through 2015 (main)			All adopters		
		ATT	%	Pre-trend	ATT	%	Pre-trend
Comprehensive sex ed	ATT $[0, 8]$	−0.033**	−3.3	−0.022	0.064*	+6.7	−0.049
		(0.015)		(0.012)	(0.039)		(0.010)
Comprehensive sex ed	ATT $[4, 8]$	−0.052***	−5.1	−0.022	0.064	+6.7	−0.049
		(0.016)		(0.012)	(0.047)		(0.010)
Traditional sex ed	ATT $[0, 8]$	−0.034	−3.3	−0.013	−0.034	−3.3	−0.013
		(0.026)		(0.012)	(0.026)		(0.012)
Traditional sex ed	ATT $[4, 8]$	−0.035	−3.4	−0.013	−0.035	−3.4	−0.013
		(0.033)		(0.012)	(0.033)		(0.012)
Treated states (comp. / trad.)		6 / 9			9 / 9		
Control states		24			24		
State-years		1147			1258		

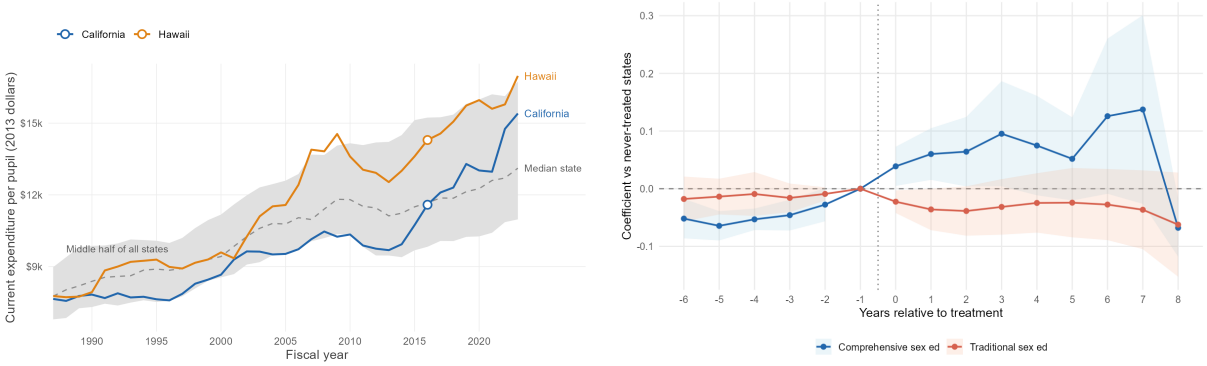
Table E.1: Fiscal placebo: ATT on current expenditure per pupil, capped and full adopter samples.

Notes: The outcome is log real current expenditure per pupil in average daily attendance, from the NPEFS state-year school finance panel for fiscal years 1987–2023, on the 51-state panel with the adoption dates and control pool of every other battery. Estimated by the Callaway–Sant'Anna group-time estimator with never-treated controls, a universal base period and no covariates, weighted by each state's mean fall enrollment, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; standard errors clustered on state. This is a placebo on the mechanism rather than an outcome of interest: a resources story requires school budgets to break at adoption. The main sample keeps adopters through 2015, which excludes California and Hawaii (2016) and Washington (2022); "All adopters" adds them, the 2016 adopters with eight years from adoption each (2016–2023) and Washington with two before the panel ends. ATT $[0, 8]$ averages the group-time effects over the first nine post-adoption years and ATT $[4, 8]$ over the last five; "%" is the ATT in percent of the pre-adoption level; "Pre-trend" is the average pre-adoption coefficient over event times −6 to −2, with its standard error beneath, the same diagnostic every other table reports. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

		Placebo ATT $[0, 6]$	Balanced $[-5, 5]$ ATT $[0, 8]$	HonestDiD $\bar{M}$ (10%)	HonestDiD $\bar{M}$ (5%)
Adopters through 2015					
Comprehensive sex ed	ATT $[0, 8]$	-0.032	-0.049^{***}	0.00	0.00
		(0.049)	(0.015)		
	ATT $[4, 8]$			0.10	0.00
Traditional sex ed	ATT $[0, 8]$	0.019	-0.030	n/a	n/a
		(0.015)	(0.036)		
	ATT $[4, 8]$			n/a	n/a
All adopters					
Comprehensive sex ed	ATT $[0, 8]$	-0.038	0.079^{**}	0.00	n/a
		(0.030)	(0.037)		
	ATT $[4, 8]$			n/a	n/a
Traditional sex ed	ATT $[0, 8]$	0.019	-0.030	n/a	n/a
		(0.015)	(0.036)		
	ATT $[4, 8]$			n/a	n/a

Table E.2: Fiscal placebo: in-time placebo, balanced window and HonestDiD sensitivity.

Notes: The outcome is log real current expenditure per pupil in average daily attendance, from the NPEFS state-year school finance panel for fiscal years 1987–2023, on the 51-state panel with the adoption dates and control pool of every other battery. Estimated by the Callaway–Sant'Anna group-time estimator with never-treated controls, a universal base period and no covariates, weighted by each state's mean fall enrollment, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; standard errors clustered on state. This is a placebo on the mechanism rather than an outcome of interest: a resources story requires school budgets to break at adoption. The in-time placebo moves every adoption date seven years earlier and re-estimates on pre-adoption data alone, averaged over $[0, 6]$; the balanced column restricts the treated states to those observed over the full $[-5, 5]$ window; the HonestDiD columns are the breakdown $\bar{M}$ at the 10 and 5 percent levels, the largest post-treatment violation of parallel trends, as a multiple of the largest pre-treatment violation, at which the estimate stays significant. "n/a" means the estimate is not significant even at $\bar{M} = 0$, so a breakdown value is not defined. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



(a) Per-pupil spending of the two 2016 adopters

(b) Event study, all adopters

Figure E.1: The 2016 adopters and the fiscal placebo on the full adopter sample.

Notes: Panel (a): real current expenditure per pupil in average daily attendance, in 2013 dollars, for California and Hawaii, both of which adopt a comprehensive mandate in 2016; open circles mark the adoption year, the grey band is the interquartile range across all 51 states and the dashed line the median. California ranks between 36th and 40th of 51 from 1995 to 2013; between 2013 and 2019 its real per-pupil spending grows by 37 percent, the fastest of any state, and Hawaii's by 26 percent, the third fastest, against a median of 8. This is why the main fiscal sample, Figure 5, caps adoptions at 2015. Panel (b): the event study of Figure 5 on the full adopter sample, keeping the two 2016 adopters; it is shown for transparency rather than relied on, because California's trajectory in panel (a) carries it. Log real current expenditure per pupil in average daily attendance, NPEFS state-year panel for fiscal years 1987–2023; Callaway–Sant'Anna with never-treated controls, a universal base period and no covariates, weighted by each state's mean fall enrollment, each arm measured against the same pool of never-mandate states. Shaded bands are 95 percent confidence intervals from standard errors clustered on state.

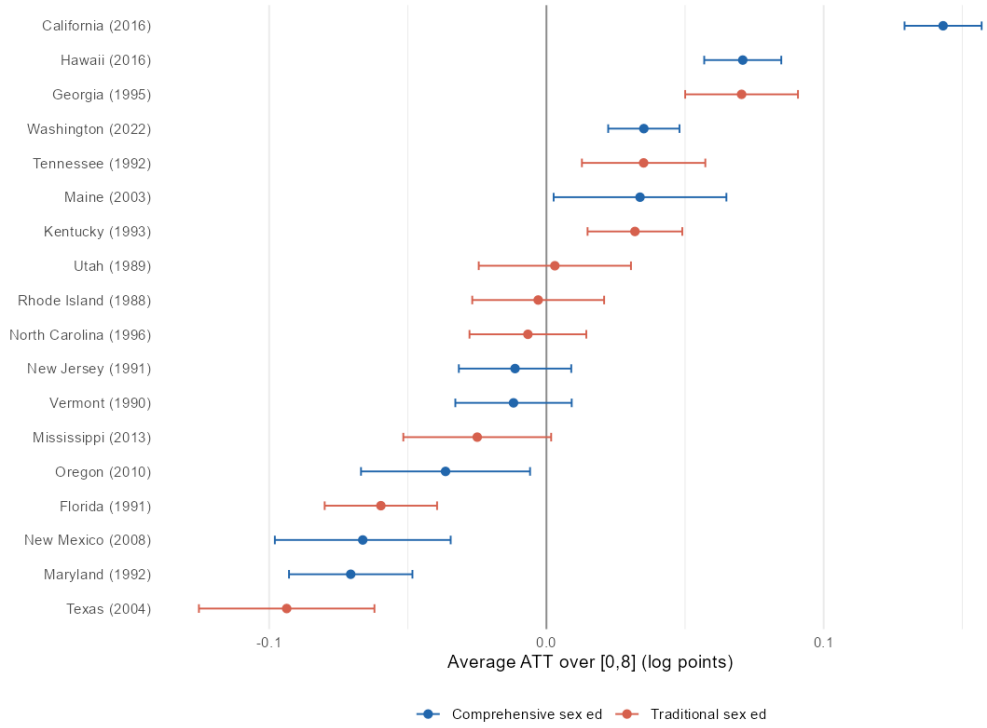


Figure E.2: Fiscal placebo: state-by-state ATT on current expenditure per pupil.

Notes: The outcome is log real current expenditure per pupil in average daily attendance, from the NPEFS state-year school finance panel for fiscal years 1987–2023, on the 51-state panel with the adoption dates and control pool of every other battery. Estimated by the Callaway–Sant'Anna group-time estimator with never-treated controls, a universal base period and no covariates, weighted by each state's mean fall enrollment, each arm—comprehensive and traditional sex education—measured against the same pool of never-mandate states; standard errors clustered on state. This is a placebo on the mechanism rather than an outcome of interest: a resources story requires school budgets to break at adoption. Each point is one adopting state's own average ATT over the first nine post-adoption years, estimated alone against the never-treated pool, with a 95 percent confidence interval from a single treated cluster (read the intervals as indicative), so that the figure shows whether the pooled estimate rests on one or two states.

E.2 Electoral cycles

Construction of the seat-share measure. Each chamber is rebuilt from its winners in the State Legislative Election Returns (Klarner et al., 2013), 1967–2016, with the Partisan Balance file (Klarner, 2013) filling the state-years the returns do not reach; over the 1,017 state-years both cover, the Democratic house seat shares correlate at 0.999. I keep one record per seat, so that a cross-endorsed winner in a fusion-voting state is counted once and by true party, and compute the two-party seat share $D/(D+R)$ of each chamber, averaged unweighted across the two; senates on staggered terms are rebuilt from each district's most recent election, and third-party and independent members, half a percent of seats, are left out. Nebraska, whose legislature is unicameral and nonpartisan, and the District of Columbia are dropped, and the panel starts in 1974, the first year Minnesota elected its legislature on a partisan ballot. Every state is placed on a common two-year grid through 2016, each chamber carried forward from its most recent election, so that $e = -1$ is the legislature seated in the two years before adoption in every state.

Estimation. Event time is $e = \lfloor (y-g)/2 \rfloor$ for grid year y and adoption year g, so $e = 0$ is the first legislature elected in or after the year the mandate takes effect. The Callaway–Sant'Anna estimator compares each adopting state's change in the aligned seat share from $e = -1$ to every other election

with the change in the never-treated states over the same calendar years, and aggregates across adoption cohorts by cohort size; the averages in Table E.3 are equal-weighted means of the event-time coefficients over the window, as elsewhere in the paper. The comprehensive arm has nine adopters and the traditional arm nine; Washington, which adopts in 2022, after the returns end, contributes no post-adoption election and is dropped, leaving eight and nine, and Wisconsin, whose policy changes twice, is left out of the 30 never-treated states as in every other battery. The in-time placebo moves every adoption date six years, three elections, earlier and averages over $e = 0$ to 2 of the placebo clock, which is the three elections before the true adoption measured against the fourth. The two-way fixed-effects version regresses the same outcome on event-time indicators with state and year fixed effects, the adopters' observations restricted to $e = -4$ to 3. No specification includes state trends: with eight or nine adopters and a pre-period of four elections, a state-specific slope would absorb most of the variation the event study is meant to show.

Levels around adoption. Table E.4 reports the raw means behind the event study. The Democratic share of comprehensive adopters rises from 62.0% at $e = -4$ to 65.0% at $e = -1$ and is 61.8% at $e = 0$, and the Republican share of traditional adopters from 35.7% to 45.0% and then 44.5%: neither steps up as the policy arrives. The comprehensive share falls to 53.9% by $e = 3$, but two things account for the fall, neither of them the mandate. California and Hawaii, among the most Democratic legislatures in the sample, adopt in 2016 and leave the table after one election, and the national swings of 1994 and 2010 lower the Democratic share of the never-treated states over the same calendar years, which the event study nets out.

	Comprehensive sex ed (Democratic share)	Traditional sex ed (Republican share)
After adoption, $e = 0$ to 3	0.6	−2.2
	(2.8)	(2.2)
In-time placebo, onsets six years earlier, $e = 0$ to 2	4.5	3.7
	(3.3)	(2.4)
Two-way fixed effects, after adoption, $e = 0$ to 3	0.0	−2.4
	(2.5)	(2.0)
Adopting states	8	9
Never-treated states	30	30
Fixed effects	State and year (Callaway–Sant'Anna group-time ATTs)	

Table E.3: Legislative composition around adoption: average effects on the aligned two-party seat share.

Notes: The outcome is the two-party seat share $D/(D+R)$, in percentage points, of the Democratic party in the comprehensive arm and of the Republican party in the traditional arm, averaged unweighted across the two chambers, 1974–2016 on a two-year election grid. The rows are equal-weighted averages, after adoption, of the event-time coefficients of Figure 6, each against the legislature elected in the two years before the mandate took effect ($e = -1$); the in-time placebo moves every adoption date six years earlier and averages over the three elections before the true adoption; the two-way fixed-effects row regresses the outcome on event-time indicators with state and year fixed effects. Callaway–Sant'Anna with never-treated controls, a universal base period, no covariates and no weights, each arm against the same never-mandate states. Standard errors clustered on state in parentheses. *, ** and *** denote significance at 10, 5 and 1 percent. Sources: Klarner et al. (2013), supplemented state-year by state-year with Klarner (2013).

	Before adoption				After adoption			
Election relative to adoption	−4	−3	−2	−1	0	1	2	3
Comprehensive adopters, Democratic share	62.0	63.7	66.0	65.0	61.8	54.8	55.3	53.9
States observed	*8*	*8*	*8*	*8*	*8*	*6*	*6*	*6*
Traditional adopters, Republican share	35.7	36.9	38.6	45.0	44.5	45.3	45.4	48.6
States observed	*9*	*9*	*9*	*9*	*9*	*9*	*8*	*8*

Table E.4: Legislative composition around adoption: mean seat share of the party aligned with the mandate among adopting states.

Notes: Descriptive means, not estimates. Each cell is the mean across adopting states, at the election indexed relative to adoption, of the two-party seat share $D/(D+R)$ in percent held by the Democratic party in comprehensive states and by the Republican party in traditional states, averaged unweighted across the two chambers; $e = -1$ is the legislature elected in the two years before the mandate took effect. The returns end in 2016, so coverage after adoption is unbalanced and the rows in italics report how many states enter each cell: California and Hawaii (comprehensive, 2016) are observed for one election after adoption, Mississippi (traditional, 2013) for two. Sources: Klarner et al. (2013), supplemented state-year by state-year with Klarner (2013).