Opt-In Revolution: A Comment*

Randy Cragun†

April 6, 2022

Abstract

A re-analysis of the evidence on effects of legal changes granting confidential early legal access (ELA) to oral contraceptives in late adolescence. Past work used a difference-in-differences framework with state-specific timing of ELA to suggest that ELA increased women's wages, annual earnings, and labor force participation. This comment shows that those conclusions misinterpret the data. The apparent effects result from a restriction on the regression model that induces substantial bias due to an empirically observable violation of the parallel trends condition. This article derives an expression for the bias, shows that the condition is violated in this case, shows that randomly assigning placebo ELA produces apparent effects that are nearly indistinguishable from the estimates of effects of actual ELA using the restricted model, and shows that the unrestricted model gives no evidence for effects of ELA.

JEL: C21, C52, J13, J18, J22, J31

*This research benefited from a grant from Birmingham-Southern College and from non-material support and feedback from Jason Lindo, Caitlin Knowles Myers, Ishita Chatterjee, James Cragun, Clinton Jenkins, Pamela A. Meyerhofer and participants in the "Impacts of access to contraception and abortion" session of the 2020 Southern Economics Association meetings, Leo Rossi, Mallika Garg, and Jacob Burgdorf. The author has no conflicts of interest to disclose.

†Birmingham-Southern College, Box 549007, 900 Arkadelphia Rd, Birmingham, AL 35254, rc@rcragun.net, +1-205-226-7748, Fax:+1-205-226-4847.

Bailey, Hershbein, and Miller (2012) (hereafter "BHM") and Bailey (2006) ("Bailey") estimate effects of confidential access to oral contraceptives ("the pill") in late adolescence on women's earnings and labor force attachment. This comment demonstrates that the effects are entirely due to bias induced by a regression model restriction.

Starting in the 1960s in the US,[1] the pill provided a safe method to reduce uncertainty about fertility without the high cost of sexual abstinence, but young people who wanted the pill needed a prescription, and the Common Law usually prevented physicians from accepting the consent of minors. Around that time, US states started lowering the age at which a person was considered a legal adult (the age of majority), allowing minors to give consent to treatment if they were mature enough to understand the consequences of their choices (mature minor doctrines), and passing medical consent and family planning laws extending the right to medical or contraceptive choice to minors. Such legal changes removed barriers for young women who wanted confidential access to the pill, a condition that Bailey terms "Early Legal Access" (or "ELA") to the pill.

The timing of ELA was state-specific, so Bailey uses a difference-in-differences (DD) framework to estimate effects, showing evidence that ELA increased women's labor force attachment in their late 20s and early 30s (but not at later ages). BHM use the same methodology to analyze effects on earnings over women's lives. They present evidence that women's earnings fell in their early 20s and rose at later ages as a result of pill access, consistent with increased formal human capital accumulation in youth that pays off later. Lindo et al. (2020) extend BHM's estimates to older ages, finding no evidence of differences in earnings near retirement age. They also estimate effects of ELA on the probability of working in a Social-Security-covered job (a measure of employment) and find positive effects on employment for young women like Bailey did.

This comment demonstrates that the effects estimated by Bailey and BHM (and, somewhat less so, Lindo et al. 2020) are artefacts of functional form. Specifically, they estimate

[1] The FDA approved Enovid for contraceptive purposes in 1960. It had been in use for regulating menses since 1957.

effects at multiple ages with DD two-way-fixed-effects regressions that interact the treatment variable with age category indicators but do not interact the state and time (birth cohort) dummies with those age categories, imposing on the model that all states share the same baseline age patterns (absent treatment with ELA) in the dependent variables and that all ages had the same rate of growth of the outcome variable (again, in the absence of treatment). I derive an expression for unbiasedness of the OLS estimator of the restricted model and show that the condition is violated in the data. I then show that the restricted model produces apparent effects of placebo contraceptive policies that are nearly indistinguishable from the estimates of effects of actual contraceptive policies and that the unrestricted model gives no evidence for effects of early confidential access to the pill.

The source of bias addressed in this article is a violation of the parallel trends condition. BHM and Bailey do not compare changes in earnings (or labor force participation) for women in their early 30s who were treated with ELA to changes in earning by other women in their early 30s but, instead, to changes in earnings by *all* women in the sample whose states did not change their laws at that time. The bias induced by this model restriction is distinct from bias in DD estimates due to endogenous effects of treatment on the rate of change of the outcome in previously-treated groups (see Goodman-Bacon 2020, for a discussion of DD models allowing effects on both levels and slopes of outcomes) and occurs even if the timing of treatment is exogenous. This comment is similar to the argument by Yzerbyt, Muller, and Judd (2004) that if the effect of a treatment is allowed to vary with some individual difference, then controls usually need to be interacted with that individual difference as well.

Myers (2017) casts doubt on claims that ELA caused delayed fertility and marriage and presents evidence that increased confidential pill access in youth induced large increases in youth sexual activity, offsetting the pill's direct effect of decreasing pregnancy risk per time of intercourse. These results suggest that increased pill use by age 21 among those with ELA (as shown by Bailey, Hershbein, and Miller 2012) is an unlikely mechanism for downstream effects like earnings and labor force participation. However, delayed fertility induced by pill use is not the only mechanism by which ELA could impact downstream outcomes. For example, increased certainty about the timing of fertility could increase the expected value of investments in human capital by reducing expected career disruptions (Bailey 2006). Thus, the effect of ELA on careers and earnings is still an open question.

Bailey and BHM estimate effects on a variety of outcomes. Because they use the same model for all estimates, the criticism in this comment applies in each case.[2] However, I estimate effects only on earnings, wages, and labor force participation for brevity.

1 The model restriction

Most states within the US granted ELA sometime during the 1960s or 1970s, and these changes may have reduced the cost of obtaining and using oral contraceptives for women who reached the affected ages after the policy change, which might have changed the incentives for early human capital accumulation, which would change the path of those women's earnings and labor force participation over their lives. We can estimate the effect of ELA on an outcome at some age a with a DD model where the first difference is between birth cohorts and the second is between states:

$$Y_{abs} = \alpha + \delta ELA_{bs} + \beta_b + \gamma_s + \varepsilon_{abs} \tag{1}$$

where Y_{abs} is the average of the outcome (labor force participation, annual earnings, hourly wage, etc.) at age a of birth cohort b in state s, β_b is a set of birth year fixed effects, γ_s is a set of state fixed effects, and ELA_{bs} is 1 if women born in state s in year b could legally consent to treatment with oral contraceptives in late adolescence and 0 otherwise.[3] δ is a

[2] The criticism does not apply to estimates by BHM of effects of ELA on the probability of using the pill before age 21.

[3] There is likely heteroskedasticity and error correlation over time, but addressing that is outside the scope of this paper (see, e.g., Bertrand, Duflo, and Mullainathan 2004; Conley and Taber 2011; Ferman and Pinto 2018; MacKinnon and Webb 2017; MacKinnon and Webb 2018; Stephen G. Donald and Kevin Lang 2007). I use state averages to avoid dealing with error correlation across individuals within states.

weighted average of the treatment effects.[4] For simplicity, all regression models in this report are written for observations at the state × birth cohort × age group level, but the empirical analysis uses individual-level data.

We can estimate effects at different ages by estimating Eq. (1) separately for each age or we can pool the data and (more efficiently) estimate one regression with *every* right-hand-side term in Eq. (1) interacted with age group dummies, which maintains the DD structure of the single-age version. The pooled (across age) "unrestricted" version of Eq. (1) is

$$Y_{abs} = \alpha_a + \sum_a \delta_a \times ELA_{bs} \times I(a) + \beta_{ab} + \gamma_{as} + \varepsilon_{abs} \tag{2}$$

where $I(a)$ is an indicator for age a and the fixed effects are now for age by state and age by birth cohort.

Bailey and BHM estimate Eq. 2 but without allowing for state and birth cohort fixed effects to differ by age (hereafter the "restricted model"):

$$Y_{abs} = \alpha_a + \sum_a \delta_a \times ELA_{bs} \times I(a) + \beta_b + \gamma_s + \varepsilon_{abs} \tag{3}$$

2 Source of bias

Suppose we had just two birth cohorts in two states, Georgia and Texas. Georgia enacted ELA between the two cohorts while neither cohort was ever treated in Texas. Suppose also that we wanted to know the effect of ELA in late adolescence on the earnings women received in their 20s and a separate effect in their 40s. Assume for simplicity that we observe each cohort twice—once in their 20s and once in their 40s. Following BHM, we would estimate

[4] Dealing with heterogeneity in treatment effects would complicate comparisons to the existing research. See Callaway and Sant'Anna (2019) for an approach to DD estimation in the presence of heterogeneous treatment effects. Readers familiar with recent advancements in econometric theory will be aware that the many-groups, many-time-periods DD model does not always estimate the average treatment effect. Rather, the coefficient of interest is a weighted average of every possible 2×2 DD comparison, and the weights can sometimes even be negative (Goodman-Bacon 2020). Dealing with this complication is beyond the scope of this paper.

(with some additional controls)

$$
\begin{aligned}
Y_{abs} = & \alpha_{20} \times I\left(a=20\right) + \delta_{20} I\left(a=20\right) \times Post_b \times GA_s + \\
& \alpha_{40} \times I\left(a=40\right) + \delta_{40} I\left(a=40\right) \times Post_b \times GA_s + \beta Post_b + \gamma GA_s + \varepsilon_{abs}
\end{aligned}
\tag{4}
$$

where Y_{abs} is the mean earnings of women in birth cohort b and state s at age a, $Post$ is an indicator for the second cohort and GA is an indicator for Georgia, the state that enacted ELA.

If Eq. (4) gives the actual data generating process, then the expected value of the change in average earnings between the two cohorts at age 20 in Georgia is $\beta + \delta_{20}$ and at age 40 is $\beta + \delta_{40}$. However, according to the restricted model, β is the change in earnings between cohorts at both age 20 *and* age 40 in Texas. In other words, the restricted model imposes that it does not matter which Texas age group we use as a comparison for either Georgia age group. We could use any linear combination (with weights summing to 1) of the changes by age group in Texas because

$$
\begin{aligned}
& E\left[Y_{TX,Post,20} - Y_{TX,Pre,20}\right] \\
& = E\left[Y_{TX,Post,40} - Y_{TX,Pre,40}\right] \\
& = E\left[\theta\left(Y_{TX,Post,20} - Y_{TX,Pre,20}\right) + (1-\theta)\left(Y_{TX,Post,40} - Y_{TX,Pre,40}\right)\right]
\end{aligned}
\tag{5}
$$

for any real-numbered θ.

The OLS estimate of δ_{20} in Eq. (4) is given by (proof in the online appendix)

$$
\begin{aligned}
\widehat{\delta_{20}} = & Y_{GA,Post,20} - \left(\frac{2}{3} Y_{GA,Pre,20} + \frac{1}{3} Y_{GA,Pre,40}\right) \\
& - \left[\frac{2}{3} Y_{TX,Post,20} + \frac{1}{3} Y_{TX,Post,40} - \left(\frac{1}{3} Y_{TX,Pre,20} + \frac{2}{3} Y_{TX,Pre,40}\right)\right]
\end{aligned}
\tag{6}
$$

This is a difference in *some* differences, but it is not the differences we might expect. Nor is the estimator intuitive like the standard 2×2 DD estimator for just people in their 20s:

$$\widehat{\delta_{20}^{2\times 2}} = Y_{GA,Post,20} - Y_{GA,Pre,20} - (Y_{TX,Post,20} - Y_{TX,Pre,20}) \tag{7}$$

We can put the OLS estimator of the restricted model $\widehat{\delta_{20}}$ in terms of the 2×2 DD estimator $\widehat{\delta_{20}^{2\times 2}}$ (proof in the online appendix):

$$\begin{aligned}\widehat{\delta_{20}} = \widehat{\delta_{20}^{2\times 2}} &+ \frac{1}{3}\left[(Y_{GA,Pre,20} - Y_{GA,Pre,40}) - (Y_{TX,Pre,20} - Y_{TX,Pre,40})\right] \\ &+ \frac{1}{3}\left[(Y_{TX,Post,20} - Y_{TX,Pre,20}) - (Y_{TX,Post,40} - Y_{TX,Pre,40})\right]\end{aligned} \tag{8}$$

If the parallel trends assumption holds within age groups, then $\widehat{\delta_{20}^{2\times 2DD}}$ will be unbiased by the usual arguments for DD. This implies that the OLS estimator of the restricted model will be unbiased if and only if

$$E\left[\begin{array}{cccc|c} & (Y_{GA,Pre,20} - Y_{GA,Pre,40}) & - & (Y_{TX,Pre,20} - Y_{TX,Pre,40}) & \\ + & (Y_{TX,Post,20} - Y_{TX,Pre,20}) & - & (Y_{TX,Post,40} - Y_{TX,Pre,40}) & \end{array} X\right] = 0 \tag{9}$$

where X is the matrix of regressors.

Each of the differences in differences in Eq. (9) shows a source of potential bias. The first term (on top) says that pre-treatment differences between age groups must be the same in the two states. The second term says that growth of the outcome variable in the comparison group must not differ by age. Both sources of bias occur regardless of whether the difference in trends (or initial levels) by age is exogenous or an effect of previous treatment. Neither source of bias depends on post-treatment outcomes for the treatment group.

For example, if educational achievement were increasing between subsequent birth cohorts, we might expect earnings for people in their 20s to grow slowly between the two cohorts because the later cohort would forego wages in their 20s to accumulate human capital (e.g. Becker 1962). Earnings at later ages would grow faster between cohorts because later cohorts have more human capital. Thus, comparing the first difference between cohorts in Georgia at age 40 to the first difference between cohorts in Texas at age 20 will indicate a positive effect of ELA even if the actual effect of ELA is zero. This bias occurs even if pill access is the reason for increasing educational attainment.

The second condition in Eq. (9) is violated in the ELA literature. Figure 1 shows average real earnings in seven age groups for women born from 1943 to 1953 in the National Longitudinal Survey of Young Women (NLSYW) and for women born from 1935 to 1960 in the Annual Social and Economic Supplement to the Current Population Survey (CPS).[5] Later cohorts earn more in middle age than earlier cohorts did, but later cohorts tend to earn about the same as earlier cohorts when they are young (and less than previous cohorts when under age 20). The growth in earnings is nearly monotonically related to age.

The differences in growth of earnings by age in Figure 1 induce bias in the restricted model estimates. The rate of growth of earnings at age 20 should be irrelevant to how we estimate effects of ELA on earnings at age 40, but the comparison group for the effect of ELA at age 40 in the restricted model is a weighted average of the earnings growth at all ages—including the less-positive growth in earnings at age 20. Thus, the effect of ELA at age 40 is overstated. Annual earnings for women at age 40 increased over birth cohorts by about $600 per year of birth, but the restricted model assumes that those earnings actually grew much slower. Similarly, the estimated effect of ELA at young ages suffers from negative bias. BHM's estimates of effects of ELA are negative at younger ages and positive at older ages—exactly the pattern we should expect from the bias imposed by the model restriction.

The problem is less severe in the case of labor force participation, as seen in Figure 2. Unlike with earnings, the relationship between growth in labor force participation and age is not monotonic. Labor force participation grew fastest for people in their late 20s and early 30s. Thus, we should expect the restricted model to overestimate effects at those ages and underestimate effects at other (particularly younger) ages. Again, positive effects of ELA

[5]Synthetic birth year for the CPS data is the sample year minus respondent's year of age.

on labor force participation for women in their late 20s and early 30s and negative effects at younger ages are exactly the pattern of estimates reported by Bailey.

3 ELA and labor force participation

Bailey uses March CPS data to estimate a restricted probit DD model[6] of effects of ELA on women's labor force participation rates for six age groups[7]. The estimates suggest that early legal access to the pill induced women in their late 20s and early 30s to participate more in the labor force. Results from Lindo et al. (2020) are similar.

In this section, I attempt to replicate the sample used by Bailey and then report three sets of estimates:

- Restricted model estimates
- Unrestricted model estimates
- Estimates of both models using randomly assigned placebo ELA

The results show that apparent effects of placebos are indistinguishable from apparent effects of actual ELA using the restricted model and that the unrestricted model estimates give no evidence of effects of ELA even though the spreads of sampling distributions of coefficient estimates are barely larger than with the restricted model.

3.1 Data for labor force participation estimates

Like Bailey, I use microdata from the 1964–2001 Annual Social and Economic Supplement to the Current Population Survey (CPS) (Flood et al. 2020). I construct synthetic birth year by subtracting the respondent's age from the survey year and limit the sample to civilian

[6] In addition to the state and birth cohort dummies (as in BHM), Bailey also controls for sample year dummies, which I follow in the initial estimates and later omit.

[7] Bailey (2006) reports the marginal effects for only five of the age groups included in the regression—possibly because there should be no effect on the youngest age group (16–20).

women aged 16–44[8] for whom the synthetic birth year is between 1935 and 1960 (inclusive) and for whom labor force participation status is not attributed. I use the ELA measure reported by Bailey.[9] I follow Bailey in estimating effects at each of the age groups 16–20, 21–25, 26–30, 31–35, 36–40, and 41–45. Bailey omits reporting estimates of marginal effects for the youngest age group, but I report estimates of marginal effects at all ages in the sample.

Unfortunately, I am not able to reproduce Bailey's exact estimates. The differences may be due to differences between the Unicon files used by Bailey (which Unicon no longer distributes) and the IPUMS files that I use. Following Bailey's sample selection rules,[10] I end up with 744,118 observations with positive sample weights, slightly more than Bailey's 733,419. However, the estimates are similar with the exception of the effect for the age 21–25 group.[11] For this group, Bailey's point estimate for the marginal effect "evaluated at the mean" is 0.003, but my estimate of the discrete marginal effect of ELA at the means of the non-Age, non-ELA variables is -0.0205. Using the means within age groups or using a linear probability model has little impact on the estimated effects (-0.0244 and -0.0238,

[8] I later extend this to ages 15–59 using later CPS samples.

[9] The CPS groups state of residence in many cases from 1968 through 1976. Thus, I follow Bailey in grouping the states into 21 consistently-identifiable groups and taking as treatment an estimate of the probability of being treated with ELA. For each state × birth year combination, ELA is 1 if the state enacted ELA before the year that is 21 years after the birth year (using the years of ELA reported by Bailey). ELA for a state *group* × birth year cell is a weighted average of those state × birth year ELA measures where each state's ELA is weighted by its population 20 years after the birth year divided by the total population of the state group 20 years after the birth year. The population estimates are from US Bureau of the Census, Population Distribution Branch (1995), US Bureau of the Census, Population Distribution Branch (1996), US Bureau of the Census, Population Distribution and Population Estimates Branches (1995), and US Bureau of the Census, Population Estimates Branch (1996).

[10] I understand the sample to be female respondents to the 1964 to 2001 Annual Social and Economic Supplement (ASEC) to the CPS who were age 16 to 44, for whom synthetic birth year (sample year minus age) is within the range [1935, 1960], who were not in the military, whose labor force participation status was not attributed, for whom state of residence (or state group) is observed, and who have positive sample weights. Bailey does not mention the last two conditions, but state is necessary to assign ELA treatment and Bailey's weighted regressions implicitly omit observations with sample weights of 0. Bailey also mentions omitting institutionalized observations, but the CPS does not sample institutionalized persons. The ASEC includes an oversample from surrounding months, but there are too many of these observations to explain the difference. At one point in the article, Bailey describes the sample as ages 18 to 44, but that is likely a typo, and dropping ages 16 and 17 would lose far more observations than the difference between our sample sizes.

[11] And possibly for the 16–20 age group, which Bailey does not report effects for.

respectively). A reviewer stated that they had replicated Bailey's estimates, but I have no information about when that was or which source files they used.

Because it seems to no longer be possible to replicate Bailey's estimates and because the probit and linear probability model estimates are so similar (trivial relative to the differences between the restricted and unrestricted model estimates), I report only linear probability estimates for brevity and ease of exposition.

3.2 Placebo procedure

Estimating the model with placebo treatments shows the bias induced by the model restriction. If treatment was randomly assigned through a known mechanism, that mechanism can be reapplied arbitrarily-many times to produce placebo treatments.[12] Because these placebo legal changes never actually occurred, they should have no effect on the outcome variable. Thus, the distribution of estimated effects of the placebos should be centered at 0 for unbiased estimators. The empirical distributions of placebo effects are estimates of the sampling distribution of the statistic of interest δ_a under the null hypothesis that there is no effect (that $\delta_a = 0$ for all a). This strategy, sometimes called "randomization inference" (RI) can also be used to calculate p-values and confidence intervals, so I report RI-based p values for the unrestricted model estimates (when those p values are less than 0.2).

Specifically, I randomly permute the state-specific years of ELA from Bailey, construct placebo ELA with the permuted years, estimate the regressions, and then repeat the process 7,200 times times to construct sampling distributions for the regression coefficients (under the null hypothesis that there is no effect of ELA).

[12] The assignment mechanism in this case is not strictly known. However, even if the placebo law dates are not drawn from the exact same data generating process as the actual laws, the fact that the placebo laws generate similar patterns of effects as the actual laws when using the restricted model is strong evidence against the validity of that model.

3.3 Effect of ELA on labor force participation

Estimates of effects of ELA on labor force participation are in Fig. 3. Like Bailey, I find a pattern of increased labor force participation in women's late 20s and early 30s and decreased participation at earlier ages with the restricted model (top panel). These estimates are consistent with theoretical predictions that ELA would cause young women to increase formal human capital accumulation through schooling (Bailey 2006; Bailey, Hershbein, and Miller 2012; Goldin and Katz 2002; Hock 2008).

However, the unrestricted model (bottom panel of Fig. 3) does not show this pattern. The two models tell almost opposite stories. If there is any effect, it is likely a *negative* effect of ELA on labor force participation at ages 26–34.

The placebo effects in violin plots estimate the sampling densities of the coefficients under the null hypothesis that ELA has no effect on labor force participation. The unrestricted model (bottom) gives estimates of effects of placebo laws that are centered at zero, whereas the restricted model (top) generates patterns that look like meaningful effects even when using placebo laws that should have no effect, suggesting that the restricted model estimates suffer from substantial bias. Furthermore, placebo laws generate effects similar to actual ELA, and none of the estimates using actual dates of ELA is extreme relative to the sampling distribution of estimates with placebo laws. The data provide no evidence for effects of ELA on labor force participation.

It is not simply that the estimates are less precise when including the full set of interactions. Neither the confidence intervals using cluster-robust standard errors (allowing for error correlation at the state level) nor the dispersions of the placebo-based sampling densities differ much between the two models.

3.4 Effect of ELA on labor force participation without sample year effects

The regressions in Fig. 3 follow Bailey in controlling for sample year fixed effects in addition to age and birth cohort effects, but these year effects are not intuitive, and omitting them helps illustrate the instability of the restricted model. Bailey constructs synthetic birth year by subtracting age from the sample year, so *year* of age and synthetic birth year dummies would be perfectly collinear with sample year dummies by construction. Thus, the year dummies are picking up variation in labor force participation within the six age *groups*, which each contain five years of age.

Fig. 4 re-estimates the effects of ELA on labor force participation but omits the year dummies from the regressions. Interestingly, the year dummies seem to matter a lot to the restricted model but not to the unrestricted model. In the DD model for a single age group (Eq. (1)), each birth cohort is observed in at most five years (the width of the age group). Thus, the coefficients on the year dummy are based on the small amount of variation that occurs within cohorts over those five years. But in the restricted model, each birth cohort is observed in many age groups, which means their observations span many years. Because there is no cohort $\times$ age interaction term, the coefficients on year dummies are no longer based only on the within-cohort, within-age-group variation and can pick up more of the variation in labor force participation.

4 ELA and earnings with NLSYW data

BHM (and Lindo et al. 2020) estimate effects of ELA on annual wage and salary earnings and hourly wages using data from the National Longitudinal Survey of Young Women (NLSYW). Their results suggest that ELA caused women to earn more in middle age and less before their 30s.

In this section, I revisit their work with three sets of estimates:

- Restricted model estimates
- Unrestricted model estimates
- Estimates of both models using randomly assigned placebo ELA

The results again show that apparent effects of placebos are indistinguishable from apparent effects of actual ELA using the restricted model and that the unrestricted model estimates give no evidence of effects of ELA even though the standard errors are barely larger than with the restricted model.

4.1 Data for earnings estimates: NLSYW

I follow BHM in estimating effects of ELA on annual wage and salary earnings and hourly wages using data from the National Longitudinal Survey of Young Women (NLSYW). The first wave of the NLSYW surveyed women aged 14 to 24 in 1968 and interviewed them 21 more times through 2003. These women were born from 1940 to 1954, and BHM limit the sample to birth years 1943 through 1953. Wherever this study uses NLSYW data, the samples and variables come from replication code from BHM. Their dependent variables are real wage and salary measures deflated to year 2000 dollars using the Personal Consumption Expenditures price level (Bureau of Economic Analysis 2021). They include only observations with positive values of the dependent variable (before taking the logarithm). They split the sample into seven age groups: 14–19, 20–24, 25–29, 30–34, 35–39, 40–44, and 45–49.[13] They omit estimates of marginal effects for the youngest age group when reporting results, but I report estimates of marginal effects at all ages in the sample.

Unfortunately, I am unable to estimate the unrestricted model with the NLSYW data because geographic (state) identifiers are limited to use at Census Research Data Centers. However, this is a minor issue. This article makes two claims: (1) the restricted model

[13]The replication code provided by BHM includes women up to age 59. They did not respond to a request for clarification. I follow their published description rather than their code and drop observations over age 49.

estimates are uninformative and (2) the unrestricted model estimates show no evidence of effects. Estimates of the unrestricted model are irrelevant to point (1), which is established by proof of the condition for unbiasedness paired with a demonstration of empirical differences in baseline trends by age and a comparison of restricted model estimates to placebos. The conclusions from BHM and Bailey should only be believed if their methods support those conclusions. If the methods are uninformative, then there is no existing evidence for effects of ELA on earnings and labor force participation. Point (2) is where unrestricted model estimates are useful, but I *do* produce estimates of the unrestricted model using CPS data in the next section and find no evidence for effects.

It is also not obvious a priori which data source should be considered more compelling evidence. The NLSYW is better for determining state of residence in late adolescence than the CPS, but the CPS has the benefit of much larger sample sizes and a wider range of birth cohorts. Considering that Bailey uses the CPS and that BHM get remarkably similar estimates with the NLSYW,[14] there seems to be no reason to conclude that we must use the NLSYW to answer questions about effects on earnings.

4.2 Placebo procedure

The permutation test procedure is nearly identical to the one used for labor force participation estimates with CPS data in section 4.2. Ideally, I would randomly permute the state-specific years for ELA from BHM and then construct placebo ELA from those years. However, the public-use NLSYW data do not include state of residence, so I instead randomly assign each person a state (with equal probability for each state) and then construct placebo ELA as if they had lived in that state at age 20. I then estimate the regressions and repeat the procedure 5,400 times.

[14]BHM report estimates of effects on labor force participation in a footnote.

4.3 Effect of ELA on earnings and wages (NLSYW)

Estimates of effects of ELA on real hourly wages and real annual earnings at various ages using the NLSYW are in Fig. 5. The points replicate estimates from BHM using their base specification with abortion controls, which imposes the restriction that state and birth year fixed effects are common across age groups. The violin plots give kernel density estimates of estimated effects of placebo laws from the randomization inference procedure discussed in the previous subsection (randomly assigning persons to states).

BHM's estimates show small declines in incomes due to ELA for women in their 20s and large positive effects on incomes after age 30 (and before age 50).[15] These estimates are consistent with theoretical predictions that ELA would cause young women to increase human capital accumulation (Bailey 2006; Bailey, Hershbein, and Miller 2012; Goldin and Katz 2002; Hock 2008) and that such investments require foregone earnings early in life that pay off later (Becker 1962; Mincer 1974).[16]

However, none of the estimates from BHM is extreme relative to the distributions of estimates of effects of placebos. When accounting for bias due to the model restriction, the data do not support the conclusions from BHM. Furthermore, the apparent effects of placebo ELA suggest that the model restriction induces substantial bias.

[15] Lindo et al. (2020) extend this work to later ages and find no effect after age 50, which could suggest that the highest-earning women treated with ELA were more likely to retire early. However, that interpretation is only reasonable if there really are effects at younger ages. Furthermore, while Lindo et al. (2020) also impose that the fixed effects are shared across ages, they use data from the Health and Retirement Study, which samples only people over age 50 and their spouses. Because the bias in the restricted model is a result of pooling ages and because their work does not pool a wide age range, their estimates likely contain little of the sort of bias addressed in this research.

[16] In some cases, the shape of work profiles over the life cycle can be compared to the shape of earnings profiles to discriminate between human capital accumulation through schooling and through on-the-job training (or learning by doing). Both sources of human capital accumulation typically increase the slope of age-earnings profiles (Becker 1962), but schooling takes time away from work in early life, while on-the-job training does not. The results on labor force participation later in the paper deal with whether ELA reduced work hours early in life.

5 ELA and earnings with CPS data

Because I cannot estimate the unrestricted model using the NLSYW data, I augment these results with estimates of effects on annual earnings using CPS data. The placebo procedure is identical to the one used above with NLSYW data except that I now leave respondents in their state of residence and randomly-permute dates of ELA across states.

5.1 Data for earnings estimates: CPS

I use microdata from the 1963–2009 Annual Social and Economic Supplement to the CPS (Flood et al. 2020).[17] I construct synthetic birth year by subtracting the respondent's age from the survey year and limit the sample to women aged 14–49 for whom the synthetic birth year is between 1935 and 1960 (inclusive). This sample is nearly identical to Bailey's; the primary differences are that I use a broader range of survey years to observe later cohorts at all ages and that I use the age range used by BHM.

The dependent variable is wage and salary earnings over the previous year (or its logarithm) deflated by the personal consumption expenditures price index (Bureau of Economic Analysis 2021) with 2000 as the base year. Because of the difficulty in determining labor earnings for the self-employed, I ignore self-employment earnings and exclude from the sample women who report that earnings at the job they worked at the longest over the previous year were not from wages or salary. The sample includes only observations with positive earnings for consistency with BHM.

5.2 Effect of ELA on earnings and wages (CPS)

Estimates of effects of ELA on real hourly wages and real annual earnings using CPS data are in Fig. 6. The figure includes estimates of both the restricted and unrestricted model. Like when using NLSYW data, the restricted model estimates show a pattern of decreased

[17] I group states using the same procedure described in the section on labor force participation.

earnings at young ages and increased earnings in the 30s and 40s. Women who could obtain the pill without parental consent in late adolescence appear to earn thousands of dollars more per year in middle age than if they could not.

However, the data do not support that conclusion. Using the restricted model with placebo ELA (the violin plots) suggests the same pattern of effects as actual ELA. The restricted-model estimates for younger ages are centered thousands of dollars below zero and the estimates for older ages are centered thousands of dollars above zero. Thus, analyses that uses Eq. (3) will severely over-reject the null of no effect. Additionally, none of the point estimates of effects of actual ELA is extreme relative to the apparent effects of placebos ($p > 0.12$ for every age).

The unrestricted model shows no evidence of effects. Unlike the restricted model, the densities for placebo estimates using the unrestricted model are centered at 0. Furthermore, adding the full set of interactions barely inflates the spread of the densities in Fig. 6 (and even reduces standard errors at younger ages), so we are not losing much by including them in this case. Once again, the data do not support the conclusions from BHM.

6 Discussion

This article revisits past research on effects of 1960s and 1970s legal changes in the US allowing young people to consent to medical treatments before age 21 and other laws that may have lowered the cost of obtaining oral contraceptives. Past research suggests that these legal changes increased earnings, hourly wages, and labor force participation for women after their 20s. This article presents evidence that the data do not support those conclusions. Past research uses difference-in-differences regressions to estimate effects at multiple ages but does not allow for differences in state fixed effects by age or (more importantly) differences in birth cohort fixed effects by age. This model restriction violates the parallel trends condition and thus induces omitted variable bias.

This article establishes six important empirical facts:

- Earnings and labor force participation grew faster from cohort to cohort at some ages than at other ages, exactly what the model restriction assumes is not true
- Apparent effects of randomly-assigned placebo ELA using the restricted model are not centered at 0
- None of the estimates of effects of actual ELA using the restricted model is extreme relative to placebo effects
- Apparent effects of placebo ELA using the *un*restricted model *are* centered at 0
- Estimates with the unrestricted model show no evidence of effects
- The precision in unrestricted model estimates is almost as good as in restricted model estimates

These facts suggest that there is no evidence that ELA caused women to earn more or participate more in the labor force at any age and that past estimates of effects are entirely due to omitted variable bias caused by a model restriction.[18]

If the evidence in this work is convincing, there are two ways to interpret that evidence: (1) ELA is a valid instrument for costs of obtaining oral contraceptives but has not yielded evidence for effects of those costs on earnings and labor force participation or (2) ELA is not a valid instrument. It may be tempting for readers who previously accepted the validity of the ELA methodology (like I did) to now reject it, but nothing in this research undermines the arguments that have been made in its favor in past research. Rather, this research shows that there is no evidence that ELA had the effects that past research suggests it had.

[18]These results do not constitute strong evidence against effects of ELA on earnings and labor force participation. The tests may lack sufficient power to identify effects or measurement error might attenuate estimates. However, this does not justify using estimates of the restricted model.

References

Bailey, Martha (2006). "More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply". In: *The Quarterly journal of Economics* 121(1).

Bailey, Martha, Brad Hershbein, and Amalia R. Miller (2012). "The Opt-In Revolution? Contraception and the Gender Gap in Wages". In: *American Economic journal: Applied Economics* 4(3).

Becker, Gary S. (Oct. 1962). "Investment in Human Capital: A Theoretical Analysis". In: *Journal of Political Economy* 70.5, pp. 9–49. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/258724. URL: https://www.journals.uchicago.edu/doi/10.1086/258724 (visited on 11/02/2019).

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (Feb. 1, 2004). "How Much Should We Trust Differences-In-Differences Estimates?" In: *The Quarterly Journal of Economics* 119.1, pp. 249–275. ISSN: 0033-5533. DOI: 10.1162/003355304772839588. URL: https://academic.oup.com/qje/article/119/1/249/1876068 (visited on 04/27/2019).

Bureau of Economic Analysis (June 24, 2021). *National Income and Product Accounts, Table 1.1.4. Price Indexes for Gross Domestic Product*. URL: https://apps.bea.gov/iTable/iTable.cfm?reqid=19&step=3&isuri=1&nipa_table_list=4&categories=survey (visited on 07/13/2021).

Callaway, Brantly and Pedro H. C. Sant'Anna (Mar. 1, 2019). *Difference-in-Differences with Multiple Time Periods*. SSRN Scholarly Paper ID 3148250. Rochester, NY: Social Science Research Network. DOI: 10.2139/ssrn.3148250. URL: https://papers.ssrn.com/abstract=3148250 (visited on 01/09/2021).

Conley, Timothy G. and Christopher R. Taber (Jan. 21, 2011). *Inference with "Difference in Differences" with a Small Number of Policy Changes*. http://dx.doi.org/10.1162/REST_a_00049. DOI: 10.1162/REST_a_00049. URL: https://www.mitpressjournals.org/doix/abs/10.1162/REST_a_00049 (visited on 08/02/2019).

Ferman, Bruno and Cristine Pinto (July 16, 2018). “Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity”. In: *The Review of Economics and Statistics*, pp. 1–16. ISSN: 0034-6535, 1530-9142. DOI: 10.1162/rest_a_00759. URL: https://www.mitpressjournals.org/doi/abs/10.1162/rest_a_00759 (visited on 05/22/2019).

Flood, Sarah et al. (2020). *Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset]*. Minneapolis, MN: IPUMS, 2020. URL: https://doi.org/10.18128/D030.V7.0.

Goldin, Claudia and Lawrence F. Katz (Aug. 2002). “The Power of the Pill: Oral Contraceptives and Women’s Career and Marriage Decisions”. In: *Journal of Political Economy* 110.4, pp. 730–770. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/340778. URL: https://www.journals.uchicago.edu/doi/10.1086/340778 (visited on 05/27/2019).

Goodman-Bacon, Andrew (Aug. 2020). *Difference-in-Differences with Variation in Treatment Timing*. Working Paper Series. National Bureau of Economic Research.

Hock, Heinrich (2008). “The Pill and the College Attainment of American Women and Men”. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.1023042. URL: http://www.ssrn.com/abstract=1023042 (visited on 05/27/2019).

Lindo, Jason M. et al. (May 1, 2020). “Legal Access to Reproductive Control Technology, Women’s Education, and Earnings Approaching Retirement”. In: *AEA Papers and Proceedings* 110, pp. 231–235. ISSN: 2574-0768, 2574-0776. DOI: 10.1257/pandp.20201108. URL: https://pubs.aeaweb.org/doi/10.1257/pandp.20201108 (visited on 05/27/2021).

MacKinnon, James G. and Matthew D. Webb (Mar. 2017). “Wild Bootstrap Inference for Wildly Different Cluster Sizes”. In: *Journal of Applied Econometrics* 32.2, pp. 233–254. ISSN: 0883-7252, 1099-1255. DOI: 10.1002/jae.2508. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2508 (visited on 09/11/2019).

— (June 1, 2018). “The wild bootstrap for few (treated) clusters”. In: *The Econometrics Journal* 21.2, pp. 114–135. ISSN: 1368-4221, 1368-423X. DOI: 10.1111/ectj.12107. URL:

https://academic.oup.com/ectj/article/21/2/114-135/5078969 (visited on 05/22/2019).

Mincer, Jacob (1974). *Schooling, experience, and earnings.* Human behavior and social institutions 2. New York: National Bureau of Economic Research; distributed by Columbia University Press. 152 pp. ISBN: 978-0-87014-265-9.

Myers, Caitlin Knowles (Dec. 2017). "The Power of Abortion Policy: Reexamining the Effects of Young Women's Access to Reproductive Control". In: *Journal of Political Economy* 125.6, pp. 2178–2224. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/694293. URL: https://www.journals.uchicago.edu/doi/10.1086/694293 (visited on 05/27/2019).

Stephen G. Donald and Kevin Lang (2007). "Inference with Difference-in-Differences and Other Panel Data". In: *The Review of Economics and Statistics* 89.2, p. 221. ISSN: 00346535. URL: http://ezproxy.bsc.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsjsr&AN=edsjsr.40043055&site=eds-live (visited on 09/11/2019).

US Bureau of the Census, Population Distribution and Population Estimates Branches (Feb. 1995). *Intercensal Estimates of the Total Resident Population of States: 1970 to 1980.* Consistent with Current Population Reports, Series P-25, No. 957, issued October 1984. Data for 4/1/70 and 4/1/80 are census counts. Population data for remaining years are intercensal estimates as of each July 1. Washington, DC. URL: https://www2.census.gov/programs-surveys/popest/tables/1980-1990/state/asrh/st7080ts.txt (visited on 07/16/2021).

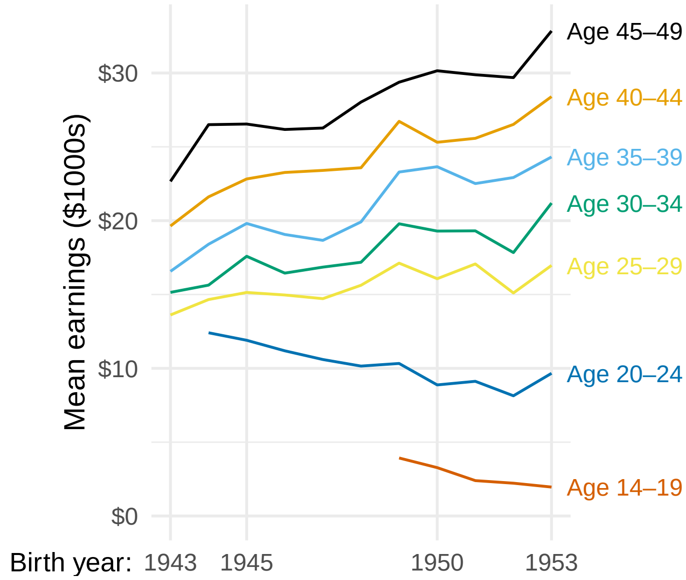
US Bureau of the Census, Population Distribution Branch (Apr. 1995). *Intercensal Estimates of the Total Resident Population of States: 1950 to 1960.* The 1950-60 estimates are consistent with those published in Current Population Reports Series P25-304 issued 4/65. P25-304 also describes the methodology used to derive the estimates. Washington, DC. URL: https://www2.census.gov/programs-surveys/popest/tables/1980-1990/state/asrh/st5060ts.txt (visited on 07/16/2021).

US Bureau of the Census, Population Distribution Branch (Aug. 1996). *Intercensal Estimates of the Total Resident Population of States: 1960 to 1970*. The 1960-70 estimates are consistent with those published in Current Population Reports Series P-25, No. 460. P-25 No. 460 also describes the methodology used to derive the estimates. To obtain a copy call (301)457-2422. Washington, DC. URL: `https://www2.census.gov/programs-surveys/popest/tables/1980-1990/state/asrh/st6070ts.txt` (visited on 07/16/2021).

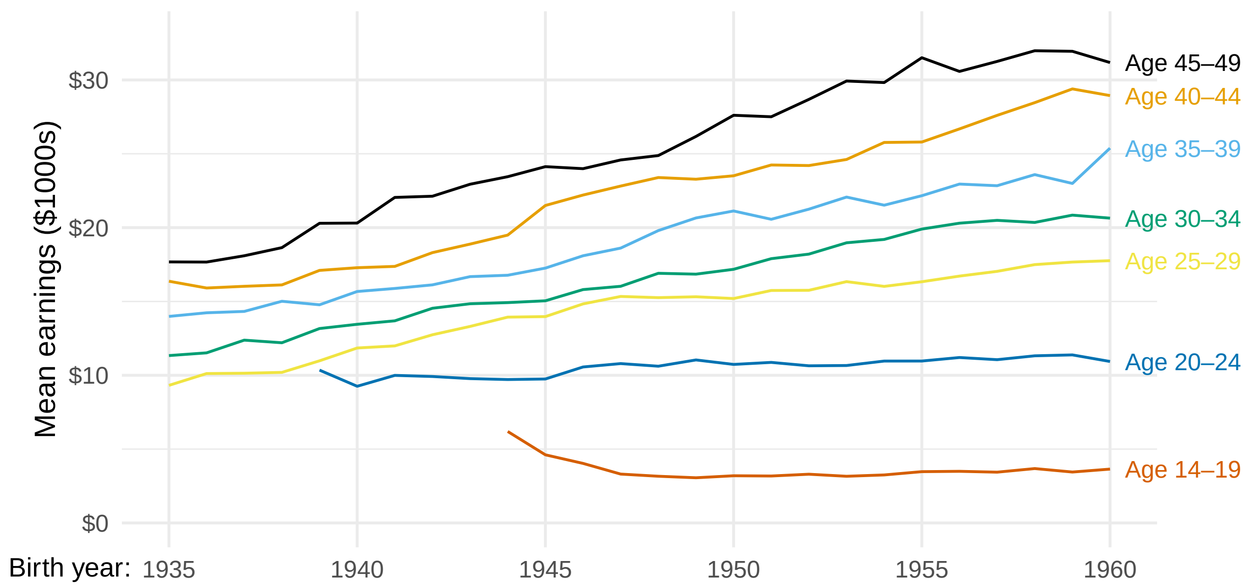
US Bureau of the Census, Population Estimates Branch (Aug. 1996). *Intercensal Estimates of the Total Resident Population of States: 1980 to 1990*. All data are consistent with the intercensal estimates shown in Table 2 of CURRENT POPULATION REPORTS Series, P25-1106, issued 11/93. Washington, DC. URL: `https://www2.census.gov/programs-surveys/popest/tables/1980-1990/state/asrh/st8090ts.txt` (visited on 07/16/2021).

Yzerbyt, Vincent Y., Dominique Muller, and Charles M. Judd (May 1, 2004). "Adjusting researchers' approach to adjustment: On the use of covariates when testing interactions". In: *Journal of Experimental Social Psychology* 40.3, pp. 424–431. ISSN: 0022-1031. DOI: 10.1016/j.jesp.2003.10.001. URL: `http://www.sciencedirect.com/science/article/pii/S0022103103001598` (visited on 04/04/2020).

Figures



(a) NLSYW



(b) CPS

Figure 1: Average annual wage and salary earnings by birth cohort and age

NLSYW sample (a): women from the National Longitudinal Survey of Young Women that were born from 1943 through 1954, had positive wage or salary earnings in the last year, and were age 14–49 at the time of observation.

CPS sample (b): civilian women in the 1963–2009 Annual Social and Economic Supplement to the Current Population Survey that were born from 1935 through 1960, were age 14–49 at the time of observation, had positive wage or salary earnings in the last year, and reported that the job they worked at the longest in the previous year paid wages or salary. Synthetic birth year is the sample year minus respondent's year of age. The data include only wage and salary earnings that are not from self-employment.

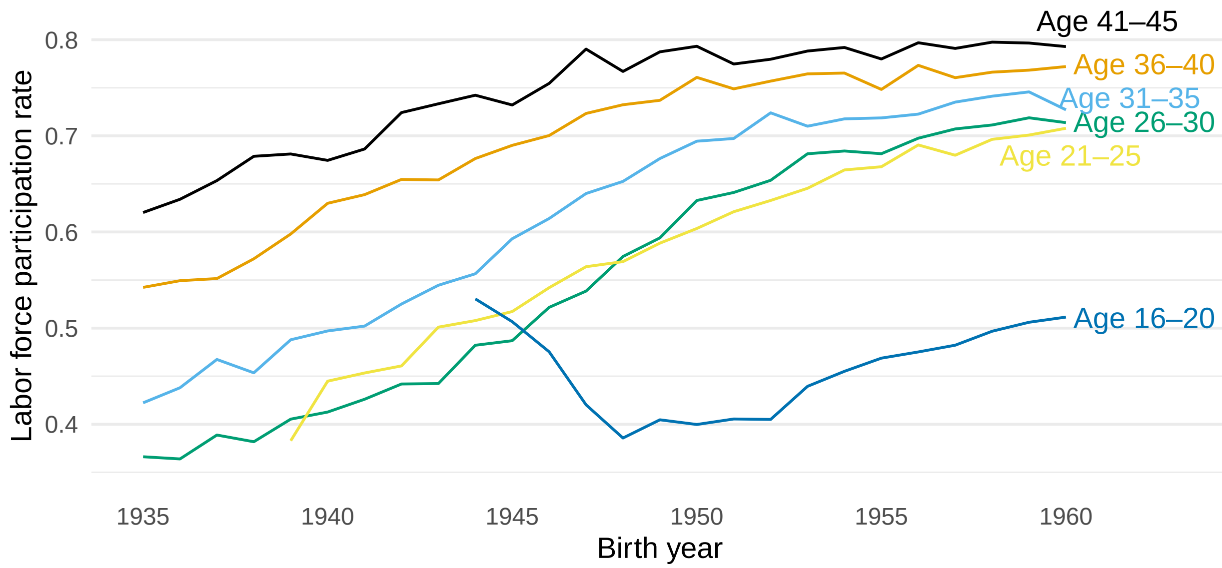


Figure 2: Labor force participation rate by birth cohort and age

Data: civilian women age 16–45 in the 1964–2001 Annual Social and Economic Supplement to the Current Population Survey that were born from 1935 through 1960 and whose labor force participation status is not attributed. Synthetic birth year is the sample year minus respondent's year of age.

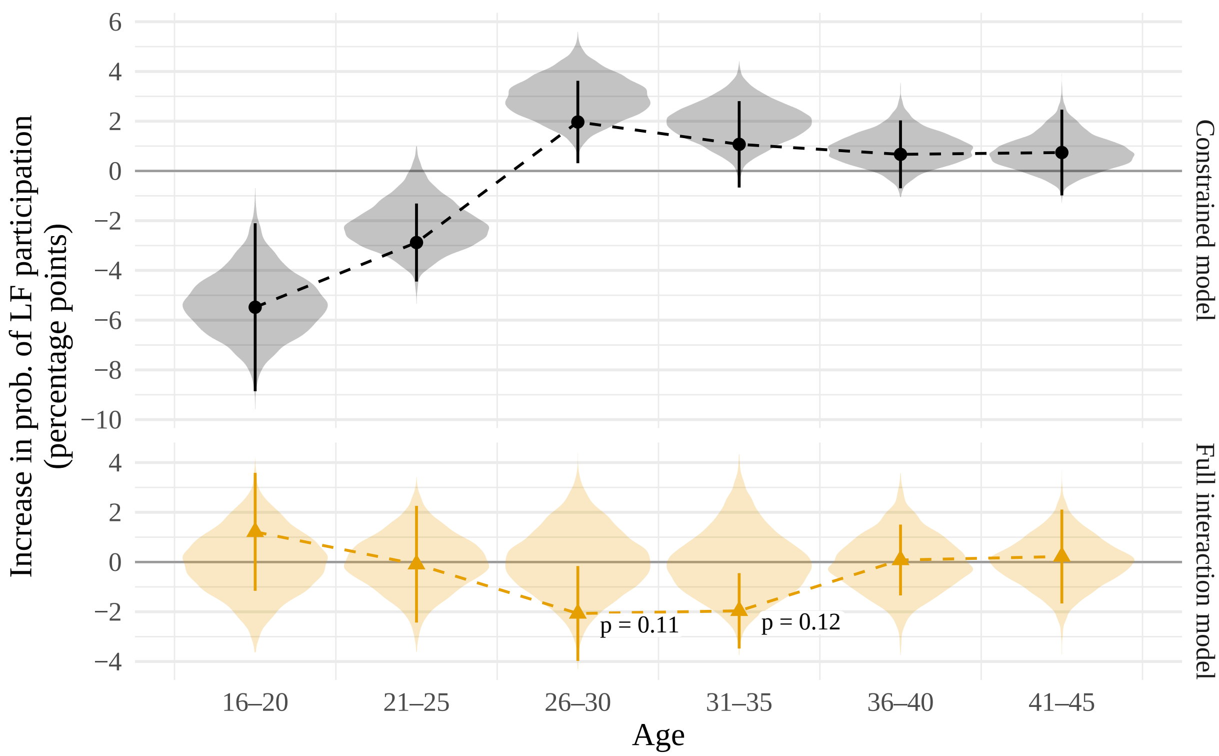


Figure 3: Effect of early legal access to oral contraceptives on female labor force participation by age

Violin plots are kernel density estimates of over 1,200 estimates (of each model) of placebo laws from random permutations of state-specific dates of ELA.

Data: 1964–2001 March CPS microdata for women aged 16 to 45 who were born from 1935 through 1960 and for whom labor force participation is not attributed.

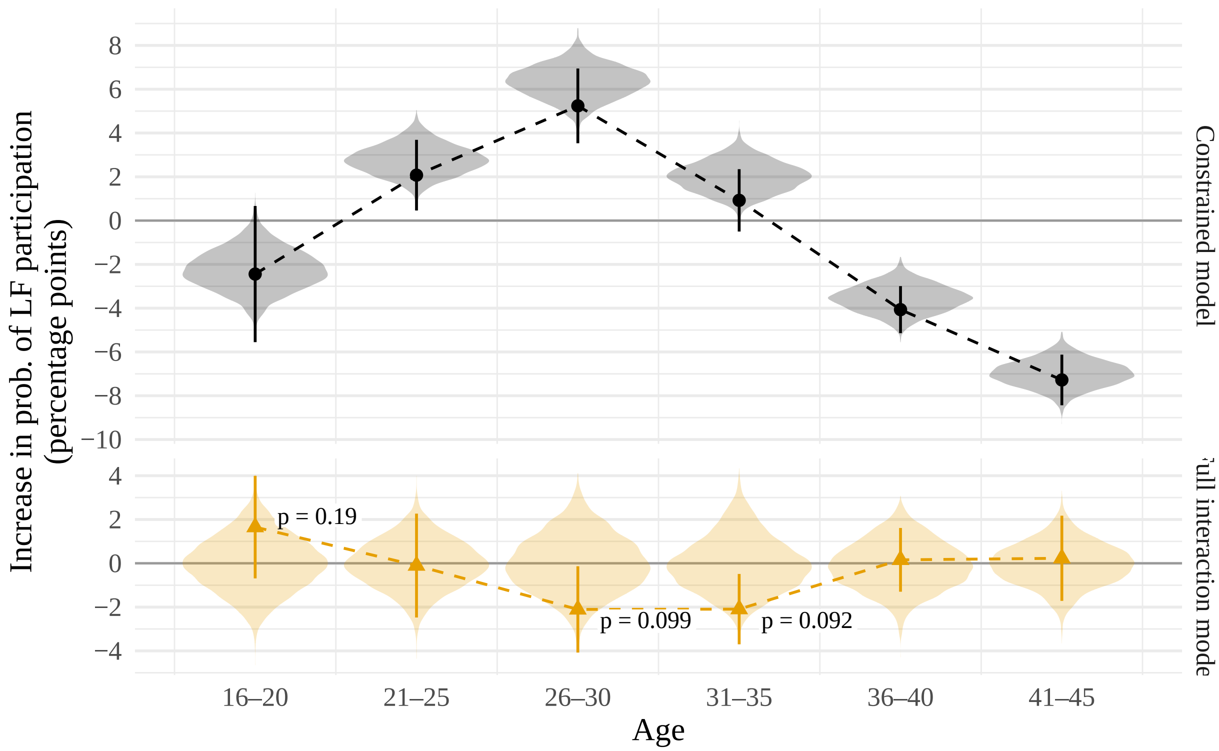


Figure 4: Estimated effect of early legal access to oral contraceptives on female labor force participation by age without sample year indicators

Data: 1964–2001 March CPS microdata for women aged 16 to 45 who were born from 1935 through 1960 and for whom labor force participation is not attributed.

Figure 5: Effect of early legal access (ELA) to oral contraceptives on hourly wages or annual wage and salary earnings by age

Points are estimates of the restricted model using actual state of residence. Vertical bars are 90% confidence intervals constructed from standard errors allowing for error correlation within states. Violin plots are kernel density estimates of 3,600 regressions with placebo ELA from randomly assigning persons to states. Wages and earnings are deflated to year 2000 dollars using the Personal Consumption Expenditures price index (using code and source files from BHM). Data: National Longitudinal Survey of Young Women (original cohort). Sample: ages 14–49 with positive values of the dependent variable (before taking the logarithm). All regressions include controls for early abortion access interacted with age group dummies; a three-way interaction between abortion access, ELA, and age dummies; sets of indicators for age group, state of residence at age 20, and birth year; and the logarithm of distance to the nearest abortion provider.

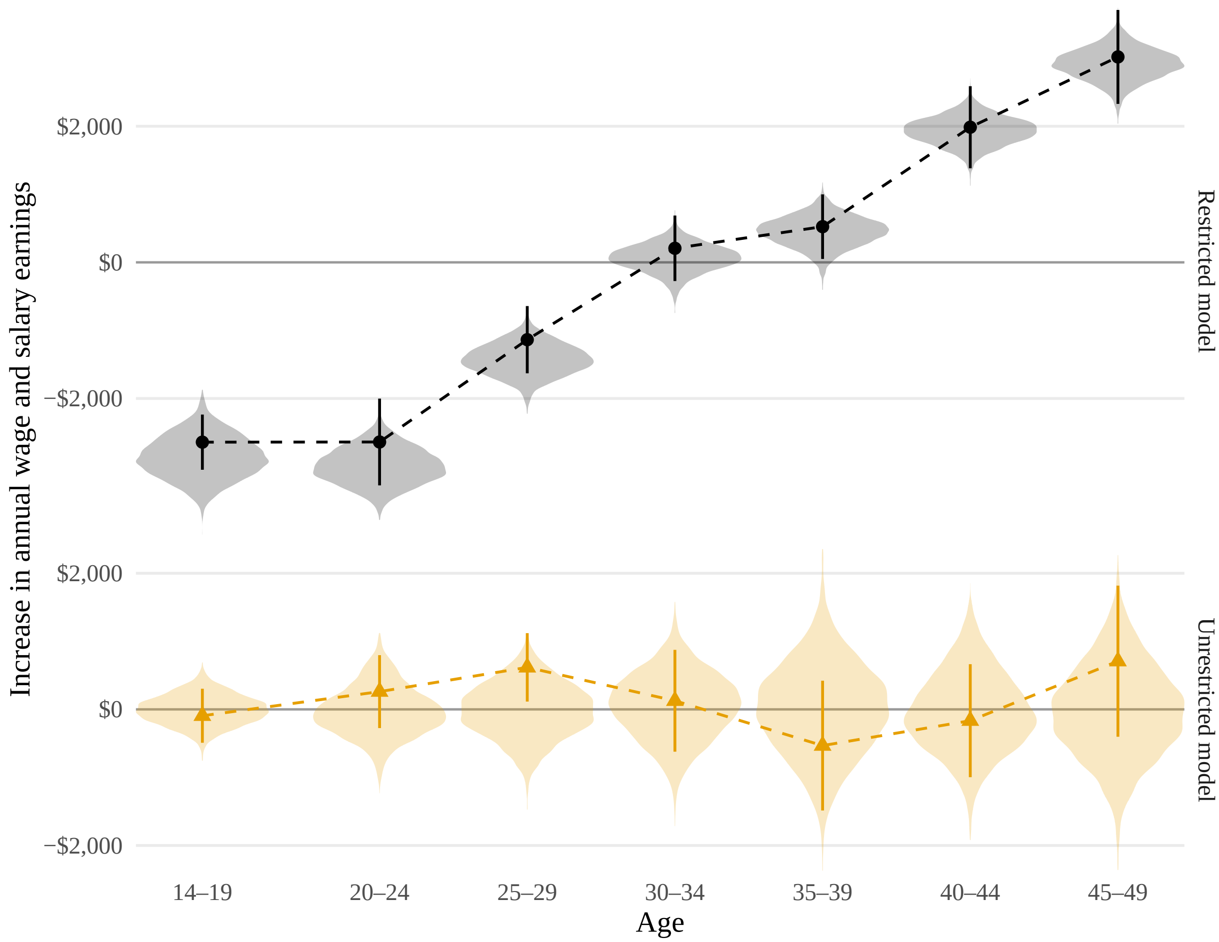


Figure 6: Effect of early legal access (ELA) to oral contraceptives on annual earnings by age

Vertical bars are 95% confidence intervals based on standard errors correcting for heteroskedasticity and error correlation within states. Violin plots are densities of estimates of effects of placebo (randomly-permuted) ELA on earnings.

Data: 1963–2009 CPS ASEC microdata (IPUMS) for women aged 14 to 49 who were born from 1935 through 1960 and had positive earnings in the last year.