The Impact of Abortion on the Long-Term Earnings Trajectory of Working Women in the United States

Yoony Lee*

4/28/2023

Department of Economics
University of California, Berkeley

Abstract

This paper investigates the immediate and residual (delayed) effects of abortion on the earnings trajectories of working women in the United States. Using the National Longitudinal Survey of Youth 1979 data set, I construct a 10-year panel data of working women born between 1957 to 1964 and run multiple Difference-in-Differences regressions to analyze the effects of abortion on their earnings over a 10-year period. I find that women who received an abortion had consistently higher earnings compared to a similar group that did not, and that the gap remained roughly constant over time. The racial wage differential between Black and White women is pronounced in these data, and my results suggest some of the inequality could be attributable to unequal access to abortion. I also examine the Net Present Value of getting an abortion. These findings suggest the presence of persistent racial wealth gaps, which I hope to address in future research by analyzing inequities in reproductive healthcare access and the industrial organization of other institutional structures at large.

*This work would not have been possible without my advisor, Professor Ryan D. Edwards, who provided unrelenting guidance and moral support. I also thank Professor of Law Emerita Kristin Luker for her counsel as my URAP mentor, whose life's work inspired me to pursue this topic. All errors and shortcomings are mine.

1 Introduction

The recent overturn of Roe v. Wade upended conditions of reproductive access across the United States, invoking the question – how does restricting and eliminating access to abortions impact women's lives? The issue of abortion and its implications for women have attained significant prominence and garnered extensive public attention in recent years. In the early 20th century, attempts to answer the question were motivated in part by growing economic interest in population control, the legitimization of obstetrics and gynecological studies in the medical field, and the gradual recognition of women's rights worldwide. In this paper, I explore the relationship between abortion and women's socioeconomic outcomes not solely through the lens of supply (abortion access), but through measurable and quantifiable factors influencing a woman's utility of obtaining an abortion.

Using the National Longitudinal Survey of Youth (NSLY79) cohort of working women born between the years 1957 to 1964, I seek to quantify the effects of abortion on a working woman's average income over time. The empirical methodology involves using a differences-in-differences regression approach and an application of the Ordinary-Least Squares results to obtain the net present value using the dimension, log(wage). I restrict the model to working women between the ages of 15-27 years, stratifying the sample into control and treatment groups characterized by whether they have given birth to a live child or have undergone an abortion and have not raised a child, respectively.

I explore the following related questions:

- How does getting an abortion affect a woman's earnings trajectory?
- What other considerations, such as the cost of an abortion and maternity leave, should be included to analyze present trends?
- Does race influence the "treatment effect" of abortion?

Assumptions used in previous literature that treated wage trajectories the same across groups of different demographics (age, ethnicity, educational attainment) used to investigate the relationship between total income and abortion also stand to be tested.

The rest of the paper proceeds in the following order. In section one, I review past macroeconomic and micro-economic literature on fertility, abortion, and women's socioeconomic performance, noting the limitations and implications of each that inform my approach in this paper. Section two is a descriptive overview of the NSLY79 data. Section three provides a comprehensive summary of my preparation of the NSLY79 data as panel data for a Difference-in-difference approach and describes my data preparation process and the results using this data. Section four describes the conceptual framework for calculating the net present value of obtaining an abortion for working women using results from the previous section. Section five summarizes my findings, and notes the limitations of the data and the modeling strategy.

2 Historical Context and Literature Review

Approaches to previous literature concerning this topic have varied over time and have often been subject to a variety of social, political, and ideological norms characterizing the time period in which they were conducted. In the early 20th century, there was little formal research on the effects of abortion, as the procedure was conducted largely covertly and underground. Historical anecdotal evidence and medical reports suggest that illegal abortions were often unsafe and associated with high rates of morbidity and mortality. Eventually, the gradual decriminalization of abortion within limited state jurisdictions in the latter half of the 20th century introduced pathways for research to conduct empirical studies examining the effects of abortion on women's health and well-being.

Early studies focused on medical complications from abortions, such as infections, internal hemorrhaging, and injury to reproductive organs. Subsequent research expanded to consider the psychological effects of abortion, including depression, anxiety, and other mental health concerns. In recent years, research has increasingly focused on the social and economic effects of abortion, including its impact on women's educational and career opportunities, family relationships, and financial stability. Overall, trends in research on the effects of abortion on women in the United States have shifted from a focus on physical complications to a broader consideration of the social, psychological, and economic impacts of the procedure. The general consensus is that access to safe and legal abortion is a critical component of women's health and reproductive rights, but equal access remains a persistent public policy issue.

Macroeconomic Approach

The establishment of Princeton's Office of Population Research in 1936 sparked population research initiatives in the United States. Several presidents, including Lyndon B. Johnson, Richard Nixon, and George W. Bush explicitly pushed for the extension of family planning services and policies regarding population control. Other countries quickly followed suit during the twentieth century by undertaking their own rendition of population control initiatives. Historically, one clear rationale for pushing for abortion and family planning legislative-related reform was empowering women to achieve desired family size, which presumably would expand female labor force participation (LFP) and thus increase national output while enhancing female earnings.

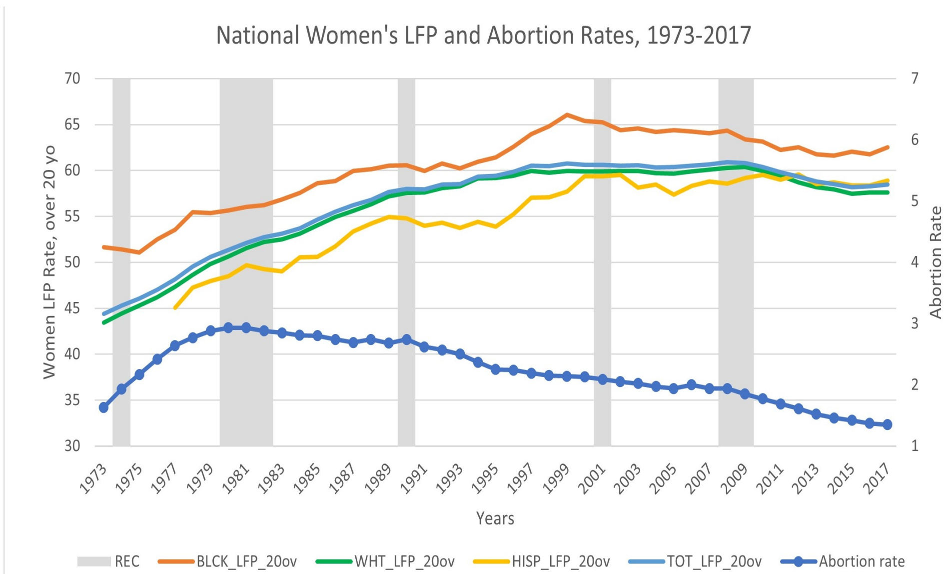
Past relevant economic literature relevant explored the relationship between pro-abortion legislation and women's formal economic participation as well as economic mobility. These methods explored hypotheses largely reliant on the indirect impacts of supply-based, federal policies restricting abortion access. In "Fertility, Female Labor Force Participation, and the Demographic Dividend", Bloom et. al. (2007), used cross-country panel data methods. They found that increasing levels of fertility decreased female labor force participation. Their model assumes Cobb-Douglas production and different productivity, capital, and age structures across countries. Under these assumptions, Bloom et al. discovered that decreasing fertility reduces population growth and shifts the average age of the working population, and increases the capital-to-labor ratio. Additionally, their findings supported their preliminary theory that "positive behavioral responses" result from increased female labor force participation. Bloom et. al treat the population as homogeneous at the macro-level, which is appropriate considering the nature of their cross-country panel study. But when placed in the context of the United States and the vast heterogeneity of its population, their approach lacks the granularity necessary to observe the micro-level effects unique to subgroups of different demographic characteristics.

The general declining trend of abortion rates and increasing rates of LFP rates across Black, White, and Hispanic women suggest, at first glance, that there is reason to expect a causal relationship between abortion and LFP, or vice versa (Figure 1). However, past literature has investigated how factors unrelated to birth control – like wars (the U.S. Civil War, WWI, WWII) and economic events such as the Great Depression – have impacted the perception and treatment of women as laborers, and that the severity of the crises pushed the economy for a more inclusive labor market, transcending barriers imposed by racial prejudice. While increased social acceptance and prevalence of abortion are attractive theories, overall increases in labor force participation rates, the entrance of women into the labor force, and/or general economic performance are more probable explanations of the upward trajectory in LFP rates shown in Figure 1. Such trends seemed to have "accentuated [the] existing trend" of "women's increasing involvement in waged labor," and "positive behavioral responses" furthered women's collective desire to work instead of giving birth (Helmbold and Schofield 1989; Bloom et. al., 2007). These findings suggest that changes in labor market conditions are responsible for changing fertility trends and the increased participation of women in the labor market. Nevertheless, the application of these findings is limited to cross-country contexts since they overlook the heterogeneity of labor market outcomes for women of different races/ethnic backgrounds.

Numerous studies have investigated the impacts of abortion access restrictions on women's fertility outcomes. Attempts to measure the pre-Roe versus post-Roe (1973) impact on fertility outcomes proved to be complex because while Roe constitutionally recognized the right to abortion via the privacy statute, the landmark case simultaneously gave states the discretion to impose gestational limits using their own criterion of fetus viability, the duration of pregnancy at which the states believed they had the right to prohibit abortion on the grounds of infanticide. According to studies performed on the CDC data on abortions from 1974 to 1977, states that mandated longer waiting periods for abortion raised the post-first-trimester percent of abortions by 2.7 percent (Bitler and Zavodny 2001). Although Mississippi's overall abortion rates among teens decreased when the state required a mandatory 24-hour waiting period, the rates of abortions among post-12-week pregnancies increased, suggesting that women from states with enforced waiting periods are more likely to travel longer distances (100 miles) to receive an abortion at a state with no such waiting periods (Joyce and Kaestner 2001). In fact, Jones and Jerman (2017)'s study, which found that women receiving abortions in states with mandated waiting periods and a professional evaluation were more likely than those who received abortions in states without such regulations to get an abortion, challenged the claim that past restrictions were significant enough to explain the observable reduction in overall abortion rates among most cohorts.

Nevertheless, there is a demonstrated need for increased abortion access. Title-X, a federal program established in 1970, was intended to provide af-

Figure 1: Women's LFP and Abortion Rates, 1973-2017



(a) Source: Aggregated annual Labor Force Participation Rates (LFP) for Black, White, and Hispanic over 20 years of age are sourced from the St. Louis Federal Economic Reserve Data (FRED). The Guttmacher Data Center public use data sets provides national abortion rates for all women in the United States. The vertical "REC" bands represent annual recession bands. I used St. Louis FRED's Real-time Sahm rule recession indicator index scores to broadly mark recessionary periods from 1973 to 2017.

fordable reproductive health services to low-income women in the United States. According to the Guttmacher Data Center, in 2016, only 17 percent of the demand for reproductive health services was met by Title X-funded centers. In the same year, an approximated 40 million women ages 13-44 expressed their demand for contraceptive services and less than 15 percent of these demands, or only 6,088,190 women in total (all ages) were met via service from publicly funded centers in the United States.

Microeconomic Approach

The direct impacts of state-mandated gestational limits on fertility and pregnancy outcomes are ambiguous. Other studies assessed the impacts of these regulations on the socioeconomic outcomes of women who had to carry unwanted pregnancies to term and had to incur costs both directly and indirectly associated with pregnancy and childbirth.

One recent study, the Turnaway Study, was UCSF's longitudinal (ten-year) study on 1000 women who experienced unwanted pregnancies and were subsequently turned away from abortions. Using this study, Miller, Wherry and Foster (2020), in "The Economic Consequences of Being Denied An Abortion," produced one of the most salient microeconomic reports on the longitudinal impacts of abortion, addressing the limitations of previous studies delineated above. They controlled for different gestational periods, indicating how controlling for different desired abortion timing may reflect differences in women's level of preparedness and their ability to care for their future children. This approach informs the understanding of women's willingness to pursue an abortion and points to behavioral implications, unlike the state-mandated waiting period analysis mentioned above. Using the Experian credit report data, the authors traced the credit report data of the participants of the Turnaway Study. By comparing the financial outcomes of women who were denied an abortion to those who were able to receive abortions, Miller et al. concluded that women who were denied abortions incurred persistent long-term financial "distresses," such as unpaid debts, evictions, bankruptcies, decrease in the women's household income, and a decreased likelihood to "have a prime credit score in the two years following the birth." Other studies also suggest a positive relationship between abortion access and infant outcomes as measured by "low birthweight babies and neonatal mortality" (Grossman and Jacobowitz 1981; Corman and Grossman 1985; Joyce 1987, and those born immediately after the legalization of Roe also experienced favorable outcomes (Gruber, Levine and Staiger 1999).

Research on the impact of abortion on intergenerational socioeconomic outcomes is limited. Donahue and Levitt's in "Abortion and Crime" (Donahue and Levitt 2001) enlisted a "back of the envelope" approach, creating a link between abortion and crime by following common associations between independently proven relationships – unwanted childbirth and unfavorable childhood outcomes, fragmented family structures, and childhood development, poor economic adjustment and crime rates, to name a few. Their work concluded that the passage of Roe reduces the chance of unintended pregnancies and eliminates the negative shock associated with women carrying unplanned or unwanted pregnancies to term. Since "the marginal children who were not born as a result of abortion legalization would have systematically been born into less favorable circumstances if the pregnancies had not been terminated" they argue that abortion access has resulted in favorable economic outcomes for the following generations, leading to fewer crime overall. Though Levitt later exhibited the novelty of these findings as a standalone chapter in Freakonomics, subsequent works relied less on correlational inference and more on rigorous econometric analyses, suggesting both economic and statistical fallacies in "Abortion and Crime." Levitt and Donohue failed to account for endogeneity in their model introduced by "age-specific crime rates"(Joyce 2009), the effects of social learning that shifts the collective demand for abortions (Akerlof et al. 1996; Lott and Whitley 2007), state-dependent characteristics, and shifts in the supply curve of abortion accessibility. Additionally, the cocaine epidemic during the 1970s may have obscured the true effects of the positive selection effect induced by increased abortion access (Joyce 2009).

Furthermore, the "Abortion as Insurance" paper hypothesized the potential effect of mandating abortion as a required insurance plan for women with unwanted and unintended pregnancies (Levine and Staiger 2002). By analyzing how the availability and affordability of abortion services affect women's fertility choices, their theoretical model treats abortion insurance analogously to fire or flood insurance. Their simple "decision-making under certainty model" hypothesizes the influence of a potential 'abortion insurance' on women's incentive for pregnancy, contraceptive use, and attitude towards abortion and childbirth. Recent works proposed alternative methods to model incentives, most notably Raute's exploitation of the causal effect of "paid leave reform on fertility decisions" via a two-group Differences-in-differences strategy, an audit study measuring the changes in the average fertility rate between women of low and high education in Germany (Raute 2017).

These past works demonstrate that demographic variables, such as age, race, and educational attainment, as well as external factors such as timing and geographic distance from the nearest healthcare facility, affect both the incentive to get an abortion. Macroeconomic policy-based studies highlight the wide scale impact of mandating reproductive health restrictions, some, like Donahue and Levitt, establish causal linkages between abortion and social phenomena separately associated with fertility. A commonality across these works is that historically, Black and BIPOC women's decisions were most sensitive to and were more impacted by changes in both national policies and micro-scale stressors compared to their White counterparts. In this paper, I will explore the 'penalty' of being non-White by measuring how the wage and earnings trajectories of working women post-Roe (1973) was impacted by a 'shock' brought about by pregnancy, and how these trends compare to that of those who obtained an abortion.

3 Introduction to the Data

The NLSY79 data set is a longitudinal study that tracks a nationally representative sample of 12,686 individuals aged between 14 to 22 in 1979. Since its first interview in 1997, the survey followed the lives of the same participants annually from 1997 to 2011 and biennially thereafter, until 2019. It contains detailed information on demographics, education, employment, income, family background, and other variables, making it a valuable resource for researchers studying labor market outcomes and intergenerational mobility. While NLSY97 is a newer version with over 90 percent of the sample still participating in the longitudinal survey to this day, the NLSY97 survey lacks crucial data, and when filtered into groups of different demographics, the resulting samples were too small to warrant its usage in empirical analysis. However, NLSY79 is an older data set and may not reflect current labor market conditions as accurately as the more recent NLSY97. Additionally, the NLSY79's sample size is relatively small compared to other major data sets which may limit the significance of its findings. The NLSY surveys have also undergone changes to their sampling methodology and survey design over the years, making it difficult to pursue direct comparisons between data from different waves. For the reasons delineated above and the simple fact that NLSY79 is an observational, not an experimental study, I proceed with caution when interpreting the data.

Figure 10 in the Appendix shows the raw data of all working women from NLSY79, with a separate table defining each variable. Each variable beginning with "inc" and ending with numbers consecutive from 79 to 94 denotes each woman's earnings from the previous year. Note that all negative values in this data set represent inapplicable, excused, omitted, etc. variables. "dob1chd" identifies the year during which the respondent gave birth to her first child, if she gave birth a live child, and "yrabort84" identifies the year the respondent received an abortion.

These two variables perform two key data-wrangling functions. First, it allows me to both subset the control and treatment groups; and second, construct panel data.

3.1 Treatment and Control Groups

By checking the presence of non-negative values, I can infer whether the individual gave birth or obtained an abortion. As such, I used the "dob1chd" variable to subset for the "control" group and the "yrabort84" variable to identify the "treatment" group or those who received an abortion. I indicate this as a dummy variable, where "treat" equals 1 for the treatment and 0 for the control. After omitting rows with inconclusive data, I extracted a data frame with a sample size of 4,085 women who consistently made earned income throughout the 10 years after their first pregnancy, aged between 15 and 27 at the time at which they got pregnant. The key variables of interest are shown in Figure 2. Note that each "wage" variable represents the total earnings from wages and salary from the preceding year. The frequency distribution in Figure 2 shows that age and year at which each woman gave birth or received an abortion, "shockyr factor", is approximately normally distributed, whereas other variables, such as race, education "hgrdc79", initial wage, and the share of those who received the treatment/abortion "treat" are heavily skewed, indicating that the sample may over-represent or under-represent the population of interest.

3.2 Constructing the Panel Data

The second function is the year. This is important because each respondent did not experience childbirth/abortion during the same year, similar to

Figure 2: Summary Statistics

```
Data Frame Summary
dfsummstat
Dimensions: 4085 x 6
Duplicates: 861

-------------------------------------------------------------------------------------------------------------------------
No   Variable          Stats / Values          Freqs (% of Valid)   Graph                   Valid      Missing
---- ----------------- ----------------------- -------------------- ----------------------- ---------- ---------
1    age               Mean (sd) : 21.3 (2.5)  13 distinct values         :                 4085       0
     [numeric]         min < med < max:                                   :                 (100.0%)   (0.0%)
                       15 < 21 < 27                                       : .
                       IQR (CV) : 4 (0.1)                               : : : : .
                                                                    . . : : : : : : : :

2    race              1. Hispanic              738 (18.1%)         III                     4085       0
     [factor]          2. Black                1134 (27.8%)         IIIII                   (100.0%)   (0.0%)
                       3. White                2213 (54.2%)         IIIIIIIIII

3    hgrdc79           1. 12                   1205 (29.8%)         IIIII                   4045       40
     [character]       2. 10                    610 (15.1%)         III                     (99.0%)    (1.0%)
                       3. 9                     600 (14.8%)         II
                       4. 11                    530 (13.1%)         II
                       5. 8                     460 (11.4%)         II
                       6. 13                    203 ( 5.0%)         I
                       7. 7                     142 ( 3.5%)
                       8. 14                    136 ( 3.4%)
                       9. 15                     64 ( 1.6%)
                       10. 6                     41 ( 1.0%)
                       [ 8 others ]              54 ( 1.3%)

4    initialwage       1. 0                     871 (24.0%)         IIII                    3636       449
     [character]       2. 3000                  102 ( 2.8%)                                 (89.0%)    (11.0%)
                       3. 7000                   89 ( 2.4%)
                       4. NA                     80 ( 2.2%)
                       5. 6000                   79 ( 2.2%)
                       6. 2000                   76 ( 2.1%)
                       7. 1000                   75 ( 2.1%)
                       8. 10000                  75 ( 2.1%)
                       9. 5000                   75 ( 2.1%)
                       10. 8000                  68 ( 1.9%)
                       [ 700 others ]          2046 (56.3%)         IIIIIIIIIII

5    shockyr_factor    1. 79                   656 (16.1%)          III                     4085       0
     [factor]          2. 80                   695 (17.0%)          III                     (100.0%)   (0.0%)
                       3. 81                   708 (17.3%)          III
                       4. 82                   753 (18.4%)          III
                       5. 83                   723 (17.7%)          III
                       6. 84                   550 (13.5%)          II

6    treat             1. 0                    3727 (91.2%)         IIIIIIIIIIIIIIIIII      4085       0
     [factor]          2. 1                     358 ( 8.8%)         I                       (100.0%)   (0.0%)
-------------------------------------------------------------------------------------------------------------------------
```

(a) Source: National Longitudinal Survey of Youth 1979 (NLSY97)

'staggered roll-out' experimental studies where experiments are conducted on a set of units over multiple time periods, hence the income observation spanning over 10 years (Figure 10 in the Appendix). The 'earliest' possible observation of earnings is from 1979 to 1989, and the 'latest' is from 1984 to 1994. Using this information and after sorting for control/treatment per the method above, I reorganized the data frame to resemble a panel data-like structure, with each observation of annual earnings following the same individuals for 10 years, or 10 periods. For example, "period 0" would denote the year 1980 woman who has given birth to a child in 1980, and "period 5" would represent 1985, 5 years after the woman has given birth.

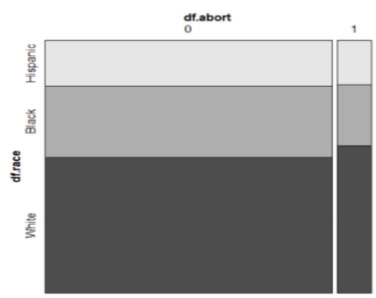
Figure 3: Contingency Table and Mosaic Plot

Cell Contents

N
Chi-square contribution
N / Row Total
N / Col Total
N / Table Total

Total Observations in Table: 4085

df$race	df$abort 0	1	Row Total
Hispanic	661	77	738
	0.007	0.057	
	0.896	0.104	0.181
	0.181	0.176	
	0.162	0.019	
Black	1028	106	1134
	0.240	1.999	
	0.907	0.093	0.278
	0.282	0.242	
	0.252	0.026	
White	1958	255	2213
	0.159	1.323	
	0.885	0.115	0.542
	0.537	0.582	
	0.479	0.062	
Column Total	3647	438	4085
	0.893	0.107	



(a) Source: National Longitudinal Survey of Youth 1979 (NLSY97)

3.3 Contingency Analysis

Figure 3 is a two-way contingency table of two dummy categories "abort", which takes on a value of 1 or 0 to indicate whether the woman got an abortion, and "race", denoting the racial characteristic of the woman. The mosaic table on the right visualizes the patterns shown in the contingency table. Within this special extract of NLSY79, 54.2 percent of women are White, 27.8 percent are Black, and 18.1 percent are Hispanic. Shares of White, Black, and Hispanic women who have gotten an abortion are 58.2 percent, 24.2 percent, and 17.6 percent, respectively. The proportional difference between the racial composition of women in the data set and the racial composition of those who received an abortion shows that white women were much more likely to undergo abortions than Black and Hispanic women. Keeping in mind the caveat that this observation takes into account only one contingency (abortion), these results suggest that a women's propensity to obtain an abortion may depend on salient individual or access-based factors commonly associated with race.

4 Difference-in-Differences

As explained in section 3, after stratifying the NLSY79 sample, I used the control and treatment-stratified group to form a panel data frame. Using wage as a proxy for utility, I will first run a differences-in-differences regression analysis on this data frame to analyze the treatment effects of obtaining an abortion. I use this information to calculate the Net Present Value as a function of log(wage), or percent change in a woman's average earnings with an arbitrary variable, C, to denote the costs of the treatment (abortion) or the benefits of the control (childbirth), such as paid maternity leave, to assess the value of obtaining an abortion in the event that a working woman got pregnant. Covariates such as age, ethnicity/race, and educational levels are controlled for as they are key essential determinants of whether the treatment effect exists.

Empirical Framework

The preliminary hypothesis is that the pre-intervention difference should be 0. In other words, all else equal, there should be no difference in the wage between working women who had an abortion and working women who gave birth; the average wage gap should be zero unless the treatment occurs. I hypothesize that the gap is greatest in the first year, where one would presumably take time off from work to give birth (the control group) or would continue to work otherwise (the treatment group). With each successive period (e.g. from year 0 to year 10), the positive wage effects of abortion should decrease with time, emulating a "promotion" effect that prevents the career setback women often experience after giving birth and raising the child. The gap in the annual earnings of both control and treatment groups should either remain the same (vertical shift in the curve) or shrink with time.

The DiD regression model measures the difference at two different states or time periods, particularly the difference in the average annual earnings of women who have gotten an abortion and those who have given birth at year 0, and a subsequent time period. As mentioned in section 3, the DiD regression data set is a rearranged version of the original NLSY79 data set, altered to form a longitudinal panel data following the average annual earnings of working women consecutively from period 0 (year at the time of birth/abortion) to period 10 (10 years after the said birth/abortion).

In this model, I assume that the effects of a non-live birth or miscarriage as similar effects to that of abortion on a woman's earnings over time. While this assumption discounts the psychological impact of miscarriages on the woman's livelihood and her capacity for gainful employment, this model controls for this by creating a dummy variable sorting women into groups who have either 1) given birth to a live child as the "control" group or 2) got an abortion and had not given birth to a live child group throughout the observed 10 year period as the "treatment" group. I impose such restrictions on the model to eliminate potential confounding variables, the details of which are thoroughly acknowledged in subsequent sections.

Figure 4 illustrates two possible wage trajectories for the treatment group - women who received an abortion. The first possibility is that the treatment group may experience an initial upward shock at period 0, followed by a diminishing marginal growth (d/dt) where the wage trajectory converges back to the same rate of change as that of the control group; or as depicted by the second possibility, the marginal growth may be constant, implying that the treatment or the recipience of abortion may alter the functional form and accelerate the growth of a woman's wage over this 10-year period. I display the actual trajectory in Figure 5 using the NLSY79 data.

Operationalizing the theory of this "widened gap", I begin the mathematical expression with the simple equation:

$$\Delta Y_i = \beta_0 + \beta_1 X_i + u_i \tag{1}$$

Where ΔY_i equals the change in the value of annual income for the ith individual.

Regressing the equation gives $\beta_1 = \beta^{DD}$, which equals the difference in the outcome variable (Y), total income.

Since β is the difference in the group means of ΔY_i, the DD estimator is equal to the OLS estimation of β.

Figure 6 is a table depicting a simple computation of a DiD estimator for two separate time periods (1 year and 10 years). In this analysis, the DD estimator provides a broad view of changes in the women's annual income and confirms the presence that there exists a noticeable wage gap between the control and treatment groups. I perform a regression analysis to analyze whether the variation in wage between the two groups is explained by the treatment itself or endogeneity in the model.

Thus, the DD estimator should equal:

Figure 4: DiD Regression Conceptualization

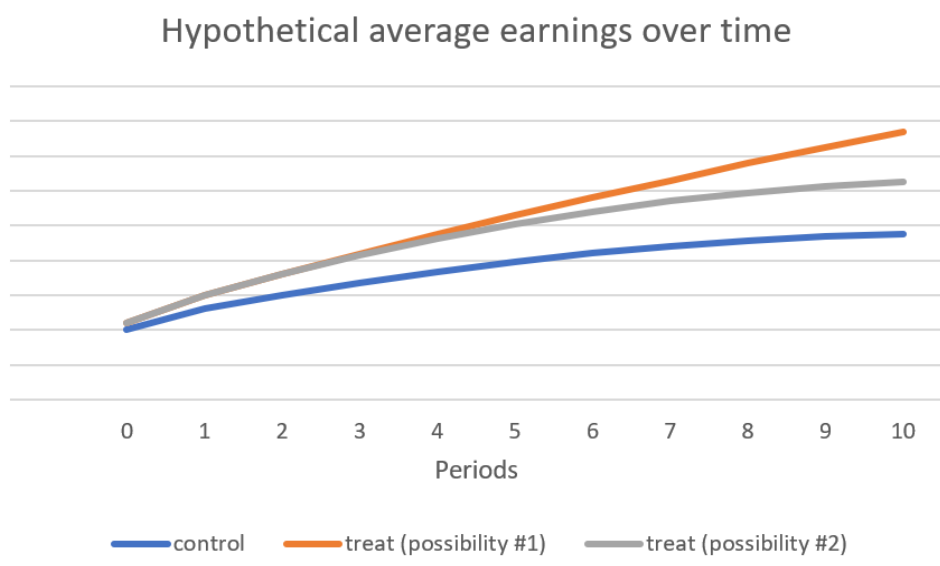
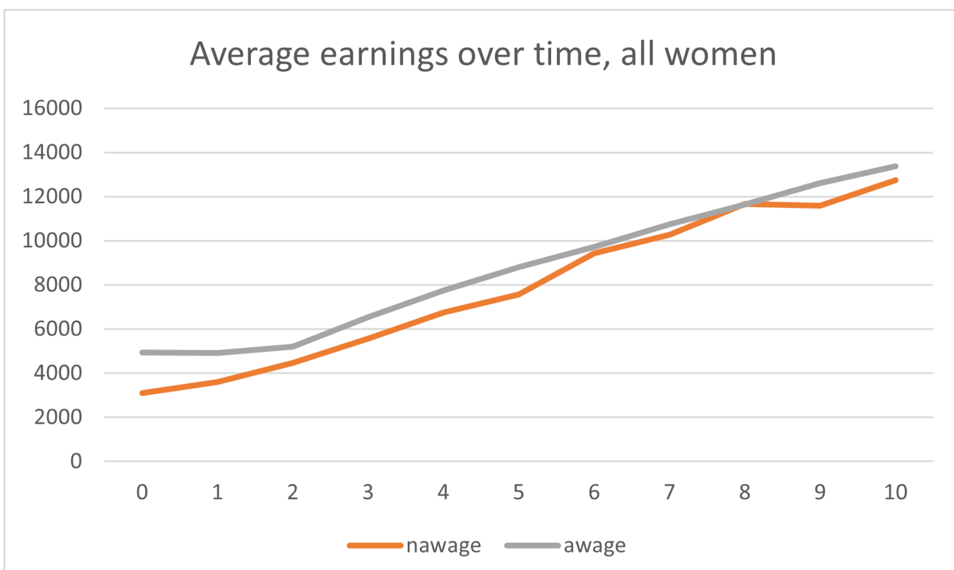


Figure 5: DiD estimator Computation



(a) Source: National Longitudinal Survey of Youth 1979 (NLSY97)

Figure 6: Average Earnings, All Women

The DiD estimator for 1 year:

	Before 80'	After 80' (89')	Change
Treatment (income, abortion)	3737	4731.8	4731.8-3737= 994.8
Control (income, birth)	3063.8	3313.7	3313.7-3063.8= 249.5
T-C	3737-3063.8 = 673.2	4731.8-3313.7= 1418.1	994.8-249.5= 745.3

The DiD estimator for 10 years:

	Before 80'	After 80' (89')	Change
Treatment (income, abortion)	3737	4731.8	4731.8-3737= 994.8
Control (income, birth)	3063.8	3313.7	3313.7-3063.8= 249.5
T-C	3737-3063.8 = 673.2	4731.8-3313.7= 1418.1	994.8-249.5= 745.3

$$\beta^{DD} = (\text{ Test }_{t,a} - \text{ Test }_{t,b}) - (\text{ Test }_{c,a} - \text{ Test }_{c,b}) = 745.3 \qquad (2)$$

Differences-in-Differences Regression

Based on the research using the NLSY79 data, the early career earnings trajectory for women generally shows a steep increase in earnings during the first few years in the labor market, followed by a gradual leveling off of the earnings growth rate. Recent literature found that race/ethnicity and education have significant effects on women's earnings trajectories (Loughran and Riehl 2018; Polachek and Xiang 2017). Polachek and Xiang (2017) found that the earnings trajectory for this group followed a downward concave relationship with respect to age, meaning that earnings growth was initially rapid and then decelerated over time. Using NSLY79, Light and Ureta (1990) used a quadratic function of age to model the early career earnings of college-educated women. Borjas and Grogger (1996) found similar implications for women of high school-educated backgrounds, but that the rate of increase slowed down after about 8-10 years. The functional forms from these studies were more or less similar in that the earnings trajectory for working women in their early careers, regardless of ethnicity, followed a concave-downward trajectory. Light and Ureta (1990)'s functional form is:

$$\log(\text{wage}) = \beta_0 + \beta_1 \text{age} + \beta_2 \text{age}^2 + \beta_3 \text{experience} + \beta_4 \text{education} + \beta_5 \text{race} + \epsilon \tag{3}$$

Where the quadratic function of age in equation (3) explains the steep initial increase in earnings during the first few years in the labor market, followed by a gradual leveling off of the earnings growth rate. Other variables in the model, such as education and race/ethnicity, capture differences in earnings trajectories across different subgroups of women.

Exploiting the panel nature of my extraction of the NLSY79 data, I modify Light and Ureta's quadratic specification and use a differences-in-differences regression to estimate the causal effect of a treatment on log wages while controlling for other factors that may affect wages and that may differ between the treatment and control groups:

$$\begin{aligned}\log(\text{ wage }) = \beta_0 + \beta_1 \text{ time\_i } + \beta_2 \text{ treat\_i } + \beta_3 \text{ time\_i } {}^* \text{ treat\_i } + \beta_4 \text{ Hispanic} \\ + \beta_5 \text{ Black } + \beta_6 \text{ age } + \beta_7 \text{ age }^2 + \beta_8 \text{ education } + \epsilon\end{aligned} \tag{4}$$

This equation models the logarithm of working women's wage as a function of several variables, including a treatment indicator (treat), an interaction term between treat and time, and other covariates, age, age^2, education, Hispanic, and Black. Time and treat are both dummy variables which take the value of 0 or 1, indicating whether the observed ith individual is at the pre or post-shock(childbirth/pregnancy) period and whether the ith individual belongs to the treatment (received abortion) or control (live childbirth) group. The interaction term between treat and time estimates the difference in the change in log wages over time between the treatment and control groups. Other covariates control for demographic factors that may affect log(wage) and that may differ between the treatment and control groups.

Conventional wisdom tells us that wages are typically higher for individuals with more education, or that there are wage differences by race/ethnicity, so including these covariates in the model allows us to isolate the effect of the treatment on log wages while holding these other factors constant. Note that the annual earnings of all women in this data frame are greater than 0, to depict their participation in the labor force.

Regression Results

The DiD regression estimation measures the immediate effect of abortion at time (t) by calculating the difference between the wage at the next time period/year (t+1) when she receives the treatment and her wage from the previous period.

Table 1 shows the regression results from period 0 to period 1, where period 1 is the time at which a woman either obtained an abortion (treatment) or gave birth (control). For example, a woman who experiences a pregnancy in the year 1979 and gets an abortion at the next year, 1980. 1979 is coded as "period 0", the wages from the NLSY79 data set reporting the total wages and income earned by each women, "wage" in the previous year. The interpretation using the regression results of the first column shows that all coefficients are statistically significant at the 1 percent level. The coefficients for time, Hispanic, age, education, time*treat(the treatment effect), and the constant are positive. The negative Treat coefficient on the time variable for the control group (the coefficient on time for individuals in the control group) suggests that the outcome variable decreases over time in the control group and that this trend should be accounted for in estimating the treatment effect. To summarize the differences in differences results for the first year, a 12.7 percent increase in the wage is attributed to a 1 unit increase in time (years); being Black is associated with a 25.4 percent decrease in wage and a 3.33 percent increase in wage for Hispanic; a one year increase in the age increases wage by about 15.9 percent; a 1 year increase in education increases wage by 4.9 percent; and the treatment effect of getting an abortion at period 0 is a 6.5 percent increase on a woman's wage, on average.

Table 2 depicts OLS results for time periods beyond the 1-year duration. For instance, the second regression compares a woman's wage one year after the treatment (which takes place in period 1). The rationale for this method is that the extracted data set is a panel data frame, allowing for a longitudinal interpretation of women's average wages, grouped either by race and/or educational attainment, over time. These separate DiD regressions from 1 to 10 years serve the purpose of testing the presence of the delayed response of the treatment, beyond the immediate 1-year duration. Based on the adjusted R-squared value alone, the increase in wage is explained less by the treatment of abortion as time passes, the effects easing off in a non-linear, potentially concave manner. While the treatment effect (time*treat) is consistently positive over time, demonstrating that the effects of abortion on women's earnings are not static. The effect of abortion in the first year on log(wage) persists beyond the first year. In fact, the treatment effect is the greatest (160 percent increase in log(wage)) in year 2 and plateaus at around 40 to 50 percent in the following years, countering possibility no. 2 references in the original hypothesis (see figure 4) that the marginal effects of getting an abortion on wage diminish over time.

Refer to R-squared values in Table 1 and Table 2. In each of my analyses, I obtained an R-squared value of less than 0.3 for all models indicating that the model and its variables explain less than half of the variance in the dependent variable, log(wage). While this suggests that my model may not provide a good fit for the data, I proceed by prioritizing the practical significance of these individual coefficients and what their statistical significance, or lack thereof, may represent in the context of my research question. It is also difficult to achieve high levels of fit due to the complexity of the NLSY79 data and the limitations my model poses as it 'coerces' an observational study to serve an experimental purpose.

Table 1: Results, Years 1 to 5

	Dependent variable:				
	log(wage)				
	1 years	2 years	3 years	4 years	5 years
	(1)	(2)	(3)	(4)	(5)
time	0.129***	−0.943***	0.457***	0.609***	0.761***
	(0.003)	(0.065)	(0.034)	(0.035)	(0.034)
treat	−0.274***	0.021	−0.311***	−0.308***	−0.330***
	(0.061)	(0.161)	(0.083)	(0.084)	(0.084)
Hispanic	0.069*	−0.018	0.060	0.068	0.082*
	(0.037)	(0.084)	(0.044)	(0.045)	(0.044)
Black	−0.226***	−0.119	−0.227***	−0.200***	−0.249***
	(0.033)	(0.076)	(0.039)	(0.040)	(0.040)
age	0.186***	0.506***	0.218***	0.218***	0.203***
	(0.007)	(0.016)	(0.009)	(0.009)	(0.009)
education	0.040***	−0.213***	0.017*	0.013	0.029***
	(0.008)	(0.019)	(0.010)	(0.010)	(0.010)
time:treat	0.059***	1.604***	0.445***	0.460***	0.500***
	(0.011)	(0.217)	(0.113)	(0.114)	(0.113)
Constant	3.767***	−0.496	3.316***	3.354***	3.527***
	(0.132)	(0.305)	(0.161)	(0.162)	(0.160)
Observations	7,384	5,330	5,212	5,228	5,191
Log Likelihood	-11,609.930	-11,893.030	-8,206.152	-8,291.774	-8,194.251
Akaike Inf. Crit.	23,235.860	23,802.060	16,428.300	16,599.550	16,404.500
R-squared	0.2836869	0.2027333	0.1988373	0.2061491	0.2266459
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 2: Results, Years 6 to 10

	Dependent variable:				
	log(wage)				
	6 years	7 years	8 years	9 years	10 years
	(1)	(2)	(3)	(4)	(5)
time	0.870***	1.017***	1.126***	1.212***	1.269***
	(0.035)	(0.034)	(0.034)	(0.034)	(0.035)
treat	−0.337***	−0.339***	−0.362***	−0.368***	−0.375***
	(0.085)	(0.081)	(0.081)	(0.081)	(0.081)
Hispanic	0.041	0.046	0.043	0.022	0.033
	(0.045)	(0.043)	(0.043)	(0.043)	(0.044)
Black	−0.213***	−0.180***	−0.249***	−0.243***	−0.254***
	(0.040)	(0.038)	(0.039)	(0.039)	(0.040)
age	0.191***	0.188***	0.175***	0.164***	0.159***
	(0.009)	(0.008)	(0.009)	(0.009)	(0.009)
education	0.027***	0.031***	0.048***	0.042***	0.049***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
time:treat	0.564***	0.534***	0.593***	0.573***	0.650***
	(0.116)	(0.111)	(0.112)	(0.114)	(0.116)
Constant	3.787***	3.806***	3.935***	4.233***	4.261***
	(0.163)	(0.156)	(0.158)	(0.160)	(0.161)
Observations	5,171	5,105	5,002	4,906	4,818
Log Likelihood	-8,255.917	-7,897.936	-7,731.189	-7,582.438	-7,457.435
Akaike Inf. Crit.	16,527.830	15,811.870	15,478.380	15,180.880	14,930.870
R-squared	0.2231157	0.2639363	0.2849242	0.2917907	0.3054179
Note:					*p<0.1; **p<0.05; ***p<0.01

5 Net Present Value

While the previous DiD method explores the implications of Bloom et al.'s "Fertility and the demographic dividend" labor force participation framework in a microeconomic setting, I develop the previously mentioned "shock" theory and supplement it with the probability framework, similar to the "abortion as insurance" model, to calculate the net present value of carrying a pregnancy to term versus obtaining an abortion.

5.1 Conceptual Framework of NPV

I also include the initial costs included in the functional form of the net present value equation. Wherry and Miller (2016)'s linear regression model examined the relationship between the role of private insurance, Medicaid, and out-of-pocket spending and the cost of obtaining an abortion. The key demographic and socioeconomic variables in their model are age, race/ethnicity, income, and education and expressed as the following:

$$\begin{aligned}\text{Total cost of abortion} &= \beta 0 + \beta 1 \text{ private insurance } + \beta 2 \text{ Medicaid} \\ &+ \beta 3 \text{ out of pocket spending} + \beta 4 \text{ age } + \beta 5 \text{ race} + \beta 6 \text{ income} \\ &+ \beta 7 \text{ education} + \varepsilon\end{aligned} \tag{5}$$

The coefficients $\beta 1$, $\beta 2$, and $\beta 3$ represent the average amount paid by patients with private insurance, Medicaid, and out-of-pocket, respectively. Wherry and Miller found that, on average, patients with private insurance paid 524 dollars for an abortion, patients with Medicaid paid 231 dollars, and patients who paid out of pocket paid 603 dollars. Out-of-pocket spending was highest for patients who were younger, had lower incomes, and had lower levels of education.

Because NLSY79 contains limited information about the insurance status of women, I treat the cost as an arbitrary variable defined by these undetermined variables, which are later interpreted to determine the maximum cost at which a woman is willing to get an abortion or alternatively, the cost which breaks even the cumulative wage benefits a woman would earn within a 10-year time span, on average.

I derived the NPV formula by exponentiating Light and Ureta (1990)'s modified equation to form the equation for the expected annual total income of women y at time t to obtain the equation for a woman's average annual income after time t as a function of age, education, race, and abortion. The latter two are controls, with race expanded as 3 dummy variables of 0 or 1 indicating whether the individual is Hispanic, White, or Black, as with whether they have gotten an abortion within the specified time frame:

$$\begin{aligned} \text{wage } &= \exp(\beta 0+\beta 1 \text{ age } +\beta 2 \text{ age }^{2}+\beta 3 \text{ education} \\ &+\beta 4 \text{ Hispanic } +\beta 5 \text{ Black } +\beta 6 \text{ abortion } +\beta 7 \text{ time } +\varepsilon) \end{aligned} \tag{6}$$

I incorporate this cost framework as an arbitrary variable in the Net Present Value equation. I express the definite integral of equation 6 as the cumulative, or the present value of a woman's total annual income from wages and earnings from the years 0 to 10 (the bounds), as:

$$\text{PV(a=1)} = \int_{1}^{10} \mathrm{w}_{a}(t) - [c(a=1)] \tag{7}$$

Where c(a=1) represents the initial cost of obtaining an abortion, such as medical fees, health insurance coverage, etc. I include the cost function in the present value framework to reflect the impact of changes in the cost of obtaining an abortion on women's decisions about when and whether to have children. and $\mathrm{w}_a(t)$ represents the expected annual income if a woman obtains an abortion. Future renditions germane to modern trends may subtract maternity leave benefits from this cost function. The present value for not obtaining an abortion is:

$$\text{PV(a=0)} = \int_{1}^{10} \mathrm{w}_{a}(t) - [c(a=1)] \tag{8}$$

Where $\mathrm{w}_c(t)$ represents the expected wage of those who did not get an abortion.

The Net Present Value (NPV) is the difference between the present value of obtaining an abortion ($a = 1$) and the present value of not obtaining an abortion ($a = 0$):

$$NPV = \int_{0}^{10} \left(w_{_a}(t)\right) - \int_{0}^{10} \left(w_c(t)\right) - [c(a=1)] \tag{9}$$

Using the coefficients derived from the differences-in-differences OLS regression results above, the net present value, or the net percent increase in women's earnings over the 10-year period varies significantly for women of different demographic characteristics. Table 3 in the Appendix shows coefficient results for wage, instead of log(wage) as previously shown in Table 2. The cumulative sum of the treatment effect (treat*time) from periods 1 to 10 on log(wage), controlling for the covariates specified above, is 5.984, or about a 598.4 percent change in the average wage of working women aged 15 to 27 who obtained an abortion compared to those who did not. Alternatively, the results from Table 3 show that the net present value in wages of getting an abortion instead of giving birth is about 15,293 dollars, on average. These values are discounted relative to the woman's initial earnings at the initial period.

Figure 8 in the Appendix illustrates the cumulative earnings and visualizes the differences in earnings of women between those who gave birth to a child and those who did not, using the raw NLSY79 data. Figure 9 uses the same raw data, but draws a comparison between Black women who obtained and did not obtain an abortion. The cumulative difference between the treatment and control groups I obtained using the wrangled and stratified data is much greater than the difference in the trajectories shown in these two graphs. The key takeaway from this discrepancy between the overall NLSY79 and my selected sample emphasizes the statistical importance of the covariates I initially controlled for when conducting the Differences-in-Differences analysis.

6 Conclusion

In this paper, I first investigate which restrictions must be imposed on the original NLSY79 data set to construct a panel data of working women born between 1957 to 1964 in the United States. A simple contingency analysis suggested that White women have a higher propensity to get an abortion than their Black and Hispanic counterparts. I then ran multiple Difference-in-differences regressions to analyze the immediate effects of abortion (1 year after the treatment) and its residual, dynamic effects (2 to 10 years after the treatment). I found the results to be consistent with my initial hypothesis that the wage trajectories of women who have gotten an abortion are greater than that of women who gave birth to a child. In the Net Present Value section, I explore the unexplained racial discrepancy of abortion recipients from the contingency analysis, including an arbitrary variable, C to encompass a variety of costs associated with getting and not getting an abortion. I presume this "Cost" variable to be a function of factors external to the model, such as limited access to abortion via state-mandated restriction or travel distance to the nearest family planning center.

6.1 Limitations

Future iterations of the Difference-in-differences model could be improved by not only investigating how getting an abortion may have a "promotional effect" on working women's earnings trajectories over time but also assessing how women might enter or exit the labor force depending on the outcome of their pregnancy (Bloom et al. 2007). The model in my extracted data frame accounts for women who have made more than "0" dollars in income.

There are many limitations to the net present value model. My model best explains the net present value of getting an abortion to the extent that the specified variables influence women's earnings from working, and that the Difference-in-differences regression results from my analysis are statistically significant. Future iterations could be focused on deriving an equation for this "Cost" function, perhaps regressing racial discrimination in abortion access and insurance plans, or even factoring in recent development to remedy the cost of childbirth and childrearing, such as paid maternity leave. For example, Raute's paper evaluates the relationship between maternity leave benefits and the baby gap between high and low-earning women (Raute 2017).

More importantly, I measure the "value" solely in monetary terms. In reality, the decision to get an abortion is influenced by factors that, suffice to say, may not be related to income at all. The woman's marital status, pre-existing poverty levels, cultural identity, geographic location, physical well-being, etc., all serve as plausible reasons to terminate or continue a pregnancy. Provided that data collection is possible, behavioral economics-based approaches may be more appropriate for this type of research, such as Upadhyay et al.'s recent qualitative work in "The effect of abortion on having and achieving aspirational one-year plans" (UUpadhyay, Biggs and Foster 2015).

6.2 Final Remarks

In summary, I find that the coefficients on the treatment (abortion), education, and race, are statistically significant. Specifically, abortion does have a positive effect on a working woman's earnings, and the Difference-in-differences regression results suggest that the residual effects seem to last beyond the first several years. There is a "wage penalty" for being Black, and a "benefit" to being White, whereas the effects for Hispanic women remain ambiguous, which is likely due to the underrepresentation of Hispanic women in the NSLY79 data. I estimate that the net present value of getting an abortion for working women aged 15-27 is a 598.4 percent increase in their cumulative earnings over a 10-year period compared to those who do not obtain an abortion. Finally, I acknowledge that a lot has changed since the mid-1900s. Today, the legal status of Roe hangs in precarious balance, even more so for equitable distribution of reproductive healthcare. When it comes to the relationship between reproductive rights and economic well-being, race and ethnicity are jointly related. In future research, I hope to gather more recent and ethnically inclusive data and address most, if not some of the key limitations I encountered throughout this paper.

References

[1] Ananat, Elizabeth Oltmans, Jonathan Gruber, Phillip B. Levine, and Douglas Staiger. 2006. “Abortion and Selection.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w12150.

[2] Bailey, Martha J. 2011. “Reexamining the Impact of Family Planning Programs on U.S. Fertility: Evidence from the War on Poverty and the Early Years of Title X.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w17343.

[3] Binsbergen, Jules H. van, and Ralph S.J. Koijen. 2010. “Predictive Regressions: A Present-Value Approach.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w16263.

[4] Bitler, M., and M. Zavodny. 2001. “The Effect of Abortion Restrictions on the Timing of Abortions.” *Journal of Health Economics.* 20 (6): 101132. https://doi.org/10.1016/s0167-6296(01)00106-0.

[5] Bloom, David E., David Canning, Günther Fink, and Jocelyn E. Finlay. 2007. “Fertility, Female Labor Force Participation, and the Demographic Dividend.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w13583.

[6] Bureau of Labor Statistics, U.S. Department of Labor. National Longitudinal Survey of Youth 1979 cohort, 1979-2016 (rounds 1-27). Produced and distributed by the Center for Human Resource Research (CHRR), The Ohio State University. Columbus, OH: 2019.

[7] Card, David, and Thomas Lemieux. 1993. “Wage Dispersion, Returns to Skill, and Black-White Wage Differentials.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w4365.

[8] Corman, H., and M. Grossman. 1985. “Determinants of Neonatal Mortality Rates in the U.S. A Reduced Form Model.” *Journal of Health Economics.* 4 (3): 213–36. https://doi.org/10.1016/0167-6296(85)90030-x.

[9] Donohue, John, and Steven Levitt. 2000. “The Impact of Legalized Abortion on Crime.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w8004.

[10] Grossman, M., and S. Jacobowitz. 1981. “Variations in Infant Mortality Rates among Counties of the United States: The Roles of Public Policies and Programs.” *Demography* 18 (4): 695–713.

[11] Guttmacher Institute. ”State Facts About Abortion: Data Tables.” *Guttmacher Institute, 2021.* https://data.guttmacher.org/states/table?topics=62dataset=data.

[12] Helmbold, Lois Rita, and Ann Schofield. 1989. “Women’s Labor History, 1790-1945.” *Reviews in American History* 17 (4): 501–18. https://doi.org/10.2307/2703424.

[13] Jones, Rachel K., and Jenna Jerman. 2017. “Population Group Abortion Rates and Lifetime Incidence of Abortion: United States, 2008-2014.” *American Journal of Public Health* 107 (12): 1904–9. https://doi.org/10.2105/AJPH.2017.304042.

[14] Joyce, Ted, and Robert Kaestner. 2001. “The Impact of Mandatory Waiting Periods and Parental Consent Laws on the Timing of Abortion and State of Occurrence among Adolescents in Mississippi and South Carolina.” *Journal of Policy Analysis and Management* 20 (2): 263–82. https://doi.org/10.1002/pam.2025.

[15] Joyce, Theodore J. 2009. “Abortion and Crime: A Review.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w15098.

[16] Lahey, Joanna N., and Marianne H. Wanamaker. 2022. “Effects of Restrictive Abortion Legislation on Cohort Mortality Evidence from 19th Century Law Variation.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w30201.

[17] Levine, Phillip B., and Douglas Staiger. “Abortion as Insurance.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* February 2002. https://doi.org/10.3386/w8813.

[18] Lott Jr, John R., and John Whitley. 2007. “Abortion and Crime: Unwanted Children and Out-of-Wedlock Births.” *Economic Inquiry* 45 (2): 304–24. https://doi.org/10.1111/j.1465-7295.2006.00040.x.

[19] Miller, Sarah, Laura R. Wherry, and Diana Greene Foster. 2020. “The Economic Consequences of Being Denied an Abortion.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w26662.

[20] Raute, Anna. 2017. “Can Financial Incentives Reduce the Baby Gap? Evidence from a Reform in Maternity Leave Benefits.” Working Paper. Working Paper Series. *National Bureau of Economic Research.* https://doi.org/10.3386/w23793.

[21] U.S. Bureau of Labor Statistics. ”Labor Force Participation Rate - 20 Yrs. and over, Black or African American Women” [LNS11300032]. *FRED, Federal Reserve Bank of St. Louis.* Accessed March 12, 2023. https://fred.stlouisfed.org/series/LNS11300032.

[22] U.S. Bureau of Labor Statistics. ”Labor Force Participation Rate - 20 Yrs. and over, Hispanic or Latino Women” [LNU01300035]. *FRED, Federal Reserve Bank of St. Louis.* Accessed March 12, 2023. https://fred.stlouisfed.org/series/LNU01300035.

[23] U.S. Bureau of Labor Statistics. ”Labor Force Participation Rate - 20 Yrs. and over, White Women” [LNS11300029]. *FRED, Federal Reserve Bank of St. Louis.* Accessed March 12, 2023. https://fred.stlouisfed.org/series/LNS11300029.

Appendix

Figure 7: Source: National Longitudinal Survey of Youth 1979

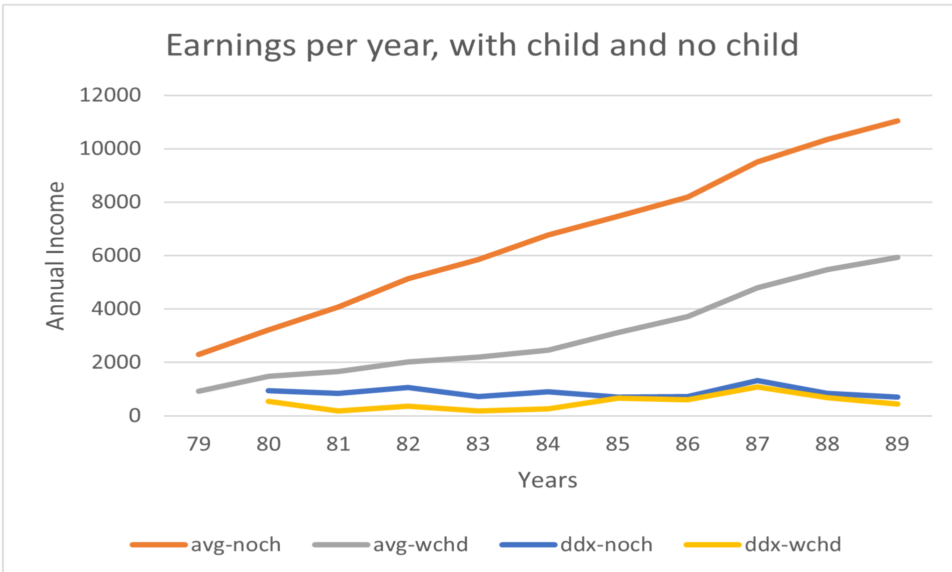


Figure 8: Source: National Longitudinal Survey of Youth 1979

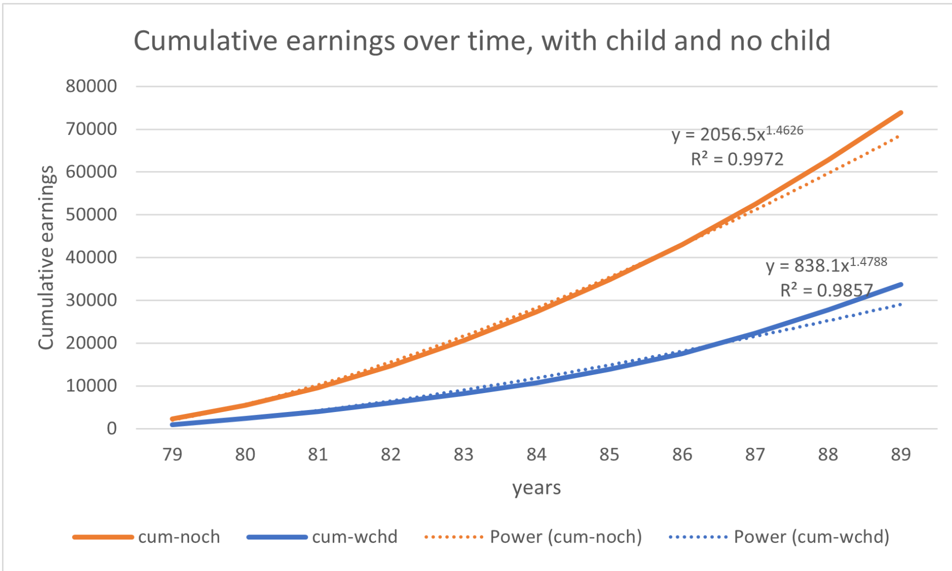
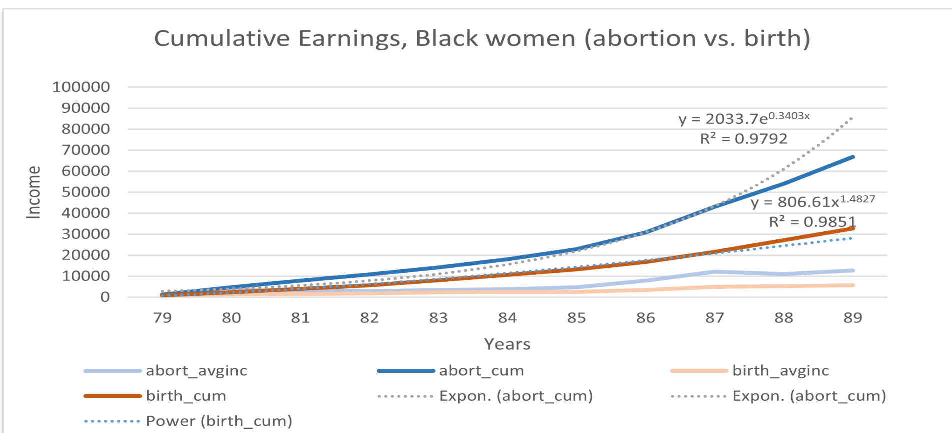


Figure 9: Source: National Longitudinal Survey of Youth 1979



Note that mortality rates among Black women who gave birth were high during this decade. The samples were filtered to only include those whose gave birth to a live child. Not included in this figure are the results for women who lost their child during pregnancy: increases to earnings for these individuals are substantial and on average, the differential effects of nonlive birth on earnings are greater than for those who obtained an abortion.

Figure 10: Source: National Longitudinal Survey of Youth 1979

```
Data Frame Summary
df
Dimensions: 12686 x 22
Duplicates: 0

----------------------------------------------------------------------------------------------------------------
No   Variable    Stats / Values               Freqs (% of Valid)     Graph                  Valid      Missing
---- ----------- ---------------------------- ---------------------- ---------------------- ---------- --------
1    id          Mean (sd) : 6343.5 (3662.3)  12686 distinct values  : : : : : : : : : :    12686      0
     [integer]   min < med < max:             (Integer sequence)     : : : : : : : : : :    (100.0%)   (0.0%)
                 1 < 6343.5 < 12686                                  : : : : : : : : : :
                 IQR (CV) : 6342.5 (0.6)                             : : : : : : : : : :
                                                                     : : : : : : : : : :

2    dob79       Mean (sd) : 60.3 (2.2)       57 : 1680 (13.2%)      II                     12686      0
     [integer]   min < med < max:             58 : 1677 (13.2%)      II                     (100.0%)   (0.0%)
                 57 < 60 < 64                 59 : 1722 (13.6%)      II
                 IQR (CV) : 4 (0)             60 : 1662 (13.1%)      II
                                              61 : 1530 (12.1%)      II
                                              62 : 1600 (12.6%)      II
                                              63 : 1550 (12.2%)      II
                                              64 : 1265 (10.0%)      I

3    hgrdc79     Mean (sd) : 10.3 (2.7)       23 distinct values       :                    12686      0
     [integer]   min < med < max:                                      :                    (100.0%)   (0.0%)
                 -4 < 11 < 95                                          :
                 IQR (CV) : 3 (0.3)                                    :
                                                                       :

4    race        Mean (sd) : 2.4 (0.7)        1 : 2002 (15.8%)       III                    12686      0
     [integer]   min < med < max:             2 : 3174 (25.0%)       IIIII                  (100.0%)   (0.0%)
                 1 < 3 < 3                    3 : 7510 (59.2%)       IIIIIIIIIII
                 IQR (CV) : 1 (0.3)

5    numchd79    Mean (sd) : 0.1 (0.4)        0 : 11503 (90.7%)      IIIIIIIIIIIIIIIIII     12686      0
     [integer]   min < med < max:             1 :   893 ( 7.0%)      I                      (100.0%)   (0.0%)
                 0 < 0 < 5                    2 :   236 ( 1.9%)
                 IQR (CV) : 0 (3.4)           3 :    45 ( 0.4%)
                                              4 :     8 ( 0.1%)
                                              5 :     1 ( 0.0%)

6    yrabort84   Mean (sd) : 0.7 (19.2)       18 distinct values     :                      12686      0
     [integer]   min < med < max:                                    :                      (100.0%)   (0.0%)
                 -5 < -4 < 84                                        :
                 IQR (CV) : 0 (28.3)                                 :
                                                                     :                 .

7    dob1chd     Mean (sd) : 1461.8 (875.6)   48 distinct values                       :    12686      0
     [integer]   min < med < max:                                                      :    (100.0%)   (0.0%)
                 -4 < 1981 < 2013                                                      :
                 IQR (CV) : 1991 (0.6)                               :                 :
                                                                     :                 :
```

Source: National Longitudinal Survey of Youth 1979

8	inc80 [integer]	Mean (sd) : 2115.9 (3775.9) min < med < max: -5 < 0 < 67500 IQR (CV) : 3004 (1.8)	1145 distinct values	: : : : : .	12686 (100.0%)	0 (0.0%)
9	inc82 [integer]	Mean (sd) : 4256.5 (5688.8) min < med < max: -5 < 1900 < 28975 IQR (CV) : 7000 (1.3)	1518 distinct values	: : : : : : : . .	12686 (100.0%)	0 (0.0%)
10	inc83 [integer]	Mean (sd) : 5139.6 (6383.8) min < med < max: -5 < 2500 < 31519 IQR (CV) : 8413.8 (1.2)	1854 distinct values	: : : : . : : : : . .	12686 (100.0%)	0 (0.0%)
11	inc84 [integer]	Mean (sd) : 6099.6 (7318.7) min < med < max: -5 < 3500 < 36574 IQR (CV) : 10000 (1.2)	1742 distinct values	: : : : . : : : : .	12686 (100.0%)	0 (0.0%)
12	inc85 [integer]	Mean (sd) : 6693.7 (8054.7) min < med < max: -5 < 3752.5 < 39389 IQR (CV) : 11000 (1.2)	1549 distinct values	: : : : . . : : : : . .	12686 (100.0%)	0 (0.0%)
13	inc86 [integer]	Mean (sd) : 7663.3 (9009) min < med < max: -5 < 4800 < 43119 IQR (CV) : 13000 (1.2)	1405 distinct values	: : : : . . : : : : . .	12686 (100.0%)	0 (0.0%)
14	inc87 [integer]	Mean (sd) : 9057.7 (10919.5) min < med < max: -5 < 6000 < 59387 IQR (CV) : 15000 (1.2)	861 distinct values	: : : : . . : : : : . .	12686 (100.0%)	0 (0.0%)
15	inc88 [integer]	Mean (sd) : 10227.7 (11608.6) min < med < max: -5 < 7300 < 57124 IQR (CV) : 17000 (1.1)	845 distinct values	: : : : . . : : : : : .	12686 (100.0%)	0 (0.0%)
16	inc89 [integer]	Mean (sd) : 11230.6 (12725.7) min < med < max: -5 < 8000 < 62836 IQR (CV) : 18900 (1.1)	916 distinct values	: : : : . . . : : : : . .	12686 (100.0%)	0 (0.0%)
17	inc90 [integer]	Mean (sd) : 12394.8 (14280.2) min < med < max: -5 < 9200 < 74283 IQR (CV) : 20000 (1.2)	852 distinct values	: : : : . : : : : : .	12686 (100.0%)	0 (0.0%)
18	inc91 [integer]	Mean (sd) : 11491.9 (15032.7) min < med < max: -5 < 5000 < 81481 IQR (CV) : 20005 (1.3)	938 distinct values	: : : : . : : : : .	12686 (100.0%)	0 (0.0%)
19	inc92 [integer]	Mean (sd) : 12077.5 (16296.1) min < med < max: -5 < 5000 < 90325 IQR (CV) : 21005 (1.3)	812 distinct values	: : : : . : : : : .	12686 (100.0%)	0 (0.0%)
20	inc93 [integer]	Mean (sd) : 12486.9 (17239.6) min < med < max: -5 < 3535 < 100948 IQR (CV) : 22005 (1.4)	784 distinct values	: : : : . : : : . .	12686 (100.0%)	0 (0.0%)
21	inc94 [integer]	Mean (sd) : 13043.5 (18008.2) min < med < max: -5 < 3500 < 101653 IQR (CV) : 22505 (1.4)	762 distinct values	: : : : . : : : . .	12686 (100.0%)	0 (0.0%)

Table 3: Results, Years 1 to 5

	Dependent variable:				
	(wage)				
	1 years	2 years	3 years	4 years	5 years
	(1)	(2)	(3)	(4)	(5)
time	$1,223.074^{***}$	$1,371.323^{***}$	$3,373.180^{***}$	$4,848.709^{***}$	$6,316.971^{***}$
	(25.035)	(176.089)	(188.095)	(208.797)	(223.875)
treat	$-1,655.430^{***}$	$-1,324.909^{***}$	$-1,499.818^{***}$	$-1,419.618^{***}$	$-1,443.384^{***}$
	(459.910)	(436.363)	(460.579)	(511.051)	(545.667)
Hispanic	327.919	83.220	9.125	84.467	197.353
	(278.296)	(228.737)	(244.300)	(270.577)	(288.517)
Black	$-1,693.144^{***}$	$-1,134.062^{***}$	$-1,354.560^{***}$	$-1,395.440^{***}$	$-1,483.181^{***}$
	(248.683)	(205.507)	(217.259)	(239.578)	(257.344)
age	$1,248.117^{***}$	$1,384.408^{***}$	$1,274.641^{***}$	$1,370.263^{***}$	$1,374.235^{***}$
	(53.928)	(44.155)	(47.692)	(52.650)	(55.990)
education	373.183^{***}	-27.771	142.503^{***}	140.002^{**}	205.021^{***}
	(64.209)	(51.698)	(55.119)	(61.427)	(65.213)
time:treat	327.460^{***}	981.667^{*}	606.690	1,040.740	656.918
	(81.567)	(587.049)	(621.029)	(689.000)	(737.135)
Constant	$-24,394.120^{***}$	$-23,060.420^{***}$	$-22,431.480^{***}$	$-24,490.450^{***}$	$-25,274.860^{***}$
	(1,001.659)	(826.298)	(888.634)	(980.273)	(1,040.306)
Observations	7,384	5,330	5,212	5,228	5,191
Log Likelihood	-77,570.180	-54,026.500	-53,110.010	-53,816.910	-53,776.320
Akaike Inf. Crit.	155,156.400	108,069.000	106,236.000	107,649.800	107,568.600

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4: Results, Years 6 to 10

	Dependent variable:				
	(wage)				
	6 years	7 years	8 years	9 years	10 years
	(1)	(2)	(3)	(4)	(5)
time	$7,507.641^{***}$	$8,694.245^{***}$	$10,103.670^{***}$	$11,174.160^{***}$	$12,233.430^{***}$
	(241.346)	(251.634)	(268.887)	(288.571)	(306.796)
treat	$-1,498.126^{**}$	$-1,514.372^{**}$	$-1,629.119^{**}$	$-1,649.473^{**}$	$-1,668.027^{**}$
	(586.777)	(606.792)	(639.834)	(679.204)	(713.796)
Hispanic	-241.762	-13.488	134.135	318.084	309.242
	(310.744)	(321.709)	(341.230)	(363.682)	(385.394)
Black	$-1,508.510^{***}$	$-1,446.748^{***}$	$-1,885.094^{***}$	$-1,770.546^{***}$	$-2,038.688^{***}$
	(275.999)	(286.963)	(305.086)	(326.611)	(346.765)
age	$1,363.261^{***}$	$1,362.844^{***}$	$1,309.051^{***}$	$1,262.087^{***}$	$1,283.497^{***}$
	(60.377)	(63.009)	(67.073)	(71.899)	(75.901)
education	278.434^{***}	354.199^{***}	467.255^{***}	457.824^{***}	500.508^{***}
	(70.225)	(73.889)	(78.634)	(85.014)	(90.210)
time:treat	$2,041.087^{**}$	$1,975.131^{**}$	$2,505.768^{***}$	$1,908.937^{**}$	$3,248.274^{***}$
	(796.656)	(830.372)	(882.814)	(954.417)	(1,017.066)
Constant	$-25,735.060^{***}$	$-26,596.840^{***}$	$-26,561.760^{***}$	$-25,500.390^{***}$	$-26,358.480^{***}$
	(1,123.685)	(1,166.773)	(1,248.945)	(1,335.500)	(1,415.395)
Observations	5,171	5,105	5,002	4,906	4,818
Log Likelihood	-53,944.780	-53,426.490	-52,613.270	-51,895.680	-51,203.500
Akaike Inf. Crit.	107,905.600	106,869.000	105,242.500	103,807.400	102,423.000
Note:					*p<0.1; **p<0.05; ***p<0.01