

ABORTION AND SELECTION

Elizabeth Oltmans Ananat, Jonathan Gruber, Phillip B. Levine, and Douglas Staiger*

Abstract—Abortion legalization in the early 1970s led to dramatic changes in fertility. Some research has suggested that it altered cohort outcomes, but this literature has been limited and controversial. In this paper, we provide a framework for understanding selection mechanisms and use that framework to both address inconsistent past methodological approaches and provide evidence on the long-run impact on cohort characteristics. Our results indicate that lower-cost abortion brought about by legalization altered young adult outcomes through selection. In particular, it increased likelihood of college graduation, lower rates of welfare use, and lower odds of being a single parent.

I. Introduction

THE legalization of abortion in the United States in the early 1970s represents one of the most important changes in American social policy in the twentieth century. This policy change had obvious implications for the likelihood of giving birth in the case of an unintended pregnancy, and resulted in a drop in birth rates. In addition, we will demonstrate that it led to a rise in pregnancy rates. As a result of both of these changes, abortion legalization may have altered the characteristics of birth cohorts. In particular, children's outcomes may have improved on average, because they were more likely to be born into a household in which they were wanted.

Two earlier papers have investigated the implications of such positive selection through abortion for the quality of cohorts born after abortion legalization, but have used different sources of variation that, implicitly, relied on different aspects of the selection process. Gruber, Levine, and Staiger (GLS, 1999) examined the living circumstances of children, and focused on the reduction in births caused by initial legalization. Donohue and Levitt (DL, 2001) examined crime among youths,¹ focusing on differences in abortion rates between states after legalization.

This paper builds on those results by considering the impact of abortion access on a broad array of cohort characteristics in adulthood. Since the relevant cohorts were born in the early 1970s, children born at that time are now in their thirties. Using the 2000 Census, we can examine adult characteristics, such as completed educational attainment, employment, and poverty status.

Received for publication June 2, 2006. Revision accepted for publication November 1, 2007.

* Departments of Public Policy and Economics, Duke University, and NBER; Department of Economics, MIT, and NBER; Department of Economics, Wellesley College, and NBER; and Department of Economics, Dartmouth College, and NBER, respectively.

We are grateful to John Donohue, Larry Katz, Steve Levitt, and Ted Joyce for helpful discussions, to an anonymous referee, seminar participants at NBER Summer Institute, and various colleges and universities for their comments; and to Steve Levitt for providing crime data. Gruber, Levine, and Staiger also gratefully acknowledge funding from NICHD (grant number R01 HD042819).

¹ Other papers addressing the crime issue followed and will be discussed subsequently.

Beyond evaluating outcomes in adulthood, this paper also makes important contributions by addressing the inconsistent methodological approaches and the implicit models of selection used in past research. The approach of GLS used, as motivation, that (i) after five states repealed anti-abortion laws in 1970, relative birth rates fell in those states; (ii) after the 1973 *Roe v. Wade* decision, birth rates converged between early legalizing states and other states. They treated these two changes as a two-part experiment in reducing the rate at which unintended pregnancies translated into births. DL, on the other hand, noted that (i) abortion rates continued to rise after 1973, and that (ii) the abortion rate remained persistently higher in the early legalizing states even after widespread legalization. They used these diverging trends in the abortion rate as a measure of variation in the overall level of selection into pregnancy and birth.

In this paper, we note that the post-1973 combination of (i) converging birth rates and (ii) persistent differences in abortion rates implies that in the latter half of the 1970s increased abortions did not map perfectly into fewer births. Instead, it must be the case that pregnancy rates rose in early legalizing states relative to other states in the late 1970s. Since the approaches used by GLS and DL rely on different changes in reproductive behavior, they are not equivalent. We introduce a comprehensive model of selection that identifies the two distinct mechanisms—increased pregnancies and decreased births—by which abortion access may alter cohort characteristics. We argue that in the first part of the decade, decreased births in early legalizing states, out of a steady number of unwanted pregnancies, were the dominant force in the selection process. In the late 1970s, on the other hand, we argue that a larger pool of pregnancies in repeal states, from which a steady number of births were selected, means that early legalizing states could continue to have more positively selected cohorts of children even after widespread legalization.

We test this model empirically by using a methodology that incorporates both the legalization variation used by GLS and the variation in “taste” for abortions across states, which DL relied on implicitly, to credibly identify selection effects. We find evidence of selection effects of abortion on young adult outcomes. In particular, lower costs of abortion led to an increase in the rate of college graduation, lower odds of welfare receipt, and lower odds of being a single parent. We are unable to sharply distinguish the mechanisms through which selection occurs, but our results are robust to the empirical methodology employed.

Our paper proceeds as follows. Section II gives background on abortion legalization and reviews previous work on its effects. Section III provides a model of the mechanisms through which abortion affects selection of birth cohorts. Section IV describes the methodology for our

study. Section V discusses the data. Section VI presents the results. Section VII concludes.

II. Background

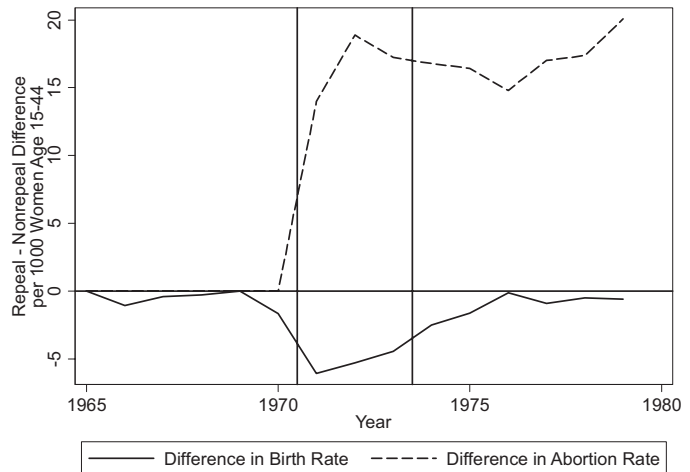
As detailed elsewhere (Garrow, 1994), abortion was first legalized in five states (Alaska, California, Hawaii, New York, and Washington) in 1970, three years ahead of the *Roe v. Wade* decision in 1973. Levine et al. (1999) used this state-level variation in the timing of legalization to show that births fell considerably in response to the law change. GLS, using this birth rate variation, determined that abortion legalization improved outcomes for those born in the early repeal states in the 1971–1973 period, relative to other states, and that this relative improvement faded afterwards. Estimating the characteristics of those children who would have been born had abortion not been legalized (the “marginal” child), they found their outcomes would have been worse than the average outcomes of those children who were born. The marginal child would have been 40% to 60% more likely to live with a single parent, to be poor, to be in a household collecting welfare, and to die during the first year of life.

DL hypothesized that, if fewer unwanted children were born when abortion rates increased, then crime may have decreased when those children would have reached adulthood.² Among other methods, they regressed state/year-of-birth arrest rates against the abortion rate in the state and year in which an individual was born. They concluded that increased abortion access in the 1970s may explain as much as half of the decline in crime observed in the 1990s.

The DL study has generated a great deal of controversy. Critiques by Joyce (2004a, 2004b) and Foote and Goetz (2008) have raised important questions regarding the empirical analysis, to which Donohue and Levitt (2004, 2008) have responded. Although it is beyond the scope of this paper to fully elaborate upon all the points raised in this debate, we do want to focus on Joyce’s criticism of their identification strategy. He argues that including the abortion rate on the right-hand side of the regression does not accurately gauge variation in unwanted births, since the abortion rate is endogenous, and only the variation in legalization is exogenous. Joyce reestimated the Donohue and Levitt model using the double quasi-experiment implemented by Levine et al. (1999) and GLS and obtained results that he argued were inconsistent with a causal interpretation.

Donohue and Levitt disputed this reading of Joyce’s findings, however, and highlighted a potential weakness of

FIGURE 1.—DIFFERENCE BETWEEN REPEAL AND NONREPEAL STATES IN THE BIRTH RATE AND ABORTION RATE PER 1,000 WOMEN AGE 15–44, 1965–1979



the GLS identification strategy: abortion rates do not correspond well to the “experiment” proposed by GLS. This is illustrated in figure 1, which plots the differences in birth rates and in abortion rates between repeal states and nonrepeal states from 1965 to 1979.³ Changes in birth rates correspond well to a natural experiment: birth rates in repeal states fell relative to nonrepeal states in 1971, but the gap disappeared by the mid-1970s once abortion was legalized nationwide. However, abortion rates do not follow this pattern. The figure imposes a zero difference in abortion rates prior to 1970, due to data limitations; in 1971, the first full year in which abortion was legal in all repeal states, abortion rates in those states jumped dramatically compared with the rest of the country (note that the abortion rate becomes nonzero in nonrepeal states as well, since women could travel to repeal states to obtain an abortion). While this first divergence in abortion rates corresponds well to the experiment proposed by GLS, it never dissipated, contrary to the assumptions of the two-part experiment. Instead, women in repeal states continued to use abortion at much higher rates than women in the rest of the country throughout the 1970s, suggesting that differential selection may have continued as well.

III. A Model of Selection

To distinguish the different ways that abortion access may lead to selection, we introduce a theoretical model that describes the potential effects of changes in the cost of abortion. This simple model of decision-making under uncertainty is closely related to those introduced by Kane and Staiger (1996) and Levine and Staiger (2002), which were designed to examine how changes in abortion cost would affect fertility decisions of a woman facing uncertain payoff to giving birth. We extend that here to explicitly incorporate heterogeneity across women in the expected payoff to

² We are aware of two other contributions to the literature on abortion and children’s outcomes other than crime. Charles and Stephens (2006), using methods similar to GLS, find that legalized abortion led to a significant reduction in drug use. Pop-Eleches (2006), exploiting a sudden ban on abortions imposed in Romania, finds that, conditional on mother characteristics, children born after the ban had worse economic outcomes as adults.

³ The data used to generate this figure are described subsequently.

giving birth, to allow for an additional margin of selection: selection into pregnancy.

A. A Simple Model

We begin by briefly summarizing our assumptions regarding the decision-making process. First, a woman makes her decision about pregnancy based on the expected payoff that she anticipates at the time of becoming pregnant; this expected payoff varies across women. Second, a woman makes her pregnancy and abortion decisions sequentially; she first chooses whether to become pregnant (*ex ante*), and then after some time has elapsed she chooses whether to abort or give birth (*ex post*). Third, a woman makes her *ex post* decision about abortion after updating her expected payoff of giving birth, because by the time of the abortion decision she is better informed about the consequences of a birth than she was at the time when she chose to become pregnant. Thus, in our model, abortion differs from other methods of avoiding a birth (such as contraception or abstinence) because the abortion decision is made with more complete information than the pregnancy decision.⁴ Finally, we assume that children's outcomes are directly linked to the payoff of giving birth. Intuitively, we assume that more "wanted" (that is, higher payoff) births have better outcomes than less wanted births.

The formal details of the model are presented in Ananat et al. (2006), but the main implications are intuitive. Within this simple model, lowering the cost of abortion affects fertility decisions on two margins. First, more women will become pregnant because the down-side risk of a pregnancy is less costly should they want an abortion after becoming pregnant. These women are essentially paying a price (the potential abortion cost) to buy the option to give birth after they have learned more about the birth payoff. The pregnancies that are added on this margin we refer to as "marginal pregnancies." The average payoff to births that result from these marginal pregnancies will be lower than that for nonmarginal pregnancies, corresponding to the fact that women with marginal pregnancies are indifferent about getting pregnant *ex ante*, while all other women strictly prefer pregnancy *ex ante*. Thus, all else equal, the addition of marginal pregnancies will increase the number of births but lower the average outcome of births.

Lowering the cost of abortion will also affect fertility decisions on the abortion margin *ex post*. Among the women who become pregnant, lowering the cost of abortion will increase the probability of abortion and reduce the probability of birth. The births that are eliminated on this margin we refer to as "marginal births." The payoff to these marginal births will be lower than that for all other births, corresponding to the fact that women with marginal births

are indifferent between having an abortion and giving birth *ex post*, while all other women strictly prefer giving birth *ex post*. Thus, all else equal, the elimination of marginal births will decrease the number of births but raise the average outcome of births.

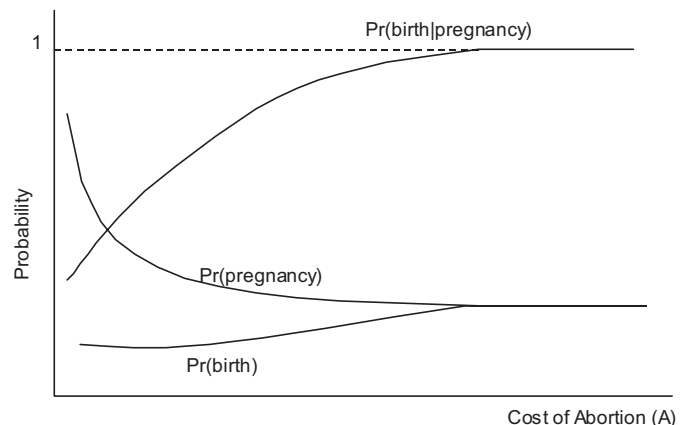
B. Empirical Implications of the Model

Figure 2 illustrates the implications of this model for pregnancy and birth rates. On the right-hand side of the diagram, where the cost of abortion is very high (for example, abortion is illegal), no women will choose to abort and all pregnancies will end in a birth; in other words, $\Pr(\text{birth}|\text{pregnancy}) = 1$ and $\Pr(\text{pregnancy}) = \Pr(\text{birth})$. Lowering the cost of abortion will increase the pregnancy rate (through the addition of marginal pregnancies) and will reduce the proportion of pregnancies that end in a birth (through the elimination of marginal births by abortion). Thus, lowering the cost of abortion will unambiguously increase the number of pregnancies and abortions.

However, since these two increases work in opposite directions, the net effect of lowering the cost of abortion on the birth rate is ambiguous. We have chosen to draw figure 2 so that lower costs result in a lower birth rate until costs become very low, at which point the birth rate flattens out. This would be the case if relatively few marginal pregnancies are induced until the cost of abortion becomes quite low; for example, women are unwilling to become pregnant in hopes of receiving positive news about the payoff to a birth *unless* abortion is very low-cost.

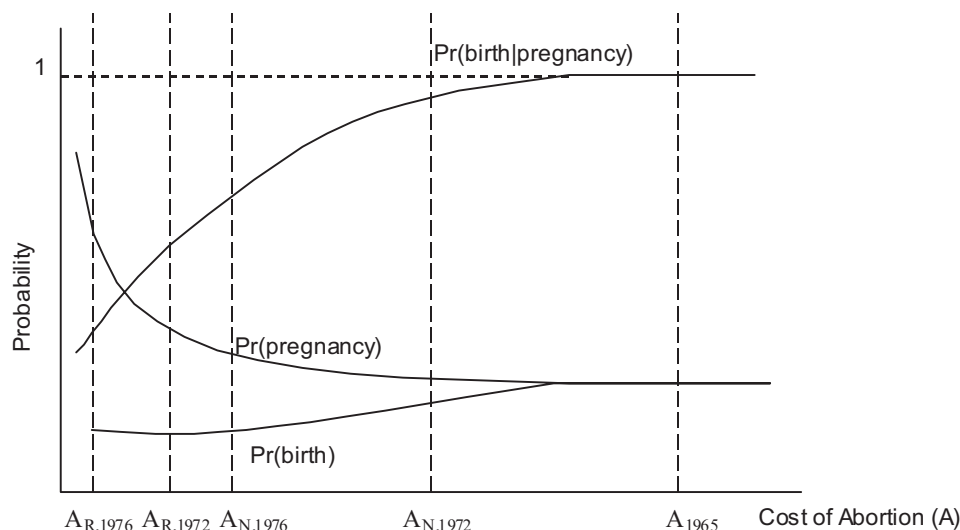
Figure 3 uses figure 2 to provide a possible explanation for why birth rates in repeal and nonrepeal states diverged from 1971 to 1973 but then converged by 1976, while abortion and pregnancy rates remained divergent throughout the 1970s. In the 1960s (we use 1965 as an example), the cost of abortion, A_{65} , was very high in all states. Then, from the 1960s to the early 1970s (such as 1972), the cost of abortion fell dramatically in the early repeal states (to $A_{R,72}$), while the costs fell only modestly in the nonrepeal

FIGURE 2.—RELATIONSHIP BETWEEN COST OF ABORTION AND PROBABILITY OF PREGNANCY AND BIRTH



⁴ In fact, there is considerable evidence that information obtained after becoming pregnant (for example, support from parents or the father, health problems of mother or fetus) is an important determinant of the abortion decision (Bankole, Singh, & Haas, 1998; Torres & Forrest, 1988).

FIGURE 3.—IMPLICATIONS FOR THE PROBABILITY OF PREGNANCY AND BIRTH, GIVEN THE LIKELY COST OF ABORTION IN REPEAL (R) AND NONREPEAL (N) STATES IN 1965, 1972, AND 1976



states (to $A_{N,72}$), partly due to travel to the repeal states for abortion. From the early to mid-1970s (such as 1976), the cost of abortion continued to fall in both sets of states, through legalization in the nonrepeal states and increased access in the repeal states. This led the costs to move to $A_{R,76}$ and $A_{N,76}$ respectively. Such a difference in abortion costs between states even when abortion is universally legal could result from differences in the social acceptability of the procedure, for example, which could lead to lower access to abortion providers or generalized stigma against the procedure.

This set of changes is consistent with the empirical evidence from figure 1, as well as other evidence that abortion costs were lower in repeal states. Birth rates post-1976 looked similar despite differences in abortion rates, and our model offers an explanation that the greater number of terminated “marginal births” in the repeal states was cancelled out by the greater number of “marginal pregnancies” that were subsequently carried to term. If this is accurate, then birth cohorts in repeal states would be made up of fewer marginal births and more births from marginal pregnancies. Since the payoffs to marginal births are lower than payoffs to all other births (including those resulting from marginal pregnancies), cohort “quality” in repeal states would then continue to improve relative to that in nonrepeal states in the mid- to late 1970s.

To summarize, our model suggests that cohort quality improves under two circumstances. If marginal births are reduced with little or no change in marginal pregnancies, then cohort quality will rise. This is what likely occurred at the time of initial abortion legalization. But cohort quality may also rise if, holding cohort size steady, pregnancy rates rise so that selection into the cohort increases. This is likely what took place in the years after abortion legalization.

IV. Empirical Methodology

Two empirical approaches have been used to document the relationship between abortion access and child outcomes in a state-year birth cohort. Both approaches run regressions in which the dependent variable is the average outcome in a birth cohort and include a comprehensive set of controls. One approach (GLS) uses an IV regression of the average outcome in a birth cohort on the log birth rate in the state and year of birth, instrumenting for the birth rate with changes in the legal status of abortion within a state. A second approach (DL) uses an OLS regression of the average outcome in a birth cohort on the ratio of abortions to births in the state and year of birth. In this section we propose a more general empirical framework that encompasses both approaches as special cases and that separately estimates the impacts of the marginal birth and the marginal pregnancy.

A. Specification

The equation estimated by GLS took the following form:

$$\ln(OUTCOME_{st}) = \alpha_1 \times \ln(BIRTHRATE_{st}) + \text{controls} + \varepsilon_{st}. \quad (1)$$

The dependent variable, $OUTCOME_{st}$, is a measure of the average outcomes of those born in state s in year t . For simplicity, we define all outcomes so that they reflect lower socioeconomic status of the cohort (for example, the proportion of the cohort living in poverty or the proportion that did not graduate from college). The key explanatory variable, $\ln(BIRTHRATE_{st})$, is the log birth rate to childbearing-age women in the cohort’s birth year and state. The equation includes generic controls for the otherwise unobservable

characteristics that differ between states or over time, including state and cohort dummies, quadratic trends by state of birth, and controls for economic and demographic conditions in the state and year of birth.

Taking the derivative of equation (1) with respect to the log birth rate shows that α_1 is an estimate of the gap between the outcome for the marginal birth and the “average birth” (defined in the model as the expected payoff to birth from an average pregnancy) in the cohort, stated in percentage terms.⁵ Our theoretical model suggests that α_1 should be positive; the marginal birth is more likely than the average birth to live in poverty, not graduate from college, and so forth.

In the context of our model, equation (1) imposes an important restriction by not acknowledging the potential role of the marginal pregnancy. An increase in pregnancy and abortion rates that leaves the birth rate unchanged, as seen in figure 1, cannot affect the average outcome of the birth cohort according to this specification. Equation (1) implicitly assumes that outcomes are the same for the marginal births that are avoided due to legalization and the marginal pregnancies that occur after legalization, while our theoretical model suggests that outcomes of marginal pregnancies will be better than those of the marginal births they are replacing. This is a potential problem for the identification strategy used by GLS.

A more general specification estimates the impacts of the marginal birth and marginal pregnancy separately:

$$\begin{aligned} \ln(OUTCOME_{st}) &= \beta_1 \times \ln(PREGRATE_{st}) \\ &+ \beta_2 \times \ln(BIRTHRATIO_{st}) \\ &+ \text{controls} + \epsilon_{st}. \end{aligned} \quad (2)$$

The variable *PREGRATE* is the number of pregnancies per childbearing-age woman in the cohort’s birth year and state, and *BIRTHRATIO* is the births to pregnancies ratio. Using the same argument that led to the interpretation of α_1 , β_1 is an estimate of the difference in outcomes between births resulting from the marginal pregnancy and births resulting from the average pregnancy, while β_2 is an estimate of the difference in outcomes between the marginal birth and the average birth.⁶ This specification illustrates that GLS imposed $\beta_1 = \beta_2$, that is, that the marginal pregnancy and the marginal birth had the same outcomes. Our theoretical model, however, suggested that outcomes from the marginal birth should be worse than the outcomes for the marginal pregnancy, that is, $0 < \beta_1 < \beta_2$ for a negative outcome such as living in poverty.

⁵ See GLS for this derivation.

⁶ Note that we measure β_2 using the birth ratio, rather than the birth rate, because this specification estimates the effect of a change in births holding the pregnancy rate constant, that is, the birth ratio. When we move to two-equation versions of this specification in equations (6) and (7), we will use the more intuitive birth rate.

Interestingly, equation (2) also includes the specification run by DL as a special case. To see this, note that one can rewrite the log of the birth ratio as

$$\begin{aligned} \ln(BIRTHRATIO) &= \ln(\text{births/pregnancies}) \\ &= -\ln(\text{pregnancies/births}) \\ &= -\ln((\text{births} + \text{abortions})/\text{births}) \\ &= -\ln(1 + \text{abortions/births}) \\ &\approx -(\text{abortions/births}). \end{aligned} \quad (3)$$

The final expression in equation (3) is the abortion ratio,⁷ the variable used by DL.⁸ Thus, equation (2) can be approximately rewritten as

$$\begin{aligned} OUTCOME_{st} &= \beta_1 \times \ln(PREGRATE_{st}) \\ &+ -\beta_2 \times ABORTRATIO_{st} \\ &+ \text{controls} + \epsilon_{st}. \end{aligned} \quad (4)$$

DL estimated this equation restricting $\beta_1 = 0$. Their restriction implicitly assumed that the birth outcome for the marginal pregnancy was the same as the outcome for the average birth.

Thus, the difference between the DL and GLS specifications rests on which restrictive assumption each made about the birth outcomes of the marginal pregnancy. In the more general specification of equation (2), the impacts of the marginal birth and marginal pregnancy are estimated separately, and the assumptions maintained in earlier papers can be tested directly.

B. Estimation

Estimates of equation (2) by OLS will be biased if much of the variation in pregnancy rates and birth ratios was not the result of changes in the cost of abortion. GLS instrumented for the birth rate using variation in the legal status of abortion across states, because variation in the birth rate was potentially driven by factors other than the cost of abortion, such as transitory economic improvements or unobserved improvements in the expected outcome of births. GLS also estimated reduced-form models using this specification:

⁷ Note that this calculation proxies pregnancies with the sum of births and abortions, implicitly making the standard assumption that miscarriages are exogenous. We continue this assumption in the empirical estimates that follow.

⁸ DL refer to the ratio of abortions to births as the abortion rate, but the abortion ratio is the conventional name for this statistic. The abortion rate typically refers to abortions performed per 1,000 women of childbearing age.

$$\begin{aligned} \ln(OUTCOME_{st}) = & \rho_1 Repeal_s \times D7173 \\ & + \rho_2 Repeal_s \times D7475 + \rho_3 Repeal_s \times D7679 \quad (5) \\ & + \text{controls} + \varepsilon_{st}. \end{aligned}$$

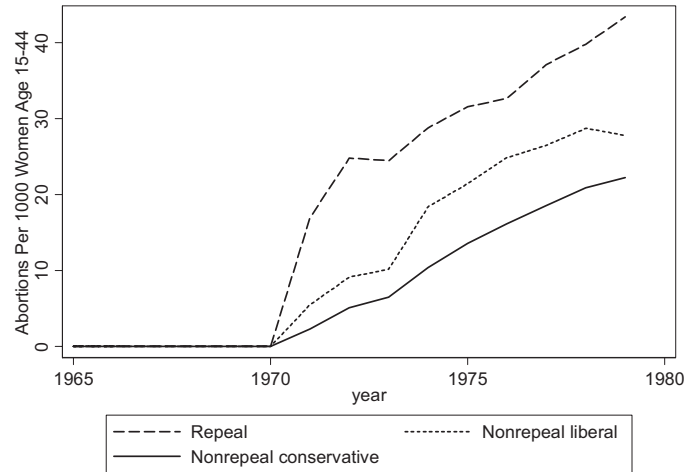
where *Repeal* is a dummy variable for states that are early repealers of abortion restrictions; D7173, D7475, and D7679 are dummies for the eras 1971–1973, 1974–1975, and 1976–1979; and controls are as above. This model serves to examine whether any gap in outcomes that emerged in 1971–1973 across these groups of states closed afterward when abortion was legal everywhere. That is, if abortion legalization improves average cohort outcomes through selection, then cohort outcomes should improve in the repeal states in 1971–1973 ($\rho_1 < 0$, since outcomes are negative), but then “catch up” in other states after 1973 ($\rho_2, \rho_3 \approx 0$).

While GLS used legalization alone as an instrument for the cost of abortion, the trends in figure 1 suggest that even after national legalization there continued to be differences in abortion cost between repeal and nonrepeal states; indeed, as we describe below, such differences exist even within the set of nonrepeal states. DL used these differences in estimating the relationship between outcomes and the abortion rate by OLS, implicitly assuming that abortion rate variation was driven by changes in the cost of abortion. Instead of relying on such an assumption, we extend the GLS instrumental variables strategy to explain heterogeneity in the cost and use of abortion across states and years beyond legalization.

We extend the set of instrumental variables used to predict abortion costs along two dimensions. The first dimension is the travel distance to the nearest state in which abortion was legal during the period after early repeal and before national legalization (1970–1973).⁹ The second is the “latent cost” of abortion in each state after national legalization. Legalization loosened a constraint on abortion demand, but there was variation across states in the extent to which the constraint had been binding; in places with high latent costs, abortions would not be heavily demanded even if legal. Such latent costs could be a function of many factors, but certainly an important one is social attitudes toward the use of abortion: higher social opprobrium on abortion use raises its psychic cost and reduces its use. In the states where abortion carries greater social stigma, we expect abortion rates to be lower after legalization.

The notion that social norms matter is strongly supported by our data. Figure 4 plots abortion rates in repeal and nonrepeal states over time, dividing the nonrepeal states into two groups: “socially liberal” states with at least 22% of the population reporting they were “liberal” in opinion polls

FIGURE 4.—TRENDS IN ABORTION RATE PER 1,000 WOMEN AGE 15–44 IN REPEAL STATE AND LIBERAL AND CONSERVATIVE NONREPEAL STATES, 1965–1979



from the 1960s, and “socially conservative” states with less than 22% “liberal.”¹⁰ Seven of the nonrepeal states fall into the liberal category. As can be seen in figure 4, repeal states had the highest abortion rates after 1973. Abortion rates were lowest in socially conservative states, while socially liberal nonrepeal states were in the middle.

We therefore include in the first-stage equation for the pregnancy rate the interactions between repeal status and year dummies, which were used as instruments by GLS, as well as additional interactions with travel distance and latent cost of abortion to capture the heterogeneity of abortion cost within nonrepeal states after legalization:

$$\begin{aligned} \ln(PREGRATE_{st}) = & D7173 \times Nonrepeal_s \times (\rho_1 + \rho_2 LC_s + \rho_3 DIST_s) \\ & + D7475 \times Nonrepeal_s \times (\rho_4 + \rho_5 LC_s) \quad (6) \\ & + D7679 \times Nonrepeal_s \times (\rho_6 + \rho_7 LC_s) \\ & + \text{controls} + \varepsilon_{st}. \end{aligned}$$

Nonrepeal_s is a dummy for a cohort born in a nonrepeal state; era dummies are defined as above; *DIST_s* is the average straight-line distance from state *s* to the nearest repeal state; and *LC_s* is a measure of latent cost of abortions (operationally defined below). Both *DIST* and *LC* have been rescaled so that they range from 0 to 1: 0 represents a state with the lowest distance (the repeal states) or the lowest latent cost (New York), and 1 represents a state with the highest distance (Louisiana) or highest latent cost (Mississippi). All of the repeal states have a distance of 0 and latent cost near 0, so there is no need to include interactions between repeal and these variables in the specification. Similarly, all states have a distance of 0 after 1974, so there

⁹ The notion that travel costs matter for abortion access is intuitive and is supported by the work of Levine et al. (1999) and Kane and Staiger (1996). Both papers find that the birth rate is affected by distance to the nearest legal state (in the 1970–1973 period) or to the nearest abortion provider (thereafter).

¹⁰ The opinion data that were used to identify conservative and liberal states are described subsequently.

is no need to include distance interacted with D7475 or D7679. The main effects of distance, latent cost, and repeal status are all absorbed by the state of birth fixed effects.

We hypothesize that the coefficients on all three interactions with D7173 should be negative ($\rho_1, \rho_2, \rho_3 < 0$), since higher costs of abortion should reduce pregnancy rates. After abortion is legalized in all states, we expect that women in states that had the highest latent costs were the least likely to have an abortion, implying that $\rho_5 < 0$ and $\rho_7 < 0$. Controlling for latent cost, we expect the difference in abortion costs between nonrepeal and repeal states to decline during the transition years 1974–1975, as abortion providers gradually entered these states following *Roe* (implying that $\rho_1 < \rho_4 < 0$). We expect no difference in abortion costs, other than that driven by latent costs, between nonrepeal and repeal states after 1976 ($\rho_6 = 0$).¹¹

Conceptually, the cost of abortion is a single variable that is being used to instrument for two different variables, leaving equation (2) unidentified. However, our model suggests that changes in abortion cost may have nonlinear effects that will differ between the two endogenous variables, and our set of instruments provides identification by capturing this nonlinearity. We hypothesize that going from illegal abortion to legal abortion with high latent cost may have a large effect on the birth ratio but little effect on the pregnancy rate (because the abortion option is so costly); thus legality is a plausible instrument for the birth ratio. In contrast, going from high latent cost to low latent cost may have a much larger relative effect on the pregnancy rate, as the down-side risk of a pregnancy is all but eliminated; thus social liberalism is a plausible instrument for the pregnancy rate.

In practice, results based on this approach were imprecise. As a result, we estimate each of the following equations (which represent restricted versions of equation (2) that are consistent with past models) separately, using OLS as well as the full set of instruments just described:

$$\begin{aligned} \ln(OUTCOME_{st}) \\ = \delta_1 \times \ln(BIRTHRATE_{st}) + \text{controls} + \epsilon_{st}. \end{aligned} \quad (7)$$

$$\begin{aligned} \ln(OUTCOME_{st}) \\ = \gamma_1 \times \ln(BIRTHRATIO_{st}) + \text{controls} + \epsilon_{st}. \end{aligned} \quad (8)$$

Equation (7) is identical to equation (1), and represents a restricted form of equation (2) in which $\beta_1 = \beta_2$ (the marginal pregnancy and the marginal birth have the same outcomes, analogous to GLS). Equation (8) is very similar to equation (4), but imposes the restriction that $\beta_1 = 0$ (the birth outcome for the marginal pregnancy is the same as the outcome for the average birth, analogous to DL).

¹¹ The first-stage equation for the birth ratio is analogous to equation (5), except that the expected signs of the coefficients in this equation are the opposite of the pregnancy-rate equation (since higher abortion costs reduce pregnancies but raise the birth ratio).

Although these restrictions make it impossible to test the full implications of the model, these specifications allow us to bound the true difference between the marginal birth and the average birth (β_2 in equation [2]). Equation (7) leads to an *overstatement* of the difference between the marginal birth averted by abortion and average birth because it attributes all cohort selection to the change in “net births”; it ignores the fact that an increase in the number of marginal pregnancies carried to term that is offset by a decrease in marginal births from unwanted pregnancies will also improve selection despite having no effect on the birth rate.¹² Equation (8), meanwhile, leads to an *understatement* of the difference between the marginal and average birth by ignoring the fact that the newly induced marginal pregnancies that are carried to term will tend to lower the average birth outcome.¹³ Thus, while we cannot estimate equation (2), we are nonetheless able to bound the true selection effect for the marginal birth.

V. Data

To estimate these models we use several different sources of data. Our main source is the 2000 decennial Census of the United States.¹⁴ We use it to measure a variety of outcomes for individuals born in a given state/year: whether the individual lives in a household that is below the poverty line, is receiving welfare, or is a single parent; whether the individual dropped out of high school, or did not graduate from college; whether the individual is incarcerated;¹⁵ and whether the individual is not employed. The Census provides two important advantages over other data sets containing similar outcome measures. First, it identifies state of birth rather than merely state of residence. If migration from one’s childhood state is related to other outcomes, then state of residence will be a biased measure of abortion availabil-

¹² This is a direct application of omitted variable bias. To see this, rewrite equation (2):

$$\begin{aligned} \text{Outcome}_{st} &= (\beta_1 - \beta_2) \ln(PREGRATE_{st}) + \beta_2 \ln(BIRTHRATE_{st}) \\ &\quad + \text{controls} + \epsilon_{st}, \\ \text{since } \ln(BIRTHRATE) &= \ln(PREGRATE) + \ln(BIRTHRATIO). \end{aligned}$$

Omitting $\ln(PREGRATE)$ from this produces equation (6). According to our model, $\beta_1 - \beta_2 < 0$ and $\text{cov}(PREGRATE, BIRTHRATE) < 0$, so the omitted variable bias is positive. Since $\beta_2 > 0$ for negative outcomes, this overstates the coefficient on $\ln(BIRTHRATE)$.

¹³ This is also a direct application of omitted variable bias. Equation (7) estimates $OUTCOME_{st} = \gamma_1 \times \ln(BIRTHRATIO_{st}) + \text{controls} + \epsilon_{st}$, omitting the pregnancy-rate term, $\ln(PREGRATE_{st})$, from equation (2). Since $\beta_1 > 0$, $\beta_2 > 0$, and $\text{cov}(PREGRATE, BIRTHRATIO) < 0$, the OLS estimate of γ_1 will be a downward-biased estimate of β_2 .

¹⁴ All Census data used were taken from the Minnesota Integrated Public Use Microdata Series (Ruggles, Sobek et al., 2003); only unallocated observations were used.

¹⁵ The Census only provides data on institutionalization, not incarceration per se. But past evidence suggests that the vast majority of institutionalized young adults are incarcerated. In the 1980 Census (the most recent Census that provides detailed institutionalization data), 68% of those aged 20 to 35 who were institutionalized were incarcerated. This rate is likely to be much higher today since incarceration rates have nearly quadrupled (U.S. Department of Justice, 2003), while the number in mental institutions declined (Grob, 2001).

ity at birth. Second, it offers large sample sizes. As noted in GLS, even large changes in the outcomes of the marginal child will result in relatively small changes in the outcome of the average child, which is our unit of observation. The 5% sample provides the precision to identify these small changes.

Our sample includes those born in the United States and observed in the 2000 Census at ages 21 to 35 in that year (born between 1965 and 1979).¹⁶ A unit of observation represents mean values for each state/year of birth cohort. The means are weighted by the person weights provided by the Census Bureau, and the cell sizes are used as weights in the regression analysis. This data structure allows us to control for both age and state of birth fixed effects in our model.

One drawback to the use of the decennial Census is that we only observe each cohort once in adulthood, in the year 2000. As a result, we cannot separate age and cohort effects within states (national age effects are captured by a full set of age dummies). To address this problem, we allow quadratic variation in age effects by state in all specifications, as in GLS. That is, we assume that any state-specific age patterns are captured by a quadratic trend in age, and that any remaining differences across states by cohort reflect true cohort effects.

We supplement Census data on cohort adult outcomes with two measures of cohort birth circumstances in 1965–1979: the percentage of all infants born to minors (under the age of 18), and the percentage born to nonwhite mothers. These measures are derived from Vital Statistics data from the natality detail files for each year between 1965 and 1979. Vital Statistics outcomes are based on all births, and tend to produce more precise estimates than outcomes based on the 5% sample from the 2000 Census.

Key explanatory variables in our model require data on abortions and births. Birth rates by year and state of residence come from Vital Statistics. Abortion rates by year and state of residence come from the Alan Guttmacher Institute (AGI) beginning in 1973. These data are generally regarded as the best available since they reflect the results of surveys of large and small abortion providers that report the counts of abortions performed at their site. We use AGI data for which an algorithm has been employed that converts data arranged by state of occurrence to one in which it is arranged by the mother's state of residence.¹⁷ We then augment these data with additional data reported by the Centers for Disease Control for the period 1970 through 1972 (Centers for Disease Control, 1971, 1972, 1974). These data, collected through the Vital Statistics system, are believed to include some undercount of the total abortions performed (cf. Saul, 1998), but they are the best data

available for this period. We impose an abortion rate of 0 for years prior to 1970.¹⁸

We use three different measures of the cost of acquiring an abortion as instrumental variables. The first is whether the cohort was born in one of the five early repeal states (California, New York, Washington, Alaska, Hawaii).¹⁹ The second is the average straight-line distance to the nearest county in which abortion is legal (calculated for each county in a state and averaged, weighting by county population). Distance has been rescaled to go from 0 to 1, and is 0 after 1973 and in repeal states.²⁰ Third, we use two different measures of the latent social cost of abortion. While the ideal variable would be a measure of tastes for abortion in the pre-legalization period that would capture the extent to which women were constrained by social costs when abortion was illegal (and therefore the extent to which they might take advantage of it upon legalization), unfortunately, such data are not available. Instead, we use two different measures of states' social climates prior to 1970 as proxies for this ideal measure.

Our first proxy instrument uses data on state political attitudes compiled from 1960s state-level voter surveys conducted by Louis Harris and Associates.²¹ In 38 states, at least one Harris poll of a representative sample of voters was conducted in 1962, 1964, or 1966 that asked the question, "What do you consider yourself—conservative, middle of the road, or liberal?" We use the fraction of self-identified liberals as one measure of latent costs: where the population is more liberal, the social costs of attaining a legal abortion post-repeal are lower.²²

Our second proxy is a measure developed by Levine (2004) to capture illegal abortion rates by state before 1970: in places where there was more demand for illegal abortion, the latent costs of abortion were lower. He employs **retrospective data from the 1982 and 1988 National Surveys of Family Growth** and estimates abortion rates by state for the 1965–1969 period (one observation for each state, aggregating over years to help overcome small sample sizes). These data suffer from recall bias and the general reluctance to report abortions (perhaps even more so if performed

¹⁸ We have also undertaken an exercise designed to test the sensitivity of our results to imposing abortion rates of 0 for the pre-1970 period. We used retrospective data on illegal abortion to fill in the missing values for the abortion rate. The results of this exercise are very similar to those reported below, with coefficient estimates generally 10%–20% larger in magnitude.

¹⁹ We have also considered other proxies for the real costs of obtaining abortion. One alternative measure is distance to the nearest state in which abortion is legal; another measure is distance to the nearest county with an abortion provider. When we replicate our analysis using either of these variables, the results are consistent with those reported here.

²⁰ Using AGI data on the location of counties with abortion providers, and calculating distance to the nearest abortion provider as in Kane and Staiger (1996), yields very similar results. Because of the potential concern about endogenous location of abortion providers, we do not report these results.

²¹ Historical Harris poll data are available from the data archive of the Odum Institute for Research in Social Science, at http://www.irss.unc.edu/data_archive/home.asp.

²² In some later polls, the option "radical" was added to the survey question. Respondents who identified themselves as "radical" are grouped with those who identified as "liberal" for the purposes of our analysis.

¹⁶ In models of college graduation and poverty status, we have restricted the sample to those between the ages of 24 and 35 in 2000 (born between 1965 and 1976), because college graduation rates are very low and poverty rates are very high for those in their early twenties, making them difficult to capture using quadratic age trends.

¹⁷ The procedure for doing so is described in Levine (2004).

TABLE 1.—IMPACT OF ABORTION LEGALIZATION ON 1965–79 VITAL STATISTICS FERTILITY OUTCOMES

	Log Birth Rate (1)	Log Pregnancy Rate (2)	Log Birth Ratio (3)	Log Survivor Rate (4)	Log % Born to Minor (5)	Log % Born to Nonwhite Mother (6)
Repeal \times 1971–1973	−0.057 (0.012)	0.129 (0.044)	−0.187 (0.035)	−0.043 (0.013)	−0.085 (0.020)	−0.053 (0.025)
Repeal \times 1974–1975	−0.018 (0.019)	0.145 (0.055)	−0.162 (0.041)	−0.000 (0.029)	−0.113 (0.019)	−0.024 (0.025)
Repeal \times 1976–1979	0.006 (0.023)	0.109 (0.064)	−0.103 (0.045)	0.039 (0.049)	−0.115 (0.026)	−0.005 (0.033)
Mean of dependent variable (before logs)	7.8%	8.8%	89.1%	5.3%	6.8%	17.8%
<i>p</i> -value for test of equality of coefficients on repeal \times 1971–1973 and repeal \times 1976–1979	0.00	0.40	0.00	0.04	0.22	0.01
<i>p</i> -value for test of joint significance of reported coefficients	0.00	0.00	0.00	0.00	0.00	0.01

Notes: Reported coefficients are for the repeal interactions from regression specifications such as equation (1), including the following other regressors: the insured unemployment rate, per capita income, crime rate, percentage of the population that is nonwhite, a full set of state and year dummies, and state-specific quadratic trends. In specifications (1)–(4), 750 observations are available representing the fifty states and fifteen years of birth cohorts (1965–1979). Specifications (5) and (6) are limited to 600 observations (1968–1979). All standard errors are clustered at the state level.

while illegal). Nevertheless, they provide a gauge of the potential differences across states that may have existed in the years prior to abortion legalization. This variable is normalized to 1 in the states where no abortions were reported during this period and 0 in the state with the most reported abortions (so that latent abortion costs rise as the index rises).²³

For each of our proxy measures, data are missing for some states. Twelve states had no survey data on political identification prior to 1970 and two small states did not have pre-1970 estimates of the abortion rate. For both measures, we replaced each missing state value with the average value for other states in its Census region (based on the nine Census divisions).²⁴

In addition, to provide some controls for the state- and year-specific environment in which the birth or pregnancy decision was made, we use data on the economic and demographic conditions in the state and year of birth. These include per capita income, the crime rate, and the percentage of the population that was white, all from the *Statistical Abstract* (various years). The insured unemployment rate comes from the U.S. Department of Labor, Employment and Training Administration (1983).

²³ Nevada has the highest reported abortion rate, but may be an outlier due to its small sample size. Among bigger states, New York had the highest rate, 7 per 1,000 women of childbearing age; its current (year 2000) value is 36. **State of residence in the survey year is not available in the public-use file of the NSFG. Researchers can access these data, however, by visiting the National Center for Health Statistics and conducting the analysis in their Research Data Center.** For the purposes of this project, we assign the state of residence in the survey year to the respondent's residence in all preceding years, as Gruber, Levine, and Staiger (1999) have done.

²⁴ We have also considered other proxies for latent costs that have the advantage of greater state coverage, but the disadvantage of being available only post-repeal. These include state attitudes toward abortion itself, measured in either the General Social Survey (which covers some states beginning in 1972) or the DDB Needham Lifestyle Survey (starting in 1985). In addition, NARAL (2003) compiles recent state rankings of access to abortion. Each of these measures is strongly correlated with both the pre-1970 abortion rate and state political attitudes, and using each as an instrument yields similar results.

VI. Census Results

A. Estimates Using Abortion Legalization Only

We begin our analysis by presenting results based on the methodology used in GLS that lays out clearly the impact of abortion legalization on key fertility outcomes, shown in table 1. The first three columns report results for the log of the birth rate, pregnancy rate, and birth ratio (births/pregnancies). Similar to Levine et al. (1999) and GLS, we estimate that birth rates declined by 5.7% in repeal states relative to nonrepeal states in the 1971–1973 period, but that the gap shrunk after 1973 and had disappeared by 1976–1979. Both the initial decline and the “catch-up” by the nonrepeal states are strongly statistically significant. Estimates for the pregnancy rate and birth ratio are quite different. As expected, pregnancy rates increased by 12.9% in repeal states relative to nonrepeal states in the 1971–1973 period, but there was no significant catch-up and the gap remained at 10.9% by 1976–1979. Similarly, the birth ratio declined by 18.7% in repeal states relative to nonrepeal states in 1971–1973, but by 1976–1979 this gap had only partially disappeared. Thus, while birth rates had converged by 1976, pregnancy rates had risen and birth ratios had fallen in repeal states relative to nonrepeal states. This pattern follows that observed in figure 1.

In column 4 of table 1 we confirm that the fertility effects associated with abortion legalization that we observe in Vital Statistics data are present in the Census data roughly thirty years later in the form of reduced cohort size. We construct a measure that we call the “survivor rate,” which represents the number of individuals in a state/year of birth cohort alive in the 2000 Census per 1,000 women of childbearing age in the state/year in which those individuals were born. If there were no mortality since birth or, more plausibly, if mortality since birth were small and roughly random, then estimates using this dependent variable should be roughly comparable to the previously estimated birth

TABLE 2.—IMPACT OF ABORTION LEGALIZATION ON ADULT OUTCOMES, 2000 CENSUS

	Log % in Poverty (1)	Log % Single Parent (2)	Log % on Welfare (3)	Log % HS Dropout (4)	Log % Not College Graduate (5)	Log % Not Employed (6)	Log Incarcerated per 1,000 (7)
Repeal \times 1971–1973	–0.036 (0.044)	–0.033 (0.038)	–0.148 (0.056)	–0.002 (0.053)	–0.027 (0.009)	0.018 (0.020)	0.007 (0.154)
Repeal \times 1974–1975	–0.085 (0.071)	–0.027 (0.058)	–0.215 (0.154)	0.014 (0.042)	–0.020 (0.013)	0.014 (0.035)	0.102 (0.264)
Repeal \times 1976–1979	–.191 (0.076)	–0.008 (0.083)	–0.177 (0.233)	–0.018 (0.050)	–0.038 (0.020)	–0.016 (0.048)	–0.295 (0.662)
Mean of dependent variable (before logs)	10.3%	9.0%	2.1%	7.4%	70.6%	22.5%	4.8
<i>p</i> -value for test of equality of coefficients on repeal \times 1971–1973 and repeal \times 1976–1979	0.00	0.64	0.88	0.83	0.44	0.32	0.58
<i>p</i> -value for test of joint significance of reported coefficients	0.01	0.39	0.04	0.69	0.01	0.24	0.50

Notes: Reported coefficients are for the repeal interactions from regression specifications such as equation (1), including the following other regressors: the insured unemployment rate, per capita income, crime rate, percentage of the population that is nonwhite, a full set of state and year dummies, and state-specific quadratic trends. All specifications include 750 observations representing the fifty states and fifteen years of birth cohorts (1965–1979), with the following exceptions: specifications (1) and (5) exclude 1977–1979 because percentage in poverty and percentage not college graduate increase sharply at these ages; specification (3) excludes one observation with 0% in welfare; and specification (7) excludes fifty observations with 0 incarcerated. All standard errors are clustered at the state level.

effect.²⁵ Indeed, our estimates indicate that these constructed survivor rates are just over 4% lower in repeal states relative to nonrepeal states in cohorts born between 1971 and 1973 relative to that from earlier cohorts, and not significantly different in later years. These results mimic rather closely those in column 1, and verify that differences in cohort size in the 2000 data reflect the earlier impact of abortion legalization.

The final two columns of table 1, which estimate equation (5), document that legalized abortion resulted in selection at the time of birth. Moreover, they provide some evidence that the resulting gap in average characteristics of birth cohorts did not completely disappear after 1973. The percentage of infants born to minors decreased by 8.5% in repeal states relative to nonrepeal states in the 1971–1973 period, but there was no significant catch-up and the gap remained at 11.5% by 1976–1979. The results for percentage born to nonwhite mothers, by contrast, are more consistent with catch-up in the nonrepeal states. Thus, despite a convergence in birth rates, some of the average characteristics of the birth cohorts did not converge, providing further evidence that changes in the birth rate alone are insufficient to identify selection into a cohort.

This conclusion is supported in table 2, which reports estimates from equation (5) of the differential patterns in repeal and nonrepeal states in children's outcomes as adults, including poverty status, welfare receipt, single parenthood, educational attainment, employment, and the likelihood of

being incarcerated.²⁶ All outcomes are defined to be negative in terms of socioeconomic status, so selection due to lower cost of abortion would predict a negative coefficient on repeal \times 1971–1973. The coefficient on repeal \times 1976–1979 should be 0 (or less negative) if there is convergence in these outcomes following widespread legalization; lack of convergence suggests continued differences in selection in the late 1970s.

In most cases, the direction of the effect based on the repeal \times 1971–1973 coefficient is that which would be predicted by the positive selection found in GLS. The results on education are perhaps most striking, with a large negative effect on the percentage of the cohort that did not graduate from college (column 5), indicating that abortion legalization shifted the distribution of education upward. But these results provide no evidence of convergence; coefficients on repeal \times 1974–1975 and repeal \times 1976–1979 are generally of the same sign and larger than the coefficient on repeal \times 1971–1973. Again, consider the results for college graduation. The coefficient on repeal \times 1971–1973 shows that the odds of not graduating from college fell by 2.7% in the early repeal states over the 1971–1973 period, relative to other states, but the coefficient on repeal \times 1976–1979 is larger yet, indicating that the odds of not graduating from college fell 3.8% in early repeal states even in the later period. This lack of convergence supports the possibility that continued growth in both the pregnancy and abortion rates in early legalization states

²⁵ In fact, we know that this assumption, taken in its strongest terms, is inaccurate. GLS show that the infant mortality rate for the marginal child following abortion legalization was 40% higher than that for the average child. But the infant mortality rate is so low (1.9% during that period) that a somewhat lower rate would be swamped by the magnitude of the impact on fertility itself.

²⁶ In some cases, we hypothesized that the effect of parental fertility control would be stronger on the outcome of a certain at-risk subpopulation. For example, women are at much higher risk of welfare receipt, while men and African Americans are at higher risk of incarceration. We therefore conducted the analysis separately by sex and by race, but the results did not differ significantly from what is presented here.

TABLE 3.—FIRST-STAGE REGRESSIONS

	Latent Cost Based on Pre-1970% Liberal			Latent Cost Based on Pre-1970 Abortion Rates		
	ln (Pregnancy Rate) (1)	ln (Birth Rate) (2)	ln (Birth Ratio) (3)	ln (Pregnancy Rate) (4)	ln (Birth Rate) (5)	ln (Birth Ratio) (6)
1971–1973 \times nonrepeal	–0.099 (0.048)	0.021 (0.014)	0.120 (0.040)	–0.109 (0.051)	0.019 (0.020)	0.127 (0.043)
Nonrepeal \times “latent cost”	–0.031 (0.031)	0.061 (0.026)	0.092 (0.023)	–0.021 (0.031)	0.035 (0.023)	0.056 (0.030)
Nonrepeal \times distance	–0.027 (0.012)	0.018 (0.016)	0.045 (0.012)	–0.010 (0.008)	0.023 (0.010)	0.033 (0.012)
1974–1975 \times nonrepeal	–0.048 (0.064)	–0.024 (0.024)	0.023 (0.050)	–0.044 (0.077)	–0.042 (0.035)	0.001 (0.062)
Nonrepeal \times “latent cost”	–0.205 (0.046)	0.090 (0.030)	0.295 (0.040)	–0.125 (0.061)	0.075 (0.037)	0.201 (0.055)
1976–1979 \times nonrepeal	0.011 (0.076)	–0.041 (0.031)	–0.053 (0.059)	0.034 (0.090)	–0.066 (0.046)	–0.101 (0.071)
Nonrepeal \times “latent cost”	–0.267 (0.060)	0.079 (0.043)	0.346 (0.054)	–0.176 (0.069)	0.076 (0.048)	0.252 (0.061)
# of observations	750	750	750	750	750	750
First-stage <i>F</i> -Statistic	14.9	32.0	37.1	9.9	33.4	13.9
<i>p</i> -value of <i>F</i> -test	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Models also include the insured unemployment rate, per capita income, crime rate, percentage of the population that is nonwhite, a full set of state and year dummies, and state-specific quadratic trends. Distance and latent cost are scaled 0–1, with 0 representing the state with the lowest cost and distance (highest pre-1970 percentage liberal or abortion rate). Standard errors and *F*-statistics account for clustering at the state level.

caused greater positive selection in the composition of births in those states, even after abortion was legalized nationally and birth rates converged.

B. Estimates Using Abortion Cost Instruments

These results motivate our updated methods for estimating selection effects based upon the model in equations (7) and (8). Before presenting OLS and IV results from these models, we first report the results of the first-stage regressions, which relate differences in proxies for abortion costs to the log birth rate and birth ratio. We also estimate first-stage models where the log pregnancy rate is the dependent variable, because of its key role in equation (2) and because we have unambiguous predictions regarding the impact of abortion costs on pregnancy rates.

Table 3 reports two sets of first-stage estimates, corresponding to our two instrumental variables strategies; the results are similar for both sets of instruments, and are consistent with our predictions. Columns 1 and 4 provide support for the prediction that higher abortion costs reduce the pregnancy rate: nonrepeal states had lower pregnancy rates during the 1971–1973 period; among nonrepeal states, those with higher latent social costs of abortion (based on pre-1970 liberalism or on the illegal abortion rate) experienced lower pregnancy rates. The same pattern continues in 1974–1975 and 1976–1979. As expected, the negative relationship between latent costs and the pregnancy rate becomes more pronounced after legalization: once the legal constraints are removed, the underlying latent costs of abortion become the primary determinant. The remainder of the table shows that higher abortion cost (as proxied by nonrepeal, latent cost of abortion, and distance) predicts higher birth ratios (or equivalently, fewer abortions) and

higher birth rates. The first-stage *F*-statistics for each of these regressions are sufficiently large to rule out weak-instrument problems.²⁷

C. OLS and IV Results

Regression results based on equations (7) and (8) are reported in tables 4 and 5, respectively. In each table, each panel represents regressions on a different outcome. There are three columns in each table, for OLS and our two IV strategies. For each IV regression, we also show the *p*-value for the Hausman test that OLS and IV outcomes are significantly different, and the *p*-value for the overidentification test of our set of instruments.

In table 4 (where the key independent variable is the log of the birth rate from equation [6]), the pattern of coefficients is fairly consistent with selection, with positive IV coefficients for living in poverty, being a single parent, receiving welfare, being a high school dropout, and not graduating from college. The IV results are wrong-signed (negative) for being unemployed and being incarcerated. The results for single parenthood and college graduation are statistically significant with either instrument, and for welfare receipt are marginally significant. The findings are similar in table 5: positive (expected sign) coefficients on education, welfare receipt, poverty rate, and single parenthood; negative (wrong-signed) for being not employed and

²⁷ IV estimation of the more general specification in equation (2) requires instruments for both the pregnancy rate and the birth ratio. A generalization of the first-stage *F*-statistic to the case of two endogenous variables, described in Staiger and Stock (1997), indicated that our instruments were too weak to reliably estimate equation (2). They do not generate sufficient independent variation to reliably estimate the coefficient on each of the endogenous variables.

TABLE 4.—ESTIMATES OF RELATIONSHIP BETWEEN LOG BIRTH RATE AND ADULT OUTCOMES
(standard errors in parentheses)

Outcome	OLS	IV 1	IV 2
% living in poverty	0.084 (0.198)	0.199 (0.354)	0.113 (0.383)
<i>p</i> -value for Hausman test		0.731	0.938
<i>p</i> -value for over-ID test		0.458	0.620
% single parent	0.161 (0.229)	0.665 (0.318)	0.689 (0.333)
<i>p</i> -value for Hausman test		0.166	0.142
<i>p</i> -value for over-ID test		0.720	0.646
% receiving welfare	0.179 (0.539)	1.939 (0.936)	1.720 (0.956)
<i>p</i> -value for Hausman test		0.043	0.075
<i>p</i> -value for over-ID test		0.666	0.293
% high school dropout	0.314 (0.325)	0.472 (0.690)	0.411 (0.740)
<i>p</i> -value for Hausman test		0.787	0.877
<i>p</i> -value for over-ID test		0.832	0.721
% not college graduate	0.219 (0.061)	0.311 (0.119)	0.289 (0.126)
<i>p</i> -value for Hausman test		0.306	0.477
<i>p</i> -value for over-ID test		0.324	0.591
% not employed	0.137 (0.166)	−0.112 (0.219)	−0.275 (0.183)
<i>p</i> -value for Hausman test		0.259	0.037
<i>p</i> -value for over-ID test		0.597	0.856
Incarcerated per 1,000	−0.648 (1.050)	−1.332 (1.490)	−2.819 (2.144)
<i>p</i> -value for Hausman test		0.624	0.252
<i>p</i> -value for over-ID test		0.587	0.059

Notes: IV 1 represents the specification that includes state pre-1970 fraction liberal as an instrument. IV 2 represents the specification that includes state pre-1970 abortion rates as an instrument. Models also include the insured unemployment rate, per capita income, crime rate, percentage of the population that is nonwhite, a full set of state and year dummies, and state-specific quadratic trends. Sample sizes are the same as those reported in tables 1 and 2. Standard errors and test statistics account for clustering at the state level. The excluded instruments in the IV models are those reported in table 3. For poverty and college graduate we use only 1965–1976, so the instruments used 1974–1976 interactions, rather than 1974–1975 and 1976–1979 interactions.

being incarcerated. Once again, the results for single parenthood, welfare receipt and college graduation are all at least marginally significant with both instruments.

These results suggest sizeable effects of selection on outcomes. Recall that the estimates from the specifications in table 4 and table 5 provide a bound on the difference between the marginal and average birth. Therefore, the estimates in the IV columns in each table suggest that the marginal birth is 23% to 69% more likely to be a single parent, 73% to 194% more likely to receive welfare, and 12% to 31% less likely to graduate college.

We have also investigated the link between abortion and crime, applying our instrumental variables strategy to the crime data gathered by DL. The main practical improvement that our approach offers relative to previous work by DL and by Foote and Goetz (2008) is that we estimate IV models where changes in the birth ratio and birth rate are instrumented using changes in abortion policy. With this approach, we found a negative but not precisely estimated effect of reduced abortion costs on crime per capita; as with our Census results, we were unable to distinguish between effects due to changes in pregnancy rates and effects due to changes in the birth ratio. Nonetheless, despite many differences in data and method, we were able to replicate the

TABLE 5.—ESTIMATES OF RELATIONSHIP BETWEEN LOG BIRTH RATIO AND ADULT OUTCOMES
(standard errors in parentheses)

Outcome	OLS	IV 1	IV 2
% living in poverty	0.125 (0.091)	0.019 (.153)	0.030 (.158)
<i>p</i> -value for Hausman test		0.201	0.328
<i>p</i> -value for over-ID test		0.426	0.618
% single parent	0.123 (0.128)	0.240 (.130)	0.232 (.123)
<i>p</i> -value for Hausman test		0.241	0.292
<i>p</i> -value for over-ID test		0.666	0.434
% receiving welfare	0.448 (0.238)	0.790 (.241)	0.733 (.243)
<i>p</i> -value for Hausman test		0.067	0.158
<i>p</i> -value for over-ID test		0.536	0.269
% high school dropout	0.260 (0.139)	0.144 (.230)	0.098 (.265)
<i>p</i> -value for Hausman test		0.392	0.324
<i>p</i> -value for over-ID test		0.739	0.671
% not college graduate	0.112 (0.023)	0.120 (.026)	0.120 (.028)
<i>p</i> -value for Hausman test		0.746	0.754
<i>p</i> -value for over-ID test		0.201	0.688
% not employed	−0.001 (0.082)	−0.062 (.082)	−0.124 (.078)
<i>p</i> -value for Hausman test		0.241	0.060
<i>p</i> -value for over-ID test		0.584	0.879
Incarcerated per 1,000	−0.490 (0.432)	−0.812 (.613)	−1.346 (.878)
<i>p</i> -value for Hausman test		0.427	0.121
<i>p</i> -value for over-ID test		0.647	0.108

Notes: IV 1 represents the specification that includes state pre-1970 fraction liberal as an instrument. IV 2 represents the specification that includes pre-1970 abortion rates as an instrument. Models also include the insured unemployment rate, per capita income, crime rate, percentage of the population that is nonwhite, a full set of state and year dummies, and state-specific quadratic trends. Sample sizes are the same as those reported in tables 1 and 2. Standard errors and test statistics account for clustering at the state level. The excluded instruments in the IV models are those reported in table 3. For poverty and college graduate we use only 1965–1976, so the instruments used 1974–1976 interactions, rather than 1974–1975 and 1976–1979 interactions.

results in DL for total crimes committed. Therefore, our results align with the previous literature in that reduced abortion costs led to reduced crime, but largely through a reduction in cohort size (total crime) rather than through selection (crime per capita).²⁸

VII. Conclusions

In this paper, we have extended past analyses of abortion and selection in several dimensions. First, we have provided a new theoretical model that provides useful insights regarding the mechanisms that may cause selection when the cost of abortion changes. The main insight of the model is that selection may still occur even if birth rates are unaffected, since the composition of births may change. Second, we have introduced an econometric methodology that encompasses the differing approaches taken by past researchers. Although we are unable to implement this methodology directly due to limited power, we can identify what restrictions need to be imposed for the past methods to be consistent with our theoretical model. Since our findings are broadly consistent across these specifications, the impact of

²⁸ Detailed results are provided in Ananat et al. (2006).

the different restrictions does not appear to be very dramatic. Third, we have updated the literature on abortion and selection to include outcomes in early adulthood. We found consistent evidence that changes in cohort composition that occurred in the 1970s that can be attributed to greater abortion access led to improved cohort outcomes, particularly in the form of higher rates of college graduation, lower rates of single parenthood, and lower rates of welfare receipt.

Most importantly, taken together with earlier results (Gruber et al., 1999), our findings suggest that the improved living circumstances experienced by the average child born after the legalization of abortion had a lasting impact on the lifelong prospects of these children. Children who were "born unwanted" prior to the legalization of abortion not only grew up in more disadvantaged households, but they also grew up to be more disadvantaged as adults. This conclusion is in line with a broad literature documenting the intergenerational correlation in income (Solon, 1999) and showing that adverse living circumstances as a child are associated with poorer outcomes as an adult (Haveman & Wolfe, 1995). Overall, our results provide further evidence that abortion is associated with positive selection and that its impact is persistent.

Thus, overall, we find evidence consistent with long-run selection effects through abortion. While the statistical significance of our findings depends on the particular outcome under consideration, it is robust to the choice of instrument, and the pattern is clear: when abortion costs are lowered, cohort outcomes improve.

REFERENCES

- Ananat, Elizabeth Oltmans, Jonathan Gruber, and Phillip B. Levine, "Abortion Legalization and Lifecycle Fertility," *Journal of Human Resources* 42:2 (2007), 375–397.
- Ananat, Elizabeth Oltmans, Jonathan Gruber, Phillip B. Levine, and Douglas Staiger, "Abortion and Selection," National Bureau of Economic Research working paper no. 12150 (2006).
- Bankole, Akinrinola, Susheela Singh, and Taylor Haas, "Reasons Why Women Have Induced Abortions: Evidence from 27 Countries," *International Family Planning Perspectives* 24:3 (May/June 1998), 117–127, 152.
- Centers for Disease Control, *Abortion Surveillance Report—Legal Abortions, United States, Annual Summary, 1970* (Atlanta: Centers for Disease Control, 1971).
- , *Abortion Surveillance Report—Legal Abortions, United States, Annual Summary, 1971* (Atlanta: Centers for Disease Control, 1972).
- , *Abortion Surveillance, Annual Summary, 1972* (Atlanta: Centers for Disease Control, 1974).
- Charles, Kerwin Kofi, and Melvin Stephens Jr., "Abortion Legalization and Adolescent Substance Use," *Journal of Law and Economics* 49:2 (2006), 481–505.
- , "Measurement Error, Legalized Abortion, and the Decline in Crime: A Response to Foote and Goetz," *Quarterly Journal of Economics* 123:1 (2008), 425–440.
- Donohue, John J. III, and Steven D. Levitt, "The Impact of Legalized Abortion on Crime," *Quarterly Journal of Economics* 116:3 (2001), 379–420.
- , "Further Evidence that Legalized Abortion Lowered Crime: A Reply to Joyce," *Journal of Human Resources* 39:1 (2004), 29–49.
- Foote, Christopher, and Christopher Goetz, "The Impact of Legalized Abortion on Crime: A Comment," *Quarterly Journal of Economics* 123:1 (2008), 407–423.
- Garrow, David J., *Liberty and Sexuality: The Right to Privacy and the Making of Roe v. Wade* (New York: Macmillan, 1994).
- Grob, Gerald N., "Mental Health Policy in 20th-Century America" (pp. 3–14), in Ronald W. Manderscheid and Marilyn J. Henderson (Eds.), *Mental Health, United States, 2000* (Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, 2001).
- Gruber, Jonathan, Phillip B. Levine, and Douglas Staiger, "Abortion Legalization and Child Living Circumstances: Who Is the 'Marginal Child'?" *Quarterly Journal of Economics* 114:1 (1999), 263–292.
- Haveman, Robert, and Barbara Wolfe, "The Determinants of Children's Attainments: A Review of Methods and Findings," *Journal of Economic Literature* 33:4 (1995), 1829–1878.
- Joyce, Theodore, "Did Legalized Abortion Lower Crime?" *Journal of Human Resources* 39:1 (2004a), 1–28.
- , "Further Tests of Abortion and Crime," National Bureau of Economic Research working paper no. 10564 (2004b).
- Kane, Thomas, and Douglas Staiger, "Teen Motherhood and Abortion Access," *Quarterly Journal of Economics* 111:2 (May 1996), 467–506.
- Levine, Phillip B., "Parental Involvement Laws and Fertility Behavior," *Journal of Health Economics* 22:5 (2003), 861–878.
- , *Sex and Consequences: Abortion, Public Policy, and the Economics of Fertility* (Princeton, NJ: Princeton University Press, 2004).
- Levine, Phillip B., and Douglas Staiger, "Abortion as Insurance," National Bureau of Economic Research working paper no. 8813 (March 2002).
- Levine, Phillip B., Douglas Staiger, Thomas J. Kane, and David J. Zimmerman, "Roe v. Wade and American Fertility," *American Journal of Public Health* 89:2 (1999), 199–203.
- NARAL, *State-by-State Report Card on Access to Abortion* (Washington, DC: NARAL Pro-Choice America Foundation, 2003).
- Pop-Eleches, Christian, "The Impact of a Change in Abortion Regime on Socio-Economic Outcomes of Children: Evidence From Romania," *Journal of Political Economy* 114:4 (2006), 744–773.
- Ruggles, Steven, Matthew Sobek, et al., *Integrated Public Use Microdata Series, Version 3.0* (Minneapolis: Historical Census Projects, University of Minnesota, 2003).
- Saul, Rebekah, "Abortion Reporting in the United States: An Examination of the Federal-State Partnership," *Family Planning Perspectives* 30:5 (1998), 244–247.
- Solon, Gary, "Intergenerational Mobility in the Labor Market" (pp. 1761–1800), in Orley Ashenfelter and David Card (Eds.), *Handbook of Labor Economics. Volume 3A* (Amsterdam: Elsevier, 1999).
- Staiger, Douglas, and James Stock, "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65:3 (1997), 557–586.
- Torres, Aida, and Jacqueline Darroch Forrest, "Why Do Women Have Abortions?" *Family Planning Perspectives* 20 (1988), 169–176.
- United States Census Bureau, *Statistical Abstract of the United States* (Washington, DC: U.S. Census Bureau, various years).
- , *Vital Statistics of the United States: Natality* (Washington, DC: U.S. Census Bureau, various years).
- U.S. Department of Justice, Bureau of Justice Statistics, *Key Facts at a Glance: Correctional Populations* (Washington, DC: U.S. Department of Justice, 2003).
- U.S. Department of Labor, Employment and Training Administration, *Unemployment Insurance Financial Data (ET Handbook 394)* (Washington, DC: Government Printing Office, 1983).