Effects of Increased Access to Infertility Treatment on Infant and Child Health: Evidence from Health Insurance Mandates[1]

Marianne P. Bitler
University of California-Irvine, NBER, and IZA

This version: February 2008
JEL classification: I3, JI
Key words: infant health, child health, fertility, insurance mandates, twins, birth outcomes, infertility, impaired fecundity

[1]I thank Reagan Baughman, Sandra Black, Kate Bundorf, Janet Currie, Ashlesha Datar, Julie DaVanzo, Will Dow, Chrissy Eibner, Susan Ettner, Deborah Haas-Wilson, Steven Haider, Janice Hauge, Vivian Ho, Hilary Hoynes, Becky Kilburn, Jacob Klerman, Brent Kreider, Darius Lakdawalla, Dave Loughran, Nicole Maestas, Terra McKinnish, Kevin Milligan, Jane Miller, Lucie Schmidt, Barbara Schone, Kosali Simon, Jennifer Ward-Batts, Madeline Zavodny, Julie Zissimopoulos, and seminar participants at RAND, PPIC, the Melbourne Institute of Applied Economic and Social Research, UC Merced, UC Irvine, Iowa State University, the All-UC Labor conference, the Economic Demography Workshop at the PAA, the PAA, and the WEA for valuable comments. I thank Aaron Yelowitz for generously providing data on Medicaid expansions and Loren Baker for generously providing data on HMO penetration. I gratefully acknowledge the financial support of the National Institute of Child Health and Human Development (R03 HD046485), the National Institute on Aging, and the RAND Corporation. This work has not been formally reviewed or edited. The views and conclusions are those of the author and do not necessarily represent those of any other organization. Correspondence to Bitler at Department of Economics, University of California-Irvine, 3151 Social Science Plaza, Irvine, CA 92697 or mbitler@uci.edu.

Abstract

Reproductive technologies have radically improved since the introduction of the first fertility drugs in the late 1960s. These technologies make conception possible for many couples who otherwise would have been unable to reproduce. Many of these technologies increase the probability of having a multiple birth, typically a more risky pregnancy. These technologies also provide women considering delaying reproduction with insurance against later infertility. In this paper, we use Detailed Natality data and U.S. Census data to look at the association between use of infertility treatments and fertility and infant and child health outcomes. Women who use infertility treatment are not randomly drawn from the population. To deal with endogeneity of use of treatment, we rely on variation in access to infertility treatments induced by state-level insurance mandates forcing insurers to cover or offer to cover infertility treatments, thus subsidizing their use.

Mandates are associated with a 10–23 percent increase in the probability of being a twin for children born to older mothers, and an even larger increase in the twin delivery rate for older women. Twin pregnancies are riskier for mothers and infants, thus this finding alone suggests a negative health effect of the mandates. Mandates are also associated with small, statistically significant negative effects on birth weight, gestation, and 5-minute Apgar score for samples of twins and, to a lesser extent, singletons. Back of the envelope calculations suggest imposing a mandate to cover infertility treatment in every state would lead to increased hospital costs of $334 million per year from the effects on twins alone. Estimates using Census data suggest that twins born to older mothers in mandate states may face higher risks as children of having a condition that limits basic physical activities.

This paper contributes to the health economics literature on mandated health insurance benefits; infertility treatment mandates affect fertility and health outcomes. It also contributes to the economics literature on birth selection.

1 Introduction

Improved infertility treatments are increasingly benefiting couples who would otherwise remain childless by allowing them to bear children while also allowing women to postpone their childbearing. But what are the costs? While we can measure the typical medical costs of the treatments themselves (estimates range from $40,000 to over $200,000 for a sucessful IVF delivery), it is important to know if there are other costs in terms of the eventual health of the children born after use of these treatments? This paper is the first to examine this question using variation in access to infertility treatments; variation induced by state-level mandates requiring private insurers to cover infertility treatment.

A key difficulty with this analysis is an inability to know the true counterfactual. Infertility treatments may affect the health of children born after use of such treatment. At the same time, infertility treatments change the characteristics of the population of women giving birth. Women using infertility treatment differ systematically from other women; they are often older and usually have been unable to conceive or carry a pregnancy to term without treatment. Age itself is associated with worse birth outcomes. It is also possible that the medical issues making it difficult to conceive or carry a pregnancy to term could directly affect infant or later child health. Thus, use of infertility treatment might be associated with worse infant health because of negative selection in who uses treatment. However, infertility treatments are time intensive and expensive, suggesting these would-be parents value their children very highly and are likely to invest positively in infant health in other ways. This suggests that if anything, the negative effects of the infertility treatment mandates as found here are a lower bound for the actual effects of increased access.

We use Detailed Natality birth certificate data and Census data to see whether the twin delivery rate and the share of births that are twins have increased with infertility treatment mandates. Because births due to use of IVF or other infertility treatments are rare in the population of singleton births, we then focus primarily on the relation between the mandates and health outcomes among twins. We use Natality data to examine the relation between mandates and birth weight, gestation, and Apgar score. We use U.S. Census data to see if there are longer-term effects of the mandates on child health.

Our identifying variation comes from state laws mandating insurance coverage of infertility treatment. A number of states passed laws in the late 1980s and early 1990s requiring health insurers to cover or offer coverage of infertility treatment. These state mandates lower the price of infertility treatment. The reduced price of treatment likely allows some women to have children who otherwise could not, but also may lead some women to delay attempts to have children or to use treatments sooner. By using exogenous variation in the cost of infertility treatment induced by these mandates, we avoid possible endogeneity bias associated with comparing users of infertility treatments with other mothers.

First, using Natality data for 1981–99, we show that infertility treatment mandates are associated with a statistically significant increase of between 10 and 23 percent in the probability that infants born to older mothers are twins and a 16–37 percent increase in the probability that infants born to older mothers are mixed-sex twins. Using Natality data and Census population data for the same period, we find similar patterns. We find that mandates are also associated with a statistically significant increase of between 24 and 40 percent in the twin delivery rate for older women. These findings establish that the mandates have affected the composition of births and rates of twin births. Given that twin pregnancies are riskier than singleton pregnancies for the infant(s) and the mother, this also suggests the mandates have had a negative effect on infant health even if this were their only impact.

We also estimate reduced-form regressions in the Natality data relating the state-level mandates to birth weight, gestation, and Apgar score separately for twins and for singletons. Our results show that twins with older mothers in states with mandates have statistically significantly lower birth weights, shorter gestation durations, and lower Apgar scores. We find some smaller in magnitude impacts for singletons as expected given the smaller share of singleton births that are due to use of ART or other infertility treatments.

With pooled Decennial Census and American Community Survey (ACS) data we then investigate whether the mandates have had longer-term health impacts on twins with older mothers (relative to other twin births), with mixed findings. Our results suggest that twin children born to older mothers in mandate states may be more likely to have a physical impairment or a condition limiting basic physical activities. There is no evidence of an association between being a twin born to an older mother in a mandate state and sensory impairments or conditions limiting the ability to learn, concentrate, or remember.

We have found that, as expected, for older women, infertility mandates are associated with higher rates of twin deliveries and a larger share of births that are twins. In addition, we find that the mandates are associated with worse birth outcomes among twins, and to some extent, among singletons.

To get a sense of these costs in financial terms, we use data on the distribution of the costs of infant and maternal hospital stays by birth weight to estimate the increased hospital costs associated with imposing an infertility mandate on all states. We estimate that additional infant and maternal hospital costs from increases in twinning and lower birth weight among twins associated with imposing a national mandate could be as large as $334 million for 1999. These additional costs are about 71 percent of the total cost for all the in vitro fertilizations performed in 2000. If the mandates affect higher-order multiple births in the same way that they affect twins, the mandate-associated cost of additional triplet and higher-order multiples is an additional $82 million per year.

The rest of the paper proceeds as follows. Section 2 provides background on infertility, infertility treatment, and infertility treatment mandates. Section 3 provides a theoretical motivation and discusses related work. Section 4 describes the data and outcomes. The empirical models are presented in Section 5 and the main results and extensions in Section 6. Section 7 discusses implications of the findings. Finally, Section 8 concludes.

2 Infertility, Infertility Treatment, and Infertility Treatment Mandates

Infertility is generally defined by clinicians as the inability to conceive after 12 months of unprotected intercourse (6 months for older women).[1] A related reproductive problem is impaired fecundity. Unlike infertility, which as commonly defined only applies to cohabiting and married women, any (not surgically sterile) woman may have impaired fecundity. A woman has impaired fecundity if she has either had problems conceiving or carrying a pregnancy to term or has been unable to conceive after a three-year period of unprotected intercourse (Chandra & Stephen (1998)).

Infertility and impaired fecundity are common. In the 2002 National Survey of Family Growth (NSFG), about 2.1 million married women (or about 7.5 percent of married women) were infertile, and numbers were similar in the 1995 NSFG (Chandra, Martinez, Mosher, Abma & Jones (2005); Abma, Chandra, Mosher, Peterson & Piccinino (1997)).

Infertility treatment can be a long, arduous, and expensive process. The first stage of infertility treatment is a thorough examination of each partner's reproductive organs and their circulatory, endocrine, and necrologic functions. An early stage of treatment for women not ovulating normally is treatment with a fertility drug to stimulate ovulation. Drugs of various potencies (e.g., clomiphene citrate; gonadotropin-releasing hormones) are available. More aggressive techniques such as in vitro fertilization (IVF) were pioneered in the early 1980s. In IVF, the oocytes (eggs) and sperm are combined in a laboratory. Once early embryos develop, they are transferred into the uterus. In the U.S., the first infant was born after use of IVF in 1981. Another frequent procedure is intrauterine insemination (IUI). In IUI, a woman is often given medication to stimulate multiple egg development and to help doctors time when egg(s) will be in the Fallopian tubes. The sperm to be used is

[1] Much of the information in this section is drawn from Office of Technology Assessment (1988). Women must also not be surgically sterile to be considered infertile.

separated from the other components of semen in the laboratory and then the concentrated sperm is placed in the cervix or high in the uterus using a catheter. Because women undergoing IUI are often superovulating, there is a non-trivial risk of multiple pregnancies with IUI. Treatments for male infertility include artificial insemination and IUI as well as other procedures. Surgery is also a treatment for some forms of infertility.

Infertility treatments are used by a large and growing number of women. The NSFG is the only data set with a long time-series of information on use of infertility treatment in the U.S. Table 1 shows the share of women from the 2002 wave of the National Survey of Family Growth who have ever gotten various types of infertility treatment as well as those who had a visit for infertility treatment during the year before the survey.[2] The NSFG asks about both medical help to get pregnant and medical help to avoid miscarriage, and considers both infertility treatment. However, only those who got medical help to get pregnant were asked what type of treatment they obtained. Table 1 shows that 13.4 percent of women 15–44 had ever had some type of infertility treatment, and 2.8 percent had some treatment during the year before the survey.

We classify infertility treatments into two types: 1) assisted reproductive technologies (ART), where both the sperm and the eggs are handled outside the body, and 2) other treatments where either the egg or sperm may be handled outside the body but not both. The most common assisted reproductive technology is IVF; other less common ART procedures include as gamete intrafallopian transfer and zygote intrafallopian transfer. Other infertility treatments that may be used simultaneously with ART but are not themselves ART are ovulation-inducing drugs, artificial insemination, IUI, testing for infertility, and surgery to repair blocked Fallopian Tubes. Infertility clinics are required by law to report data on ART procedures and success rates to the CDC (but need not report on use of other infertility treatments). However, a large share of infertility treatment does not involve ART. Table 1 also shows that of the 9.3 percent of women who ever got medical help to get pregnant, most did not use IVF or other ART but used other less invasive treatments. 4.2 percent of all women or almost half of those who ever had treatment to get pregnant were prescribed ovulation-inducing drugs and 1.2 percent of all women or 13 percent of those getting help to get pregnant used artificial insemination. Only 0.003 percent of all women or 3 percent of women who had help to get pregnant used IVF.

There is more comprehensive data available from providers on use of IVF and other assisted reproductive technologies. Starting in the mid-1980s, the Society of Assisted Alternative Reproductive Technology (SART) began a voluntary registry of clinic data on IVF, embryo transfers, and other ART treatments and their outcomes. In 1992, Congress passed a law requiring the CDC to publish clinic-specific success rates for ART procedures.

[2]The numbers in Table 1 are tabulations for women 15–44 in the NSFG who reported having had sex after menarche.

As a result, the total number of pregnancies and live births resulting from ART as well as the number and type of ART cycles are available from the registry and/or the CDC data. Unfortunately, these registry data do not span the introduction of the infertility treatment mandates.

Figure 1 shows the total number of pregnancies, deliveries, and live births due to use of ART over the period. While there are no registry data for the period before 1985, it is unlikely that there were many ART-related pregnancies before that; recall that the first live birth in the U.S. from IVF was in 1981. Figure 1 shows that the number of pregnancies, deliveries, and live infants born after use of ART has increased considerably over the 1980s and 1990s, with much of the increase occurring in the late 1980s, shortly after the first state infertility treatment mandates would have impacted deliveries. It also shows an increase in the share of deliveries that were multiple births (if all births were singletons, then the number of deliveries and live infants would be the same).

Infertility services are not covered by all insurance plans. Table 1 shows that in 2002, only 76 percent of women who obtained medical help to get pregnant had it paid for by private health insurance. Treatment can be quite expensive. In 1988, a simple diagnosis of scanty or infrequent menstruation followed by treatment with fertility drugs would have cost about $3,668 (Office of Technology Assessment (1988)). At this time, about 30 percent of the couples would have successfully conceived via this treatment. A comprehensive evaluation for non-ovulatory causes of infertility would have cost a further $2,905. Tubal surgery to deal with blocked or damaged Fallopian tubes would have cost $7,118, and IVF $9,376 for two cycles (two sets of fertilized eggs implanted in the woman). 1988 estimates suggest that 70 percent of infertile women would have become pregnant by one of these treatments. Neumann, Gharib & Weinstein (1994) calculate that the cost of a successful delivery via IVF ranged from $44,000 to $211,940 in 1992. While these figures are dated, they show that infertility treatment can be quite costly. Ignoring the costs of unsuccessful IVF cycles and assuming costs are similar today, these estimates suggest that the 33,141 live deliveries via IVF and similar technologies in 2002 (Wright, Schieve, Reynolds & Jeng (2005)) cost between $1.46 and $7.02 billion without accounting for post-delivery hospital stays. This suggests that having access to insurance coverage for infertility treatment might be quite valuable.

A number of states have passed laws requiring insurance plans to cover or offer to cover infertility treatment. Table 2 contains a list of states with mandates and indicates the year the state passed the law, whether the mandate specifically excludes IVF or does not, whether the law mandates coverage or that the insurer offer coverage, and how the mandate applies to HMOs.[3]

[3]Information about the state laws was taken from National Conference of State Legislatures (2002) and cross-checked with law

Several papers from the medical literature use cross-sectional data to examine the impact of insurance mandates for infertility treatment on use of IVF and pregnancy outcomes. Massachusetts imposed a broad mandate to cover infertility treatment in 1987. Griffin & Panak (1998) use administrative data from insurance companies and find that the Massachusetts mandate is associated with increased use of ART. Jain, Harlow & Hornstein (2002) compare IVF cycles and live births for states with no coverage of IVF, partial coverage of IVF, and full coverage of IVF. They find that women at clinics in full-coverage states underwent more cycles (per capita) than did women in other states. However, women in full-coverage states also transferred fewer embryos, suggesting that insurance coverage of IVF might be associated with a decrease in multiple births.

A number of economists have also examined the effects of these mandates. Bitler & Schmidt (2007), using NSFG data, find that for older highly educated women, living in a state with an infertility mandate is associated with a statistically significant and economically meaningful increase in the probability of having used infertility treatment. Schmidt (2007) uses Natality data from 1981–2000 and differences-in-differences techniques, finding that first birth rates are significantly higher for women 35 and over in mandate states. Bundorf, Henne & Baker (2007), using Natality data, find that mandates are associated with an increased probability of having a multiple birth for older women. They also use clinic data on use of assisted reproductive technologies such as IVF and, using cross-sectional differences in the presence of mandates, find that mandates lead to more use of such technologies among women of low and high fecundity.[4] Buckles (2006) uses Natality data, finding mandates are associated with higher average age at first birth and more deliveries per first birth. Buckles also considers effects of mandates on labor market participation and wages. Bart Hamilton and Brian McManus model the behavior of infertility clinics in several papers. Hamilton & McManus (2005) use clinic data on use of IVF, the number of cycles undergone, the number of embryos transferred, the birth rate, and the rate of multiple births, finding that insurance mandates which include IVF increase cycles of IVF but cut the number of embryos transferred. Their clinic data only report the number of cycles of ART, so they cannot tell whether this means more women are gaining access to ART or women are undergoing more cycles. The clinic data do not contain information about use of other infertility treatments, and thus cannot capture all impacts of the insurance mandates unless the mandates had no effect on use of any infertility treatments but ART.

information available at `http://www.asrm.org/Patients/insur.html`. These laws do not cover self-insured companies' health insurance plans; as this is preempted by ERISA.

[4]Henne & Bundorf (2007) use similar data and find that use of ART was higher in states that had comprehensive mandates in place while the number of births per cycle of IVF and multiples per birth were lower in such states.

3 Theoretical Perspectives and Literature Review

In this section, we present the theoretical framework behind our analysis and provide a literature review.

3.1 Theoretical perspectives

Improvements in infertility treatment lower the cost to infertile women of having their own children. This may increase fertility among couples who would have been unable to conceive before these new technologies were widely available. In effect, these technologies may have shrunk the tail of the fecundity distribution.

Much of the rhetoric about infertility treatment implicitly presumes that children born via use of infertility treatment are the same as other children. Yet, there are reasons why this might not be the case. Children born via infertility treatment could have worse outcomes than other children. If women at the lower tail of the fecundity distribution are particularly likely to have infants at the lower tail of the infant health distribution, all other things being equal, subsidized infertility treatment might be associated with worse infant and child health because of this birth selection. The existence of infertility treatment might also provide some women with insurance against being infertile, enabling them to delay childbearing. If there is an important age gradient in birth outcomes, these women's infants may have worse health than if the women had not delayed childbearing.

Outcomes for children born from infertility treatment could also be better than those of other children. Any woman who goes through the more invasive medical interventions described above in Section 2 clearly wants any resulting children very much. Thus, one expects that women who respond to the insurance subsidies or who delay childbearing will invest more in fetal health during their pregnancy than will other women. This investment may offset the possibly negative selection touched on above. If treatment allows some women to postpone childbearing to obtain human capital and/or achieve higher socioeconomic status, these women may have more resources to invest.

Certain infertility problems may have little or no correlation with infant health (e.g., blocked Fallopian tubes). Some types of treatments themselves may affect infant health. Because of these potentially countervailing influences, it is impossible to predict from theory alone the effect of infertility treatment on infant health, motivating an empirical examination. Next we discuss the existing empirical evidence.

3.2 Infertility, infertility treatment, and infant and child health

Many papers in medical and public health journals examine trends in infertility and infant health and in the use of infertility treatments and infant health.[5] Findings on the association between infertility and birth outcomes are mixed. Some studies find that infertility is associated with having small-for-gestational-age and preterm births while other studies find no such association. Many studies note one difficulty in identifying an association between infertility (and infertility treatment) and birth outcomes: Infertile women who become pregnant are often older and age is independently associated with some negative birth outcomes.

Another strand of literature compares the health of infants conceived via use of ART (usually) or other infertility treatments to the health of other infants. Schieve, Meikle, Ferre, Peterson, Jeng & Wilcox (2002) find that ART infants born in 1997 are of lower birth weight than the general infant population. Bergh, Ericson, Hillensjo, Nygren & Wennerholm (1999) find an increased risk of multiple births and an increased risk of prematurity and low birth weight for all IVF births relative to non-IVF births. They also find some increase in the frequency of congenital malformations among IVF singletons as compared to other singletons. Dhont, Sutter, Martens & Bekaert (1999) find perinatal outcomes of singleton pregnancies conceived via use of ART are significantly worse than those of singleton pregnancies conceived spontaneously. Lastly, several recent studies examine the longer-term health impacts of ART on children. Sutcliffe (2004) summarizes this literature, concluding that when born at term, ART children are as healthy as other children with the exception that ART children have higher rates of congenital anomalies.

This clinical literature on effects of infertility treatment relies almost exclusively on cross-sectional comparisons between women using ART and other women. These cross-sectional differences may not be totally due to use of ART. By contrast, we use data spanning the period before and after use of ART was common. The clinical literature also pays less attention to possible selection than the birth selection literature (discussed below) suggests is warranted. Our approach deals with any selection in who uses infertility treatment by using exogenous variation in the cost of treatment induced by the insurance mandates for infertility treatment.

3.3 Insurance mandates

Absent asymmetric information, economic theory provides clear predictions for the impact of employer mandates on wages, employment, and insurance coverage; namely that they should all decrease (e.g., Summers

[5]Papers on this topic include Varma & Patel (1987), Sheiner, Shoham-Vardi, Hershkovitz, Katz & Mazor (2001), Levi, Raynault, Bergh, Drews, Miller & Scott (2001), and Tan, Doyle, Campbell, Beral, Rizk, Brinsden, Mason & Edwards (1992).

(1989)). With homogeneous workers, wages should fall until (almost) all of the cost is covered by the worker's lower wages. With heterogeneous workers, firms may respond by cutting insurance availability or shifting employment. All of these predictions hold to some extent for the case of group-specific mandated benefits (e.g., Gruber (1994)). However, in the presence of informational asymmetries, theoretical predictions for the effect of employer mandates on coverage and health are less clear. For example, with adverse selection, even if individuals value a particular form of coverage above their own actuarially fair price, they may not be able to obtain coverage in the individual market without a mandate.[6] If a state has passed a law mandating that health insurers cover or offer to cover infertility treatment, then the state is forcing insurers to subsidize infertility treatment. In themselves, these mandates do not automatically represent a subsidy. If the insurance market worked perfectly and there were no information problems or other inefficiencies, absence of such coverage would merely imply people were unwilling to pay for such coverage. However, even in an efficient market, if such mandates forced insurers to pool infertility treatment with other treatments, it could result in infertility treatment being subsidized. In the case of adverse selection where insurers cannot identify who likely users of infertility treatment are, mandates may also have real effects.

The empirical literature on the effects of mandated benefits on use of health care and health is mixed, with some papers finding little or no significant impact of mandates on utilization or health (e.g., Klick & Markowitz (2006), Pacula & Sturm (2000), and Bao & Sturm (2004) on mental health mandates), while other papers find some significant impacts on health care utilization (e.g., Liu, Dow & Norton (2004) on mandates about maternal postpartum hospital stays).

3.4 Birth outcomes and selection

Our goal as well as that of the studies discussed above is to identify the causal impacts of infertility treatment. However, use of ART or other infertility treatments is not randomly assigned. In fact, higher income women, women with private insurance, white women, and highly educated women are more likely to obtain medical help to get pregnant than other women (Chandra & Stephen (1998), Bitler & Schmidt (2006)). This suggests selection in who uses infertility treatment may color the findings discussed in Section 3.2.

The economics literature on birth production functions, surveyed in Wolpin (1997), focuses more attention on possible selection bias (e.g., Rosenzweig & Schultz (1983)). A frequent question of interest is whether health

[6]Practically, the Employment Retirement Income Security Act (ERISA) of 1974 exempts self-insured firms from state-level mandates. Bitler & Schmidt (2007) estimate that a large share of private sector workers are covered by infertility treatment mandates despite ERISA.

inputs such as prenatal care are associated with infant health. Another is the effect of access to new technologies on health outcomes.

Selection is an important issue when evaluating the impacts of various inputs on infant health. There can be selection both in which women use health inputs and in the health distribution of infants who are actually born. Both types of selection can bias findings. Women who anticipate adverse birth outcomes may be more likely to invest in these inputs, biasing down estimates of the positive effects of prenatal care, for example. An alternate hypothesis posits positive selection; women who initiate care early may engage in other forms of healthy behavior, leading researchers to overestimate the positive impacts of prenatal care. The resolution of the pregnancy itself may also be characterized by selection (e.g., women who think they have very unhealthy fetuses may choose to abort these fetuses). Grossman & Joyce (1990) find evidence of positive selection among black but not white women choosing to carry their pregnancies to term. Ananat, Gruber, Levine & Staiger (2006) and Gruber, Levine & Staiger (1999) show that average outcomes of children born after the legalization of abortion were better, suggesting that access to legal abortion led to positive selection among children who were born. Donohue & Levitt (2001) find that the introduction of legalized abortion is associated with a decrease in crime, a result questioned by Joyce (2004) and defended by Donohue & Levitt (2004). Pop-Eleches (2006) finds that abolition of abortion in Romania is associated with more child schooling and better labor market outcomes for children born after the ban. He finds that this is due to more births to highly educated urban women. Controlling for these observables, children born after the ban had worse outcomes.

This existing literature on birth selection suffers from an inability to separate out heterogeneity or selection among mothers from the impact of investment by mothers. In this literature, heterogeneity and investment are expected to affect birth outcomes in the same direction. Our analysis does not suffer from this drawback. Users of ART and other infertility treatment are infertile and may have more difficulty supporting a fetus through an entire pregnancy than do other mothers despite having invested heavily in their pregnancies. The selection biases from these two factors (heterogeneity and investment) are expected to go in opposite directions for users of infertility treatment.

The demographic literature documents that women in the U.S. are delaying childbearing, perhaps to accumulate more human capital or to advance in the labor market. This delay in fertility may bear costs if maternal age is independently associated with infant health, confounding the selection bias discussed above. Estimates of the negative impact of age on birth outcomes from Royer (2003), using panel data and correlated random effects to deal with maternal heterogeneity, are similar to those from the medical literature (e.g., Cnattingius, Forman, Berendes & Isotalo (1992)).

The existing clinical literature on effects of infertility treatment tends to rely on cross-sectional data. It also pays less attention to possible selection than the birth selection literature suggests is warranted. The economics literature on these mandates focuses mostly on fertility behavior or on use of ART and infertility treatment. We augment this literature by examining the health impacts of the mandates. Next, we turn to a more detailed discussion of the data sets we use.

4 Data and Trends

This paper uses two large individual-level data sets, pooled birth certificate data from 1981–99 and pooled 2000 Decennial PUMS data and 2001–02 ACS data. It also uses data on twin delivery rates, constructed by aggregating data from birth certificates and Census population estimates.[7]

4.1 Detailed Natality data and Natality/Census birth rate data

We use extracts from the Natality data for the years 1981–99 (Division of Vital Statistics, NCHS (Various years)) to look at the effects of mandates on fertility and infant health. These data cover the universe of births in the U.S., including demographic information about the parents such as education, age, and marital status. The data also include many pregnancy outcomes. For the analysis of the share of births which are twins, our sample is a one in fifty random subset of singleton and twin births. For the analysis of infant health outcomes, our sample includes all twin births or a one in fifty random subset of singletons. We also combine counts of pregnancies with live twin births from the Natality data for mothers by age, race, state, and year with Census population estimates for women by age, race, state, and year to construct twin delivery rates.

The first task is to establish that there is a program to evaluate, answering the following question: Have the mandates affected fertility? We consider whether mandates have led to more twins. Many infertility treatments explicitly involve fertilizing and implanting more than one egg in order to increase the chances of a successful pregnancy. Use of some ovulation-inducing drugs can cause multiple eggs to be released, also leading to an increase in the probability of a multiple birth (Callahan, Hall, Ettner, Christiansen, Greene & Crowley (1994)). Thus, a finding of more twinning associated with mandates plausibly means that mandates have affected fertility. This paper is not the first or only paper to consider whether mandates have affected the twin birth or delivery

[7]For more detail about the Natality and PUMS/ACS data and the construction of twin delivery rates, see Appendix A.

rate (Bundorf et al. (2007)) or whether mandates have affected the share of live births that are twins (Buckles (2006)). Nonetheless, we consider these outcomes to establish that mandates are associated with more twins for older women in our data with our preferred specification and controls.

The first outcome is an indicator for whether the infant is a twin (rather than a singleton). If mandates are associated with an increase in the share of births that are twins, it could mean mandates led to more twins without changing the number of deliveries or it could mean mandates caused some women who would not have had children at all to have twins. Thus, we also look at the number of twin deliveries per 1000 women (number of pregnancies resulting in twin live births per 1000 women).[8]

Our third outcome measure is an indicator for whether a twin is part of a mixed-sex twin pair (we also examine whether a twin is part of a same-sex boy or same-sex girl twin pair). Mixed-sex twins must be fraternal (dizygotic) while same-sex twins can be either identical (monozygotic) or fraternal (dizygotic). Wimalasundera, Trew & Fisk (2003) report that while ART is associated with higher rates of both monozygotic and dizygotic twinning, the increase in dizygotic twinning rates with ART is much, much larger. Thus, if mandates lead to a larger increase in the share of mixed-sex twins than in the share of single-sex twins, it would suggest that this increase is the result of more infertility treatment. We create our mixed-sex or same-sex twin indicators by probabilistically matching twins in each pair. Matched twins are those with the same maternal and paternal characteristics, same birth outcomes, and same date and place of birth. We are able to match the bulk of our twin records in this fashion (88 percent).

For the analysis of infant health we focus on three birth outcomes: Birth weight, gestation, and Apgar score.[9] All three of these measures have been linked to infant and child health and other outcomes. We have also considered various cutoffs in the birth weight and gestational age distributions such as low birth weight ($<$ 2500 grams) or premature birth (gestation $<$ 37 weeks); these results are presented in Bitler (2005).

A large medical literature links low birth weight with adverse infant health outcomes (e.g., McCormick (1985), Rees, Lederman & Kiely (1996)). A less settled literature examines the impacts of low birth weight on child morbidity (e.g., McCormick, Brooks-Gunn, Workman-Daniels, Turner & Peckham (1992)). Wolpin

[8]One reason to look at the share of births that are twins in the Natality data rather than just the twin delivery rate is that the individual birth data contain many controls such as race/ethnicity, completed education, and previous births that might themselves be associated with fertility, so including them may improve precision. It is also impossible to get population data on women by year/state/age/race as well as completed education and parity in order to get denominators for twin deliveries per 1000 women at the more disaggregate level.

[9]The 5-minute Apgar score is used to evaluate the condition of the newborn infant at 5 minutes after birth (Division of Vital Statistics, NCHS (2000)). It is a summary measure of the infant's condition based on heart rate, respiratory effort, muscle tone, reflex irritability, and color. Each of these is given a score of 0, 1, or 2; their sum is the Apgar score. A higher score is better. The Apgar score is a good predictor of surviving the first year.

(1997) surveys the literature on the observed relation between birth weight and infant mortality, while Almond, Chay & Lee (2005) suggest the relationship between birth weight and short term health may not be as strong as the previous literature suggests. Numerous studies suggest that birth weight is correlated with later outcomes such as self-reported health status, earnings, and educational attainment (e.g., Behrman & Rosenzweig (2001), Currie & Hyson (1998), Black, Devereux & Salvanes (2007)).

Evidence also suggests that premature infants are at risk of worse health outcomes later in life (e.g., McCarton, Wallace, Divon & Vaughan (1996), ACOG Committee on Practice Bulletins—Obstetrics (2002)). For example, nearly one-fifth of very preterm births (gestation under 32 weeks) in 1999 did not survive the first year (Martin, Hamilton, Ventura, Menacker & Park (2002)).

The Apgar score is an alternate measure of infant health at delivery. Almond, Chay & Lee (2004) discuss the existing literature on the importance of birth weight and offer evidence that the 5-minute Apgar score may be a better predictor of infant mortality among twins.

4.1.1 Means

Figure 2 shows the share of infants that are multiples for each state during 1981–99 from the Natality data. There has been an increase over this period in the share of live infants who are multiple births, and the upward trend seems to pick up around the late 1980s. Figure 3 shows that the average age of mothers of twins has risen considerably over this period. The share of twins that are mixed-sex increased from 29 percent of twins in 1981–88 to 33 percent in 1995–99. Average birth weight and gestation for twins have both gone down over this period, perhaps due in part to the dissemination of and tremendous improvements in technology for keeping infants alive.[10]

Panel A of Table 3 contains means for the share of births that are twins (as a share of twins and singletons) and for the share of births that are mixed-sex and same-sex (girl or boy) twins for two periods: One when there was no state mandate for infertility treatment in effect during the year before the infant was born (column 1) and the other when there was such a mandate in effect (column 2). Twin records which matched no other twin or more than one twin are dropped from the calculations involving matched twins. Twins are a larger share of births when mandates were in effect (this difference is highly significant, with an F statistic of 70). The increase is about twice as large for the mixed-sex twins as the same-sex twins. Panel C of Table 3 contains the mean

[10]For a discussion of the benefits of these new technologies, see Cutler & Meara (2000). For a discussion of factors affecting their dissemination, see Baker & Phibbs (2002).

number of twin deliveries per 1000 women (pregnancies resulting in twin live births per 1000 women) for each time period. Twin delivery rates are also considerably higher when mandates were in effect—about 16 percent higher.

Panel B of Table 3 presents average birth outcomes for the two time periods (mandates/no mandates) for the sample of twins. All of the means in panel B of Table 3 are statistically significantly different across the two periods except for birth weight and low birth weight. Average twin gestation is slightly lower when mandates are in effect while twin birth weight is about the same in both periods and twin Apgar scores are higher.

These means show a substantial increase in the share of births that are multiples over the past 20 years. Twin births are more likely and the twin delivery rate per 1000 women is higher in states with mandates. Twin mothers have grown older and more highly educated over time (latter not shown). Health outcomes for twins vary by the presence of mandates. These differences may be the result of the mandates and increased use of infertility treatment or they may merely reflect changing trends.

4.2 PUMS and ACS data

We use PUMS and ACS data to look at longer-term health impacts of the mandates. The PUMS/ACS data include information on mother's age, state of residence at birth, race and ethnicity, disability status, educational attainment, and a host of other variables. We pool the PUMS and ACS to maximize sample size. Our final sample is children aged 5–17 (health outcomes are unavailable for children 0–4). This is a panel of children still alive and residing in the U.S. in 2000–02 who were born between 1982 and 1997. Note that the most sickly children may have died before they appear in our PUMS/ACS sample. This should bias us against finding long-term effects of the mandates.

Our first outcome of interest is whether a child is a "twin." We identify twins in the PUMS/ACS data by using information about children's age, relationship to the householder, and the presence of their mother. For example, two children of the same age in the primary family with the same relationship to the householder would be identified as twins. Because the PUMS/ACS data only report children's age in years, this procedure will incorrectly identify as twins some sets of siblings who were born only 9–11 months apart. However, this measurement error should also attenuate our coefficients.

Our main sample is restricted to children or grandchildren of the householder, in order to be more certain the "twins" are biological siblings. We also link the children's records to their likely mother. Then, mother's age at the child's birth is mother's current age minus child's age. Unfortunately, our only education measure for the mother is her current education, which is not necessarily the same as her education at the time of the birth(s). We link the child record to characteristics of the state where the child was born in the approximate year the child was conceived ($2000 - age - 1$).[11]

Our longer-term health outcomes are measures of disability status. They are (1) does the person have long-lasting blindness, deafness, or a severe vision or hearing impairment?; (2) does the person have a long-lasting condition that substantially limits one or more basic physical activities?; (3) because of a physical, mental, or emotional condition lasting 6 months or more, does the person have difficulty learning, remembering, or concentrating?; and (4) because of a physical, mental, or emotional condition lasting 6 months or more, does the person have difficulty dressing, bathing, or getting around inside the home? The clinical literature associates a higher risk of having sensory and physical impairments and physical and mental limitations with being born very preterm or of extremely low birth weight.[12]

4.2.1 Means

Panel A of Table 4 contains means for children 5–17 who are singletons or twins and in our final PUMS/ACS sample. Column 1 contains the average probability of being a twin for twin and singleton children born in states with no infertility treatment mandate the year before birth (approximate year of conception) and column 2 for children born in states with mandates the year before birth. We see that in the PUMS/ACS data, as in the Natality data, children are more likely to be twins if born in mandate states. The rate of twinning in the PUMS/ACS data is a bit higher than that found in the Natality data. This is partially due to our miscoding some siblings as twins but could also be due to compositional differences between twins we can identify (who must still be alive and coresident with their parents in the PUMS/ACS) and all twins born. Panel B of Table 4 contains summary statistics for the disability measures, for the twin sample only, by mandate status during the year of conception. Here we see that the long lasting impairments and conditions limiting the ability to concentrate or learn or the ability to dress or get around the house appear more common for twins born in states without mandates in place at conception.

[11] For more details on PUMS/ACS data, see Appendix A.

[12] Vohr, Wright, Dusick, Mele, Verter, Steichen, Simon, Wilson, Boryles, Bauer, Delaney-Black, Yolton, Fleisher, Papile & Kaplan (2000) report neurodevelopmental, neurosensory, and functional outcomes of extremely low birth weight (401–1000 gram) survivors being tracked by the NICHD Neonatal Research Network. Extremely low birth weight children born in 1993 and 1994 were assessed at age 18–22 months. Only 51 percent of the children had normal neurodevelopmental and sensory assessments, and outcomes worsened as birth weight decreased. Colvin, McGuire & Fowlie (2004) review the evidence about neurodevelopmental outcomes after preterm birth, noting that very preterm infants (many of whom are also low birth weight) are at increased risk of vision and hearing problems, even though most preterm infants have good neurodevelopmental outcomes.

5 Empirical Model

We would like to know the direct impact of infertility treatment on fertility and health. However, use of treatment is both potentially endogenous and rarely recorded in public-use data sets. Thus, we compare women who are likely to have easy access to infertility treatment (live in mandate states) to women who do not. Infants born via use of ART or other infertility treatment, while growing in number over the 1980s and 1990s, still represent a small share of total births. We focus on a population where they are likely to be a larger share of overall births, namely twins. We further focus on older women (30 and older) as the group most likely to have demand for and be using infertility treatments.[13] This is an intent to treat estimator, since even among older twin mothers, most women are not infertile and have not used any infertility treatment.

We run ordinary least squares regressions (OLS) of outcome measures on demographic covariates; state-level controls; state and year fixed effects; policy variables related to infertility mandates; and interactions of the policy variables with maternal age indicators. Results for birth outcomes that are binary variables are generally robust to running the regressions as probits and calculating appropriate marginal effects. Where the probit results differ, we discuss it in the text.

The regressions have the following form:

$$y_{ist} = X_{ist}\delta + S_{st}\alpha + IT_{st}\beta_1 + A_{ist} \cdot IT_{st}\beta_2 + \gamma_s + \nu_t + \varepsilon_{ist}. \tag{1}$$

In the Natality fertility regressions, y_{ist} is an indicator for being a twin in a pooled sample of twins and singletons from the Natality data or the PUMS/ACS, or for being a mixed-sex or same-sex twin in the Natality data. In the twin delivery rate regressions, y_{ist} is the number of pregnancies yielding twins per 1000 women in a given age-race-state-year cell.

In the health regressions using Natality data, y_{ist} is birth weight, length of gestation, or the 5-minute Apgar score. We examine different birth outcomes to test whether our findings are robust. We present results for birth weight first, as it is both very reliably measured (gestation is dated from the last menstrual period in much of the data) and often the focus of the literature. For each outcome in the Natality data, we only use the balanced panel of states and years where the outcome was reported by the state for every year. This maximizes sample for each

[13] Table 1, which presents summary statistics from the 2002 NSFG, suggests that 18.2 percent of women 30–44 reported ever having obtained any infertility treatment (including lower cost treatments such as use of ovulation-inducing drugs) while only 6.3 percent of women 15–29 had gotten such treatment. Women over 30 also represented the lion's share of users of IVF and other ART (Society for Assisted Reproductive Technology, American Society for Reproductive Medicine (2002)).

outcome.[14]

In the health regressions using PUMS/ACS data, y_{ist} is any of the four disability measures: (1) being blind or deaf or having a severe vision/hearing impairment; (2) having a long-lasting condition that substantially limits physical activity; (3) having a condition lasting at least 6 months that causes difficulty learning, remembering, or concentrating; and (4) having a condition lasting at least 6 months that causes difficulty dressing, bathing, or getting around inside the house.

X_{ist} is a vector of demographic characteristics associated with fertility and health outcomes. All regressions include controls for mother's age and race. All individual regressions also control for infant/child sex. The Natality and PUMS/ACS individual-level regressions also control for Hispanic ethnicity (or an indicator for Hispanic ethnicity not being reported) and education (high-school graduate, some college, four-year college degree, or unreported). Regressions using the Natality data additionally control for the number of previous live births.

S_{st} is a vector of state-level demographic, labor market, and public assistance controls that may be associated with fertility or the level of health care available to women. S_{st} includes the unemployment and aggregate employment growth rates, real median income for a family of four in the state, real annual AFDC/TANF benefits for a family of 4, the cutoff for Medicaid eligibility for a pregnant woman in the state as a share of the federal poverty level, the share of the state population that is under the poverty level, the share of births to unmarried mothers in the state, and the percent of the total state population that is black or Hispanic.[15]

The γ_s and ν_t terms represent state and year fixed effects. The state (year) fixed effects control for unobserved factors that differ across states and not over time (over time and not across states). Regressions using the Natality data also contain month of birth fixed effects. Unobservable determinants are captured by ε_{ist}.[16] All regressions and summary statistics are weighted to be population representative of children (share of births that are twins or health outcomes) or women (twin delivery rates). Standard errors are clustered at the state level. This accounts for the fact that the key policies—infertility mandates—only vary at the state-by-year level. It also addresses concerns raised by Bertrand, Duflo & Mullainathan (2004) and Kezdi (2002) about over-rejection of the null in the presence of serial correlation with differences-in-differences methodology using state policy reforms.[17]

[14] These restrictions are more onerous for some outcomes than others. New Mexico is the only state excluded from the gestation regressions. California, Delaware, and Oklahoma are excluded from the Apgar score regressions.

[15] Means for these state-level variables for the sample of twins from the Natality data are presented in Appendix Table 1. Sources for all state-level variables are discussed in Appendix B.

[16] Robustness tests discussed in Section 6.5 below also include state-specific linear time-trends or controls for neonatal intensive care availability at the state-by-year level.

[17] We have also estimated the models with state-by-year level clustering. While the standard errors for some coefficients on state-by-

Our main focus is on the coefficients of interactions of the mandates with a dummy for the mother being 30 and older. We interact mandates with age because older women are those most likely to be suffering from infertility or impaired fecundity, and thus more likely to be affected by the mandates. We have included main age and mandate effects. The coefficients on the age-mandate interactions represent estimates of the effects of the mandates on older women. To the extent that younger women also are affected by the mandates, the age-mandate interactions may understate the effect on the older women. If younger women were entirely unaffected, then the main mandate variable is an additional control for differential fertility trends in mandate states, and our age-mandate interaction would be a differences-in-differences-in-differences estimate. However, we know that a small share of the younger women do obtain treatment.

Effects of mandates on some (mostly older) women who know they are infertile should be immediate—some women will take up treatment who previously did not and others will intensify their use. It is these effects which we will pick up with our strategy. The effects of mandates on other (mostly younger women) is to increase their incentives to delay childbirth. This delay will take time, and likely cannot be picked up here. We will return to discussion of this issue in the results section.

We investigate three different codings of the insurance mandate variables. The first is simply an indicator for whether the state has any mandate concerning infertility treatment. The second splits the "any law" indicator into two variables, 1) an indicator for whether the state law mandates that health insurers cover or offer to cover infertility treatment but specifically excludes IVF and 2) an indicator for whether the state law mandates that insurers cover or offer to cover infertility treatment but does not exclude IVF. The third coding splits the "any law" indicator by whether the mandate requires insurance companies to cover treatment or merely offer coverage to employers purchasing insurance policies.[18] The indicator variables for the mandates are set to one for every year after the year the law passed (treatment and a successful pregnancy usually take at least 9 months).[19] Table 2 shows that there is considerable cross-state over-time variation in when and where the laws were passed.

If states adopt mandates in response to differential trends in twinning or birth outcomes among twins, it raises important questions about our empirical strategy. A key condition for our model is that pre-existing trends be similar in the treatment (older women in states with mandates) and control groups (older women in states

year level variables are larger with state than with state-by-year level clustering, many are smaller.

[18]Many of the other papers on effects of insurance mandates focus on "cover" mandates or mandates that include IVF. We have examined all three codings for robustness sake and because we think it is plausible that all types of mandates could have impacts. For example, mandates excluding IVF itself may still cover expensive aspects of an IVF cycle such as ovulation-inducing drugs. We discuss this further in the results section.

[19]Robustness regressions also include two-year lags of the laws, with no substantive differences in findings.

without mandates). We test for this by looking at trends in the share of twin plus singleton births that are twins and in birth weight, gestation, and Apgar score for twins, during the period 1981–84 for all states but West Virginia, whose mandate was passed in 1977. We regress state-year level means on a trend for the states that ever pass a mandate after 1984 and states that do not, as well as a dummy for being a mandate state. We then test whether coefficients on the trend variables are the same. We cannot reject that they are equal. We also test whether leads of the mandate variables are important in the regressions, and find that they are not.

We now turn to results.

6 Results

6.1 Mandates, the probability of twinning, the twin delivery rate, and the share of mixed-sex twins

Figure 4 shows the share of births that are twins (among twins and singletons) for the treatment group (women 30 and older) and the control group (women 29 and under) during the four years before a mandate was passed and the 7 years after.[20] Year 0 is the year after the year when the mandate was passed (approximate year of conception) and other years are dated similarly with respect to the policies. The solid line shows the share of births that are twins for women 30 and older. It appears that there is a trend break in this series at year 0. For comparison, the dashed line shows the share that are twins for women 29 and under. While this line bounces around some, there is no trend break visible at year 0.

Next we turn to the regression results. First we present results for the regressions predicting the share of twin and singleton births that are twins. Table 5 reports coefficients on our key mandate variables and their interactions with the mother being at least 30 for these regressions, using a one in fifty random sample of singleton and twin births from the 1981–99 Natality data. Each column presents coefficients from a different regression including a different set of mandate variables and all of the other controls discussed above.

First, note that some of the main mandate effects are negative, and that for cover mandates is statistically significant. This suggests a smaller share of births to younger women are twins after cover mandates are in effect. This is also true for some main effects in the regressions predicting twin delivery rates reported in Table 6. If the younger women were unaffected by the mandates, then we could interpret the age-mandate interactions

[20]The data points for each age group are the coefficients on dummies for the number of years since a mandate was passed from weighted regressions of the share of births that are twins for that age group on the year dummies and state fixed effects.

as differences-in-differences-in-differences estimates; in this case the main effects would hopefully have been a precisely estimated zero. However, we know some share of the younger women are using the treatments. We have explored the negative effect further by interacting age group with mandate, and find that it is driven by the youngest women (women under 25) who are indeed unlikely to be obtaining treatments. We also note that the analogous main effects in the regressions predicting health outcomes are almost all insignificant, and additionally are often the same sign as the age-mandate interactions.

Turning to the mandate-age interaction effects, the first column shows that children born to older mothers in mandate states are more likely to be twins. The coefficient of 0.00308 represents about a 10 percent increase in the probability of being a twin, compared to the pre-mandate mean of 0.0295 for children with mothers 30 and older. Running this regression as a probit and calculating the average marginal effect for the interaction (accounting for the main effects) leads to a average marginal effect of the mandates for children of women 30 and older of 0.00225 (standard error of 0.0014 and p-value of 0.096, with the standard error of the average marginal effect and p-value calculated using the delta method).

Turning to column 2, we see that whether or not the state excludes IVF from the mandate, children of women 30 and older in mandate states are more likely to be twins, although the coefficients are insignificant. Finally, column 3 presents results with the mandate variable allowed to vary by whether it is a mandate to cover or offer to cover treatment. As one might expect, the age interaction with a mandate to cover has a large, statistically significant coefficient, representing about a 23 percent increase relative to the pre-reform mean. Here the marginal probability from a probit regression is about the same size and is significant.

Table 6 shows coefficients from regressions predicting the twin delivery rate per 1000 women.[21] Mandates are associated with a significant increase in the twin delivery rate for older women. The effects are largest for cover mandates, then mandates that exclude IVF, and then any mandate. Overall, the mandates are associated with an increase in the twin delivery rate for women 30 and older of between 24 and 40 percent. These effects may seem large. However, nearly 10 percent of all twin births are due to use of ART and an unknown number due to use of other infertility treatments. Bitler & Schmidt (2007) find evidence that mandates are associated with a large (relative to the baseline) increase in ever having used infertility treatment for women 15–44.

Table 7 presents results from regressions predicting the probability that an infant is a mixed-sex twin. Again the effects of mandates for children of older mothers are statistically significant and large. Being born to a

[21]Note that these are OLS regressions weighted by female population, so the effect of mandates on twin delivery rates is simply one-half the effect on twin birth rates, since the number of twin births per 1000 women is twice the number of twin deliveries per 1000 women.

mother 30 or older in a mandate state is associated with about a 20 percent increase in the probability of being a mixed-sex twin. The analogous impact of being born to a mother 30 or older in a mandate state on the probability of being a same-sex boy or girl twin is about 10 percent. We expected a considerably larger impact of mandates on the share of births that are fraternal twins than on the share that are identical births. All of the mixed-sex twins are fraternal while a large share of the same-sex ones are identical. Thus, the fact that the impact on the share of births that are mixed-sex twins is much larger than that on the single-sex twin birth share supports our interpretation that this is indeed a result of the mandates.

Lastly, we note that we can replicate our twinning results from the Natality data in the PUMS/ACS data. Appendix Table 2 shows results from regressions predicting twin births among children 5–17 using PUMS/ACS data. Here, the mandate variables are interacted with mother's age being at least 30 at birth. While the age-mandate interactions are insignificant for the "any mandate" and IVF excluded/included variables (likely due to attenuation from measurement error), column 3 shows that being born to an older mother in a cover-mandate state is associated with a statistically significant increase of about 17 percent in the probability of being a twin in the PUMS/ACS data. For brevity, we will only report regression results for the disability outcomes controlling for the cover and offer mandate variables.

Because twin births are higher risk than singleton births, the fact that the mandates are associated with an increase in twinning and in twin delivery rates indicates that the mandates have had a negative effect on infant health even if this were their *only* impact. We next describe our findings about the effects of the mandates on infant health.

6.2 Mandates and infant health outcomes for twins: Detailed Natality data

The infant-health results for twins are reported in Tables 8–10 for the Natality data. Each table has a similar format and contains selected coefficients from three separate regressions predicting each outcome.The only coefficients shown are those on the main mandate variables and the interaction of the mandate variables with the mother being 30 and older. The first column presents results from regressions with a dummy variable for the state having any mandate and its interaction with the mother being 30 and older. The second column presents results from regressions with two separate dummy variables for the type of insurance mandate the state has (including or excluding IVF) and their interactions with the mother being at least 30. The third column regression instead splits the law according to whether the mandate is to cover or offer to cover treatment and includes age interactions with those variables. We focus the discussion on the age-mandate interactions.

Table 8 present results for regressions explaining birth weight for twins. First, note that the main mandate effects here are the same sign as the age-interactions and are never significant, in contrast to the results for twinning. Column 1 of Table 8 shows that for twins of women 30 and older, being born in a state with any infertility mandate is associated with a statistically significant decrease in birth weight of around 16 grams (the mean at baseline for women 30 and older is 2475 grams). This is not a large effect. However, the only women who should be impacted by this law are those using infertility treatment; thus a zero impact on other women is being averaged with a larger effect on women using treatment, leading to a small overall effect.

A simple back of the envelope calculation gives a sense of the possible overall magnitude of these results for the "treated" group (twins with older mothers in mandate states). Society for Assisted Reproductive Technology, American Society for Reproductive Medicine (1999) suggests that around 6,360 twins were born after use of IVF or other ART from pregnancies initiated during 1996. Suppose twice as many women had twins due to any use of infertility treatment as did due to use of ART and that two-thirds of the women using infertility treatment are at least 30 (as opposed to about 46 percent of all twin mothers). Suppose further that all twin births to women 30 and older were evenly split between insurance-mandate states and other states. Then, the mandate's effect on birth weight for twins with mothers 30 and older in the mandate states who were actually impacted by the mandate would be about $(-16*104,000*0.46*0.5)/(6360*2*0.67*0.5) = 90$ grams, if all the twins were born in 1997 (104,000 twins were born in 1997), or about a 3.6 percent decrease.

This calculation assumes the mandates have no effect on women not using infertility treatment. It could be an overestimate of the effect of mandates on the treated group if the share of twins born after infertility treatment as a fraction of total twins is larger in mandate states than in other states. It could also be an underestimate if the number of twins born to older women using ART is less than half the number born to older women using other infertility treatments.

Column 2 shows that the impact is significant only for women in mandate-but-no-IVF states, although the point estimate for living in a mandate-and-IVF-covered state is also negative. For twins with mothers of at least 30, being born in a state with a mandate excluding IVF is associated with being 23 grams lighter. Column 3 shows that for twins with older mothers, cover and offer mandates are both associated with being lighter. Older women in cover-mandate states have twins that are on average 12 grams lighter, and in offer-mandate states have twins that are on average 20 grams lighter.[22]

[22]Results for low birth weight, very low birth weight, and "extremely low" birth weight (being under 1000 grams) are reported in Bitler (2005). The pattern of the coefficients is similar for regressions predicting low birth weight, very low birth weight, or extremely low birth weight. Any mandate, mandates excluding IVF, and offer mandates are associated with increases in the share of all measures

Column 1 of Table 9 shows that twins born to older mothers in mandate states are shorter gestation infants but the coefficient is not significant. Column 2 of Table 9 shows that forcing insurers to cover infertility treatment but not IVF is negatively associated with length of gestation for older twin mothers. The coefficient suggests a decrease of around 0.089 weeks on a baseline of 36.1 weeks for women 30 and older. A rough calculation similar to the one above for birth weight suggests that this translates to an impact of about -0.50 weeks for the women actually affected by the mandate. Finally, column 3 suggests that for older mothers, offer mandates are associated with statistically significantly shorter gestation periods.[23]

Table 10 presents results for regressions predicting the 5-minute Apgar score. For Apgar score, an important measure of infant health, any mandate, mandates that include or exclude IVF, and mandates to cover or offer coverage are all associated with statistically significantly lower scores for children of older mothers. Again, the effects here are small, but a back of the envelope calculation suggests a cover mandate is associated with a 1.6 percent decrease in Apgar score for the group affected.

In sum, the results for the Natality data suggest a statistically significant but small negative association between important health outcomes for twins—birth weight, gestation, and the 5-minute Apgar score—and being born to an older mother in a mandate state. Effects are negative and significant for birth weight and Apgar scores for older mothers in states with any mandates. Effects are negative and significant for Apgar scores for older mothers in states with mandates that include IVF and for all three outcomes for older mothers in states with mandates that exclude IVF. Effects are also negative for older mothers for all three outcomes for offer mandates, and for birth weight and Apgar score for cover mandates.

One might expect the effects of insurance mandates to be larger in states with mandates covering IVF than in states with mandates excluding IVF. However, there are alternate hypotheses that are consistent with both our findings and with the findings from the medical literature. Jain et al. (2002) suggest that women in mandate-but-no-IVF states may be transferring more embryos since they may be able to afford fewer cycles of IVF than women in mandate-and-IVF-covered states although Bundorf et al. (2007) suggest that Jain, Harlow, and Hornstein's finding is due to differences in the composition of who obtains treatment across states. Alternatively, if women in mandate-but-no-IVF states substitute away from IVF and toward intrauterine insemination or other procedures that also increase the multiple-birth rate, then there may be more infertility treatment related twins

of lighter twins among twins with older mothers. Cover and IVF-allowed mandates are both associated with a higher probability of being a low birth weight twin among twins with older mothers.

[23]Twins with older mothers are more likely to be premature (gestation < 37 weeks) or very premature (gestation < 32 weeks) in states with any mandate, with a no-IVF mandate, and with an offer mandate.

born in the mandate-but-no-IVF states than in the mandate-and-IVF-covered states.[24] Regardless, these findings suggest that mandating insurance coverage of infertility treatment may have had unexpected consequences.

Another natural question is why we would find any effects of offer mandates. One explanation is that, at least in the late 1990s and early 2000s, many states with any sort of mandate also cover infertility treatment for state employees (perhaps on equity grounds).[25] This might lead to effects even in offer states. Another possible explanation is that in states with offer mandates, firms choose to cover infertility treatment as a way of recruiting or retaining highly productive employees.

We anticipated finding impacts of the mandates among twins if there indeed were any impacts of the mandates. Still, one might expect the twins results to reflect predominately use of ART. It is possible that findings for singletons would instead pick up use of less aggressive infertility treatment such as ovulation-inducing drugs. The estimates should also be smaller in magnitude for singletons, given the smaller share of women in mandate states actually using infertility treatment. Thus, looking at singletons should provide a specification check for our twin results.

6.3 Mandates and infant health outcomes for singletons: Detailed Natality data

Table 11 presents results for the one in fifty subset of singletons from 1981–99 for regressions predicting birth weight. Results for gestation and Apgar score are in Appendix Tables 3 and 4. Most of the mandate-age interactions are negative as in the twin regressions, and several are statistically significant. Further, the coefficients are considerably smaller in magnitude as would be expected given the (likely) much smaller share of singletons born after use of infertility treatment. The negative effects for birth weight are only statistically significant for older mothers in the mandate-but-no-IVF states. The negative effects for women 30 and older for gestation are only significant for mandates which include IVF. Curiously, effects on Apgar score are significant and positive for the age interaction in offer states.

Next, we examine the longer-term health impacts of these mandates.

[24] We thank Vivian Ho for suggesting this hypothesis.

[25] Historical data on state coverage of infertility treatment for public employees is unavailable, else controls for this would also be included in the regressions.

6.4 Mandates and child health outcomes for twins: PUMS/ACS data

Table 12 presents results from OLS regressions predicting longer-term health impacts for twins born from 1982–97, with cover and offer mandates included. Here the mandates are interacted with an indicator for the mother being at least 30 at conception. Most of the point estimates for the mandate-age interactions are positive (being a twin born in a mandate state to an older mother is associated with a higher probability of having the disabilities), and one is statistically significant. The OLS regressions suggest that cover mandates are associated a higher prevalence of a condition limiting physical activity (column 4, coefficient of 0.0092, standard error of 0.0046). Coefficients from regressions with mandates interacted with the mother being at least 35 are 0.0082 but no longer significant. However, the mandate-age interactions for physical impairments are now a marginally significant 0.0270 (standard error of 0.0144).

We have also estimated probits for the disability outcomes. Some of the coefficients on the age-mandate interactions are imprecisely estimated but none suggest any positive effects of being born in a mandate state to an older mother. Like the OLS regressions, they suggest more conditions limiting physical activities and more long-lasting physical impairments among twins born in cover-mandate states to older mothers. The average marginal effect on a condition limiting physical activity of being born to a mother 30 and older in a cover-mandate state is 0.0076 (standard error of 0.0035, *p*-value of 0.029). This is relative to a baseline mean of 0.0091, so it represents a very large effect, perhaps too large to be entirely credible. Taken as a whole, we interpret these findings as suggesting that long-lasting physical impairments and conditions limiting daily physical activities may be more likely for twins born to older mothers in mandate states.

6.5 Robustness

We experimented with a number of subsamples and specifications to test the robustness of the results. Below, we discuss the impact of these changes on the main results. More detail on the robustness regressions is available on request.

Restricting the Natality samples to women more likely to have used infertility treatment—women with no previous live deliveries before this pregnancy or college graduates with no previous live deliveries before this pregnancy lead to larger (in magnitude) estimated coefficients and similar patterns of significance. Including single year of age dummy variables did not change key coefficients of interest nor did estimating regressions with the unbalanced panel of all observations available for each outcome. Results were also robust to including age by year, age by state, and year by state interactions (in this case, only effects for older women are identified). Regressions including leads of the policy variables showed that the leads were not significant, addressing concerns about policy endogeneity.

One concern about our findings is that they are merely picking up the effects of advanced equipment to save babies in mandate states. We test this by including measures of neonatal care technology dissemination in the regressions as well as by including state-specific linear time trends.[26] Our measures of access to these technologies are the total number of obstetric hospital beds, neonatal intensive care hospital beds, and neonatal intermediate care hospital beds per 1000 women aged 15–44 in each state for various years from American Hospital Association Surveys. Results from regressions controlling for neonatal care dissemination or trends were substantively similar to the main results.

Another concern may be that our current mandate variables are not the right measure of which women are affected by the mandates. We have tried to address this concern by testing the effects of any mandate, mandates which include or exclude IVF, and mandates to cover or offer to cover infertility treatment. There is, however, another relevant issue, namely that mandates do not apply to all the women in a given state. Ideally, we would have information for each woman on whether she had private insurance that was covered by state mandates (not exempt through ERISA). But no such information is available in our data (or others we are aware of). Instead we use March CPS data from 1980–2000 to calculate the share of women aged 15–44 who are covered by private insurance. We replace our mandate dummies with the share of women at risk of being affected by the mandates—the share of women 15–44 covered by private insurance in mandate states. Thus, the new mandate-age interaction is the interaction of the share privately covered with the mother's being 30 or older in mandate states. We also include the share of women covered by private insurance, and the interaction of the privately insured share with the mother's age being 30 and older. The results are quite robust to this. For example, in the regression predicting birth weight, the coefficient on the interaction of "share of women with private insurance in mandate states" with the mother being 30 and older in the birth weight regressions is -17 grams and it is significant at the 1 percent level.

A second robustness test considers whether the mandate only applies to HMOs, excludes HMOs, or covers all private insurers. Here, our measure of the share of women likely affected by the mandate is one if the mandate applies to all private insurers, the HMO penetration rate in the state if the mandate is only for HMOs, and one minus the HMO penetration rate if the mandate excludes HMOs. Again, the regressions include the level of

[26]Dissemination of advanced neonatal care technology may reflect demand for these technologies from women seeking infertility treatment, suggesting these coefficients may be biased. Thus, we exclude these measures from our main results.

HMO penetration and its interaction with mother's age. The results here are also robust. The interaction of this mandate variable and the woman being 30 and older is negative and statistically significant at the 1 percent level for birth weight (-25 grams).

Given the incentives firms face to self-insure when facing mandates due to ERISA and the fact that self-insurance is cheapest for the largest firms, one might expect different effects according to the size of firm offering private insurance to women. We constructed from March and May CPS data the share of women who worked or whose husbands worked at firms with ≤ 24 employees, 25–99 employees, and ≥ 100 employees. We interact these employment size variables with mandate, age of mother, and mandate times age of mother, and include the main employment size variables as well. These models have a large number of interactions and are rather hard to interpret, and generally, imprecisely estimated.

Overall, these various exercises suggest that the results for gestation, birth weight, and the 5-minute Apgar score are quite robust.

The findings from the PUMS/ACS data of longer-term impacts of the mandates for twin children of older women are more mixed in general. However, we have performed many of the same robustness checks as above (including single year of age dummies, controlling for neonatal care availability, controlling for the share of the population affected by the mandate). Given the ambiguity of assigning the maternal characteristics (i.e., some of the "mothers" may be step-mothers), we also explored a variety of other samples. These results are fairly consistent with the ones we report in the text, namely the mandates for older mothers are generally associated with higher rates of physical impairment or conditions limiting physical activity.

7 Discussion

We have shown that mandates to private insurers to cover infertility treatment have increased both the twin delivery rate and the share of births that are twins among women 30 and older. Even if the mandates had no other impacts, this would imply that they have had negative effects on infant health. Complications for the mother such as pre-eclampsia are higher for mothers carrying twins, as is maternal morbidity (Sibai, Hauth, Caritis, Lindheimer, MacPherson, Klebanoff, VanDorsten, Landon, Miodovnik, Paul, Meis, Thursau, Dombrowski, Roberts & McNellis (2000)). Prematurity and intrauterine growth restriction, which are the main causes of neonatal morbidity and mortality for multiples, are also much more common for multiples than for singletons. Infant mortality rates for twins in 1995–97 were 4.8 times those for singletons (Martin & Park (1999)). Among surviving twins, the risk of severe or overall handicap is also significantly higher than for singletons (Luke & Keith (1992)).

Among twins, being born to a mother of 30 or older in a mandate state is associated with being a lower birth weight, shorter gestation infant with a lower Apgar score, and some of these relationships hold for singletons as well. As discussed above in section 4.1, the clinical literature shows that preterm and low birth weight infants are at higher risk of a range of adverse neonatal outcomes. There is some evidence in Census data that there may be longer-term negative health impacts of these mandates. At the very least, the negative health outcomes for infants suggest that for these women, negative selection or possibly effects of the treatments themselves outweigh any positive effects of investment by these women.

One important rationale behind provision of insurance is to compensate persons who experience a negative health draw. These mandates may be compensating women with low fecundity so they do not need to bear the full costs of their poor health. However, much of the rhetoric about these mandates has been based on claims about their low social costs. If these social costs are large, then mandates may drive up costs of providing insurance. Mandate-associated increases in twin births and lower birth weight, shorter gestation, and lower Apgar scores among twins are likely associated with higher public costs (some states have programs to help pay for very high cost deliveries). If women are delaying childbearing based on the assumption that infertility treatment (subsidized by these mandates) will permit them to have children at older ages, and if this is associated with worse outcomes, this may represent a type of moral hazard (as pointed out by Bundorf et al. (2007)). It is important to remember that there are benefits associated with families being able to have children due to the mandates who otherwise could not have. However, quantifying these costs is still useful, since a large share of them will be borne by the public sector.

We use information about the distribution of hospital costs for live births by birth weight (from Schmitt, Sneed & Phibbs (2006)) to estimate the additional costs of imposing these mandates relative to a no-mandate world. There are two sources of increased costs of mandates. The first stems from the increase in the share of births to older mothers that are twins. The second is because twins born to the older mothers in the mandate states are of lower birth weight. We use information on the actual distribution of birth weight for twins and singletons and the distribution of costs by birth weight category to estimate the hospital cost of an additional twin born to an older mother and the cost of the mandate-associated decrease in birth weight for older mothers. For example, our estimate of the cost associated with increased twinning of imposing a mandate in every state is the product of the number of additional twins and the cost per additional twin.[27]

Table 13 presents the components for calculations of the cost of imposing various mandates from increased twins (rows 5 and 7), from lower birth weights among twins (row 10), or from both (row 11). It also presents the increase in the probability that a twin will die before first hospital discharge associated with the decreased birth weight from mandates (row 9).[28] For example, we found that children born to older women in states mandating coverage were about 23 percent more likely to be twins. This suggests that if every state had a cover mandate, it would mean another 12,600 twins per year (relative to a no-mandate world), using the number of twins born to women over 30 in 1999 ($54820 * 0.22944 = 12,578$, column 4 of row 2). This translates into added costs per year of $320 million for imposing a national cover mandate if the twins were all born to women who otherwise would not have had children. If, instead, the twins were all born to women who otherwise would have had singletons, the added costs per year of a cover mandate would be $228 million.[29] The fact that being born to an older mother in a cover mandate state lowered average birth weight for twins by 12 grams suggests additional costs of about $14 million in 1999 of implementing a cover mandate in every state.[30]

Taken together, these estimates suggest that imposing a mandate to cover infertility treatment in every state would lead to an overall increase in hospital costs for women and infants of $334 million per year if the new twins were all born to women who otherwise would not have had children and an increase of $242 million if the new twins were all born to women who otherwise would have had singleton births. To put these cost figures in context, the national cost of IVF cycles in the U.S. in 2000 was about $470 million and the overall cost of ovulation stimulation in 2000 was about $257 million (Collins & Graves (2000)). This suggests that the hospital related cost of imposing a national mandate to cover would be 71 percent of the total annual cost of IVF and 130 percent of the total annual cost of ovulation stimulation. IVF and ovulation stimulation are

[27] Schmitt et al. (2006) link birth certificates, death certificates, and infant and maternal hospital discharge records to generate cost estimates for live singleton births in California by birth weight and gestation category (as well as probabilities of death by birth weight category). They were able to match more than 99 percent of the 518,704 records for the 2000 birth cohort. Hospital charges were converted to actual costs using hospital specific ratios from OSHPD Hospital Financial Reporting data. Infant costs include hospital expenses from birth until death or first discharge. Maternal costs include prenatal and delivery hospitalizations. These cost figures include only prenatal care resulting in a hospitalization and exclude return visits of the mother or infant. Amounts are in 2003 dollars. If California hospital costs exceed those of other states, these may be an overestimate of national costs. However, this is the only recent study we could find with information on the distribution of costs by birth weight and gestation, and births in California represent approximately one-eight of all births in the U.S.

[28] Appendix Table 5 presents the results of similar calculations for the costs of added triplet and higher order births, assuming the triplet and higher order birth rate increase from mandates is the same as that for twins.

[29] Luke, Bigger, Leurgans & Sietsema (1996) suggest that infant and maternal hospital charges for births in a large urban medical center for twins and singletons of similar gestational age, payer, and race did not differ significantly. Infant costs for these twins and matched singletons were from 13 to 33 times higher than those for normal singletons.

[30] The distribution of costs by gestational age is only available quite coarsely. Thus, we focus on cost estimates by birth weight.

the two largest components of ART expenditures, and thus represent a large share of the expenses targeted by the mandates. Another useful benchmark for the costs of these mandates is other health care expenses for infants. 2000 Healthcare Cost and Utilization Project data on child hospitalizations for the entire U.S. suggest that the aggregate national bill for the top diagnosis related group for all infants 1 year old or younger (extreme immaturity or respiratory distress syndrome, neonates) was $6.8 billion (including re-admissions after initial discharge.) The cost of imposing a mandate to cover infertility treatment nationally would be about 4.0 percent of the $6.8 billion.

The final exercise we perform with the data from Schmitt et al. (2006) uses the distribution of mortality rates by birth weight to calculate the increase in the probability that a twin will die suggested by the effect of mandates on birth weight. The cover mandate estimate suggests an increase in the probability of death for a twin born to an older mother of 0.0010 (on a baseline of 0.0257), or an increase of 4.0 percent (column 4 of row 9).

These estimates of the financial and other costs of imposing a national mandate do not take into account the added cost of additional triplets or higher-order births, which are larger on average than the costs of additional twins. Assuming that the effect of imposing a mandate to cover on the triplet and higher-order birth rate is about the same as the effects for twins, this suggests an additional cost of $82 million of additional triplets and higher-order births if they are born to women who otherwise would not have had children, and $73 million if they replace singletons. The rough cost estimates reported above also ignore longer-term health care and other costs for the additional multiple births and for the increased number of low birth weight or short gestation births.[31]

8 Conclusion

Reproductive technologies have improved radically since the introduction of the first fertility drugs in the late 1960s. These technologies make conception possible for many couples who otherwise would be unable to reproduce. Many of these technologies increase the probability of having a multiple birth, typically a more risky pregnancy. These technologies also provide women considering delaying reproduction with insurance against later infertility.

This study examines the effect of subsidies for use of infertility treatments on fertility outcomes (twin births

[31] For example, Petrou, Mehta, Hockley, Cook-Mozaffari, Henderson & Goldacre (2003) find that the cost of hospital re-admissions after birth during the first five years of life for premature infants range from 2.04–2.28 times the costs of the initial birth admissions. If these estimates (from UK data) are accurate for the U.S., this could suggest a much larger total long-term cost of imposing these mandates.

as a share of twin plus singleton births and the twin delivery rate), birth outcomes (for twins and singletons), and longer-term child health (for twins). The subsidies are state mandates that health insurers cover or offer to cover infertility treatment. Children born in mandate states to older mothers are 10 percent more likely to be twins in the birth certificate data, and 20 percent more likely to be mixed-sex twins. The mandates lead to an increase in the probability of being a twin in the PUMS/ACS data as well. The mandates also lead to a large and significant increase in the number of twin deliveries per 1000 women.

This study contributes to the literature on health insurance mandates. Our findings suggest that the infertility mandates are having real and substantive effects on fertility. Because twin births are riskier for infants and mothers, they also suggest the mandates have affected infant health. With existing data, we are unable to determine what share of the increase in twin deliveries is due to women waiting because of the availability of new technology and what share is women having children who otherwise could not.

Subsidized infertility treatment and increased use of this treatment by older women (as proxied by interactions between age indicators and the insurance mandates) are associated with small, statistically significant, negative effects on birth outcomes such as birth weight, gestation, and the 5-minute Apgar score for twins; and on birth weight and gestation for singletons. Findings for longer-term impacts of being born in a mandate state for twins are more mixed, with some evidence that mandates may be associated with more conditions limiting physical activities and more physical impairments among twins born to older mothers.

Imposing a national mandate to cover infertility treatment in every state would increase hospital costs both because it would increase the number of twins and because it would lower birth weight among twins. Estimates of these additional hospital costs of a national mandate to cover infertility treatment are $334 million for 2000. These additional costs are 71 percent of the total cost of all the in vitro fertilizations performed in 2000.

This paper also contributes to the ongoing debate about birth selection in health economics. In previous work that examines birth selection, heterogeneity or selection among mothers and the impact of maternal investment on birth outcomes are expected to affect birth outcomes in the same direction. Here, the taxing nature of infertility treatments suggests that women aided by ART or other infertility treatment may invest more in their pregnancies than do other women. Thus, the negative impacts of insurance mandates suggest that infertility treatment itself may be associated with adverse birth outcomes or that selection into childbearing of previously infertile women may lead to worse birth outcomes. Our findings for birth outcomes and more preliminary findings for child health suggest further investigation of longer-term impacts of mandates on child health is warranted.

References

Abma, J., Chandra, A., Mosher, W., Peterson, L. & Piccinino, L. (1997), 'Fertility, family planning, and women's health: New data from the 1995 National Survey of Family Growth', *Vital Health Statistics* **23**(19).

ACOG Committee on Practice Bulletins—Obstetrics (2002), 'ACOG practice bulletin no. 38: Perinatal care at the threshold of viability', *Obstetrics and Gynecology* **100**(3), 618–23.

Almond, D., Chay, K. Y. & Lee, D. S. (2004), The costs of low birth weight, Working Paper W10552, NBER.

Almond, D., Chay, K. Y. & Lee, D. S. (2005), 'The costs of low birth weight', *Quarterly Journal of Economics* **120**(3), 1031–83.

Ananat, E. O., Gruber, J., Levine, P. & Staiger, D. (2006), Abortion and selection, Working Paper W12150, NBER.

Baker, L. C. & Phibbs, C. S. (2002), 'Managed care, technology adoption, and health care: The adoption of neonatal intensive care', *RAND Journal of Economics* **33**(3), 524–48.

Bao, Y. & Sturm, R. (2004), 'The effects of state mental health parity legislation on perceived quality of insurance coverage, perceived access to care, and use of mental health speciality care', *Health Services Research* **39**(5), 1361–1377.

Behrman, J. R. & Rosenzweig, M. R. (2001), The returns to increasing body weight. Pier Working Paper 01-052.

Bergh, T., Ericson, A., Hillensjo, T., Nygren, K.-G. & Wennerholm, U.-B. (1999), 'Deliveries and children born after in-vitro fertilization in Sweden 1982–95: A retrospective cohort study', *Lancet* **354**(5), 1579–85.

Bertrand, M., Duflo, E. & Mullainathan, S. (2004), 'How much should we trust differences-in-differences estimates?', *Quarterly Journal of Economics* **119**(1), 249–275.

Bitler, M. P. (2005), Effects of increased access to infertility treatment on infant and child health outcomes: Evidence from health insurance mandates, Working Paper 2005.06, Public Policy Institute of California.

Bitler, M. P. & Schmidt, L. G. (2006), 'Health disparities and infertility: Impacts of state-level insurance mandates', *Fertility and Sterility* **85**(4), 858–65.

Bitler, M. P. & Schmidt, L. G. (2007), Who do health insurance mandates affect? The case of infertility treatment.

Black, S. E., Devereux, P. J. & Salvanes, K. G. (2007), 'From the cradle to the labor market? The effect of birth weight on adult outcomes', *Quarterly Journal of Economics* **122**(1), 409–439.

Buckles, K. (2006), Stopping the biological clock: Infertility treatments and the career-family tradeoff.

Bundorf, M. K., Henne, M. & Baker, L. (2007), Mandated health insurance benefits and the utilization and outcomes of infertility treatments. Mimeo.

Callahan, T. L., Hall, J. E., Ettner, S. L., Christiansen, C. L., Greene, M. F. & Crowley, W. F. (1994), 'The economic impact of multiple-gestation pregnancies and the contribution of assisted-reproduction techniques to their incidence', *New England Journal of Medicine* **331**(4), 244–49.

Chandra, A., Martinez, A., Mosher, W., Abma, J. & Jones, J. (2005), 'Fertility, family planning, and reproductive health of U.S. women: Data from the 2002 National Survey of Family Growth', *Vital Health Statistics* **23**(25).

Chandra, A. & Stephen, E. H. (1998), 'Impaired fecundity in the United States: 1982–1995', *Family Planning Perspectives* **30**(1), 34–42.

Cnattingius, S., Forman, M. R., Berendes, H. W. & Isotalo, L. (1992), 'Delayed childbearing and risk of adverse perinatal outcome: A population based study', *Journal of the American Medical Association* **268**(7), 886–90.

Collins, J. & Graves, G. (2000), 'The economic consequences of multiple gestation pregnancy in assisted conception cycles', *Human Fertility* **3**(4), 274–283(9).

Colvin, M., McGuire, W. & Fowlie, P. (2004), 'Neurodevelopmental outcomes after preterm birth', *BMJ* **329**, 1390–1393.

Currie, J. & Hyson, R. (1998), Is the impact of health shocks cushioned by socio-economic status? The case of low birthweight.

Cutler, D. M. & Meara, E. (2000), The technology of birth: Is it worth it?, *in* A. M. Garber, ed., 'Frontiers in Health Policy Research', Vol. 3, MIT Press, Cambridge, MA, pp. 33–67.

Dhont, M., Sutter, P. D., Martens, G. R. . G. & Bekaert, A. (1999), 'Perinatal outcome of pregnancies after assisted reproduction: A case-control study', *American Journal of Obstetrics and Gynecology* **181**(3), 688–695.

Division of Vital Statistics, NCHS (2000), 'Technical appendix from Vital Statistics of the United States 1998 Natality'.

Division of Vital Statistics, NCHS (Various years), 'Natality detail'.

Donohue, J. J. & Levitt, S. (2001), 'The impact of legalized abortion on crime', *Quarterly Journal of Economics* **116**(2), 379–420.

Donohue, J. J. & Levitt, S. (2004), 'Further evidence that legalized abortion lowered crime: A reply to Joyce', *Journal of Human Resources* **39**(1), 29–49.

Griffin, M. & Panak, W. F. (1998), 'The economic cost of infertility-related services: An examination of the Massachussetts infertility insurance mandate', *Fertility and Sterility* **70**(1), 22–29.

Grossman, M. & Joyce, T. J. (1990), 'Unobservables, pregnancy resolutions, and birth weight production functions in New York City', *Journal of Political Economy* **98**(5, Part 1), 983–1007.

Gruber, J. (1994), 'The incidence of mandated maternity benefits', *American Economic Review* **84**(3), 622–641.

Gruber, J., Levine, P. & Staiger, D. (1999), 'Abortion legalization and child living circumstances: Who is the "marginal" child?', *Quarterly Journal of Economics* **115**(1), 263–292.

Hamilton, B. H. & McManus, B. (2005), Infertility treatment markets: The effects of competition and policy.

Henne, M. & Bundorf, M. K. (2007), 'Insurance mandates and trends in infertility treatments', *Fertility & Sterility* .

Jain, T., Harlow, B. L. & Hornstein, M. D. (2002), 'Insurance coverage and outcomes of in vitro fertilization', *New England Journal of Medicine* **347**(9), 661–666.

Joyce, T. (2004), 'Did legalized abortion lower crime?', *Journal of Human Resources* **39**(1), 1–28.

Kezdi, G. (2002), Robust standard error estimation in fixed-effects panel models.

Klick, J. & Markowitz, S. (2006), 'Are mental health insurance mandates effective?: Evidence from suicides', *Health Economics* **15**(1), 83–97.

Levi, A. J., Raynault, M. F., Bergh, P. A., Drews, M. R., Miller, B. T. & Scott, R. T. (2001), 'Reproductive outcome in patients with diminished ovarian reserve', *Fertility and Sterility* **76**(4), 666–669.

Liu, Z., Dow, W. H. & Norton, E. C. (2004), 'Effect of drive-through delivery laws on postpartum length of stay and hospital charges', *Journal of Health Economics* **23**(1), 129–155.

Luke, B., Bigger, H. R., Leurgans, S. & Sietsema, D. (1996), 'The cost of prematurity: A case-control study of twins vs. singletons', *American Journal of Public Health* **86**(6), 809–14.

Luke, B. & Keith, L. G. (1992), 'The contribution of singletons, twins and triplets to low birth weight, infant mortality and handicap in the United States', *Journal of Reproductive Medicine* **37**(8), 661–666.

Martin, J. A., Hamilton, B. E., Ventura, S. J., Menacker, F. & Park, M. M. (2002), 'Births: Final data for 2000', *National Vital Statistics Reports* **50**(5).

Martin, J. A. & Park, M. M. (1999), 'Trends in twin and triplet births: 1980–1997', *National Vital Statistics Reports* **47**(24).

McCarton, C. M., Wallace, I. F., Divon, M. & Vaughan, H. G. (1996), 'Cognitive and neurologic development of the premature, small for gestational age infant through age 6: Comparison by birth weight and gestational age', *Pediatrics* **98**(6), 1167–78.

McCormick, M. C. (1985), 'The contribution of low birth weight to infant mortality and childhood morbidity', *New England Journal of Medicine* **312**(2), 82–90.

McCormick, M. C., Brooks-Gunn, J., Workman-Daniels, K., Turner, J. & Peckham, G. J. (1992), 'The health and developmental status of very low-birth-weight children at school age', *Journal of the American Medical Association* **267**(16), 2204–08.

National Conference of State Legislatures (2002), Insurance coverage for infertility therapy. Downloaded 10/11/2002.

Neumann, P. J., Gharib, S. D. & Weinstein, M. C. (1994), 'The cost of a sucessful delivery with in vitro fertilization', *New England Journal of Medicine* **331**(4), 239–43.

Office of Technology Assessment (1988), *Infertility: Medical and Social Choices*, U.S. Government Printing Office, Washington, DC.

Pacula, R. & Sturm, R. (2000), 'Mental health parity: Much ado about nothing?', *Health Services Research* **35**(1 Part 2), 263–275.

Petrou, S., Mehta, Z., Hockley, C., Cook-Mozaffari, P., Henderson, J. & Goldacre, M. (2003), 'The impact of preterm birth on hospital inpatient admissions and costs during the first 5 years of life', *Pediatrics* **112**(6 Part 1), 1290–1297.

Pop-Eleches, C. (2006), 'The impact of an abortion ban on socioeconomic outcomes of children: Evidence from Romania', *Journal of Political Economy* **114**(4), 744–773.

Rees, J. M., Lederman, S. A. & Kiely, J. L. (1996), 'Birth weight associated with lowest neonatal mortality: Infants of adolescent and adult mothers', *Pediatrics* **98**(6), 1161–15.

Rosenzweig, M. R. & Schultz, T. P. (1983), 'Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight', *Journal of Political Economy* **91**(5), 723–746.

Royer, H. (2003), What all women (and some men) want to know: Does maternal age affect infant health?

Schieve, L. A., Meikle, S. F., Ferre, C., Peterson, H. B., Jeng, G. & Wilcox, L. S. (2002), 'Low and very low birth weight in infants conceived with use of assisted reproductive technology', *New England Journal of Medicine* **346**(10), 731–737.

Schmidt, L. (2007), 'Effects of infertility insurance mandates on fertility', *Journal of Health Economics* **26**(3), 431–446.

Schmitt, S. S., Sneed, L. & Phibbs, C. S. (2006), 'The costs of newborn care in California: A population-based study', *Pediatrics* **117**(1), 154–160.

Sheiner, E., Shoham-Vardi, I., Hershkovitz, R., Katz, M. & Mazor, M. (2001), 'Infertility treatment is an independent risk factor for caesarean section among nulliparous women aged 40 and above', *Obstetrics and Gynecology* **185**(4), 888–891.

Sibai, B. M., Hauth, J., Caritis, S., Lindheimer, M. D., MacPherson, C., Klebanoff, M., VanDorsten, J. P., Landon, M., Miodovnik, M., Paul, R., Meis, P., Thursau, G., Dombrowski, M., Roberts, J. & McNellis, D. (2000), 'Hypertensive disorders in twin versus singleton gestations', *American Journal of Obstetrics and Gynecology* **192**(4), 938–942.

Society for Assisted Reproductive Technology, American Society for Reproductive Medicine (1999), 'Assisted reproductive technology in the United States: 1996 results generated from the American Society for Reproductive Medicine/Society for Assisted Reproductive Technology Registry', *Fertility and Sterility* **71**(5).

Society for Assisted Reproductive Technology, American Society for Reproductive Medicine (2002), 'Assisted reproductive technology in the United States: 1998 results generated from the American Society for Reproductive Medicine/Society for Assisted Reproductive Technology Registry', *Fertility and Sterility* **77**(1).

Summers, L. (1989), 'Some simple economics of mandated benefits', *American Economic Review Papers and Proceedings* **79**(2), 177–183.

Sutcliffe, A. G. (2004), 'The long-term paediatric outcomes of assisted reproductive technologies', *International Congress Series* **1266**, 359–366.

Tan, S.-L., Doyle, P., Campbell, S., Beral, V., Rizk, B., Brinsden, P., Mason, B. & Edwards, R. G. (1992), 'Obstetric outcome of in vitro fertilization pregnancies compared with normally conceived pregnancies', *American Journal of Obstetrics and Gynecology* **167**(3), 778–784.

Varma, T. R. & Patel, R. H. (1987), 'Outcome of pregnancy following investigation and treatment of infertility', *International Journal of Gynaecology and Obstetrics* **25**(2), 113–20.

Vohr, B., Wright, L., Dusick, A., Mele, L., Verter, J., Steichen, J., Simon, N., Wilson, D., Boryles, S., Bauer, C., Delaney-Black, V., Yolton, K., Fleisher, B., Papile, L.-A. & Kaplan, M. (2000), 'Neurodevelopmental and functional outcomes of extremely low birth weight infants in the National Institute for Child Health and Human Development Neonatal Research Network, 1993–1994', *Pediatrics* **105**(6), 1216–1226.

Wimalasundera, R. C., Trew, G. & Fisk, N. M. (2003), 'Reducing the incidence of twins and triplets', *Best Practice and Research Clinical Obstetrics and Gynaecology* **17**(2), 309–329.

Wolpin, K. I. (1997), Determinants and consequences of the mortality and health of infants and children, *in* M. Rosenzweig & O. Stark, eds, 'Handbook of Population and Family Economics', Vol. 1 A, Elsevier Science, New York, NY, pp. 483–557.

Wright, V. C., Schieve, L. A., Reynolds, M. A. & Jeng, G. (2005), 'Assisted reproductive technology surveillance—United States, 2002', *Morbidity and Mortality Weekly Report* **54**(SS02), 1–24.

Figure 1: Pregnancies, Deliveries, and Live Infants from ART, SART Registry

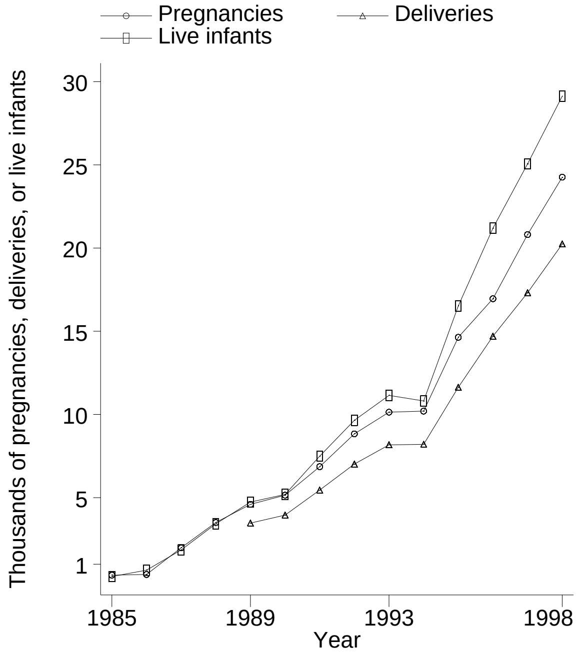

Figure 2: Share of Births That Are Multiples by State and Year, Detailed Natality Data

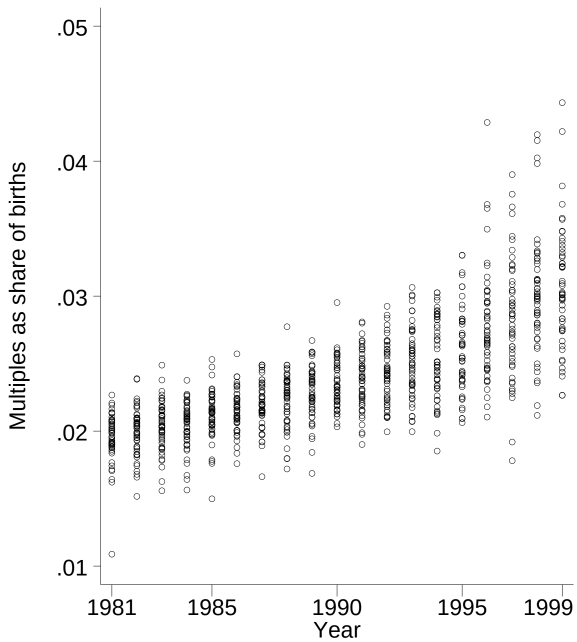

Figure 3: Average Age of Twin Mothers by State and Year, Detailed Natality Data

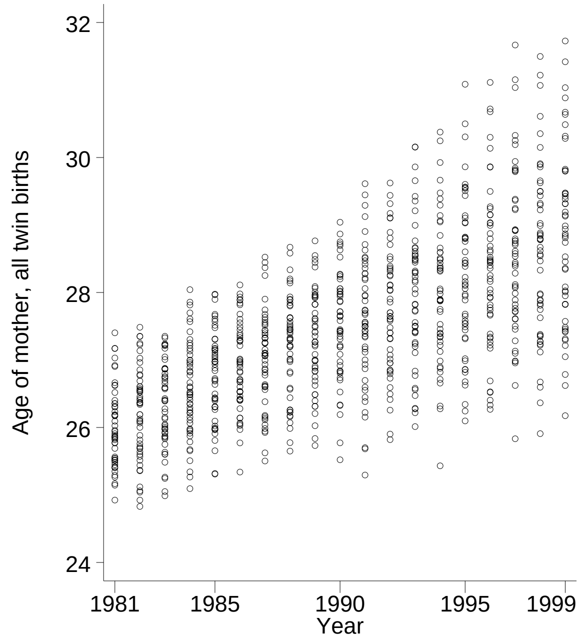

Figure 4: Share of Twins among Twins and Singletons by Years Since Infertility Mandate Was Adopted, by Age of Mother

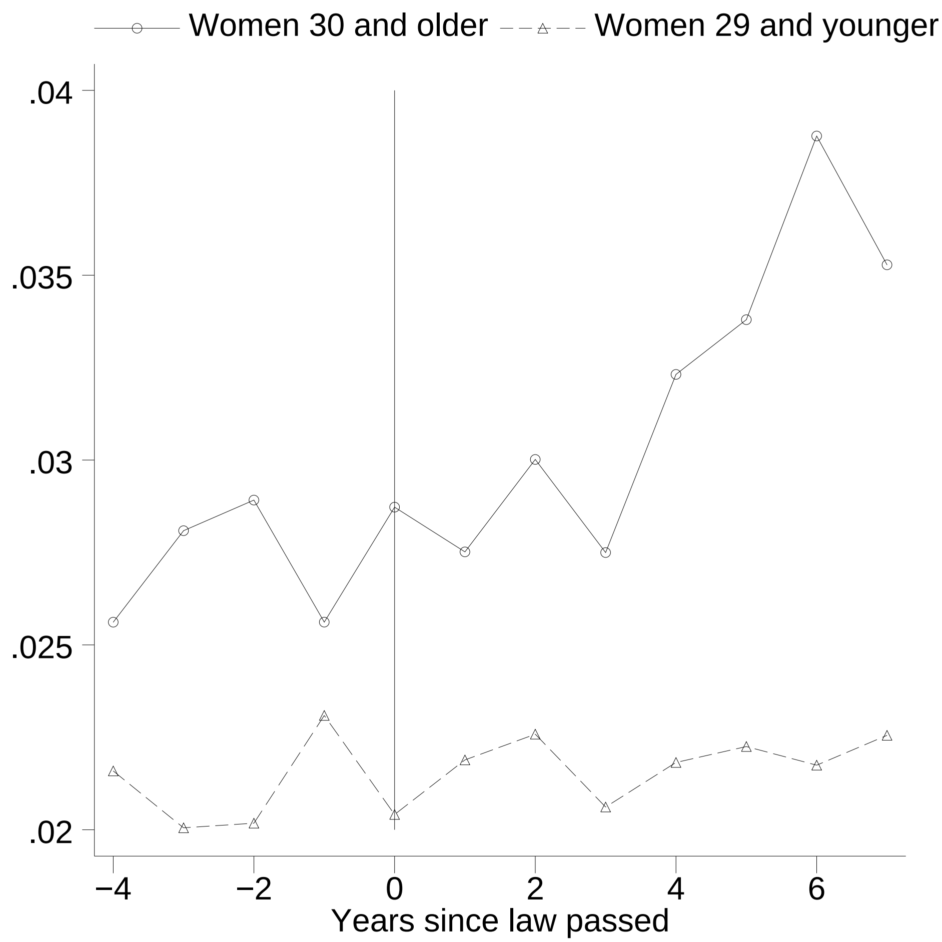

Table 1: Use of Various Infertility Treatments by Women, 2002 National Survey of Family Growth

	All women	Age group 15–29	Age group 30–44
Any infertility treatment, year before interview (to get pregnant or prevent miscarriage)	0.028	0.020	0.033
Ever had any infertility treatment (to get pregnant or prevent miscarriage)	0.134	0.063	0.182
Ever had treatment to prevent miscarriage	0.062	0.034	0.081
Ever had treatment to help get pregnant	0.093	0.036	0.133
Source paying for treatment to help get pregnant			
Private health insurance	0.071	0.027	0.100
Other source	0.023	0.008	0.032
Type of treatment to help get pregnant (not mutually exclusive)			
Advice	0.068	0.027	0.097
Woman tested for infertility	0.049	0.016	0.072
Man tested for infertility	0.041	0.013	0.060
Ovulation-inducing drugs	0.042	0.014	0.061
Artifical insemination	0.012	0.002	0.019
Surgery to correct blocked Fallopian tubes	0.008	0.001	0.013
In vitro fertilization	0.003	0	0.005
Some other treatment	0.007	0.002	0.010

Columns 1–3 show the share of women 15–44 (column 1), 15–29 (column 2), and 30–44 (column 3) reporting any infertility treatment during the year before the interview or ever having used various types of infertility treatment, from the 2002 National Survey of Family Growth (NSFG). Any infertility treatment is defined by the NSFG as having obtained treatment to help get pregnant or to prevent miscarriage or both. Question about whether private insurance paid for treatment or type of treatment are only asked of women reporting they obtained treatment to help get pregnant. Statistics are weighted. Sample is women of each age group who have ever had intercourse after menarche. Rounding of sources paying for treatment done independently, thus numbers do not necessarily sum to total ever obtaining treatment to get pregnant.

Table 2: States Enacting Laws Mandating That Health Insurers Cover or Offer to Cover Infertility Treatment, 1981–2001

State	Year law	*IVF is*		*Mandate to insurers to*		*Law applies to*		
	Passed	Not excluded	Excluded	Cover	Offer	All firms	Non-HMOs	Only HMOs
Arkansas	1987	1	0	1	0	0	1	0
California	1989	0	1	0	1	1	0	0
Connecticut	1989	1	0	0	1	0	1	0
Hawaii	1987	1	0	1	0	1	0	0
Illinois	1991	1	0	1	0	1	0	0
Louisiana	2001	0	1	1	0	1	0	0
Maryland	1985	1	0	1	0	1	0	0
Massachusetts	1987	1	0	1	0	1	0	0
Montana	1987	0	1	1	0	0	0	1
New York	1990	0	1	1	0	0	1	0
New Jersey	2001	1	0	1	0	1	0	0
Ohio†	1991	Before 1997	1997 on	1	0	0	0	1
Rhode Island	1989	1	0	1	0	1	0	0
Texas	1987	1	0	0	1	1	0	0
West Virginia	1977	0	1	1	0	0	0	1

Shown are the year each state passed a law related to health insurance coverage of infertility (column 1), whether the coverage mandate includes IVF or does not specify (column 2), whether the coverage mandate excludes IVF (column 3), whether the coverage is mandated (column 4), whether the insurer is required to offer it to the customer (column 5), whether all firms are covered (column 6), whether HMOs are excluded from the law (column 7), and whether the law applies only to HMOs (column 8).

† In Ohio, the insurance commissioner issued a statement that IVF would not be covered in 1997; thus the state is coded as a "no IVF" state from 1997 on.

Table 3: Share of Twin and Singleton Births That Are Twins, Twin Birth Outcomes, and Twin Deliveries per 1000 Women, by Presence of Infertility Treatment Mandate, Natality and Census Data

	No mandate for Infertility treatment	Mandate for Infertility treatment
A. Share of twin and singleton births that are twins		
Twin	0.0227	0.0251
Mixed-sex twin	0.0058	0.0070
Same-sex female twin	0.0067	0.0078
Same-sex male twin	0.0067	0.0077
B. Birth outcomes for twins		
Birth weight (grams)	2394.4	2394.9
Low BW ($<$ 2500 grams)	0.5187	0.5179
Very low BW ($<$ 1500 grams)	0.1027	0.0985
Extremely low BW ($<$ 1000 grams)	0.0490	0.0458
Gestation (weeks)	35.915	35.740
Premature ($<$ 37 weeks)	0.4813	0.5079
Very premature ($<$ 32 weeks)	0.1139	0.1098
Five minute Apgar score (0–10)	8.486	8.569
C. Twin deliveries per 1000 women		
Twin deliveries per 1000 women	0.5795	0.6748

Summary statistics from Detailed Natality files (panels A and B) or combined Natality data and Census population data (panel C) for 1981–99. Panel A contains the share of births that are twins or share of births that are mixed-sex or single-sex twins. Sample for first row is one in fifty random subsample of twins and singletons and sample for other rows is one in fifty random subsample of singletons and of twins where twins could be matched to one another. Panel B contains mean birth outcomes for sample of all twins. Panel C contains the twin delivery rate (number of pregnancies resulting in twin live births) per 1000 women in 5-year age by race by state by year cells over the period 1981–1999 from combined Natality data (numerator) and Census population data (denominator). For each panel, column 1 contains means during years when no infertility treatment mandate was in place during the year before birth and column 2 contains means during years when any infertility treatment mandate was in place during the year before birth. Statistics in panels A and B weighted to account for 50% sampling in some states before 1985. Counts for numerator for panel C multiplied by two if records are from states with 50% sampling before 1985. Statistics in panel C weighted by number of women in age/race/state/year cell.

Table 4: Share of Twin and Singleton Children That Are Twins and Twin Health Outcomes, by Presence of Infertility Treatment Mandate, PUMS and ACS

	No mandate for Infertility treatment	Mandate for Infertility treatment
A. Share of twin and singleton children that are twins		
Twin	0.0305	0.0317
B. Health outcomes for twins		
Blindness/deafness/vision or hearing impairment	0.0122	0.0100
Limited in basic physical activities	0.0132	0.0129
Physical, mental, or emot. cond. lasting $\geq$ 6 months causes		
Difficulty learning, remembering, or concentrating	0.0570	0.0379
Difficulty dressing, bathing, or getting around house	0.0106	0.0098

Summary statistics from PUMS/ACS. Panel A contains the share of children that are twins out of a sample of singletons and twins. Sample is all children aged 5–17 in combined 2000 5% and 1% PUMS whose mother is in the household, who are one of two children their age, both of whom are a child or grandchild of the householder and share the same mother, and whose age and sex and relationship to the head are not allocated. Twins are identified as children of the same age, whose mother lives in the household, who share the same mother, and for whom the householder is their parent or grandparent. Panel B contains mean health outcomes for the sample of twins. Column 1 contains means in years with no mandate for infertility treatment the year before birth and and column 2 contains means in years with a mandate for infertility treatment the year before birth. Statistics weighted with child weight.

Table 5: Determinants of Being a Twin, Coefficients on Mandates and Their Interactions with Mother's Age ≥ 30, Sample of Singletons and Twins, Natality Data

Controls for	Any mandate and Any mandate $*$ age ≥ 30	IVF/no IVF and IVF/no IVF $*$ age ≥ 30	Cover/offer and Cover/offer $*$ age ≥ 30
HI for infert.	-0.00062		
	(0.00090)		
HI for infert. $* \geq 30$	0.00308*		
	(0.00180)		
HI for infert. may incl. IVF		-0.00125	
		(0.00094)	
HI for infert. may incl. IVF $* \geq 30$		0.00354	
		(0.00229)	
HI for infert. excl. IVF		0.00047	
		(0.00130)	
HI for infert. excl. IVF $* \geq 30$		0.00256	
		(0.00244)	
HI must cover infert.			-0.00208**
			(0.00091)
HI must cover infert. $* \geq 30$			0.00677***
			(0.00114)
HI must offer infert. coverage			0.00078
			(0.00105)
HI must offer infert. $* \geq 30$			-0.00027
			(0.00076)

Table presents coefficients on insurance mandate variables in regressions of determinants of birth being a twin. Sample is one in fifty random subsample of all singleton and twin births for 1981–99. Each column represents one regression. The coefficient in column 1 is for an indicator for any mandate that insurers cover/offer to cover infertility treatment and its interaction with the mother being ≥ 30, those in column 2 for indicators for infertility treatment mandates for insurers that exclude IVF treatment or may include IVF and their interactions with the mother being ≥ 30, and those in column 3 for indicators for mandates for insurers to cover or offer to cover infertility treatment and their interactions with the mother being ≥ 30. Regressions include state of residence at birth fixed effects, year fixed effects, and month of birth fixed effects. Regressions also include indicators for the mother's education or for education missing or unreported; for the mother's age in 5-year intervals; for the mother being black, Asian/Pacific Islander, or other non-white or race missing; for the mother being Hispanic or ethnicity missing or unreported; for parity or parity missing; for the child's gender; and for state-level economic, demographic, and public assistance variables. Standard errors clustered at the state level. Regressions are weighted to account for the 50% sampling in some states before 1985. N is 1,439,525. Pre-reform mean of dependent variable is 0.0227 and that for infants with mothers 30 or older is 0.0295. $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .10$.

Table 6: Determinants of Twin Delivery Rate, Coefficients on Mandates and Their Interactions with Mother's Age ≥ 30, Twin Deliveries per 1000 Women, Combined Natality and Census Data)

Controls for	Any mandate and Any mandate $*$ age ≥ 30	IVF/no IVF and IVF/no IVF $*$ age ≥ 30	Cover/offer and Cover/offer $*$ age ≥ 30
HI for infert.	-0.05037*		
	(0.02541)		
HI for infert. $* \geq 30$	0.09040*		
	(0.04952)		
HI for infert. may incl. IVF		-0.03639	
		(0.03946)	
HI for infert. may incl. IVF $* \geq 30$		0.05715	
		(0.08064)	
HI for infert. excl. IVF		-0.06298*	
		(0.03362)	
HI for infert. excl. IVF $* \geq 30$		0.12812***	
		(0.04508)	
HI must cover infert.			-0.07564**
			(0.03455)
HI must cover infert. $* \geq 30$			0.14729**
			(0.06067)
HI must offer infert. coverage			-0.03976*
			(0.02312)
HI must offer infert. $* \geq 30$			0.03112
			(0.06596)

Table presents coefficients on insurance mandate variables in regressions of determinants of twin delivery rates (twin deliveries per 1000 women). Denominator from Census population counts per women in 5-year-age-group by race (white-black-Asian-other) cells. Numerator is count of deliveries per 5-year-age-group by race (white-black-Asian-other) cell (= twin births/2). Each column represents one regression. The coefficient in column 1 is for an indicator for any mandate that insurers cover/offer to cover infertility treatment and its interaction with the mother being ≥ 30, those in column 2 for indicators for infertility treatment mandates for insurers that exclude IVF treatment or may include IVF and their interactions with the mother being ≥ 30, and those in column 3 for indicators for mandates for insurers to cover or offer to cover infertility treatment and their interactions with the mother being ≥ 30. Regressions include state of residence at birth fixed effects and year fixed effects. Regressions also include indicators for the mother's age in 5-year intervals; for the mother being black, Asian, or other non-white; and for state-level economic, demographic, and public assistance variables. Standard errors clustered at the state level. Counts of births are weighted to account for the 50% sampling in some states before 1985. The number of cells is 27132. Pre-reform mean of dependent variable is 0.58 twin deliveries per 1000 women and that for women 30 and older is 0.37 twin deliveries per 1000 women. $^{***}p < .01,^{**} p < .05,^{*} p < .10$.

Table 7: Determinants of Being a Mixed-Sex Twin, Coefficients on Mandates and Their Interactions with Mother's Age ≥ 30, Sample of Singletons and Twins, Natality Data

Controls for	Any mandate and Any mandate $*$ age ≥ 30	IVF/no IVF and IVF/no IVF $*$ age ≥ 30	Cover/offer and Cover/offer $*$ age ≥ 30
HI for infert.	-0.00063		
	(0.00048)		
HI for infert. $* \geq 30$	0.00160**		
	(0.00079)		
HI for infert. may incl. IVF		-0.00087	
		(0.00052)	
HI for infert. may incl. IVF $* \geq 30$		0.00135*	
		(0.00070)	
HI for infert. excl. IVF		-0.00009	
		(0.00076)	
HI for infert. excl. IVF $* \geq 30$		0.00184	
		(0.00137)	
HI must cover infert.			-0.00099*
			(0.00050)
HI must cover infert. $* \geq 30$			0.00305***
			(0.00077)
HI must offer infert. coverage			-0.00063
			(0.00082)
HI must offer infert. $* \geq 30$			0.00028
			(0.00034)

Table presents coefficients on insurance mandate variables in regressions of determinants of birth being a mixed-sex twin. Sample is one in fifty random subsample of all singleton and matched twin births for 1981–99. Each column represents one regression. The coefficient in column 1 is for an indicator for any mandate that insurers cover/offer to cover infertility treatment and its interaction with the mother being ≥ 30, those in column 2 for indicators for infertility treatment mandates for insurers that exclude IVF treatment or may include IVF and their interactions with the mother being ≥ 30, and those in column 3 for indicators for mandates for insurers to cover or offer to cover infertility treatment and their interactions with the mother being ≥ 30. Regressions include state of residence at birth fixed effects, year fixed effects, and month of birth fixed effects. Regressions also include indicators for the mother's education or for education missing or unreported; for the mother's age in 5-year intervals; for the mother being black, Asian/Pacific Islander, or other non-white or race missing; for the mother being Hispanic or ethnicity missing or unreported; for parity or parity missing; for the child's gender; and for state-level economic, demographic, and public assistance variables. Standard errors clustered at the state level. Regressions are weighted to account for the 50% sampling in some states before 1985. N is 1,435,000. Pre-reform mean of dependent variable is 0.0058 and that for infants with mothers 30 or older is 0.0082. $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .10$.

Table 8: Determinants of Birth Weight (Grams), Coefficients on Mandates and Their Interactions with Mother's Age ≥ 30, Sample of Twins, Natality Data

Controls for	Any mandate and Any mandate $*$ age ≥ 30	IVF/no IVF and IVF/no IVF $*$ age ≥ 30	Cover/offer and Cover/offer $*$ age ≥ 30
HI mandate for infert.	-3.750		
	(6.610)		
HI for infert. $* \geq 30$	-15.765**		
	(6.305)		
HI for infert., may incl. IVF		-1.877	
		(6.826)	
HI for infert., may incl. IVF $* \geq 30$		-8.571	
		(5.606)	
HI for infert., excl. IVF		-9.694	
		(8.761)	
HI for infert., excl. IVF $*$ mother ≥ 30		-23.160***	
		(5.349)	
HI must cover infert.			0.966
			(6.289)
HI must cover infert. $* \geq 30$			-12.142*
			(6.774)
HI must offer infert. coverage			-20.661
			(14.654)
HI must offer infert. $* \geq 30$			-20.220***
			(7.050)

Table presents coefficients on insurance mandate variables. Sample is twin births in states reporting outcome for 1981–99. Each column represents one regression. The coefficient in column 1 is for an indicator for any mandate that insurers cover/offer to cover infertility treatment and its interaction with the mother being ≥ 30, those in column 2 for indicators for infertility treatment mandates for insurers that exclude IVF treatment or may include IVF and their interactions with the mother being ≥ 30, and those in column 3 for indicators for mandates for insurers to cover or offer to cover treatment and their interactions with the mother being ≥ 30. All regressions include state of residence fixed effects, year fixed effects, and month of birth fixed effects. Regressions also include indicators for the mother's education or for education missing or unreported; for the mother's age in 5-year intervals; for the mother being black, Asian/Pacific Islander, or other non-white or race missing; for the mother being Hispanic or ethnicity missing or unreported; for parity or parity missing; for the child's gender; and for state-level economic, demographic, and public assistance variables. Standard errors clustered at the state level. Regressions are weighted to account for 50% sampling in some states before 1985. N is 1,676,147. Pre-reform mean of dependent variable is 2394 and that for infants with mothers 30 or older is 2475. $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .10$.

Table 9: Determinants of Gestation (Weeks), Coefficients on Mandates and Their Interactions with Mother's Age ≥ 30, Sample of Twins, Natality Data

Controls for	Any mandate and Any mandate $*$ age ≥ 30	IVF/no IVF and IVF/no IVF $*$ age ≥ 30	Cover/offer and Cover/offer $*$ age ≥ 30
HI mandate for infert.	-0.033		
	(0.038)		
HI for infert. $* \geq 30$	-0.041		
	(0.030)		
HI for infert., may incl. IVF		-0.038	
		(0.043)	
HI for infert., may incl. IVF $* \geq 30$		0.004	
		(0.033)	
HI for infert., excl. IVF		-0.037	
		(0.049)	
HI for infert., excl. IVF $*$ mother ≥ 30		-0.089***	
		(0.022)	
HI must cover infert.			0.008
			(0.037)
HI must cover infert. $* \geq 30$			-0.000
			(0.038)
HI must offer infert. coverage			-0.198**
			(0.082)
HI must offer infert. $* \geq 30$			-0.091***
			(0.028)

Table presents coefficients on insurance mandate variables. Sample is twin births in states reporting outcome for 1981–99. Each column represents one regression. The coefficient in column 1 is for an indicator for any mandate that insurers cover/offer to cover infertility treatment and its interaction with the mother being ≥ 30, those in column 2 for indicators for infertility treatment mandates for insurers that exclude IVF treatment or may include IVF and their interactions with the mother being ≥ 30, and those in column 3 for indicators for mandates for insurers to cover or offer to cover treatment and their interactions with the mother being ≥ 30. Regressions include state of residence at birth fixed effects, year fixed effects, and month of birth fixed effects. Regressions also include indicators for the mother's education or for education missing or unreported; for the mother's age in 5-year intervals; for the mother being black, Asian/Pacific Islander, or other non-white or race missing; for the mother being Hispanic or ethnicity missing or unreported; for parity or parity missing; for the child's gender; and for state-level economic, demographic, and public assistance variables. Standard errors clustered at the state level. Regressions are weighted to account for 50% sampling in some states before 1985. N is 1,636,418. Pre-reform mean of dependent variable is 35.9 and that for infants with mothers 30 or older is 36.1. $^{***}p < .01,^{**}p < .05,^{*}p < .10$.

Table 10: Determinants of 5-Minute Apgar Score, Coefficients on Mandates and Their Interactions with Mother's Age ≥ 30, Sample of Twins, Natality Data

Controls for	Any mandate and Any mandate $*$ age ≥ 30	IVF/no IVF and IVF/no IVF $*$ age ≥ 30	Cover/offer and Cover/offer $*$ age ≥ 30
HI mandate for infert.	-0.020		
	(0.033)		
HI for infert. $* \geq 30$	-0.027**		
	(0.012)		
HI for infert., may incl. IVF		0.009	
		(0.027)	
HI for infert., may incl. IVF $* \geq 30$		-0.033**	
		(0.016)	
HI for infert., excl. IVF		-0.079**	
		(0.035)	
HI for infert., excl. IVF $*$ mother ≥ 30		-0.015*	
		(0.008)	
HI must cover infert.			-0.023
			(0.034)
HI must cover infert. $* \geq 30$			-0.025**
			(0.012)
HI must offer infert. coverage			0.039
			(0.024)
HI must offer infert. $* \geq 30$			-0.061***
			(0.008)

Table presents coefficients on insurance mandate variables. Sample is twin births in states reporting outcome for 1981–99. Each column represents one regression. The coefficient in column 1 is for an indicator for any mandate that insurers cover/offer to cover infertility treatment and its interaction with the mother being ≥ 30, those in column 2 for indicators for infertility treatment mandates for insurers that exclude IVF treatment or may include IVF and their interactions with the mother being ≥ 30, and those in column 3 for indicators for mandates for insurers to cover or offer to cover treatment and their interactions with the mother being ≥ 30. All regressions include state of residence fixed effects, year fixed effects, and month of birth fixed effects. Regressions also include indicators for the mother's education or for education missing or unreported; for the mother's age in 5-year intervals; for the mother being black, Asian/Pacific Islander, or other non-white or race missing; for the mother being Hispanic or ethnicity missing or unreported; for parity or parity missing; for the child's gender; and for state-level economic, demographic, and public assistance variables. Standard errors clustered at the state level. Regressions are weighted to account for 50% sampling in some states before 1985. N is 1,313,105. Pre-reform mean of dependent variable is 8.49 and that for infants with mothers 30 or older is 8.60. $^{***}p < .01, ^{**}p < .05, ^{*}p < .10$.

Table 11: Determinants of Birth Weight (Grams), Coefficients on Mandates and Their Interactions with Mother's Age ≥ 30, Sample of Singletons, Natality Data

Controls for	Any mandate and Any mandate $*$ age ≥ 30	IVF/no IVF and IVF/no IVF $*$ age ≥ 30	Cover/offer and Cover/offer $*$ age ≥ 30
HI mandate for infert.	-2.745		
	(4.137)		
HI for infert. $* \geq 30$	-7.550		
	(5.796)		
HI for infert., may incl. IVF		-5.469	
		(4.018)	
HI for infert., may incl. IVF $* \geq 30$		-1.936	
		(4.790)	
HI for infert., excl. IVF		1.059	
		(5.897)	
HI for infert., excl. IVF $*$ mother ≥ 30		-13.383^{***}	
		(4.687)	
HI must cover infert.			1.181
			(3.949)
HI must cover infert. $* \geq 30$			-6.522
			(5.358)
HI must offer infert. coverage			-13.613^{*}
			(7.926)
HI must offer infert. $* \geq 30$			-8.909
			(7.001)

Table presents coefficients on insurance mandate variables. Sample is one in fifty random sample of singleton births in states reporting outcome for 1981–99. Each column represents one regression. The coefficient in column 1 is for an indicator for any mandate that insurers cover/offer to cover infertility treatment and its interaction with the mother being ≥ 30, those in column 2 for indicators for infertility treatment mandates for insurers that exclude IVF treatment or may include IVF and their interactions with the mother being ≥ 30, and those in column 3 for indicators for mandates for insurers to cover or offer to cover treatment and their interactions with the mother being ≥ 30. All regressions include state of residence fixed effects, year fixed effects, and month of birth fixed effects. Regressions also include indicators for the mother's education or for education missing or unreported; for the mother's age in 5-year intervals; for the mother being black, Asian/Pacific Islander, or other non-white or race missing; for the mother being Hispanic or ethnicity missing or unreported; for parity or parity missing; for the child's gender; and for state-level economic, demographic, and public assistance variables. Standard errors clustered at the state level. Regressions are weighted to account for 50% sampling in some states before 1985. N is 1,404,233. Pre-reform mean of dependent variable is 3361 and that for infants with mothers 30 or older is 3422. $^{***}p < .01, ^{**}p < .05, ^{*}p < .10$.

Table 12: Determinants of Impairments or Long-Lasting Conditions, Coefficients on Mandates and Their Interactions with Mother's Age ≥ 30, Sample of Twins, PUMS and ACS

			Condition limiting	
	Sensory Impairment	Physical Impairment	Mental Activities	Physical Activities
HI must cover infert.	-0.00925	-0.00165	-0.00770	-0.00666**
	(0.00741)	(0.00687)	(0.00841)	(0.00332)
HI must cover infert. $* \geq 30$	0.00752	0.01184	-0.00034	0.00920**
	(0.00607)	(0.00910)	(0.00941)	(0.00459)
HI must offer to cover infert.	-0.00898	-0.01363***	-0.01426	-0.00574*
	(0.00689)	(0.00508)	(0.01285)	(0.00308)
HI must offer infert. $* \geq 30$	0.00494	0.00421	0.00133	0.00781
	(0.00668)	(0.00277)	(0.00635)	(0.00499)

Table presents coefficients on indicators for mandates for insurers to cover or offer to cover infertility treatment and their interactions with the mother being ≥ 30. Dependent variable in column 1 is having long-lasting blindness, deafness, or a long-lasting severe vision or hearing impairment. Dependent variable in column 2 is having a long-lasting condition that severely limits one or more basic physical activities such as walking, climbing stairs, reaching, lifting, or carrying. Dependent variable in column 3 is having a physical, mental, or emotional condition lasting at least 6 months that causes difficulty learning, remembering, or concentrating. Dependent variable in column 4 is having a physical, mental, or emotional condition lasting at least 6 months that causes difficulty dressing, bathing, or getting around inside the house. Sample is twin children aged 5–17 in combined 2000 5% and 1% PUMS and 2001–2002 ACS whose mother is in the household, who are one of two children their age, both of whom are a child or grandchild of the householder and share the same mother, and whose age and sex and relationship to the head are not allocated. Twins are identified as children of the same age, whose mother lives in the household, who share the same mother, and for whom the householder is their parent or grandparent. Each column represents one regression. Regressions include state of residence at birth fixed effects (for the child), and year fixed effects. Regressions also include indicators for the mother's education; for the mother's age in 5-year intervals; for the child being black, Asian, American Indian, or other non-white; for the child being Hispanic; for the child's gender; and for state-level economic, demographic, and public assistance variables. Standard errors clustered at the state level. Regressions are weighted. N is 82,875. Pre-reform mean of column 1 dependent variable is 0.0121 and that for children of mothers 30 or older is 0.0091. Pre-reform mean of column 2 dependent variable is 0.0135 and that for children of mothers 30 or older is 0.0109. Pre-reform mean of column 3 dependent variable is 0.0594 and that for children of mothers 30 or older is 0.0620. Pre-reform mean of column 4 dependent variable is 0.0100 and that for children of mothers 30 or older is 0.0090. $^{***}p < .01$, $^{**}p < .05$, $^{*}p < .10$.

Table 13: Estimates of Costs of Imposing Various Types of Mandates Regarding Infertility Treatment Nationally Due to Additional Twins Born to Older Women and Lower Birth Weight among Twins Born to Older Women

Mandate type	Any mandate	IVF	No IVF	Cover	Offer
(1) Effect of mandates on share of births which are twins to women ≥ 30 (percent, as share of baseline of 0.0295)	0.104			0.229	
(2) # of additional twins born to women ≥ 30 (= total twins to women $\geq 30 * (1)$)	5,716			12,578	
(3) Average total cost of twin birth ($)	25,472				
(4) Average total cost of singleton birth ($)	7,306				
(5) National costs of extra twins, if not born without mandate ($ millions, $= (2) * (3)$)	146			320	
(6) National costs of displaced singletons ($ millions, $= (2) * (4)$)	42			92	
(7) National costs of extra twins, if they displace singletons ($ millions, $= (5) - (6)$)	104			228	
(8) Effect on twin birth weight for twins born to women ≥ 30 (grams)	-15.765		-23.160	-12.142	-20.220
(9) Effect of (8) on probability of death for twins born to women ≥ 30 (percent, as share of baseline of 0.0247)	0.055		0.071	0.040	0.062
(10) National costs from decrease in twin birth weight ($ millions, (8) applied to twin birth weight distribution)	25		44	14	29
(11) National costs of additional twins & lower birth weight if not born without mandate ($ millions, $= (5) + (10)$)	170			334	

Shown are the components of the cost estimates for the added cost due to mandates of extra twins born to older women (rows (5) and (7)), the added cost due to mandates of reduced birth weight among twins born to older women (row (10)), and their sum (row (11)). All costs are those due to imposing national mandates of various types (estimates from Tables 5 and 8). The table also shows the impacts of reduced birth weight on the probability of infant mortality in row (9). Each column represents estimates for the type of mandate in the column label. Calculations use the distribution of costs by birth weight category and the actual birth weight distribution for twins and the one in fifty subset of singletons from the Detailed Natality data for 1999. Two estimates are presented for the cost of the extra twins born after mandates, one assuming that the new twins are born to women who otherwise would not have had live births (row (5)), and one assuming that the new twins are born to women who otherwise would have had singletons (row (7)). Estimates are all for imposing a national mandate of the relevant type relative to a no-mandate baseline. Estimates only shown where effects of mandates are significant at the 10 percent level. Estimates are in constant 2003 dollars.