NBER WORKING PAPER SERIES

MANDATED HEALTH INSURANCE BENEFITS AND THE UTILIZATION AND OUTCOMES OF INFERTILITY TREATMENTS

M. Kate Bundorf
Melinda Henne
Laurence Baker

Working Paper 12820
http://www.nber.org/papers/w12820

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2007

We thank the Iris M. Litt Fund from the Institute for Research on Women and Gender at Stanford University for financial support. We thank Marianne Bitler as well as participants at various conferences and seminars for helpful comments. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Mandated Health Insurance Benefits and the Utilization and Outcomes of Infertility Treatments
M. Kate Bundorf, Melinda Henne, and Laurence Baker
NBER Working Paper No. 12820
January 2007, Revised April 2008
JEL No. I1

ABSTRACT

During the last two decades, the treatment of infertility has improved dramatically. These treatments, however, are expensive and rarely covered by insurance, leading many states to adopt regulations mandating that health insurers cover them. In this paper, we explore the effects of benefit mandates on the utilization and outcomes of infertility treatments. We find that use of infertility treatments is significantly greater in states adopting comprehensive versions of these mandates. While greater utilization had little impact on the number of deliveries, mandated coverage was associated with a relatively large increase in the probability of a multiple birth. For relatively low fertility patients who responded to the expanded insurance coverage, treatment was often unsuccessful and did not result in a live birth. For relatively high fertility patients, in contrast, treatment often led to a multiple, rather than a singleton, birth. We also find evidence that the beneficial effects on the intensive treatment margin that have been proposed in other studies are relatively small. We conclude that, while benefit mandates potentially solve a problem of adverse selection in this market, these benefits must be weighed against the costs of the significant moral hazard in utilization they induce.

M. Kate Bundorf
Health Research and Policy
Stanford University
HRP T108
Stanford, CA 94305-5405
and NBER
bundorf@stanford.edu

Laurence Baker
Department of Health Research & Policy
HRP Redwood Bldg, Rm T110
Stanford University
Stanford, CA 94305-5405
and NBER
laurence.baker@stanford.edu

Melinda Henne
Walter Reed Army Medical Center
mhenne@stanford.edu

I. Introduction

Mandating that employers provide certain benefits to their workers can be an attractive way for governments to achieve policy goals. In the case of health insurance, mandates exist primarily as regulations requiring insurers to cover specified services in the policies they sell. Because most people with health insurance obtain it through an employer, mandating that private insurers cover particular benefits compels employers who offer health insurance to purchase coverage for those services.[1] Over time, state governments and, on occasion, the Federal government have adopted mandates covering a wide range of services (Laugesen, Paul et al. 2006).

The most compelling economic argument for health insurance benefit mandates is that they mitigate problems of adverse selection for particular types of services (McGuire and Montgomery 1982; Summers 1989). When asymmetric information exists between insurers and consumers, a situation frequently encountered in health care, competitive insurance markets may fail to provide optimal levels of coverage (Rothschild and Stiglitz 1976). Mandated provision can remedy this. Against the potential welfare gains of benefit mandates, however, weigh the possible negative welfare effects. In the case of health insurance, mandates may exacerbate moral hazard, leading to inefficiently high consumption of the mandated service (Gruber 1992).

Existing empirical literature provides mixed evidence on the efficiency implications of health insurance benefit mandates. For maternity benefits, state and federal mandates appear to have addressed a market failure due to adverse selection without generating significant moral hazard (Gruber 1994). The mandates dramatically

---

[1] The well known exception is that the Employee Retirement Income Security Act of 1974 (ERISA) exempts employers who self-insure from this type of regulation.

expanded insurance coverage, with employers passing on the costs of the mandated benefit to the workers most likely to benefit from expanded coverage. The wage offset, in turn, had no impact on the labor supply of affected workers, indicating that their value of the incremental coverage exceeded the cost.

Other studies provide a less positive picture of the efficiency implications of benefit mandates, consistent with the view that mandate passage may represent the preferences of particular interest groups (e.g. Stigler 1971) rather than the attempts of regulators to increase social welfare. Some studies have found that benefit mandates negatively affect both insurance and labor markets (Jensen and Gabel 1992; Sloan and Conover 1998; Jensen and Morrisey 1999), suggesting that they generate moral hazard that offsets any welfare gains. A third group of studies, however, reports that mandates of even high cost benefits have *not* had negative effects on either health insurance or labor markets. These studies suggest that employers often provide mandated benefits even in the absence of a mandate (Gruber 1992; Kaestner and Simon 2002). In this case, mandates neither alleviate adverse selection nor create moral hazard.

In this paper, we investigate the impact of mandates requiring health insurers to provide coverage for infertility treatments. Over the last three decades, new technologies have transformed the treatment of infertility. Among the most significant advances has been the development of assisted reproductive therapies (ART) in which a physician surgically removes an egg from a woman, combines the egg with sperm in a laboratory, and returns the developing embryo to a woman's body.[2] In vitro fertilization (IVF), in

---

[2] The development of drugs that stimulate ovulation in women was also an important advance in the treatment of infertility. Patients often use these drugs in ART cycles as well as on their own or with intrauterine (or artificial) insemination.

which the embryo is returned to the woman's uterus, is the most common type of ART.[3] The use of infertility treatments, particularly ART, increased rapidly during the last two decades. The first delivery using ART in the U.S. occurred in 1981. By 2004, 411 clinics performed 127,977 cycles leading to 36,760 births of 49,458 infants (CDC 2006).

An important concern in the treatment of infertility is its association with high rates of multiple births. In IVF, patients and physicians decide how many embryos to return to the uterus. While transferring more embryos increases the likelihood of a pregnancy, it also increases the likelihood of a multiple birth (Thurin, Hausken et al. 2004). More than 90% of patients transfer more than one embryo in given cycle, and approximately one-third of ART deliveries are multiple births compared to about 3% of all births (CDC 2005; Martin, Hamilton et al. 2005). The use of ovulation-inducing drugs, even in the absence of ART, generates similar concerns. Approximately one-third of the dramatic increase in multiple birth rates during the last three decades was due to increased use of reproductive technologies (CDC 2000).

Infertility treatment offers an intriguing case of the potential the trade-off between solving a problem of adverse selection and creating a problem of moral hazard as a consequence of a benefit mandate. Adverse selection is likely to be a problem in this market. Treatment, particularly ART, is expensive. A single cycle of IVF costs approximately $10,000 (Neumann and Johannesson 1994; Collins 2001). Because many cycles do not result in a live birth, the average cost per delivery is over $50,000 (Collins 2001). The uncertain incidence of infertility and the high cost of treatment make it a

---

[3] Other types of ART include gamete intrafallopian transfer (GIFT) and zygote intrafallopian transfer (ZIFT), in which either unfertilized eggs and sperm (gametes) or fertilized eggs (zygotes) are placed in the fallopian tubes. While GIFT and ZIFT were used in the early 1990s, they represented less than 1% of ART cycles in 2003 (CDC 2005).

significant financial risk, and experimental evidence suggests that people highly value insurance against this risk (Neumann and Johannesson 1994). However, people are likely to have private information about both their fertility and their desire for children that is highly predictive of their utilization. Because health insurance enrollment decisions are generally made annually, individuals are able to use this information when purchasing insurance, potentially leading to adverse selection.

The fact that infertility is rarely covered by insurance is consistent with, although not direct evidence of, adverse selection in this market. Although data on insurance coverage of infertility treatments are limited, a 1997 Mercer Consulting employer survey found that 65% of employers did not cover any infertility services and, of those that covered some, less than half provided coverage for the most expensive treatments like IVF. If the absence of insurance coverage for infertility treatments signals the existence of adverse selection in this market, a benefit mandate may be a useful regulatory solution. In practice, a number of states enacted mandates requiring insurers to cover infertility treatments during the late 1980s and early 1990s, and Louisiana, New Jersey, and New York passed legislation as recently as 2001.

In our analysis, we focus on the moral hazard effects of mandated coverage for infertility treatments. Our key contribution is that we demonstrate that, not only did expanding insurance coverage increase utilization of ART, but that the coverage expansions had heterogeneous effects on the population of women of child-bearing age. We demonstrate that, in theory, both relatively low and relatively high fertility patients may respond to a reduction in the price of infertility treatments. The effects on outcomes, however, are likely to differ between these groups. For relatively low fertility couples, greater utilization is likely to lead to unsuccessful cycles. For relatively high fertility couples, in contrast, greater utilization of fertility treatments may lead to higher rates of multiple births without a corresponding increase in birth rates. Using mother's age as a proxy for fertility, we provide empirical evidence consistent with these types of effects.

Our focus on the heterogeneous treatment effects both differentiates our study from other research on the effects of these mandates and provides important insights into the interpretation of the findings of existing studies. Hamilton and McManus (2005) examine the impact of benefit mandates as well as competition among fertility clinics on utilization of ART and multiple birth risk. While they conclude that mandates reduce multiple birth risk by changing the financial incentives influencing embryo transfer decisions, our results provide evidence of an alternative mechanism for their findings – that the reduction in multiple birth rates associated with expanded coverage reflects changes in the composition of the treated population. While Bitler and Schmidt (2007) document higher rates of utilization of reproductive therapies in states with insurance mandates, particularly among older, highly-educated women, they do not link increased utilization to outcomes, which is necessary for evaluating the welfare implications of greater utilization.

Bitler (2006), who examines differences in health outcomes among twins in states with and without mandates, documents that mandates lead to higher rates of twinning and that twins born in mandated states have slightly worse health outcomes, suggesting a negative effect of either ART treatment or patient selection into ART on infant health. Our results provide evidence on how patient selection is likely to affect the interpretation of results based on this identification strategy. In particular, we propose that the marginal patient in an analysis of the effects of expanded insurance coverage on rates of twinning is likely to represent a relatively fertile patient (among ART births), suggesting that any negative effects of patient selection into ART on health outcomes could be larger among the broader population.

Schmidt (2007) documents higher rates of first births among women 35 and older in states adopting infertility mandates. As an analysis of the moral hazard associated with this type of benefit mandate, however, this study has significant limitations. We demonstrate that, because infertility treatments are often used for couples experiencing secondary infertility, restricting the analysis to first births misses important effects of utilization on outcomes, particularly among older women. More importantly, because the study does not examine rates of multiple births, it does not capture a key dimension of the heterogeneity of the treatment effect – whether insurance coverage leads some women to have a multiple rather than a singleton birth. Using a separate data source, we also analyze outcomes among those treated to determine whether mandates result in incremental utilization that does not ultimately result in a birth and, thus, would not be captured by the population birth data.

Finally, while not the focus of our study, our results provide some indirect evidence on whether benefit mandates cause women to delay child birth by reducing the price of infertility treatment (Buckles 2006).

While each of these studies examines an important aspect of extending insurance coverage for the treatment of infertility, none provides evidence on how the mandates affect the composition of the treated population. Yet, we demonstrate that understanding these heterogeneous treatment effects is essential for evaluating the welfare effects of expanded coverage. In our discussion, we reconcile our findings with those of the existing literature and identify their implications for research and policy.

II. Theoretical Effects of Insurance Coverage for Infertility Treatment

Effect of Insurance Coverage on Utilization along the Extensive Margin

By reducing the price of treatment, insurance coverage may increase demand for infertility services, resulting in greater utilization among those for whom the expected benefits of treatment are lower. Patients responding to the lower price may include both relatively low and relatively high fertility patients. Figure 1 plots a hypothetical relationship between a couple's fertility and birth probabilities with and without treatment. Treatment has the largest effect on the probability of a birth for couples in the middle of the fertility distribution. Lower fertility patients are unlikely to have a birth without treatment and treatment generates a relatively small increase in that probability. Although the treatment effect is also relatively small for high fertility patients, they are likely to have a birth even in the absence of treatment.

In Figure 1, L and H represent the fertility thresholds for patients seeking treatment in the absence of insurance coverage. The expected benefits of treatment exceed the expected costs only for couples with fertility greater than L and less than H. A reduction in the out-of-pocket price of treatment shifts these thresholds from L to L' (a shift toward relatively low fertility patients) and from H to H' (a shift toward relatively high fertility patients).[4] In both cases, the lower price results in utilization by patients with lower expected benefits, and utilization increases birth rates within each group.

---

[4] This result is formalized in a theoretical model developed by Hamilton and McManus (2005).

While the effect of insurance coverage on the probability of a birth is similar for the marginal low and high fertility patient, the effect on the probability of a multiple birth differs dramatically. In Figure 1, we plot the relationship between the probability of a multiple birth and fertility, with and without treatment, assuming that treatment has a constant, positive effect on the probability of a multiple birth, conditional on a birth. Under this assumption, holding embryo transfer practices constant, the probability of a multiple birth increases with fertility due primarily to its relationship with the probability of any birth. Figure 1 demonstrates, however, that a portion of both high and low fertility patients undergoing infertility treatment in response to more generous insurance coverage would have given birth even in the absence of treatment. Because multiple birth rates are high among those treated with ART, for this subset of women, treatment had no effect on whether they gave birth, but a large effect on their likelihood of a multiple birth.

A key distinction between low and high fertility patients in this effect is the number of patients for whom it applies. Because many high fertility patients would have given birth in the absence of treatment, insurance coverage may generate a relatively large increase in the number of multiple births relative to the increase in the number of births. Among low fertility couples, in contrast, relatively few would have given birth in the absence of treatment. As a result, the magnitude of the effect of treatment on the number of multiple births is likely to be smaller. Thus, a potentially important form of moral hazard in response to insurance coverage for the treatment of infertility is an increase in multiple birth rates without a corresponding increase in birth rates among relatively high fertility patients.

Table 1 provides a hypothetical numerical example to demonstrate this effect. High and low fertility patients differ in their probability of a birth without treatment (0.8 versus 0.2). We assume that treatment has the same effect on the probability of a birth (a 10 percentage point increase) for both high and low fertility patients, and that, for both groups, infertility treatment increases the probability of a multiple birth conditional on a birth from 0.03 to 0.30. The implication is that treatment has a larger (in absolute value) incremental effect on the probability of a multiple birth for high (0.246) than low (0.084) fertility patients. In addition, the ratio of the incremental probability of a multiple birth to the incremental probability of a birth is much higher for high than low fertility patients. For low fertility patients, each incremental birth is associated with 0.8 incremental multiple births. For high fertility patients, in contrast, each incremental birth is associated with 2.46 incremental multiple births.

The assumptions underlying this stylized model are generally consistent with fertility rates and infertility treatment. Rates of multiple births among those conceiving with ART are high - approximately 30% of ART deliveries are multiple deliveries, compared to approximately 3% of births in the population as a whole. In our empirical work, we use female age as a proxy for fertility. While female fertility declines significantly with age (Menken, Trussell et al. 1986), the proportion of births that are multiples varies relatively little with age, either with or without treatment. Prior to the widespread use of infertility treatments, differences by age in rates of twin and triplet births were relatively small. In 1980, rates of multiple births per 1,000 births were 20.5, 23.5, 25.3, and 23.0 for women aged 25-29, 30-34, 35-39, and 40-44, respectively (Martin and Park 1999). In the case of ART, both birth rates and multiple birth rates conditional on a birth decline with age, particularly after age 35 (CDC 2005). These facts suggest that age is likely to be a good proxy for fertility in this context.

Effects Insurance Coverage on Utilization along the Intensive Margin

Patients seeking treatment for infertility face a variety of treatment alternatives, and generous insurance coverage may cause them to progress more rapidly from inexpensive to more expensive treatment options. For example, artificial insemination (AI) is thought to be more cost-effective than IVF, which has a higher success rate but also a much higher cost, so that it would typically be optimal for patients to progress to IVF only after AI has failed (Van Voorhis, Sotovall et al. 1997). Because patients with generous insurance coverage are protected from the out-of-pocket costs of IVF, they may progress to IVF more quickly.

Other research, however, suggests insurance coverage of ART may lead to *more* efficient utilization of ART by influencing treatment patterns in ways that ultimately lead to lower rates of multiple births (Jain, Harlow et al. 2002). Because multiple births are associated with low birth weight and other comorbidities for babies as well as greater risk of complications for mothers, many consider them a negative outcome of infertility treatment (Adashi, Barri et al. 2003).[5] Infants born as part of a multiple rather than a singleton birth are more likely to have low or very low birth weight (Schieve, S.F. et al. 2002), and low birth weight is highly correlated with infant mortality as well as a variety

---

[5] Although the evidence of complications associated with the use of ART, conditional on the number of births, is less conclusive, some studies point to the existence of negative effects on infant health. Population-based studies have found that infants conceived using reproductive technologies have worse health outcomes relative to those conceived naturally (Schieve et. al. 2002; Bitler 2006), and reviews of the medical literature conclude that the possibility that infants conceived using ART have an increased risk of birth defects relative to those conceived naturally cannot be ruled out (Hansen et. al 2005; Kurinzuk et. al. 2006).

of longer-term poor health outcomes (Hack, Klein et al. 1995; CDC 2002). Medical care costs are also higher for multiple than for singleton births. A study conducted during the late 1980s and early 1990s found that the average *per baby* cost for a singleton delivery was $9,845 compared to $18,974 and $36,588 for twin and triplet deliveries, respectively (Callahan, J.E. et al. 1994).[6]

Due to the high per cycle cost of IVF, patients paying out-of-pocket face strong financial pressure to minimize the number of cycles. Thus, they are willing to increase their chances of success at the increased risk of a multiple birth by transferring more embryos in a given cycle. Patients with comprehensive health insurance covering multiple rounds may be more conservative in their decisions regarding embryo transfer since their out of pocket costs will be much lower for an additional cycle if the first is unsuccessful.[7]

A potential mechanism for these positive efficiency effects is that insurance coverage of ART corrects a distortion in utilization created by differential coverage of complementary services. Patients usually do not bear the high incremental medical care costs associated with the multiple births because they are generally covered by insurance. Thus, when they make embryo transfer decisions, they have no incentive to consider the incremental medical care costs associated with multiple births. Insurance coverage of ART essentially corrects this distortion by subsidizing the cost of additional cycles of

---

[6] Almond, Chay and Lee (2005), however, propose that these studies, which are largely cross-sectional analyses, overstate the magnitude of both the negative health consequences and medical care costs of low birth weight.

[7] In fact, some clinical evidence is consistent with this view, finding that fewer embryos are transferred per cycle and that the percentages of cycles resulting in live births and multiple births are lower in states with mandates for comprehensive health insurance coverage of ART than in those without (Reynolds, 2001; Jain, 2002). However, it is not clear that these studies have adjusted well for the underlying characteristics of treated patients.

ART. Alternatively, Hamilton and McManus (2005) propose that this effect can be attributed to intertemporal income effects that result in patients pursuing ART earlier due to the availability of greater wealth in subsequent periods, conditional on treatment in the earlier period. Insurance essentially generates moral hazard by creating incentives for couples to pursue treatment earlier.

Effects of Insurance Coverage on the Timing of Childbearing

Insurance coverage of the treatment of infertility may also cause women to delay childbirth (Buckles 2006). Women pursuing a career face a trade-off with respect to the timing of births. While delaying child birth may lead to greater lifetime earnings, because female fecundity declines with age, it increases the risk of infertility. By enabling women to bear children at older ages, effective infertility treatment reduces this risk. Insurance coverage, in turn, reduces the medical care costs of expected future use of infertility treatment, creating stronger incentives for women to delay.

III. Data Sources

We analyze separately data from two sources in order to examine the effects of the mandates. Using population birth data, we examine birth rates, independent of treatment. By not restricting our analyses to the treated population, we are able to test whether the incremental utilization associated with mandates indeed leads to incremental births or multiple births, essentially accounting for births that would likely have taken place in the absence of more generous insurance coverage. Then, using clinic registry data, we examine utilization and outcomes among those treated with ART. These data allow us to assess the relationship between state mandate status and utilization of ART as well as outcomes among those treated.

The source of population-level data on birth rates for the year 1981 to 1999 is the National Vital Statistics System of the National Center for Health Statistics (NCHS 1981-1999). These data are abstracted from birth certificates filed each year in vital statistics offices of each state. While the data include 100% of birth certificates for most states for the years of our study, for a small number of states during the earlier years, the data include only a 50% sample of birth certificates, and we adjust the total numbers accordingly. From these data, we calculated the total number of births, singleton births, multiple births (any order), and triplet or higher order births in each state in each year by 5-year age categories (25-29, 30-34, 35-39, 40-44, and 45-49). We restrict our analyses to women age 25-49 because women in this age group represent the vast majority of ART cycles (CDC 2005). We transform the birth data to represent deliveries, rather than births, because a delivery is the appropriate unit of analysis for most of our models. Because the use of infertility treatments is associated with high rates of multiple births, analyses that examine the number of infants born, as opposed to the number of women giving birth, may overstate the effects of the mandates on fertility rates. We divide the number of births which were either twin or higher order by the relevant number of infants in order to estimate the number of deliveries for a given age group-state-year cell.[8]

We examine utilization and outcomes among those treated using registry data from the Society for Assisted Reproductive Technologies (SART). In 1989, SART, an

---

[8] Prior to 1989, quadruplets and quintuplets were characterized as triplet or higher order births. For these years, we estimated the distribution of triplet, quadruplet, and quintuplet births based on the average of the national age-specific distribution for 1989 and 1990. Because the numbers of quadruplet and quintuplet births are very small, we believe this estimate is unlikely to significantly bias our results.

affiliate of the American Society for Reproductive Medicine (ASRM), began a voluntary reporting system to collect clinic-level information about the utilization and outcome of ART services. The Fertility Clinic Success Rate and Certification Act of 1992 federally mandated participation in the system for clinics providing ART. The results are compiled annually by the Centers for Disease Control (CDC), and the first national report under this law was based on data collected for utilization in 1995. We obtained hard copies of the 1989-1994 published reports from SART. The clinic-level data from subsequent years are publicly available on the CDC website. We use SART data from 1991 to 2001. We exclude the earlier years in order to restrict our analyses to the post-adoption period since the vast majority of states adopted mandates prior to 1989.

The reports include the number of various types of ART cycles performed (in vitro fertilization (IVF), gamete intrafallopian transfer (GIFT), zygote intrafallopian transfer (ZIFT), transfer of cryopreserved embryos, and combinations of the above procedures) and the number of births and multiple births for each procedure. Clinics do not report utilization of non-ART therapies such as artificial insemination or the use of ovulation inducing drugs without ART. We identified the state of each clinic based on its ZIP code and calculated the numbers of cycles, live deliveries, and multiple live deliveries for each state in each year. From this information, we calculated rates of cycles per 1,000 women age 25-44, live deliveries per cycle, and multiple deliveries per live delivery.[9] While the reports include information on cycles and outcomes by age, the age groupings vary over time. Thus, prior to 1995, we are only able to analyze utilization

---

[9] The SART reports include the number of multiple deliveries (twins and higher) for all cycles from 1990-1994 and only for cycles involving fresh embryos from non-donor eggs for later years. As a result, we analyze only multiple deliveries resulting from fresh cycles for all years. In addition, in 1991 and 1992, the number of reported cycles is based on the number of retrievals as opposed to the number of cycles initiated. This will cause us to under-estimate the number of cycles in these years.

and outcomes aggregated across age groups. From 1995 to 2001, we analyze separately women under 35 and 35 and over, although we limit our analyses to cycles using fresh eggs due to a change in the reporting of cycles using donor or frozen eggs in 1999.

Insurance Mandates

We obtained information on which states had adopted insurance mandates from RESOLVE (www.resolve.org), a consumer organization for people experiencing infertility. We then verified and expanded on this information by reviewing the actual legislation for each mandate. Twelve states passed legislation mandating insurance coverage of ART between 1985 and 1999 (Table 2).[10] The conditions of the mandates, however, vary significantly across states based on the types of plans affected, the number and types of treatments covered, the cost-sharing associated with treatment, and the population to which the mandate applies. Of course, firms that self-insure are exempt from mandates in all states.

We classified the laws into three categories based on their likely effectiveness in changing treatment patterns. We first differentiate between mandates that require insurance companies only to offer coverage for treatment and those that require insurance companies to cover infertility treatment. It is not clear why mandating that insurers offer a policy but not mandating that consumers purchase the coverage would affect rates of coverage of the mandated benefit. Insurers could have offered this type of policy at the chosen premium in the absences of a mandate. While this treatment is consistent with most of the existing literature on benefit mandates that does not even consider offer-only

---

[10] Louisiana and New Jersey each adopted a mandate and New York updated its mandate in 2001.

mandates (e.g. Gruber 1992), we explicitly test the effects of this type of mandate by analyzing them separately.

Among mandates that require insurers to cover infertility treatment, we distinguish between those that require generous coverage of IVF from those that do not. We defined a "comprehensive" mandate (three states) as a requirement that insurance companies, including health maintenance organizations (HMOs), provide coverage for the cost of diagnosis and treatment of infertility, including ART of at least four cycles, with few exclusions on the population covered by the mandate. A "limited" mandate (seven states) refers to mandating that insurers cover fewer than 4 cycles of IVF. For the most part, these mandates require insurers to cover either a single cycle or exclude IVF from the mandated coverage. An exception is the state of Maryland. Although Maryland's mandate required coverage of up to 3 cycles when it was adopted in 1985, coverage was limited to those with a 5-year history of infertility.[11] This restriction on the covered population, combined with fewer covered cycles, potentially makes the mandate less generous than those we define as comprehensive. However, it is unclear how strictly the restriction based on pre-existing infertility was enforced. In addition, the mandated three cycles of IVF is more generous that the coverage required in other states adopting limited mandates. Due to this ambiguity, we test the sensitivity of our results to reclassifying Maryland's mandate as comprehensive. The states adopting mandates, the year of adoption, and their classification are listed in Table 2. We also note the specific restrictions on coverage that differentiate comprehensive from limited mandates.

---

[11] In 1994, the mandate was extended to couples with a 2-year history of infertility as well as those with particular diagnoses with no waiting period.

A natural concern for our analysis is whether mandate passage was exogeneous with respect to demand for infertility treatment. Existing papers in this literature provide strong evidence that this is not the case. Hamilton and McManus (2005) find little difference between states adopting and those not adopting mandates with IVF-specific coverage regulations in demographic characteristics potentially reflecting demand for treatment, including female labor force participation rates, female educational attainment, average family size, and median household income. Instead, they find substantive differences between the types of states in their regulatory environments as proxied by the extent to which they tend to mandate other types of services as well as the voting behavior of their citizens. From this, they conclude that mandate adoption appears to reflect governing tastes rather than preferences for children or other life-style factors. Finally, they demonstrate that neither clinic size nor the number of clinics in a market varied by future state mandate status prior to the implementation of most mandates. Bitler (2006) demonstrates that trends in neither rates of twin births nor twin outcomes during the period prior to the adoption of mandates (1981-1984) differed by state mandate status. These findings indicate that mandate adoption is unlikely to reflect state-specific trends in demand for infertility treatment.

Control variables

We obtained data on the number of women in each state by age from 1981-2001, which are the basis of our per capita calculations, from the U.S. Census Bureau. Data on time-varying characteristics of states, which we use as control variables in the empirical models, are from a variety of sources. The estimate of per capita income is from the regional economic accounts produced by the U.S. Bureau of Economic Analysis. We obtained annual state-level unemployment rates from the Bureau of Labor Statistics. From the Current Population Survey, we calculated state-year level female labor force participation rates, rates of education among women of child-bearing age, the distribution of the population based on family income relative to poverty, and rates of minorities and Hispanics. Finally, using the County Business Patterns data produced by the U.S. Census Bureau, we calculate the proportion of workers in each state employed in small firms (<100 workers) in each year.

IV. Insurance Mandates and Population Delivery Rates

In Figures 2-4, we present trends in rates of deliveries, multiple deliveries (including twin and higher order births), and triplet or higher order deliveries per women 25-49 by whether the state ever adopted each type of mandate. The vertical lines indicate the earliest and latest years in which mandates of any type were adopted.

Although the baseline levels differ, there are not obvious differences in the trends for the total number of deliveries by state mandate status (Figure 2).[12] There are, however, more noticeable differences in trends in the rates of multiple births (Figure 3). In the post-mandate period, rates of multiple deliveries for women 25-49 increased more rapidly in states that adopted comprehensive mandates than in states that did not adopt mandates or that adopted offer-only or limited mandates. These trends are even more pronounced for triplet and higher order deliveries (Figure 4). The case of triplet and higher order births is notable because both the levels and trends in triplet or higher order

---

[12] The peak in birth rates in 1991 is well documented in the literature and is driven primarily by birth rates for women 25 to 29.

deliveries were similar by state mandate status prior to the adoption of the mandates, but clearly diverged in the post-mandate period.

We identify the effects of mandates by examining changes in delivery rates in states adopting laws before and after their implementation compared to changes in states that did not adopt mandates. The following model provides the basic framework for our difference-in-difference estimates:

$$y_{i,t} = \alpha + M_{i.t-2}\lambda + X_{i,t}\beta + S_i\gamma + Y_t\delta + \varepsilon_{i,t}\,. \tag{1}$$

The data are aggregated to the state (i)-year (t) level. Y represents a measure of the delivery rate among women of a particular 5-year age group, ranging from 25-49. We estimate all models separately by age group because we expect the effects of the mandates to differ by age group. This specification also allows the effects of state characteristics and time trends to vary across age groups and avoids grouping populations with very different baseline fertility rates. M is a series of indicators representing whether the state had adopted a comprehensive, limited, or offer-only mandate in place two years prior to the indicated year. We lag the mandate indicators by two years for two reasons. First, the lag between successful treatment and a birth is approximately nine months, indicating that the effect should be lagged by at least one year. In addition, in preliminary analyses, we found that the mandates appear to affect birth rates with an additional year lag, potentially reflecting either lags in technology diffusion or the use of multiple cycles prior to achieving a live birth. We include state (S) and year (Y) fixed effects as well as a set of time-varying state characteristics (X) described above. Table 3 includes definitions and summary statistics for the independent variables.

We estimate the model for each age group using two types of dependent variables. We first examine deliveries per capita to test the relationship between mandates and fertility rates. We also examine rates of multiple deliveries (twin or higher order) and triplet or higher order deliveries per 1,000 deliveries to test whether mandates are associated with a shift in the distribution of deliveries toward multiples. We estimate linear models using least squares, weighting by the size of the female population in each age group in a given state and year. We allow for clustering by state when calculating the standard errors in order to obtain estimates that are unbiased by within-state serial correlation in the error terms (Bertrand, Duflo et al. 2004).

The differences-in-differences estimates indicate that comprehensive mandates were associated with a decrease in delivery rates for women 25-29 and an increase in delivery rates for women 35-39 (Table 4 – Panel A for each age group). The decrease of 3.192 deliveries per 1,000 women represents a 3% decline for women 25-29 and the increase of 1.208 deliveries per 1,000 women represents a 4% increase for women 35-39. While the negative effect for younger women is consistent with a delay in child birth in response to more generous insurance coverage of infertility, we also find a similar negative effect in states adopting offer-only mandates, which are, in theory, less likely to effect insurance coverage and treatment patterns. For women 30-34 and 40-44, the estimates of the effects of comprehensive mandates are positive, but not statistically significant at conventional levels. For each group, the estimate represents a 1% increase in delivery rates, and the estimate borders on statistical significance for women 30-34 (p=0.14) We find no evidence that limited mandates affected delivery rates for any age group, and offer-only mandates are associated with delivery rates only for women 25-29.

For twin or higher order deliveries, the implementation of a comprehensive mandate led to a statistically significant increase in the proportion of deliveries that were multiples among women 30-34 and 35-39 (Table 4 – Panel B), and, for each of these age groups, the point estimate of the number of incremental multiple deliveries represents a 10% increase in multiple delivery rates. For women 40-44 and 45-49, the estimates are relatively large, representing a 16% and 17% increase, respectively, but imprecise. We find no evidence that either limited or offer-only mandates increased rates of twin or higher-order deliveries among women of any age.

The results for triplet or higher order multiple deliveries are similar to those for twin or higher order deliveries, although we find that, in addition to women 30-34 and 35-39, comprehensive mandates increased rates of triplet or higher order deliveries among the youngest age group we study – women 25-29. In general, the point estimates from these models represent a large percentage increase due to the relatively low rate of triplet or higher order deliveries in the population as a whole. For example, a comprehensive mandate led to 0.304 incremental triplet or higher order deliveries per 1,000 deliveries among women 30-34, representing a 54% increase. Once again, these results provide no evidence that either limited or offer-only mandates were associated with higher rates of multiple deliveries.

Our results for the effects of mandates on fertility rates differ somewhat from those of Schmidt (2007) who estimates that mandates increased first birth rates by 5% among women 35 and older but did not affect first births rates among younger women. One important methodological difference is that we do not restrict the measure of fertility to first births.[13] In 2002, 43% of women who had one or more visits for medical help to get pregnant had already given birth at least once.[14] Because infertility treatments are often used for couples experiencing secondary infertility, restricting the analysis to first births may miss important effects of utilization on outcomes, particularly among older women. In Table 5, we present the results of re-estimating our delivery rate models, restricting the dependent variable to first deliveries. While we continue to find little evidence that either limited or offer-only mandates had an effect on delivery rates, the results for comprehensive mandates differ somewhat from those presented in Table 4. In particular, we find the strongest evidence of an increase in first birth rates among women 30-34. For these women, the estimate of an incremental 1.303 deliveries per 1,000 women represents a 7% increase in the first birth rate. While the estimates for older women are large in percentage terms (8% and 9% for women 35-39 and 40-44, respectively), they are imprecisely estimated. Comparing the results from Tables 4 and 5, we conclude that the effects of comprehensive mandates on delivery rates were not restricted to first deliveries and that they were larger for first deliveries for women 30-34 and for subsequent deliveries for women 35-39.

We also test the sensitivity of the results to re-classifying Maryland as having adopted a comprehensive, rather than a limited, mandate in 1985 (results are presented in Appendix Table 1). Because few states adopted comprehensive versions of the laws and Maryland's mandate was more comprehensive than most regulations that we classified as

---

[13] Other differences include specifying the types of mandates into mutually exclusive groups, measuring fertility based on the number of women giving birth, rather than the number of infants born, a potentially important adjustment given the high rates of multiple births associated with infertility treatments, not log transforming the dependent variables, and not including state-specific quadratic time trends.

[14] Author's calculations using data from the National Survey of Family Growth published by the CDC (2005).

limited (although less comprehensive than those we classified as comprehensive), this potentially represents an important test of the robustness of our results. The results are not particularly sensitive to this reclassification. Although the magnitudes of most of the estimates of the effects of comprehensive mandates are slightly smaller (with the exception of the effect of the comprehensive mandate on delivery rates for women 30-34), they remain statistically significant. In addition, we continue to find no evidence that limited mandates affected either fertility or multiple birth rates.

In summary, we find that the effects of mandates on both fertility and multiple delivery rates were concentrated among states adopting the most comprehensive versions of the laws. For states adopting comprehensive laws, mandates had the largest percentage effect on delivery rates for women 35-39 – a 4% increase for these women. For slightly younger (30-34) and older (40-44) women, the estimates indicate that mandates increased delivery rates by approximately 1%, although these estimates are not statistically significant at conventional levels. For multiple deliveries, in contrast, we find strong evidence that mandates increased the proportion of deliveries that were multiples for women 30-34 and 35-39. In the case of triplet or higher order deliveries, these effects extended to women 25-29.

V. Insurance Mandates and the Utilization and Outcomes of ART

In this section, we examine the relationship between state mandate status and utilization and outcomes of ART. These analyses supplement those from the birth data in three ways. First, a relationship between state mandate status and utilization of ART provides important evidence of the mechanism underlying the effects we observe in the population birth data. Second, we directly test the relationship between state mandate status and changes in treatment patterns along the intensive margin by examining embryo transfer decisions. Finally, we examine the relationship between state mandate status and the outcomes of ART patients.

ART utilization increased dramatically from 1991-2001. The number of cycles per 1,000 women of reproductive age (25-44) more than doubled from 0.79 to 1.83.[15] During the same time period, births per cycle increased from 0.17 to 0.25 and multiples per birth increased from 0.31 to 0.38. Cycles per capita, however, increased much more quickly in states that adopted comprehensive mandates than in states either without mandates or with other types of mandates (Figure 5). Consistent with findings from the medical literature, deliveries per cycle (Figure 6) and multiples per ART delivery were lower in states with comprehensive mandates than in other states (Figure 7).

We examine these trends using multivariate models. Because the registry data do not span the period prior to the implementation of mandates, we are unable to control for characteristics of states that are fixed over time using state fixed effects. Thus, these results are necessarily more tentative than those using the birth data. The analyses are based on the following basic model:

$$y_{i,t} = \alpha + M_i \lambda + Y_t \delta + X_{i,t} \beta + \varepsilon_{i,t} \qquad (3)$$

Once again, we aggregate the data to the state(i)-year(t) level, and the dependent variables include cycles per 1,000 women, live deliveries per cycle, and multiple deliveries per live delivery. M is a set of dummy variables indicating whether the state had a comprehensive, limited, or offer-only mandate in place. In some models, we

---

[15] We restrict the denominator to women 25-44 in this analysis because women 45-49 represent relatively few ART cycles. While we were able to analyze this group separately using the birth data, we are unable to do so using the SART data.

include dummy variables indicating the year (Y) to control for trends in treatments and outcomes that were common across states over time. Because utilization demonstrated a strong linear time trend, we also estimate models in which we instead include a linear time trend and the interaction of the trend with state mandate status. This allows us to differentiate between a one-time shift in the level of outcome variables and a change in the rate of growth associated with the implementation of the mandate. X includes time-varying state-level characteristics that potentially affect treatments or outcomes. Table 6 includes summary statistics for the independent variables in these models.

We estimate one set of models, pooling data for patients of all ages covering the period 1991-2001 as well as separate models by age group (under 35 and 35 and over) for the subset of the years for which we have consistent age breakdowns for the dependent variables (1995-2001). We estimate the models using least squares, weighting by the corresponding denominator in the dependent variable. We estimate the standard errors allowing for within-state correlation in the error terms.

Comprehensive mandates were associated with an increase in the rate of growth, rather than a one-time shift, in utilization of ART relative to states not adopting mandates. Although the point estimate of the effect of a comprehensive mandate is large in the model with only the main effects of the mandates (Table 7 – Column 1), it is not statistically significant. However, when we interact a linear time trend with the mandate indicators, we find that the cycles per capita rose more quickly in states with comprehensive mandates than in states without mandates (Table 7 – Column 2). We find no evidence that either limited or offer-only mandates were associated with greater utilization of ART. When we re-estimate the model and restrict the sample to 1995-2001, we obtain a similar result, although the effect of the comprehensive mandate is only weakly statistically significant (p=0.107), likely due to the smaller sample size (Table 7 – Column 3). When estimating the model separately for women under 35 years and 35 and over, we find little evidence of differences in the effect of the comprehensive mandate on utilization for the two groups (Table 7 – Columns 5 and 6). Although the point estimate is not statistically significant at conventional levels for the older group, the magnitude of the effect is similar to that for the younger group. The estimates indicate that utilization of ART in states with comprehensive mandates is nearly double that of states without mandates.

Consistent with research from the medical literature, we find that rates of deliveries per cycle and multiples per delivery were lower in states that adopted comprehensive mandates than in states without mandates. The magnitudes of the effects are similar for women under 35 (Table 8 - Columns 1 and 2) and 35 and over (Table 8 - Columns 6 and 7), and the effects are large – representing an approximately 25% reduction in births per cycle and an eight percent reduction in multiples per birth. Although fewer embryos are transferred per cycles in states with comprehensive mandates for patients in both age groups, these effects are not statistically significant (Table 8 - Columns 3 and 8). More importantly, we also find no evidence that differences in embryo transfer rates explain the lower rates of deliveries and multiple deliveries in state with comprehensive mandates. When we control for the average number of embryos transferred (Table 8 – Columns 4, 5, 9 and 10), the point estimates of the effects of a comprehensive mandate on rates of deliveries per cycle and multiples deliveries per delivery are nearly identical to those from the models without the control.

While existing literature has attributed the relationship between mandates and lower rates of births per cycle and multiples per birth to the ability of insurance to reduce incentives to transfer more embryos in a given cycle (Jain, Harlow et al. 2002), our results imply that the relationship may have more to do with the composition of the treated population, which has typically been poorly controlled in previous work.

While we find no evidence of an association between ART outcomes and state mandate status for states adopting limited mandates, we do find similar effects for states adopting offer-only mandates to those for states adopting comprehensive mandates. However, in the analysis of utilization, we found no evidence that utilization was greater in states with offer-only mandates, suggesting that these effects represent differences in the populations seeking treatment between the two types of states.

VI. Discussion

We find that mandates requiring insurers to provide comprehensive coverage of infertility treatments affect the utilization and outcomes of these services. These effects, however, are concentrated within states adopting the most comprehensive versions of the laws – those that place fewer exclusions on the population to which the mandate applies and that require coverage of 3 or more cycles of IVF as well as less expensive treatments. We do not find evidence that more limited mandates affect either ART treatment patterns or population-level delivery or multiple delivery rates for women of any age.

The effects of comprehensive mandates on population delivery rates are largest in percentage terms for women 35-39. For this age group, comprehensive mandates resulted in an additional 1.208 deliveries per 1,000 women, a 4% increase. For women

30-34, our estimates indicate that comprehensive mandates increased delivery rates by 1% (p=0.14) and that this increase was concentrated among first births. We do not find evidence that comprehensive mandates increased delivery rates among women of other ages, although for women 40-44, both the magnitude and standard error of the estimate are large.

The estimates from the models of population delivery rates are consistent with those from the SART data. In particular, by multiplying the point estimates of the effects of comprehensive mandates on delivery rates for women 30-34, 35-39, and 40-44 (Table 4) by the number of women in each age group in comprehensive mandate states, we calculate that comprehensive mandates resulted in 1,867 incremental deliveries annually. From the SART data, we multiply the estimate of incremental annual utilization in states with comprehensive mandates (1.536 cycles per capita) by the number of women 25-44 in these states to estimate the number of incremental cycles – 4,824. Dividing the two, we obtain a ratio of incremental deliveries to incremental cycles of 0.39, close to the 1996 rate of births per ART cycle of 0.23. This calculation, however, assumes that all incremental births associated with the comprehensive mandate are due to ART. If they are driven in part by increased utilization of other reproductive therapies such as ovulation-inducing drugs or artificial insemination, which seems likely, the estimated rate of incremental deliveries per ART cycle would be lower.

While our results do not differ entirely from those of Schmidt (2007), who studies the effects of the mandates on rates of first birth among women of different ages using national natality data, they do provide a different picture of the effects of mandates on fertility rates. First, while Schmidt finds that mandates increased first birth rates by 5%

among women 35 and older but did not affect first births rates among younger women, our analysis indicates that the largest effects were among women 35-39. The percentage increases were smaller for both younger and older women. In addition, our results indicate that limiting the analysis to first births may fail to capture important effects of on delivery rates. In particular, we find that mandates had a larger effect on second or higher parity deliveries than on first deliveries for women 35-39. This difference is potentially important when interpreting our findings because it is more consistent with an increase in the overall number of births than with a shift in the timing of births due to the implementation of a mandate. Finally, Schmidt's analysis provides no evidence on the corresponding effects on rates of multiple births among women of any age.

Do younger women delay childbirth in response to mandated insurance coverage? While we find that delivery rates for women 25-29 declined after the adoption of comprehensive mandate in our main model, this result is sensitive to the classification of the Maryland law as comprehensive. In addition, we find a similar effect for offer-only mandates, but not for limited mandates. Thus, we view our findings on this issue as inconclusive because they are sensitive to model specification, although we note that other research using different methods and datasets has documented this type of effect (Buckles 2006).

We also find that comprehensive mandates had relatively large effects on rates of multiple deliveries for women 25 to 39. For women 30-34 and 35-39, comprehensive mandates are associated with a 10% increase in rates of multiple births. The magnitudes of the effects of comprehensive mandates on rates of triplet or higher order deliveries range from a 29% increase for women 25-29 to a 54% increase for women 35-39.

While our results are not inconsistent with those of Bitler (2006) who documents an increase in the proportion of births that are twins among women 30 and over, they provide a more complete picture of the effects of mandates on rates of multiple births. In particular, by not limiting our analysis to twin births, we find that comprehensive mandates increased rates of triplet and higher order births as well, and that this increase was not limited to women 30 and over, but extended to younger women aged 25-29. Because triplet and higher order births are both more costly and more risky than twin births, this effect has potentially important public health implications.

Do comprehensive mandates increase the risk of multiple births among the relatively high fertility group - women who would have been likely to give birth even in the absence of treatment? We find evidence of this type of effect among women 25-29 and possibly among women 30-34. Among the youngest women (25-29), we find no evidence that infertility mandates increased the number of deliveries although they did increase the rate of multiple deliveries. Because comprehensive mandates resulted in an increase in both deliveries and multiple deliveries among women 30 to 34 and 35 to 39, whether they had this type of effect depends on the ratio of incremental multiple deliveries to incremental deliveries induced by the mandate. We examine this in Table 9 by re-estimating the model of multiple deliveries based on multiple deliveries per capita rather than per delivery and calculating the ratio of incremental multiple deliveries to incremental deliveries associated with a comprehensive mandate. This type of effect would result in the ratio of incremental multiple deliveries to incremental deliveries in the population exceeding the ratio of multiple births among those treated with ART.

For women 35-39, we find that proportion of incremental deliveries associated with a comprehensive mandate that were multiples was 0.11 with a 95% confidence interval from 0.06 to 0.18. This is somewhat lower than 30%, the approximate average rate of multiple births among those conceiving with ART, providing little evidence that the mandate led disproportionately to incremental utilization among those who would have given birth in the absence of treatment. For women 30-34, in contrast, the ratio of the estimates is 0.18 with a 95% confidence interval from -0.412 to 1.589. Although the imprecision of the estimates makes it possible that the true ratio is either higher or lower than 0.3, the potential for higher values suggests that it is possible that many women in this age group conceiving with ART may have delivered a baby in the absence of treatment.

Did comprehensive mandates increase utilization among women for whom treatment was not successful? The SART data provides evidence of this type of an effect. We find that utilization of ART was higher in states adopting comprehensive mandates for women both under 35 and 35 and over, and rates of deliveries per cycle and multiples per delivery were lower for both age groups. This is consistent with an increase in utilization among poor prognosis patients. These patients were less likely to conceive in a given cycle and may have produced fewer viable embryos to transfer, resulting in a fewer multiples per delivery.

The effects of mandates on the composition of the population using ART as well as other infertility treatments may have important implications for evaluating the effects of infertility treatment on infant health. In particular, we demonstrate that the marginal patient induced to use infertility treatment in response to a mandate is likely to differ from the average patient using these treatments. For example, Bitler (2006) finds that twins born in mandate states have slightly worse health outcomes than those born in non-mandate states, indicating the existence of negative effects of either selection of patients into treatment or the use of infertility treatments on infant health. While our analysis indicates that the women undergoing treatment for infertility in response to a benefit mandate may represent those from either the lower or higher end of the fertility distribution relative to those receiving treatment in the absence of a mandate, the relatively high fertility patients are those most likely to have a multiple birth. Thus, the relatively small effects observed by Bilter may be representative of the most fertile portion of the distribution of patients receiving treatment. Based on this reasoning, the negative effects on infant health among the treated population in the absence of a mandate could be larger.

While the results from our analyses of utilization are necessarily more tentative due to the absence of data from the pre-mandate period, we do find evidence that utilization of ART grew more rapidly after the adoption of mandates for both younger and older women of reproductive age. Consistent with other research, we find that births per cycle, multiples per birth, and the number of embryos transferred per cycle are lower in states with comprehensive insurance mandates than in states without mandates. While other studies have attributed these relationships to less aggressive embryo transfer decisions caused by more generous insurance coverage, our results point to an alternative explanation – that, on net, comprehensive insurance coverage extended treatment to a relatively low fertility population for whom treatment was often not successful. While we cannot rule out the possibility that in states with comprehensive mandates, patients transferred fewer embryos per cycle – the estimate is negative, but imprecisely estimated - the lower rates of deliveries per cycle and multiples per delivery in mandated states is not explained by embryo transfer rates. Thus, we conclude that these relationships are more likely to have been driven by a worse case mix in states adopting comprehensive mandates than more conservative embryo transfer decisions due to expanded insurance coverage.

An important limitation of our study is that the data on utilization and outcomes of ART do not span the period prior to the adoption of most insurance mandates. In addition, we are unable to assess the effects of mandates on treatments other than ART such as AI or ovulation-inducing drugs. In the case of the models using the population birth data, the effects we observe may be driven by increased utilization of any type of infertility treatment, not just ART. In addition, the mandates may have affected utilization of these other treatment in ways we do not observe. While the existence of endogeneity in the passage of the laws represents is also possible, other researchers have documented that the states passing laws were very similar to those not passing laws in the time period proceeding the adoption of the laws (Hamilton and McManus 2005; Bitler 2006), somewhat alleviating this concern.

VII. Conclusion

In summary, we find that the adoption of mandates that insurers cover the treatment of infertility significantly increased the utilization of these treatments. For many patients, the outcome was favorable in the sense that the incremental utilization resulted in a birth which likely would not have taken place in the absence of a mandate.

For others, however, the welfare implications are less clear. While both relatively high and low fertility patients responded to the price reduction created by the mandate, the consequences of greater utilization of these technologies differed between the two groups. For some high fertility patients seeking treatment in response to the availability of more generous insurance coverage, utilization led to a multiple rather than a singleton delivery. This shift in the birth distribution toward multiples is important because multiple births are not only high cost, but are also high risk for both mothers and infants, primarily due to complications associated with low birth weight. For low fertility women, in contrast, the adverse consequences of greater utilization were primarily in the form of unsuccessful outcomes. For many of these couples, treatment was costly and did not ultimately lead to a live birth.

These findings indicate that, even if infertility mandates increase welfare by reducing adverse selection in the coverage of infertility treatments, these benefits must be weighed against the significant moral hazard that they induce. However, they also point to an alternative explanation for the absence of insurance coverage for infertility treatment. Perhaps the moral hazard associated with generous insurance coverage itself reduces the extent to which insurers cover these services. While our analyses cannot differentiate between these two explanations, their implications for policy differ. If the market for insurance coverage for infertility treatment is limited due to adverse selection, a benefit mandate intended to solve this problem would need to be accompanied by mechanisms to minimize the resulting moral hazard. These types of mechanisms may include restrictions on the population covered by the mandate and controls on utilization through either supply side mechanisms or demand side cost-sharing. If the market for coverage is limited primarily by the extent of moral hazard generated by more generous insurance coverage, in contrast, a benefit mandate will not solve this problem. A more effective solution would be the development of more sophisticated mechanisms to target utilization to those for whom treatment the expected benefits of treatment exceed the expected costs.

References:

Adashi, E. Y., P. N. Barri, et al. (2003). "Infertility Therapy-Associated Multiple Pregnancies (Births): an Ongoing Epidemic." Reproductive BioMedicine Online 7(5): 151-542.

Bertrand, M., E. Duflo, et al. (2004). "How Much Should We Trust Differences-in-Differences Estimates?" Quarterly Journal of Economics February: 249-275.

Bitler, M. and L. Schmidt (2006). "Health Disparities and Infertility: Impacts of State-Level Insurance Mandates." Fertility and Sterility 85: 858-865.

Bitler, M. and L. Schmidt (2007). Who do Health Insurance Mandates Affect? The Case of Infertility Treatment: 28.

Bitler, M. P. (2006). Effects of Increased Access to Infertility Treatment on Infant and Child Health Outcomes: Evidence from Health Insurance Mandates.

Buckles, K. (2006). Stopping the Biological Clock: Infertility Treatments and the Career-Family Tradeoff: 60.

Callahan, T. L., H. J.E., et al. (1994). "The economic impact of multiple-gestation preganancies and the contribution of assisted reproduction techniques to their incidence." New England Journal of Medicine 331(4): 244-249.

CDC (2000). "Contribution of Assisted Reproductive Technology and Ovulation-Inducing Drugs to Triplet and Higher-Order Multiple Births--United States, 1980-1997." MMWR 49(24): 535-8.

CDC (2002). "Infant Mortality and Low Birth Weight among Black and White Infants -- United States, 1980-2002." Morbidity and Mortality Weekly 51(27): 589-592.

CDC (2005). Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports, U.S. Department of Health and Human Services, Centers for Disease Control and Prevention: 1-84.

CDC (2006). 2004 Assisted Reproductive Technology Success Rates: National Summary and Fertility Clinic Reports. Atlanta, GA, Centers for Disease Control and Prevention.

Collins, J. (2001). "Cost-effectiveness of in vitro fertilization." Semin Reprod Med 19(3): 279-89.

Gruber, J. (1992). "State-mandated benefits and employer-provided health insurance." Journal of Public Economics 55: 499-464.

Gruber, J. (1994). "The Incidence of Mandated Maternity Benefits." American Economic Review 84(3): 622-641.

Hack, M., N. K. Klein, et al. (1995). "Long-Term Developmental Outcomes of Low Birth Weight Infants." The Future of Children 5(1).

Hamilton, B. H. and B. McManus (2005). Infertility Treatment Markets: The Effects of Competition and Policy: 42.

Hansen, M., C. Bower, et al. (2004). "Assisted reproductive technologies and the risk of birth defects - a systematic review." Human Reproduction 20(2): 328-338.

Jain, T., B. L. Harlow, et al. (2002). "Insurance Coverage and Outcomes of in Vitro Fertilization." The New England Journal of Medicine 347(9): 661-666.

Jensen, G. A. and J. R. Gabel (1992). "State Mandated Benefits and the Small Firm's Decision to Offer Insurance." Journal of Regulatory Economics 4(4): 379-404.

Jensen, G. A. and M. A. Morrisey (1999). "Employer-Sponsored Health Insurance and Mandated Benefit Laws." The Milbank Quarterly 77(4): 425-459.

Kaestner, R. and K. Simon (2002). "Labor Market Consequences of State Health Insurance Regulation." Industrial and Labor Relations Review 56(1): 136-159.

Kurinczuk, J. J., M. Hansen, et al. (2004). "The risk of birth defects in children born after assisted reproductive technologies." Current Opinion in Obstetrics and Gynecology 16: 201-209.

Laugesen, M. J., R. R. Paul, et al. (2006). "A Compariative Analysis of Mandated Benefit Laws, 1949-2002." Health Services Research 41(3): 1082-1103.

Martin, J., B. Hamilton, et al. (2005). Births: Final data for 2003. Hyattsville, MD, National Center for Health Statistics.

Martin, J. A. and M. M. Park (1999). "Trends in Twin and Triplet Births: 1980-97." National Vital Statistics Reports 47(24): 1-16.

McGuire, T. G. and J. T. Montgomery (1982). "Mandated Mental Health Benefits in Private Health Insurance." Journal Of Health Politics, Policy, and Law 7(2): 380-399.

Menken, J., J. Trussell, et al. (1986). "Age and infertility." Science 233: 1389-94.

NCHS. (1981-1999). "Data File Documentations, Natality, 1981-1999 (machine readable data file and documentation, CD-ROM Series)."

Neumann, P. J. and M. Johannesson (1994). "The Willingness to Pay for In Vitro Fertilization: A Pilot Study using Contigent Valuation." Medical Care 32: 686 - 99.

Reynolds, M. A., L. A. Schieve, et al. (2001). "Risk of multiple birth associated with in vitro fertilization using donor eggs." American Journal of Epidemiology 154(11): 1043-1050.

Rothschild, M. and J. Stiglitz (1976). "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." Quarterly Journal of Economics 90: 629-49.

Schieve, L. A., M. S.F., et al. (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: 421-446.

Sloan, F. A. and C. J. Conover (1998). "Effects of State Reforms on Health Insurance Coverage of Adults." Inquiry 35(3): 280-293.

Stigler, G. (1971). "The Theory of Economic Regulation." The Bell Journal of Economics and Management Science 2(1): 3-21.

Summers, L. H. (1989). "Some Simple Economics of Mandated Benefits." American Economic Review 79(2): 177-183.

Thurin, A., J. Hausken, et al. (2004). "Elective Single-Embryo Transfer versus Double-Embryo Transfer in in Vitro Fertilization." New England Journal of Medicine 351: 2392-402.

Van Voorhis, B. J., D. W. Sotovall, et al. (1997). "Cost-effectiveness of infertility treatments: a cohort study." Fertility and Sterility 67(6): 830-836.

Figure 1: The Effect of Treatment on Birth and Multiple Birth Rates by Patient Fertility

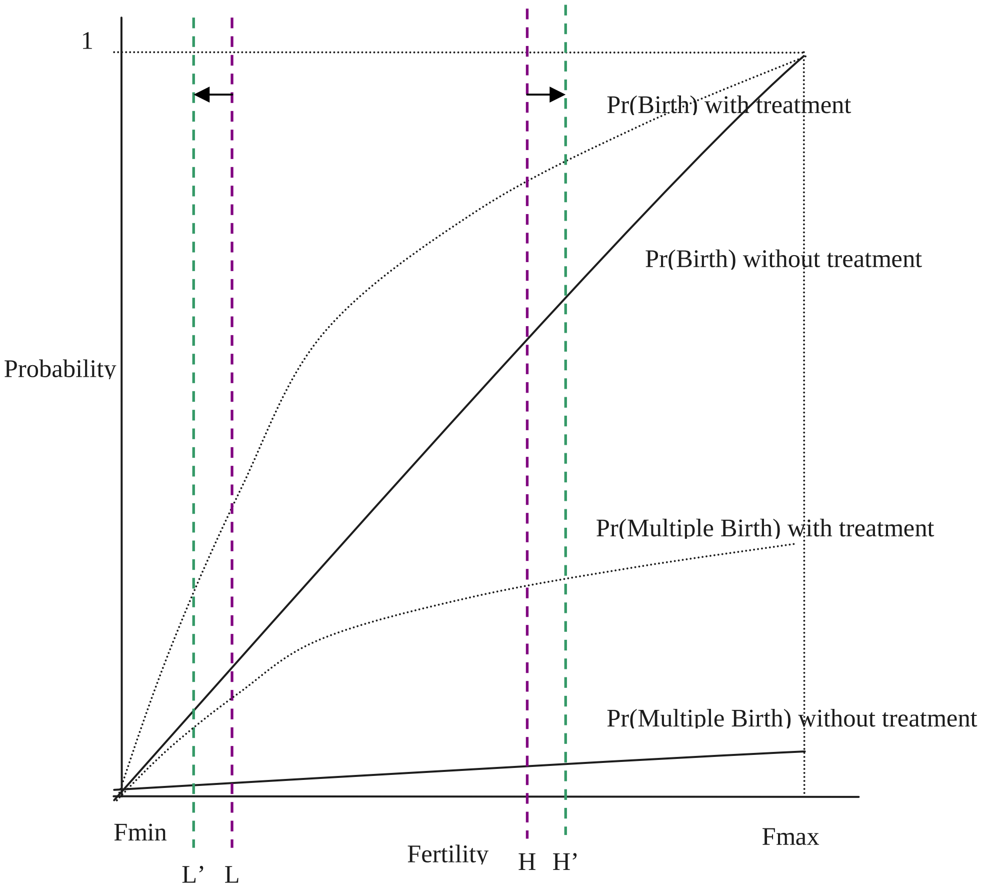

Figure 2: Deliveries per 1,000 Women 25-49 by State Mandate Status

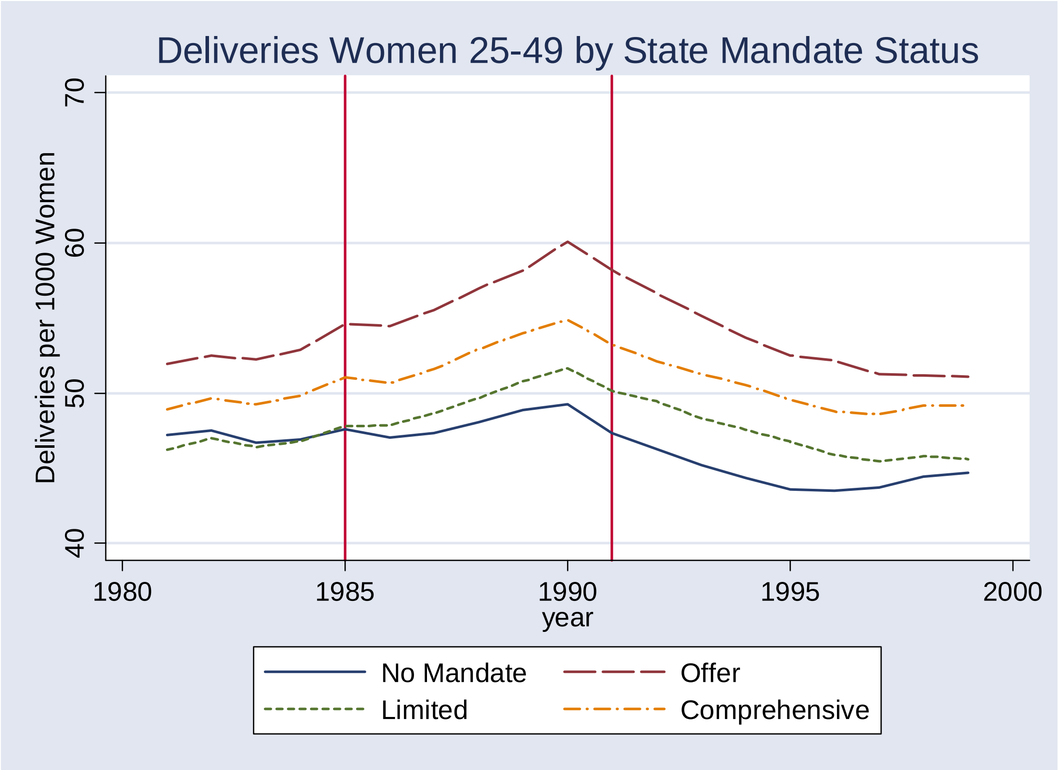

Figure 3: Multiple Deliveries per 1,000 Women 25-49 by State Mandate Status

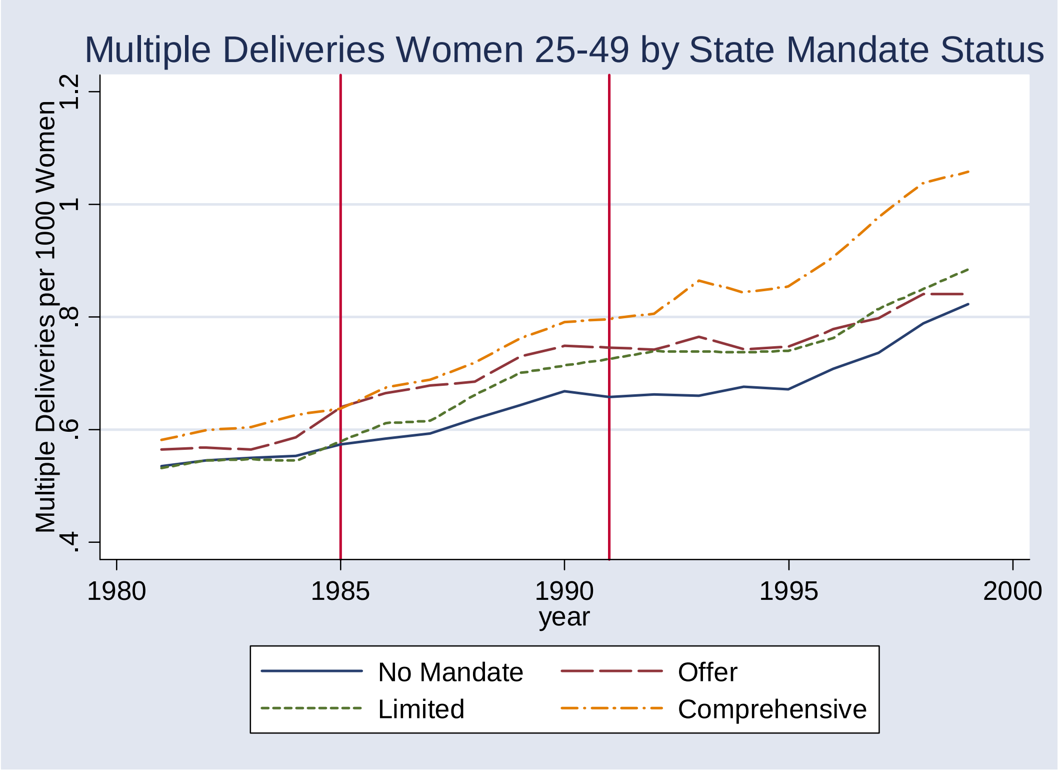

Figure 4: Triplet + Deliveries per 1,000 Women 25-49 by State Mandate Status

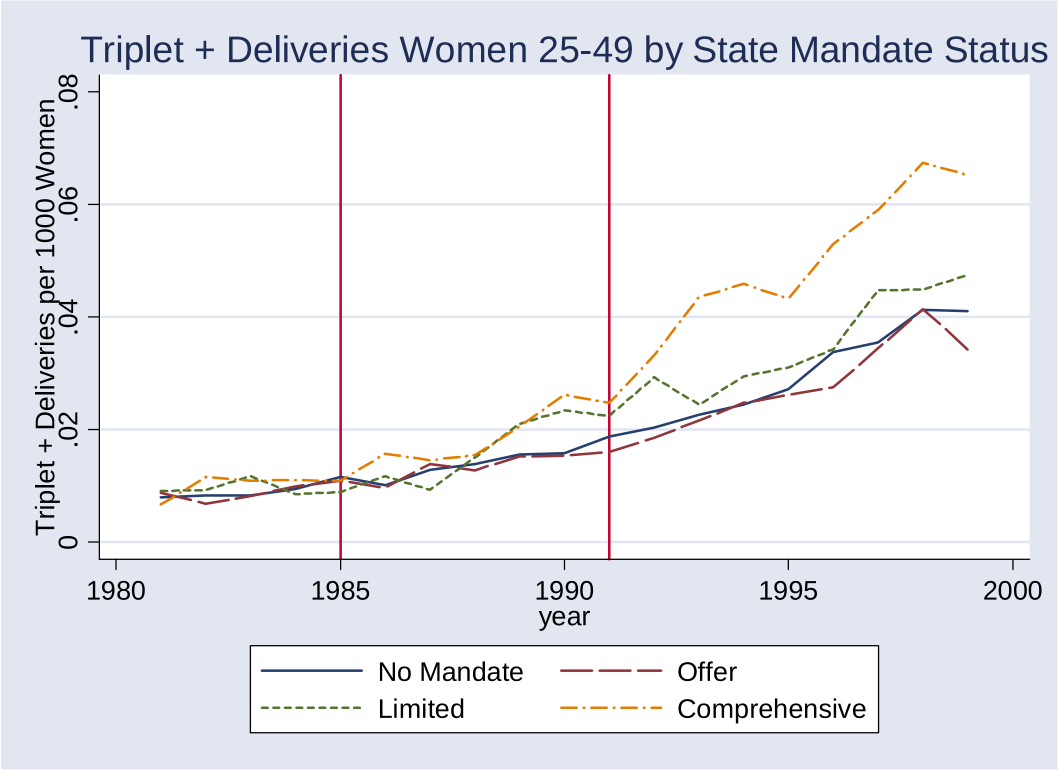

Figure 5: ART Cycles per 1000 Women 25-44 by State Mandate Status

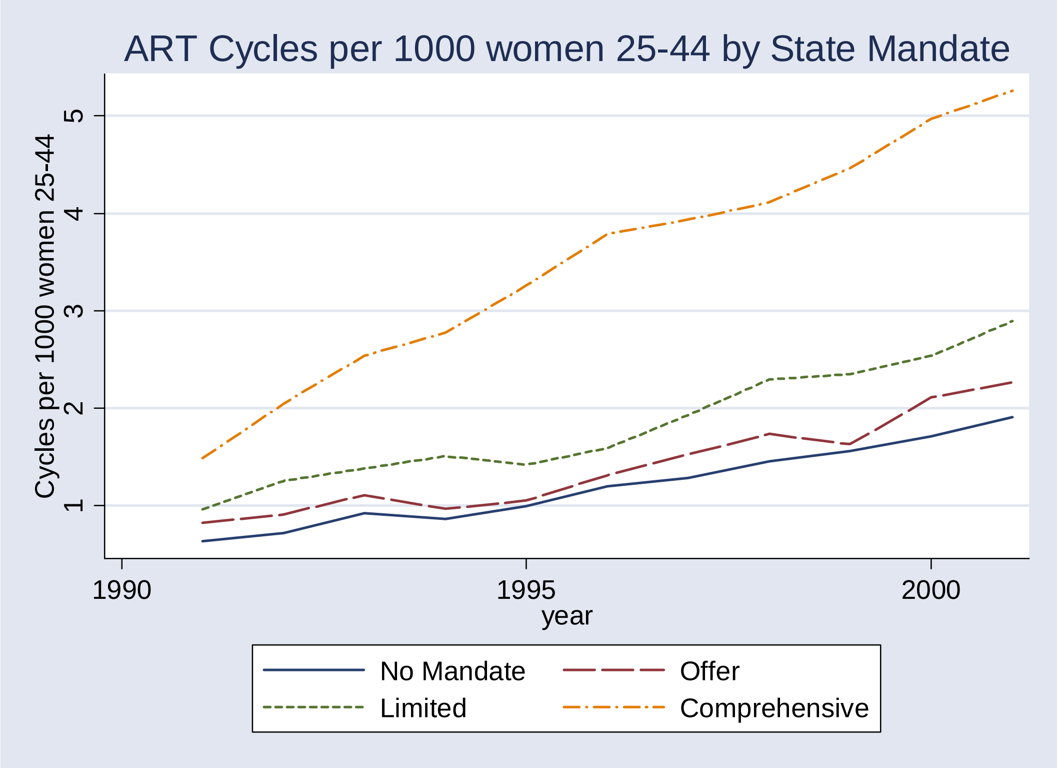

Figure 6: ART Deliveries per Cycle by State Mandate Status

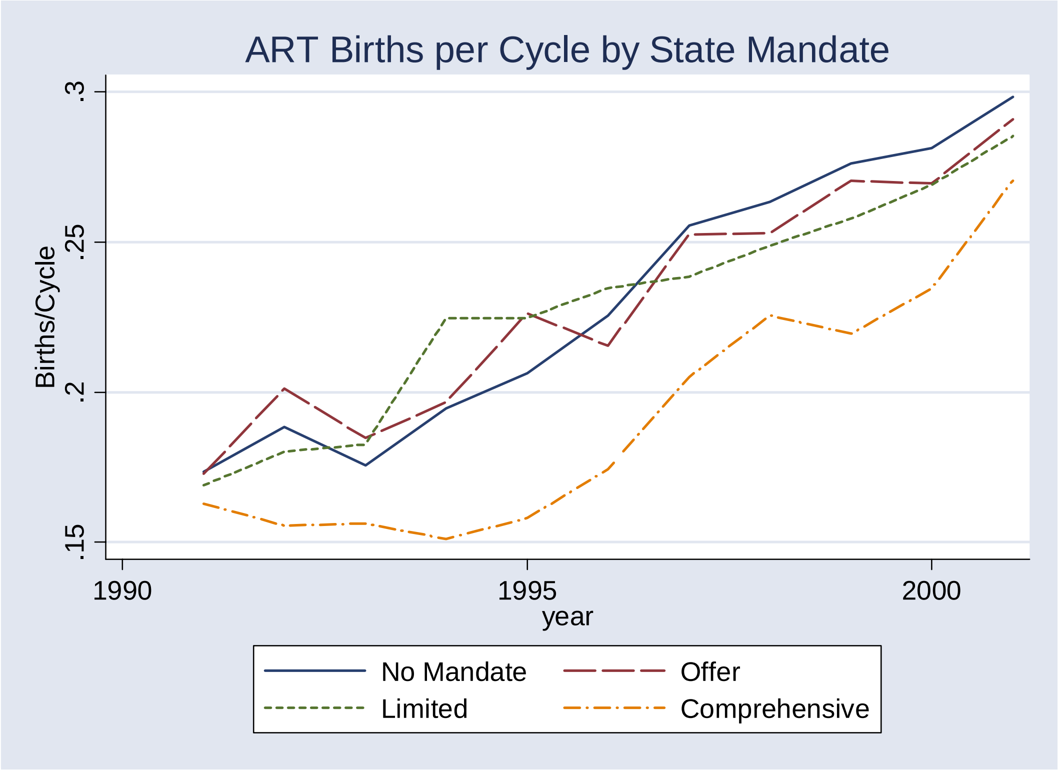

Figure 7: Multiples Deliveries per ART Delivery by State Mandate Status

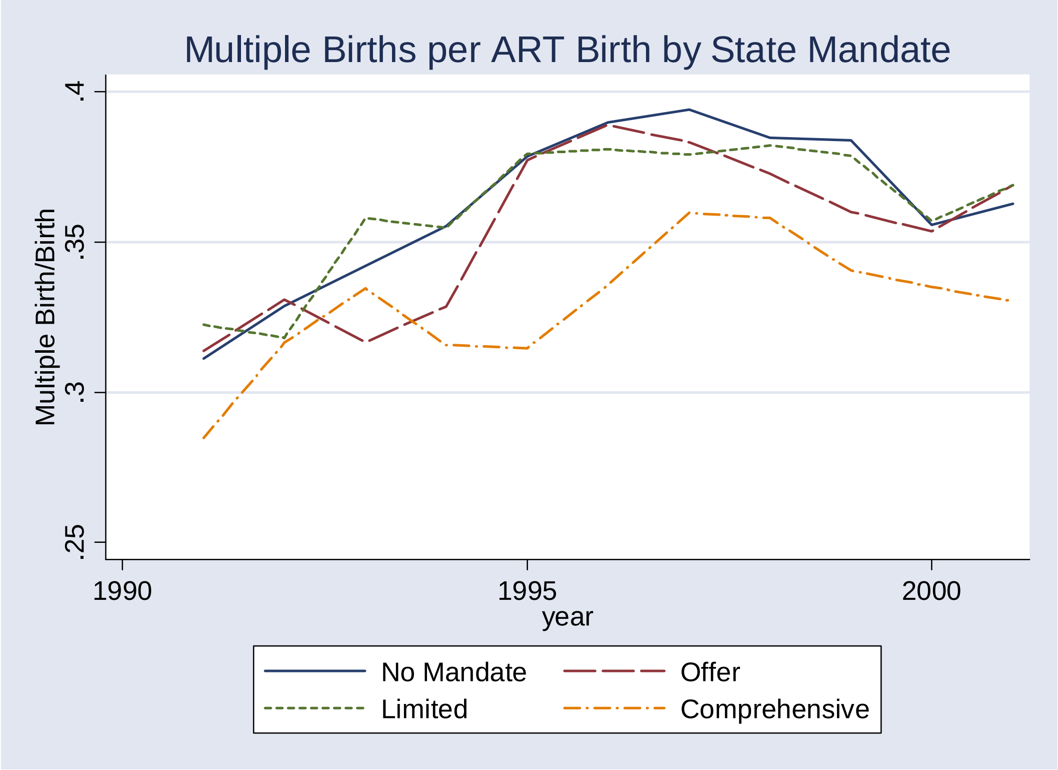

Table 1: Hypothetical Example of the Effects of Treatment on Birth and Multiple Birth Rates for Low and High Fertility Patients

	No Treatment			Treatment			Incremental Effect		
Fertility	Pr(Birth)	Pr(Multiple Birth|Birth)	Pr(Multiple Birth)	Pr(Birth)	Pr(Multiple Birth|Birth)	Pr(Multiple Birth)	Pr(Birth)	Pr(Multiple Birth)	Incremental Births /Incremental Multiple Births
Low	0.200	0.030	0.006	0.300	0.300	0.090	0.100	0.084	0.84
High	0.800	0.030	0.024	0.900	0.300	0.270	0.100	0.246	2.46

Table 2: Summary of the Timing and Type of Infertility Mandates Adopted by States

State	Year of Adoption	Type	Restrictions
Arkansas	1987	Limited	Insurers may limit the benefit to $15,000; applies to non-HMOs only
California	1989	Offer-only	
Connecticut	1989	Offer-only	
Hawaii	1987	Limited	Insurers must cover only one cycle of IVF
Illinois	1991	Comprehensive	Insurers must cover up to 4 complete oocyte retrievals for a first birth and up to 2 complete oocyte retrievals for a second birth.
Maryland	1985	Limited[1]	Limited to 3 IVF attempts per live birth and $100,000 maximum lifetime benefit. When the law was adopted, coverage was limited to patients with a 5-year history of infertility. In 1994, this requirement was changed to limit coverage to couples with a 2-year history of infertility or infertility due to particular causes with no waiting period.
Massachusetts	1987	Comprehensive	
Montana	1987	Limited	IVF excluded; applies to HMOs only
New York	1990	Limited	IVF excluded; applies to non-HMOs only
Ohio	1991	Limited	IVF excluded; applies to HMOs only
Rhode Island	1989	Comprehensive	
Texas	1987	Offer-only	
West Virginia	1977	Limited	IVF excluded; applies to HMOs only

[1] For our main results, we classify Maryland's law as limited due to its limitation to 3 cycles and its requirement of a 5-year history of infertility. The law, however, is more comprehensive than those classified as limited in our analysis. As a result, we present results re-estimating the models with Maryland classified as a comprehensive mandate.

Table 3: Descriptive Statistics for Control Variables in Models of Population Birth Rates

N=969

Variable	Mean	Std. Dev.	Min	Max
Offer-only Mandate	0.113	0.317	0.000	1.000
Limited Mandate	0.082	0.274	0.000	1.000
Comprehensive Mandate	0.035	0.185	0.000	1.000
Female Labor Force Participation Rate	0.686	0.053	0.465	0.838
Years of Education Women 20-49: 12-15	0.650	0.047	0.411	0.779
Years of Education Women 20-49: 16-17	0.162	0.035	0.069	0.284
Years of Education Women 20-49: 18+	0.052	0.020	0.006	0.197
Per Capita Income (000s)	24.473	3.733	14.222	38.332
Family Income 2 to <4 times Poverty Level	0.346	0.036	0.213	0.466
Family Income 4+ times Poverty Level	0.317	0.066	0.123	0.579
Unemployment Rate	6.400	2.032	2.200	18.000
Proportion of Workers in Firms with < 100 Workers	0.553	0.046	0.431	0.846
Proportion of Population Minorities (Non-White)	0.164	0.085	0.008	0.714
Proportion of Population Hispanic	0.094	0.098	0.004	0.417

Estimates are weighted by the size of the female population 25-49

Table 4 - The Effects of Benefit Mandates on Population Delivery Rates

Panel A	Deliveries per 1,000 women by age group				
	25-29	30-34	35-39	40-44	45-49
Comprehensive Mandate	-3.192*	0.999	1.208*	0.073	-0.019
	[1.222]	[0.663]	[0.455]	[0.128]	[0.014]
Limited Mandate	0.317	-0.806	-0.074	0.009	0.015
	[2.006]	[1.467]	[0.649]	[0.128]	[0.016]
Offer Only Mandate	-2.362*	0.37	-0.419	-0.116	-0.003
	[0.987]	[0.739]	[0.637]	[0.231]	[0.023]
Mean of dependent variable[1]	110.69	74.87	29.64	5.517	0.262
Observations	969	969	969	969	969
R-squared	0.91	0.97	0.98	0.98	0.88

Panel B	Twin or higher order deliveries per 1,000 deliveries by age group				
	25-29	30-34	35-39	40-44	45-49
Comprehensive Mandate	-0.018	1.504**	1.680*	2.512	7.043
	[0.181]	[0.359]	[0.738]	[2.053]	[5.770]
Limited Mandate	-0.038	-0.46	-0.292	-1.276	-1.322
	[0.223]	[0.474]	[0.599]	[0.857]	[5.339]
Offer Only Mandate	-0.115	-0.548	-0.475	0.831	3.418
	[0.198]	[0.478]	[0.448]	[1.156]	[5.130]
Mean of dependent variable[1]	12.29	15.2	17.58	16.139	41.664
Observations	969	969	969	969	963
R-squared	0.78	0.82	0.81	0.66	0.27

Panel C	Triplet or higher order deliveries per 1,000 deliveries by age group				
	25-29	30-34	35-39	40-44	45-49
Comprehensive Mandate	0.083**	0.304**	0.292**	0.196	-0.097
	[0.029]	[0.059]	[0.066]	[0.151]	[1.022]
Limited Mandate	0.027	-0.046	0.038	-0.208	-2.066
	[0.026]	[0.075]	[0.078]	[0.149]	[1.380]
Offer Only Mandate	-0.058*	-0.088	0.057	0.075	-0.911
	[0.027]	[0.061]	[0.084]	[0.104]	[1.568]
Mean of dependent variable[1]	0.291	0.565	0.730	0.666	3.267
Observations	969	969	969	969	963
R-squared	0.58	0.75	0.66	0.41	0.27

Standard errors are adjusted for clustering by state.
+ significant at 10%; * significant at 5%; ** significant at 1%

Note: Models include state and year fixed effects and time varying state level characteristics including female labor force participation rates, educational attainment, per capita income, distribution of the population by family income relative to poverty level, the unemployment rate, the minority rate, Hispanic ethnicity rate, and the proportion of workers employed in small firms.

[1] Mean rate over all states and all years

Table 5 - The Effects of Benefit Mandates on First Delivery Rates

Dependent Variable: First Deliveries per 1,000 women of indicated age group

	25-29	30-34	35-39	40-44	45-49
Comprehensive Mandate	-0.546	1.303**	0.458	0.093	-0.002
	[0.435]	[0.339]	[0.342]	[0.108]	[0.004]
Limited Mandate	-0.909	-0.797+	-0.175	-0.04	0.002
	[0.833]	[0.468]	[0.199]	[0.035]	[0.007]
Offer Only Mandate	-0.542	0.228	-0.147	-0.021	0.000
	[0.586]	[0.401]	[0.191]	[0.042]	[0.010]
Mean of dependent variable[1]	40.097	19.641	6.123	1.022	0.043
Observations	969	969	969	969	969
R-squared	0.89	0.98	0.98	0.97	0.81

Standard errors are adjusted for clustering by state.

+ significant at 10%; * significant at 5%; ** significant at 1%

Note: Models include state and year fixed effects and time varying state level characteristics including female labor force participation rates, educational attainment, per capita income, distribution of the population by family income relative to poverty level, the unemployment rate, the minority rate, Hispanic ethnicity rate, and the proportion of workers employed in small firms.

[1] Mean rate over all states and all years

Table 6: SART Descriptive Statistics

N=561

Variable	Mean	Std. Dev.	Min	Max
Comprehensive Mandate	0.07	0.26	0.00	1.00
Limited Mandate	0.16	0.36	0.00	1.00
Offer-only Mandate	0.21	0.41	0.00	1.00
Proportion of Women 30-34	0.26	0.01	0.21	0.28
Proportion of Women 30-34	0.26	0.01	0.22	0.30
Proportion of Women 40-44	0.24	0.01	0.20	0.31
Female Labor Force Participation Rate	0.71	0.04	0.55	0.84
Years of Education Women 20-49: 12-15	0.65	0.05	0.41	0.77
Years of Education Women 20-49: 16-17	0.17	0.03	0.08	0.28
Years of Education Women 20-49: 18+	0.06	0.02	0.01	0.20
Per Capita Income (000s)	25.51	3.31	16.76	37.64
Family Income 2 to <4 times Poverty Level	0.34	0.03	0.23	0.46
Family Income 4+ times Poverty Level	0.33	0.06	0.14	0.49
Proportion of Population Hispanic	0.11	0.11	0.00	0.41
Proportion of Population Minorities (Non-White)	0.17	0.08	0.01	0.70
Proportion of Workers in Firms with >= 100 Workers	0.57	0.04	0.39	0.68
Unemployment Rate	6.00	1.53	2.50	11.40

Estimates weighted by the size of the female population 25-44

Table 7: Relationship between State Mandate Status and Utilization of ART

	Cycles per 1,000 Women of Indicated Age				
	1991-2001		1995-2001		
	25-44	25-44	25-44	<35	>=35
Comprehensive Mandate	1.29	-0.079	1.536	1.571*	1.545
	[0.822]	[0.555]	[0.935]	[0.691]	[1.150]
Limited Mandate	-0.127	-0.344	-0.125	-0.033	-0.142
	[0.197]	[0.251]	[0.294]	[0.236]	[0.331]
Offer Only Mandate	-0.079	-0.206	-0.091	-0.037	-0.035
	[0.312]	[0.343]	[0.378]	[0.362]	[0.379]
Linear Year		0.080*			
		[0.037]			
Comprehensive*Linear Year		0.222**			
		[0.060]			
Limited*Linear Year		0.034			
		[0.033]			
Offer Only*Linear Year		0.027			
		[0.042]			
Year Fixed Effects	X		X	X	X
Linear Year		X			
Mean of dependent variable[1]	1.513	1.513	1.8	1.64	1.77
Observations	561	561	357	357	357
R-squared	0.74	0.75	0.77	0.74	0.76

Standard errors adjusted for clustering by state.

+ significant at 10%; * significant at 5%; ** significant at 1%

All models include the control variables listed in Table 8 including the age distribution of the female population of reproductive age, female labor force participation rate, education, per capita income, distribution of family income based on poverty thresholds, the unemployment rates, the proportion of workers employed in large firms, proportion hispanic and proportion minority. Standard errors are adjusted for clustering by state.

[1] Mean rate over all states and all years. Age specific rates are lower than the pooled rate due to the inability to identify donor cycles by age in the later years.

Table 8: Relationship between State Mandate Status and Outcomes of ART

	Women Under 35					Women 35 and Over				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Births per Fresh Cycle	Multiples per Birth (Fresh Cycles)	Number of Embryos Transferred per Cycle	Births per Fresh Cycle	Multiples per Birth (Fresh Cycles)	Births per Fresh Cycle	Multiples per Birth (Fresh Cycles)	Number of Embryos Transferred per Cycle	Births per Fresh Cycle	Multiples per Birth (Fresh Cycles)
Comprehensive Mandate	-0.071**	-0.033**	-0.125	-0.074**	-0.031**	-0.056**	-0.027**	-0.107	-0.056**	-0.026**
	[0.011]	[0.010]	[0.108]	[0.010]	[0.009]	[0.008]	[0.008]	[0.102]	[0.008]	[0.008]
Limited Mandate	0.012	0.009	-0.159+	0.008	0.011	0.021**	0.01	0.004	0.021**	0.01
	[0.012]	[0.007]	[0.080]	[0.012]	[0.007]	[0.007]	[0.009]	[0.092]	[0.007]	[0.008]
Offer Only Mandate	-0.039*	-0.025+	-0.047	-0.040**	-0.025+	-0.027*	-0.031*	-0.124	-0.026*	-0.030*
	[0.015]	[0.013]	[0.142]	[0.014]	[0.014]	[0.011]	[0.012]	[0.175]	[0.011]	[0.012]
Embryo Transfer Rate				-0.023*	0.017				0.002	0.008
				[0.011]	[0.011]				[0.002]	[0.005]
Mean of dependent variable[1]	0.306	0.408	3.191	0.306	0.408	0.195	0.314	3.234	0.195	0.314
Observations	324	322	324	324	322	324	323	324	324	323
R-squared	0.55	0.24	0.78	0.57	0.25	0.57	0.15	0.41	0.57	0.16

Standard errors adjusted for clustering by state.

+ significant at 10%; * significant at 5%; ** significant at 1%

Note: All models include controls for the age distribution of the female population of reproductive age, female labor force participation rate, education, per capita income, distribution of family income based on poverty thresholds, the unemployment rate, the proportion of workers employed in large firms, proportion hispanic and proportion minority.

[1] Mean rate over all states and all years.

Table 9: The Ratio of Incremental Multple Deliveries per Incremental Delivery

	30-34	35-39
Incremental deliveries per 1000[1]	0.999	1.208
	(-0.333 , 2.332)	(0.294 , 2.121)
Incremental multiple deliveries per 1000[2]	0.183	0.116
	(0.103, 0.264)	(0.042 , 0.190)
Incremental multiple deliveries per incremental delivery[3]	0.184	0.096
	(-0.412 ,1.589)	(0.056 , 0.168)

[1] Point estimates and 95 percent confidence internvals of the effect of a comprehensive mandate on deliveries per 1,000 women of indicated age. These estimates are identical to those presented in Table 4.

[2] Point estimates and 95 percent confidence intervals of the effect of a comprehensive mandate on multiple deliveries per 1,000 women of indicated age. These estimates are derived from models identical to those used to calculate incremental deliveries with the exception of the change in the dependent variable from deliveries (any order) to multiple deliveries.

[3] Estimates of the ratio of incremental multiple deliveries to incremental deliveries. 95 percent confidence intervals are estimated using the bootstrap method.

Appendix Table 1 - The Effects of Benefit Mandates on Population Delivery Rates
(reclassifying Maryland as having a comprehensive mandate)

Panel A	Deliveries per 1,000 women by age group				
	25-29	30-34	35-39	40-44	45-49
Comprehensive Mandate	-1.323	1.013+	1.099*	0.08	-0.023+
	[1.715]	[0.599]	[0.461]	[0.112]	[0.012]
Limited Mandate	-0.688	-1.012	-0.126	-0.002	0.02
	[2.184]	[1.575]	[0.682]	[0.128]	[0.014]
Offer Only Mandate	-2.136*	0.321	-0.46	-0.117	-0.003
	[1.059]	[0.750]	[0.636]	[0.229]	[0.023]
Mean of dependent variable[1]	110.69	74.87	29.64	5.517	0.262
Observations	969	969	969	969	969
R-squared	0.91	0.97	0.98	0.98	0.88

Panel B	Twin or higher order deliveries per 1,000 deliveries by age group				
	25-29	30-34	35-39	40-44	45-49
Comprehensive Mandate	0.032	1.105*	1.381*	1.922	6.381
	[0.176]	[0.504]	[0.640]	[1.791]	[5.558]
Limited Mandate	-0.077	-0.344	-0.228	-1.102	-1.344
	[0.239]	[0.470]	[0.609]	[0.807]	[5.215]
Offer Only Mandate	-0.112	-0.621	-0.532	0.74	3.223
	[0.197]	[0.490]	[0.466]	[1.175]	[5.164]
Mean of dependent variable[1]	12.29	15.2	17.58	16.139	41.664
Observations	969	969	969	969	963
R-squared	0.78	0.82	0.81	0.66	0.27

Panel C	Triplet or higher order deliveries per 1,000 deliveries by age group				
	25-29	30-34	35-39	40-44	45-49
Comprehensive Mandate	0.075*	0.249**	0.238**	0.145	-0.586
	[0.029]	[0.073]	[0.077]	[0.141]	[1.136]
Limited Mandate	0.027	-0.037	0.058	-0.198	-1.858
	[0.026]	[0.075]	[0.078]	[0.148]	[1.416]
Offer Only Mandate	-0.060*	-0.100	0.049	0.066	-0.968
	[0.026]	[0.063]	[0.086]	[0.105]	[1.581]
Mean of dependent variable[1]	0.291	0.565	0.730	0.666	3.267
Observations	969	969	969	969	963
R-squared	0.58	0.75	0.66	0.41	0.27

Standard errors are adjusted for clustering by state.

+ significant at 10%; * significant at 5%; ** significant at 1%

Note: Models include state and year fixed effects and time varying state level characteristics including female labor force participation rates, educational attainment, per capita income, distribution of the population by family income relative to poverty level, the unemployment rate, the minority rate, Hispanic ethnicity rate, and the proportion of workers employed in small firms.

[1] Mean rate over all states and all years