Mandated Health Insurance Benefits and the Utilization and Outcomes of Infertility Treatments

M. Kate Bundorf
Melinda Henne
Laurence Baker

January, 2007

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.

We thank the Iris M. Litt Fund from the Institute for Research on Women and Gender at Stanford University for financial support and participants at the 2005 AHEC conference and 2006 NBER Summer Institute for helpful comments.

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 in the form of regulations that require 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 their workers to purchase coverage for those services.[1] Over time, mandates covering a wide range of services have been adopted by state governments and, on occasion, the Federal government.

The most common economic argument in favor of 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, in turn, can remedy this. Against the potential welfare benefits, however, weigh the possible negative welfare effects of mandates. One important concern, particularly in the case of health insurance, is that mandates may exacerbate moral hazard, leading to inefficiently high levels of consumption of the mandated benefit (Gruber 1992).

The existing empirical literature provides mixed evidence on the efficiency implications of the health insurance mandates that have been adopted in the U.S. State and federal regulations mandating that insurers cover maternity benefits appear to have corrected a market failure due to adverse selection in the coverage of these services without generating significant moral hazard

---

[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.

(Gruber 1994). The mandate dramatically expanded insurance coverage, and employers were able to fully pass on the costs of the mandated benefit to the workers most likely to benefit from expanded coverage with no impact on their labor supply, indicating that the value of the incremental coverage for these workers exceeded the cost.

Other studies provide a less positive picture of the efficiency implications of benefit mandates, emphasizing the possibility that mandate passage may represent the preferences and power 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 significant moral hazard in the consumption of medical care that offsets any welfare gains. A third group of studies, however, report that mandates of even high cost benefits have *not* had these negative effects on either health insurance or labor markets (Gruber 1992; Kaestner and Simon 2002). These authors propose that mandated benefits are often provided by firms even in the absence of a mandate. In this case, mandates neither increase welfare by solving a problem of adverse selection nor decrease welfare by generating inefficiently high utilization of the mandated service.

In this paper, we investigate the impact of mandates that require health insurers to provide coverage for infertility treatments. During the last three decades, the development of new technologies has transformed the treatment of infertility, enabling people who would have otherwise been unable to conceive to bear children. 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 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, has increased rapidly during the last two decades. The first delivery using ART in the U.S. occurred in 1981. By 2003, 399 infertility clinics reported that they performed 122,872 cycles leading to 35,785 live births and 48,756 infants (CDC 2005).

Benefit mandates requiring coverage for infertility treatment offer 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 mandate. Adverse selection is likely to be a problem in this market. Infertility treatment, particularly ART, is costly for patients. 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). In addition, most patients pay out-of-pocket for these therapies. 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.

The uncertain incidence of infertility combined with the high cost of treatment make it a significant financial risk for many people, and experimental evidence suggests that individuals place a high value on insurance against this risk (Neumann and Johannesson 1994). Like prenatal and maternity services, however, asymmetric information is likely to exist between insurers and consumers in the expected use of infertility treatments. People are likely to have

---

[2] The development of drugs that stimulate ovulation in women was another important advance in the treatment of infertility. In addition to facilitating ART, these drugs are often used either on their own or with intrauterine (or artificial) insemenination in which sperm are injected into the woman's uterus at the time of ovulation and fertilization takes place in the uterus.

[3] Other approaches to ART include gamete intrafallopian transfer (GIFT) and zygote intrafallopian transfer (ZIFT), in which either unfertilized eggs and sperm (gametes) or fertilized eggs (zygotes) are returned to the fallopian tubes. While GIFT and ZIFT were used more widely in the early 1990s, their use has declined over time. In 2003, less than 1% of ART cycles involved either GIFT or ZIFT (CDC 2005).

private information about both their fertility and their desire for children that is highly predictive of their utilization but is unobservable to insurers. Because health insurance enrollment decisions are generally made annually, individuals are able to use this information when purchasing insurance. The fact that infertility is rarely covered by insurance is consistent with, although not direct evidence of, the existence of adverse selection in this market. Thus, a benefit mandate may be a useful regulatory solution to inefficiently low levels of insurance coverage of infertility treatment.

While a mandate may solve a problem of adverse selection in this market, there may be important offsetting moral hazard effects. By reducing the out-of-pocket price of treatment, expanding health insurance coverage may generate demand among consumers for whom the benefits are relatively low. While the prediction that expanding coverage will increase the size of treated population is straight-forward, the resulting effects on outcomes are potentially more subtle. To illustrate, suppose we can characterize people by their underlying fertility level. The benefits of infertility treatments are likely to be largest for those somewhere in the middle of fertility distribution. These patients have some difficulty getting pregnant, but respond well to infertility treatments so that using the treatments significantly increases the likelihood of a birth.

The expected benefits of treatment are lower for those at both the low and high ends of the fertility distribution. People at the lower end have difficulty getting pregnant and do not respond well even to advanced infertility treatment. For these patients, the expected benefits of treatment are low because they are unlikely to get pregnant either with or without treatment. Relatively high fertility people, in contrast, may have some difficulty, but would often eventually become pregnant in the absence of treatment. For example, infertility treatments may allow some women to become pregnant more quickly, even though it is also likely that they would have become pregnant eventually without them. Like the relatively low fertility patients, the expected effect of treatment on the probability of a birth is thus small. In this group, however, there is an important side effect. Since the use of advanced infertility treatments dramatically increases the chance of twins, triplets, or other higher order multiple births, these individuals become much more likely to have a multiple birth than they would have in the absence of treatment. The welfare effect of an increased probability of a multiple births is not clear a priori – some people may prefer multiple births and others may not – but there are certainly non-negligible reductions in health status and increases in health care (and lifetime) costs associated with multiple births, particularly high order multiples.

Expanding insurance coverage through a benefit mandate may also have interesting effects along the intensive treatment margin. Although infertility treatment is controversial for a variety of reasons (Spar 2006), one important concern is the high rate of multiple births among those conceiving with ART. In IVF, for example, patients and physicians must decide how many embryos to return to the uterus. While transferring a greater number of embryos increases the likelihood of a pregnancy, it also increases the likelihood of a multiple birth. More than 90% of patients transfer more than one embryo in given cycle (CDC 2005), and, in 2003, approximately one-third of deliveries resulting from ART cycles were multiple births compared to a rate of about 3% among all births (CDC 2005; Martin, Hamilton et al. 2005). The dramatic increase in multiple birth rates observed during the last three decades has been largely attributed to the utilization of reproductive technologies (CDC 2000).

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).[4] 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 of longer-term poor health outcomes (Hack, Klein et al. 1995; CDC 2002). Medical care costs are also higher for multiple than 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).

In response to these concerns, some clinical studies have proposed that more generous insurance coverage may promote *more* efficient utilization of ART by reducing the proportion of cycles that result in multiple births (Jain, Harlow et al. 2002). Because the cost per ART cycle is high, patients may try to minimize total treatment costs by maximizing the probability that a given cycle results in a birth through embryo transfer decisions. While transferring a greater number of embryos increases the probability that a cycle results in a pregnancy, it also dramatically increases the likelihood of a multiple birth. By reducing the cost of additional cycles, more generous insurance coverage may cause patients to make less aggressive embryo transfer decisions, ultimately leading to fewer multiple births among all treated patients.

Our investigation focuses on the moral hazard effects of expanding insurance coverage for infertility treatment through a benefit mandate. The high cost of treatment, combined with the lack of insurance coverage for most patients, has generated support for laws mandating that health insurance cover the treatment of infertility. Correspondingly, a number of states enacted mandates during the late 1980s and early 1990s, and Louisiana, New Jersey, and New York

---

[4] The evidence of complications associated with the use of ART, conditional on the number of births, is less conclusive. 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).

passed legislation as recently as 2001. We examine the impacts of the earlier mandates using two approaches. First, we study the relationship between the passage of mandates and the numbers of births and multiple births in the population using population birth data from 1981 to 1999, a period encompassing the passage of most of the mandates. We find that mandated coverage is associated with little change in the population-level probability of a birth, but a relatively large change in the probability of a multiple birth. We interpret this as evidence that mandates induce significant moral hazard along the extensive treatment margin among relatively high fertility patients. For this population, moral hazard results in higher rates of multiple births without a corresponding increase in birth rates.

We also study the relationship between state mandate status and rates of utilization and outcomes of ART using from longitudinal data from the vast majority of clinics operating in the U.S. We find that the potential beneficial effects on the intensive treatment margin are relatively small. While rates of births per cycle and multiples per birth are indeed lower in states adopting mandates, these differences do not appear to be driven by embryo transfer practices. Rather, these reductions in multiple birth rates appear to be associated with the selection of patients treated in the presence of expanded insurance coverage. These findings also suggest that mandates result in significant moral hazard among relatively low fertility patients as well.

II. Theoretical Effects of Expanding Insurance Coverage

In this section, we describe the theoretical relationship between expanded insurance coverage of ART and the utilization and outcomes of these treatments.

Moral Hazard along the Extensive Treatment Margin

By reducing the marginal cost of treatment, insurance coverage may increase demand for infertility services, extending treatment to those for whom it has lower expected benefits. Patients responding to the lower price may include those characterized by both relatively low and relatively high fertility. This is demonstrated in Figure 1, which plots a hypothetical relationship between a couple's underlying fertility and birth probabilities with and without treatment. Treatment is most effective for couples in the middle of the fertility distribution because it has the largest incremental effect on the probability of a birth for these couples. Lower fertility patients (L) are unlikely to have a birth without treatment and treatment generates a relatively small increase in that probability. While the treatment effect is also relatively small for high fertility patients (H), they are likely to have a birth even in the absence of treatment. A reduction in the out-of-pocket price increases utilization by both relatively low fertility couples, which is illustrated by the shift from L to L', and high fertility couples, which is illustrated by the shift from H to H'.[5] In both cases, the lower price results in patients with lower expected benefits seeking treatment, and greater utilization results in higher birth rates by increasing the probability of a birth for both groups.

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 between the two groups. Rates of multiple births among those conceiving with ART are high - approximately 30% of ART deliveries are multiple deliveries, compared to approximately 3% in the population as a whole. The proportion of births that are multiples varies relatively little with age, which can be considered a proxy for female fertility (Menken, Trussell et al. 1986), either with or without treatment. In the 1980s, prior to the widespread use of infertility treatments, differences by age for women 25-39 in rates of twin and triplet births

---

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

were small (Martin and Park 1999). In the case of ART, although birth rates among women using ART decline dramatically with age, particularly after age 35, the proportion of ART births that are multiples does not vary much by age of the mother up to age 40, when it begins to decline (CDC 2005). The implication is that, holding embryo transfer practices constant, the probability of a multiple birth will increase with fertility due primarily to its relationship with the probability of any birth. Correspondingly, in Figure 1, we plot the relationship of a multiple birth by fertility, with and without treatment, assuming that treatment has a constant, positive effect on the probability of a multiple birth, conditional on a birth.

As demonstrated in Figure 1, 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 among relatively high fertility patients without a corresponding increase in birth rates.

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.

Moral Hazard along the Intensive Treatment 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 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). Due to the high per cycle cost of IVF, individuals 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. Individuals 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 round in the event of a failure to achieve a birth.[6]

Insurance coverage may have these positive efficiency effects on treatment patterns because it *reduces* moral hazard in embryo transfer decisions among those using ART. Because the costs of delivery and neonatal care are generally covered by insurance, patients usually do not bear the high incremental medical care costs associated with the multiple births. Thus, in the absence of insurance for ART, they have incentives to make embryo transfer decisions that result in inefficiently high levels of multiple births. Insurance coverage of ART essentially corrects the distortion in utilization created by differential coverage of complementary services. 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.

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,

---

[6] 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.

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 certifications 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, the data include only a 50% sample of birth certificates during the earlier years, 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, and 40-44). We restrict our analyses to women age 25-44 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. We divide the number of twin births by two to calculate the number of deliveries of twins, and perform a similar transformation for higher-order births.

We examine utilization and outcomes among those treated using registry data from the Society for Assisted Reproductive Technologies (SART). In 1989, SART, an 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. For the data from 1989-1994, we obtained hard copies of 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 are not required to 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.[7] While the reports include information on cycles and outcomes by age, the groupings vary over time. As a result, prior to 1995, we are only able to analyze utilization and outcomes aggregated across age groups. From 1995 to 2001, we analyze separately women 35 and under and over 35, although we limit our analyses to cycles from fresh eggs due to a change in the reporting of donor cycles and cycles using frozen eggs in 1999.

Insurance Mandates

Our primary source of data for information on insurance mandates was RESOLVE (www.resolve.org), a consumer organization for people experiencing infertility. We then

---

[7] 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 the earlier years.

verified and expanded on the information from RESOLVE by reviewing the actual legislation for the mandate in each state. Twelve states passed legislation mandating insurance coverage of ART between 1985 and 1991 (Table 2).[8] 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. For example, while in Arkansas, HMOs are exempt from the mandate, in Montana, the mandate applies only to HMOs. In Illinois, insurance carriers must cover up to 4 cycles for a first birth and up to 2 for a second birth, while in Hawaii, insurers must cover only one cycle of IVF. In Rhode Island, insurers may impose cost sharing up to 20% and Maryland and Arkansas limit the lifetime maximum benefit to $100,000 and $15,000, respectively. In Maryland and Hawaii, patients must have a 5 year history of infertility, while in Arkansas, patients with a two-year history of infertility are covered by the mandate. Finally, in some states, insurers are required only to offer a policy covering infertility treatments rather than to cover infertility treatment in all policies they sell. 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 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. A “limited” mandate (seven states) refers to mandating much more limited insurance such as applying the mandate only to HMOs and either greater limits on the amount of ART coverage or exclusion of ART from the infertility services covered. States that require that insurance companies to make available for purchase a policy that covers infertility treatments,

---

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

but do not require that employers purchase these policies are considered “offer only” (three states). The states adopting mandates, the year of adoption, and their classification are listed in Table 2.

Control variables

We obtained data on the number of women in each state by age from 1981-2001 from the U.S. Census Bureau, which are the basis of our per capita calculations. We obtained information on time-varying characteristics of states, which we use as control variables in the empirical models, 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-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 employed in small firms (<100 workers) for each year.

IV. Insurance Mandates and Population Delivery Rates

IV.A. Trends in Rates of Deliveries and Multiple Deliveries

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 25-44 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).[9] 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 25-44 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 deliveries were similar by state mandate status prior to the adoption of the mandates, but clearly diverge in the post-mandate period

Important differences exist by age group in both the level and trend in delivery rates. Delivery rates (any order) decline dramatically with age (Figure 5). For older women, delivery rates increased throughout the period of the study. Though the effect is minimized by the scale of the figure, there are noticeable increases for women 40 to 44 as well as those 30 to 34 and 35 to 39. Delivery rates increased from 3.76 to 7.23 per 1,000 women 40 to 44 between 1981 and 1999.

In the period preceding the adoption of mandates, the proportion of deliveries that were multiples (twin or higher order) among women giving birth were somewhat similar across age groups (Figure 6). While rates of multiple deliveries increased for women of all ages during the study period, they increased more quickly for older than for younger women, with the most dramatic rise among women 40 to 44. The trends differ, however, when restricting the analysis to triplet or higher order multiple deliveries (Figure 7). While these rates rose throughout the study period for women of all ages, particularly in the post-mandate period, here the increase

---

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

was most dramatic for women 30 to 34 and women 35 to 39. The shift in the delivery distribution toward triplet or higher order deliveries was least dramatic for women 40 to 44.

Due to differences by age group in both the baseline levels and pre-mandate trends in delivery and multiple delivery rates, we conduct all analyses separately by age group and carefully account for pre-mandate trends.

IV.B. Mandates and Delivery Rates

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

$$y_{a,i,t} = \alpha + M_{i.t}\lambda + X_{i,t}\beta + S_i\gamma + Y_t\delta + \varepsilon_{a,i,t}\,. \qquad (1)$$

The data are aggregated to the age group (a)-state (i)-year (t) level. Y represents a measure of the delivery rate among women of a particular age group (a). M is a series of indicators representing whether the state had a comprehensive, limited, or offer-only mandate in place in a given year. We include state (S) and year (Y) fixed effects as well as a set of time-varying state characteristics (X), which were described above. Table 3 includes definitions and summary statistics for the independent variables.

We also estimate a model that augments this difference-in-difference approach to track differential trends in adopting states in the period prior to the adoption of the mandate and to capture lagged effects of mandates:

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

In this model, the superscripts on M refer to time periods relative to the year in which the state adopted the mandate: M^{-1} and M^{-2} are indicators equal to 1 for the periods 2 to 3 years and 4 or more years, respectively, prior to the adoption of a particular type of mandate. M^{+1} and M^{+2} indicate the periods 1-2 and 3 or more years after adoption. The omitted time period is then the year of adoption and the year prior to adoption.

We estimate models 1 and 2 for each age group using two different dependent variables, allowing us to examine different potential sources of bias related to policy adoption. Examining per capita delivery rates (deliveries per 1,000 women) allows us to test for differences in fertility trends across states relative to the timing of mandate adoption. We propose that this test is most sensitive to differential trends in the timing of fertility. We also examine rates of multiple deliveries per 1,000 deliveries in order to test more directly for pre-adoption differences in the utilization of infertility treatments. Our assumption is that this measure is more sensitive to trends in the use of infertility treatments, conditional on child-bearing decisions, due to the large effect of infertility treatments on rates of multiple births. Comparing the results of models 1 and 2 allows us to examine whether the estimates from the basic difference-in-difference specification (Model 1) are biased by differences in trends in delivery rates between states adopting and not adopting mandates prior to mandate adoption.

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 serial correlation in the error terms (Bertrand, Duflo et al. 2004).

The basic difference-in-difference estimates (Equation 1) indicate that comprehensive mandates were associated with a decline in delivery rates for women 25-29 and an increase in delivery rates for women 35-39 (Table 4 – Model 1 for each age group). For women 25-29, however, the negative effect of the benefit mandate appears to be driven by both relatively high rates of deliveries among women in this age group in states adopting comprehensive mandates prior to their adoption and relatively low rates in the period subsequent to their adoption. While it is possible that the lagged negative effect represents younger women delaying childbirth in response to the availability of more generous insurance coverage of infertility treatments, the positive effect in the pre-mandate period suggests that the decline is more likely driven by trends that were in place prior to mandate adoption.

For women 35-39, the estimates of pre-adoption effects in states adopting comprehensive mandates are small and not statistically significant, indicating that the difference-in-difference estimate from model 1 is not likely biased by these types of effects. While neither the immediate nor the lagged effects of the mandates are statistically significant in model 2, the lagged effect is much larger. Combining the results from models 1 and 2 for women 35 to 39, we conclude that comprehensive mandates may have increased delivery rates for these women, and that the increase occurred with a lag from the time of adoption.

Although the estimate of the effect of a comprehensive mandate on delivery rates for women 30-34 is not statistically significant in the basic difference-in-difference specification (Model 1), its magnitude is similar to that for women 35-39, creating the possibility that the comprehensive mandates did indeed have an effect on delivery rates for this population that we have inadequate power to detect. The magnitude of the estimate, however, is sensitive to the inclusion of the pre-mandate indicators, providing evidence that the effect we observe in model 1 more likely represents pre-adoption trends in delivery rates among this group than a true positive effect of comprehensive mandates. In addition, we find no evidence of a lagged effect as we did for women 30 to 34. For women 40-44, we find no evidence either of an effect of comprehensive mandates on delivery rates or of differential trends in delivery rates related to subsequent adoption of comprehensive mandates in the pre-adoption period.

We find no evidence that either limited or offer only mandates affected delivery rates among women of any age (Table 4). For limited mandates, the estimates of neither the effect of the mandate in the basic difference-in-difference specification (Model 1) nor the pre-adoption trends (Model 2) are statistically significant for women of any age group. While offer only mandates are positively associated with delivery rates for women 30-34, the large and statistically significant negative effects of the pre-mandate period indicators demonstrate that this effect is largely driven by pre-adoption trends in delivery rates for this group. In addition, the positive effect quickly dissipates in the post-adoption period.

In summary, we do not find evidence that even comprehensive mandates had dramatic effects on delivery rates for women of reproductive age. Our results indicate that any positive effects were concentrated among women age 35 to 39. For women of other ages, any effects we observe in the basic difference-in-difference specification appear to be driven by trends in delivery rates in the period preceding the adoption of the mandates. This finding differs from that of a study which documents evidence of a relatively dramatic increase in rates of first births in response to infertility benefit mandates for women age 35 and over (Schmidt 2005). This relatively large response, however, may be an artifact of the empirical specification which uses birth rates for younger women (under 35) as a control for those of older women (35 and over) to identify the effects of the mandates. Our findings indicate that trends in delivery rates in the pre-mandate period were very different both across age groups and across states adopting and not adopting these mandates. These differences suggest that using younger women as a control group may generate biased estimates of the effects of the laws.

For multiple deliveries, we find that the implementation of a comprehensive mandate was associated with a shift in the distribution of deliveries toward multiples among women 30-34 and 35-39 (Table 5). For women 30-34, there is little evidence that pre-adoption trends bias this estimate. The positive effects on rates of multiple deliveries, however, were concentrated primarily in the period three or more years after the implementation of the mandate. For women 35-39, multiple delivery rates were higher in states adopting comprehensive mandates prior to their adoption. Controlling for these trends, however, does not explain the positive effect we observe for women 35-39. The positive effect of comprehensive mandates on rates of multiple deliveries occurred earlier for women 35 to 39, although it also increased over time. We do not find evidence of effects of comprehensive mandates on rates of multiple deliveries for women 25 to 29 or women 40-44. In addition, these models provide no evidence that either limited or offer only mandates affected rates of multiple deliveries. This provides support for our interpretation of the effects of these types of mandates that we observed in Table 4 as driven by general trends in fertility rather than by mandates.

The results from Tables 4 and 5 indicate that the effects of mandates on delivery rates were limited to states adopting comprehensive versions of the laws and were concentrated among women 30 to 34 and 35 to 39. We find strong evidence that the adoption of comprehensive mandates was associated with a shift in the distribution of deliveries toward multiples, but less evidence that the laws led to more deliveries. In Table 6, we examine the relationship between mandates, focusing on the comprehensive versions of the laws among women 30 to 34 and 35 to 39, and rates of multiple deliveries per capita (rather than per delivery). This allows us to compare the number of incremental multiple deliveries to the number of incremental deliveries. Consistent with the evidence presented in Table 4, for women 30 to 34, an increase in the number of multiple and triplet or higher order deliveries was not accompanied by an increase in the number of deliveries. The primary effect of the greater utilization associated with the law was to convert singleton to multiple deliveries. For women 35 to 39, in contrast, we cannot rule out the possibility that the mandates increased the numbers of both deliveries and multiple deliveries. Unlike the case of women 30 to 34, the effect of comprehensive mandates on the number of deliveries is large, although not statistically significant. Dividing the lagged incremental multiple deliveries by the lagged incremental deliveries for women 35 to 39, we find that multiple deliveries represented approximately 15% of incremental deliveries – close to the rate of multiple deliveries among those treated with ART – approximately 33%.

IV.C. Mandates and Rates of Low Birth Weight Births

As discussed earlier, much of the concern with multiple births is due to its association with low birth weight. While our estimates of the relationship between adopting a comprehensive mandate and rates of multiple deliveries imply a corresponding increase in low birth weight infants, we formally test this by estimating a model similar to that of equation 1 with the exception that births, rather than the deliveries are the unit of analysis. Multiple infants born as part of a single delivery may have different birth weights.

Consistent with our findings of the relationship between state mandates and rates of multiple deliveries, we find evidence that comprehensive mandates were associated with an increase in the number of low and very low birth weight infants, particularly among women 35 to 39. For women 30-34, the point estimates imply that the implementation of the mandates was associated with 15 additional low birth weight and 5 additional very low birth weight births per 100,000 women, although the estimates are not statistically significant. The shift in the distribution of births toward very low birth weight births associated with a comprehensive mandates, however, is statistically significant for this group. For women 35-39, the implementation of the mandate was associated with 18 additional low birth weight and 3 additional very low birth weight births per 100,000 women.

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 allow us to supplement the results 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 using the 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 in order to determine the net effect of greater utilization on outcomes. This allows us to determine whether the non-findings for older women from the birth data are because the mandate had no effect on their utilization or because they responded to the mandate by increasing their utilization, but primarily in the form of unsuccessful cycles that did not lead to incremental deliveries.

From 1991-2001, use of ART increased dramatically. The number of cycles per 1,000 women of reproductive age (25-44) more than doubled from 0.79 to 1.83. 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 8). Consistent with findings from the medical literature, births per cycle (Figure 9) and multiples per ART birth were lower in states with comprehensive mandates than in other states (Figure 10).

We examine these trends in greater detail using multivariate models. Because the registry data do not span the period prior to the implementation of mandates, we are unable to control for either characteristics of states that are fixed over time or differences by state mandate status in ART utilization prior to the adoption of the mandate. Thus, these results are necessarily more tentative 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, the data are aggregated to the state(i)-year(t) level, and the dependent variables include cycles per 1,000 women, live deliveries per cycle, and multiples 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 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 8 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 a 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 correlation within states in the error terms.

V.A. State Mandate Status and Utilization of ART

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 9 – 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 9 – Column 2). We find no evidence that either limited or offer only mandates were associated with greater utilization of ART. When we reestimate the model and restrict the sample to 1995-1998, we obtain a similar result, although the effect of the comprehensive mandate is smaller in magnitude and only weakly statistically significant, likely due to the smaller sample size. 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. Although the point estimate is not statistically significant at conventional levels for the older group, the magnitude of the effect is similar to that of the younger group.

V.B. State Mandate Status and Outcomes of ART

Consistent with research from the medical literature, we find that rates of deliveries per cycle and multiple deliveries per delivery were lower in states that adopted comprehensive mandates than in states without mandates (Table 10). The magnitudes of the effects are similar for women under 35 (Columns 1 and 2) and 35 and over (Columns 6 and 7). Once again, the effects appear to be concentrated primarily within states adopting comprehensive mandates. We find no evidence that comprehensive mandates are associated with fewer embryos transferred (Columns 3 and 8). In addition, when we control for the average number of embryos transferred (Columns 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. In summary, we find no evidence that a comprehensive mandate leads to less aggressive embryo transfer decisions. While existing literature has attributed the relationship between mandates and lower embryo transfer rates to the ability of insurance to reduce incentives for high transfer rates (Reynolds, Schieve et al. 2001; 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.

VI. Conclusions

We find that mandates requiring insurers to provide comprehensive coverage of infertility treatments generate significant moral hazard in the utilization of these services. In theory, both relatively low and high fertility patients may respond to the reduction in the price of treatment generated by insurance coverage, and our empirical analyses provide evidence consistent with a utilization response from both groups. Using age as a proxy for fertility, we find that, among women 30-34, comprehensive mandates are associated with higher rates of multiple deliveries but have no effect on rates of deliveries. We conclude that, among this age group, more comprehensive insurance coverage of infertility treatment due to a benefit mandate led to many multiple deliveries among women who would have likely had singleton deliveries in the absence of a mandate.

Among women 35-39, we also find that multiple delivery rates were higher in states adopting comprehensive mandates. However, we cannot rule out the possibility that delivery rates also increased for these women. When we compare the point estimate for the number of incremental deliveries associated with a comprehensive mandate among women 35-39 (1.034) to the estimate of the number of incremental multiple deliveries (0.153), we find a multiple birth rate (15%) similar, although slightly lower, to that of published ART statistics (approximately 1/3). However, the estimate of incremental deliveries is noisy (indeed, not statistically significant at conventional levels in our preferred specification). As a result it is not possible to draw definitive conclusions regarding whether delivery rates did indeed increase for this group. Analyzing utilization data, we find that, in states with comprehensive mandates, utilization of ART is higher among both women under 35 and those 35 and over, providing support for our findings of effects on population delivery rates for both groups.

Among women 40 and over, we find no evidence that expanded insurance coverage increased either delivery or multiple delivery rates. Two potential explanations exist for this non-finding. One possibility is that utilization among older women is not price elastic. In this case, the implementation of a benefit mandate would have generated transfers to those who would have sought treatment in the absence of a mandate, but would not have resulted in higher rates of utilization. Alternatively, the mandate may have generated a utilization response among this age group, with relatively low success rates among those seeking treatment in response to more generous coverage. We believe that the analyses of the outcomes of ART indicate that the explanation is more likely to be the latter. If increased utilization in response to the mandate were concentrated among relatively high fertility populations, we would have expected to observe an increase in success rates, primarily deliveries per cycle. Instead, we observed the opposite, implying that the net effect of increased utilization was a decline in outcomes. Because our analyses are based on the number of cycles, rather than the number of patients seeking treatment, we cannot disentangle whether this is due to fewer patients undergoing more unsuccessful cycles or many patients seeking treatment but pursuing few cycles.

We also find that the effects of the mandates were concentrated among states adopting the most comprehensive versions of these laws and that the comprehensive mandates generally affected delivery rates with a lag. The effects on delivery rates were larger in 3 years subsequent to the adoption of the mandate than they were in the period immediately following adoption. This result is consistent with our findings from the utilization data which indicate that the adoption of a comprehensive mandate was not associated with a one-time increase in the utilization of ART, but rather an increase in the rate of growth of utilization.

While our analyses of the utilization data 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 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 comprehensive insurance coverage extended treatment to a relatively low fertility population for whom treatment was often not successful. We find no evidence of a relationship between the existence of a comprehensive mandate and the number of embryos transferred in a cycle. In addition, the lower rates of deliveries per cycle and multiples per delivery observed among this group in mandated states was not associated with the 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. Our results from the birth data suggest that this is likely to bias our findings in the direction of overstating the effects of the mandates because utilization of infertility treatments may have been increasing in states adopting mandates prior to their adoption. 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.

We conclude that the adoption of mandates requiring insurers to comprehensively cover the treatment of infertility resulted in significant moral hazard along the extensive treatment margin. While this moral hazard took the form of both relatively high and low fertility patients seeking treatment, the consequences of greater utilization of these technologies differed between the two groups. For 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. Indeed, we find direct evidence of a shift toward low-birth weight births among women 35 to 39 and very low birth weight births among women 30-34. 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 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 generated by more generous insurance coverage, in contrast, a benefit mandate will not necessarily solve this problem. A more effective solution would be the development of more sophisticated mechanisms to target utilization to those for whom treatment provides the greatest expected benefits.

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

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.

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.

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), National Center for Health Statistics, Hyattsville, Maryland.

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. (2005). "Infertility Insurance Mandates and Fertility." American Economic Review 95(2): 204-213.

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

Spar, D. (2006). "Where Babies Come From: Supply and Demand in an Infant Marketplace." Harvard Business Review February: 133-142.

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.

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

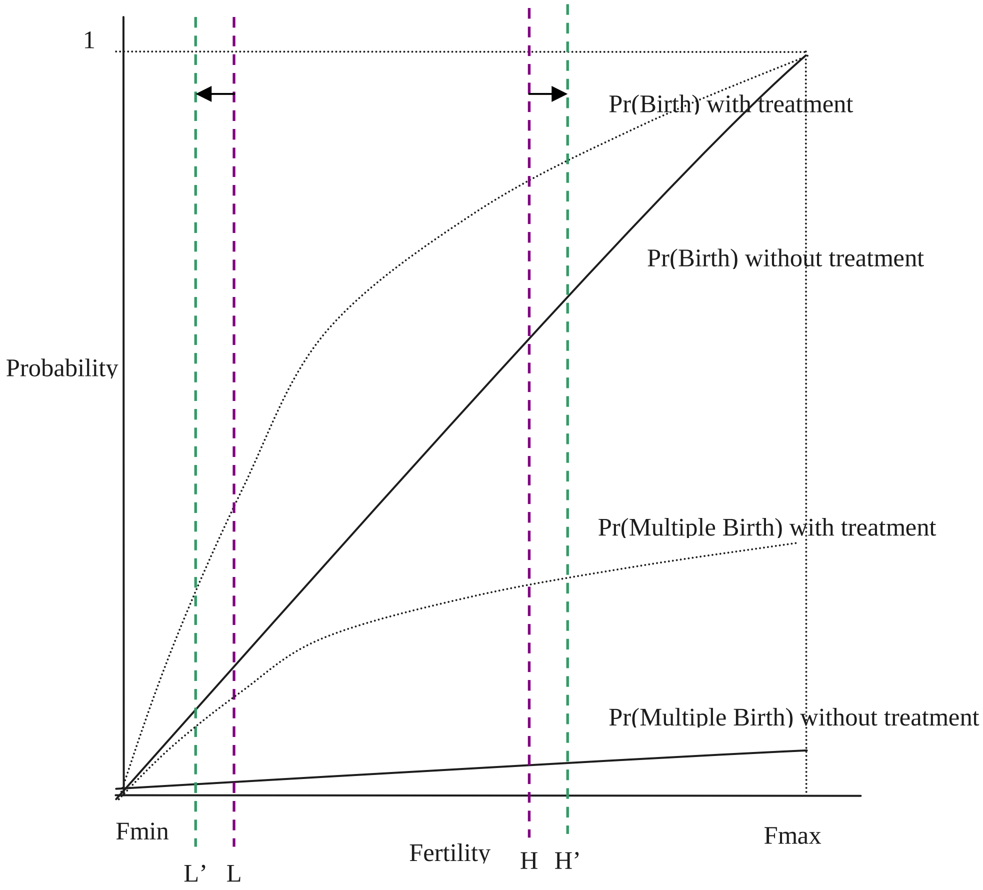

Figure 2: Deliveries per 1,000 Women 25-44 by State Mandate Status

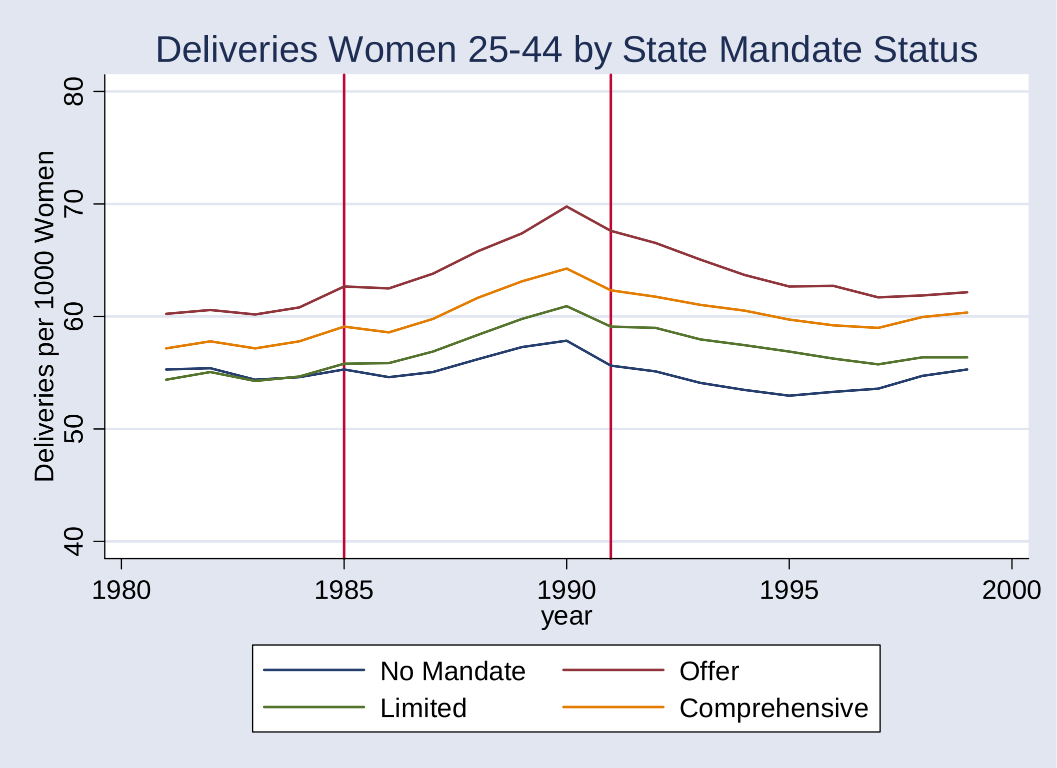

Figure 3: Multiple Deliveries per 1,000 Women 25-44 by State Mandate Status

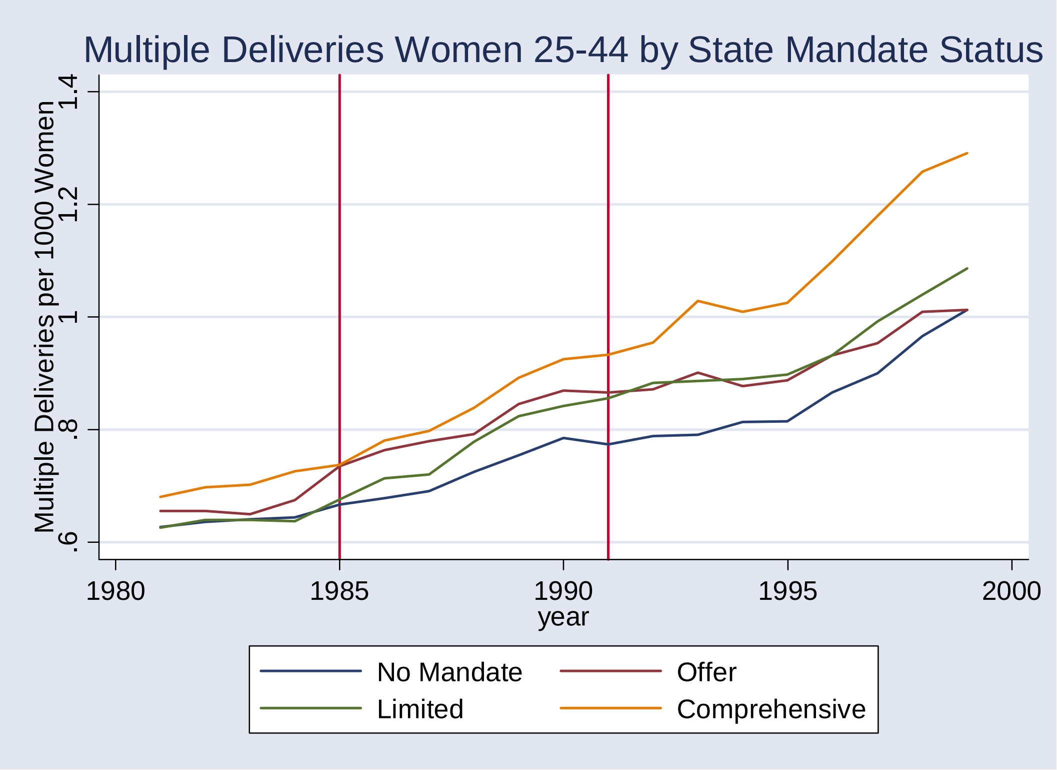

Figure 4: Triplet + Deliveries per 1,000 Women 25-44 by State Mandate Status

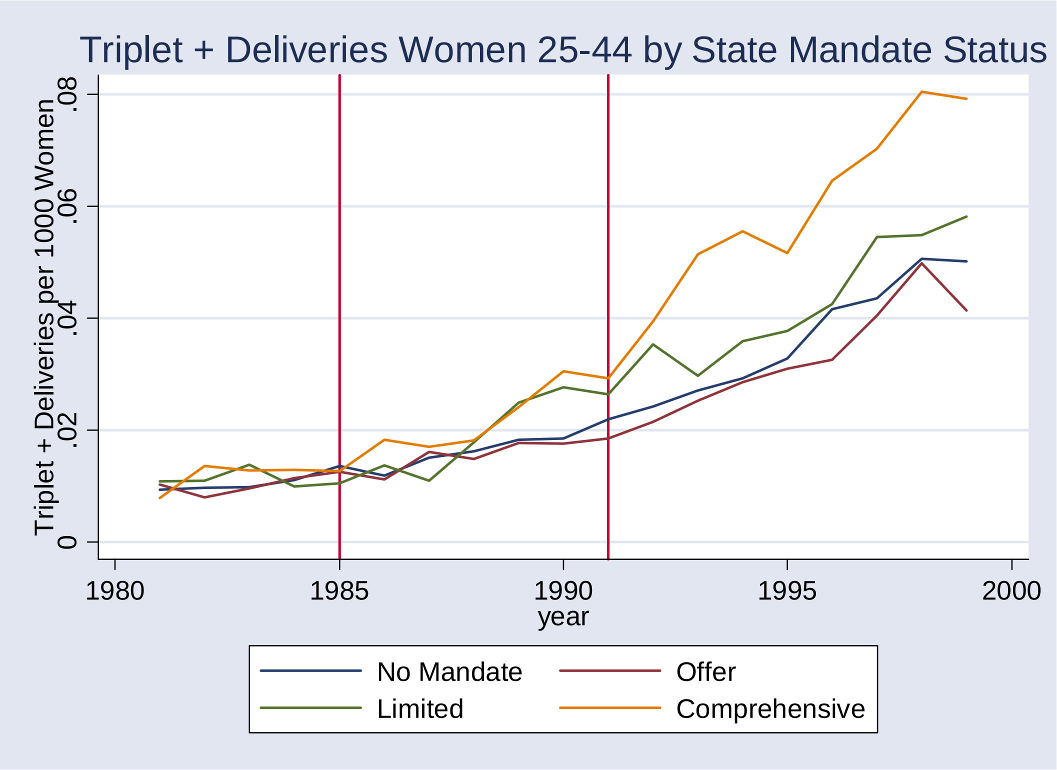

Figure 5: Deliveries per 1,000 Women by Age Group

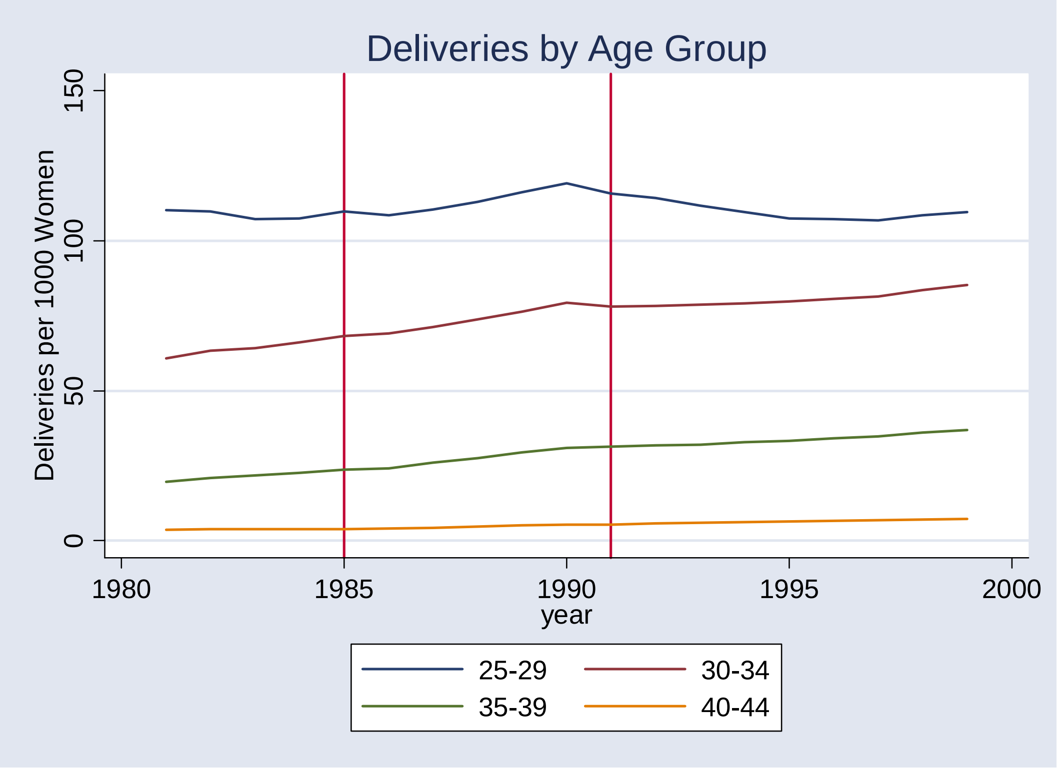

Figure 6: Multiple Deliveries per 1,000 Deliveries by Age Group

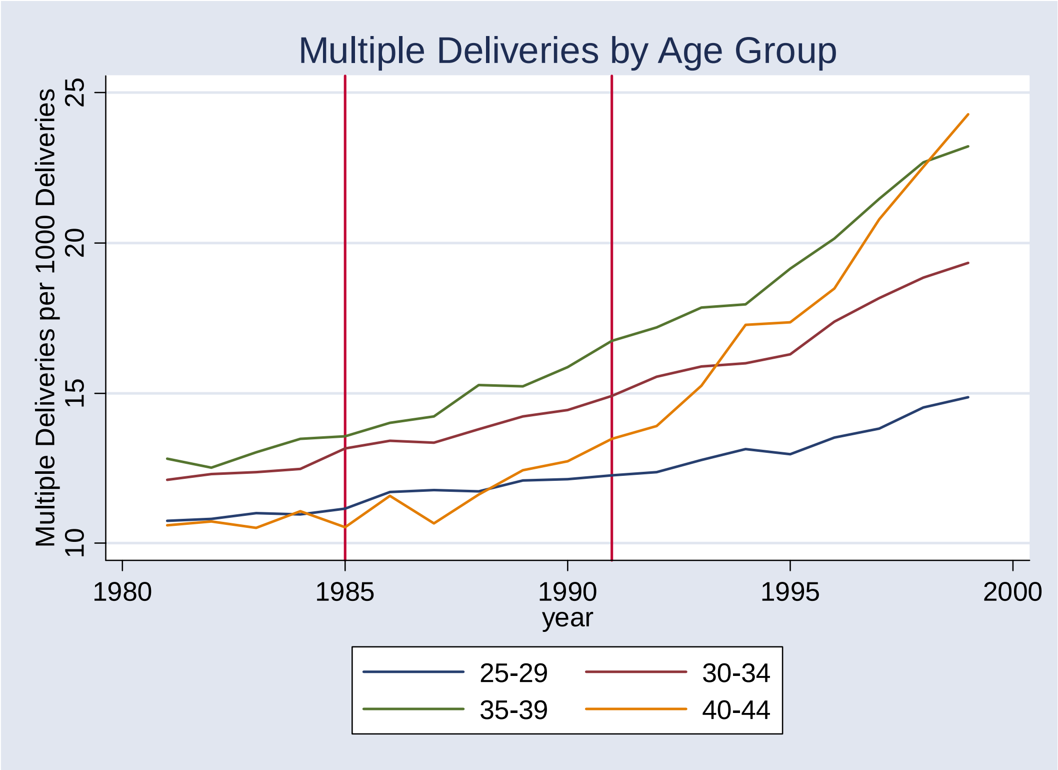

Figure 7: Triplet or Higher Order Deliveries per 1,000 Deliveries by Age Group

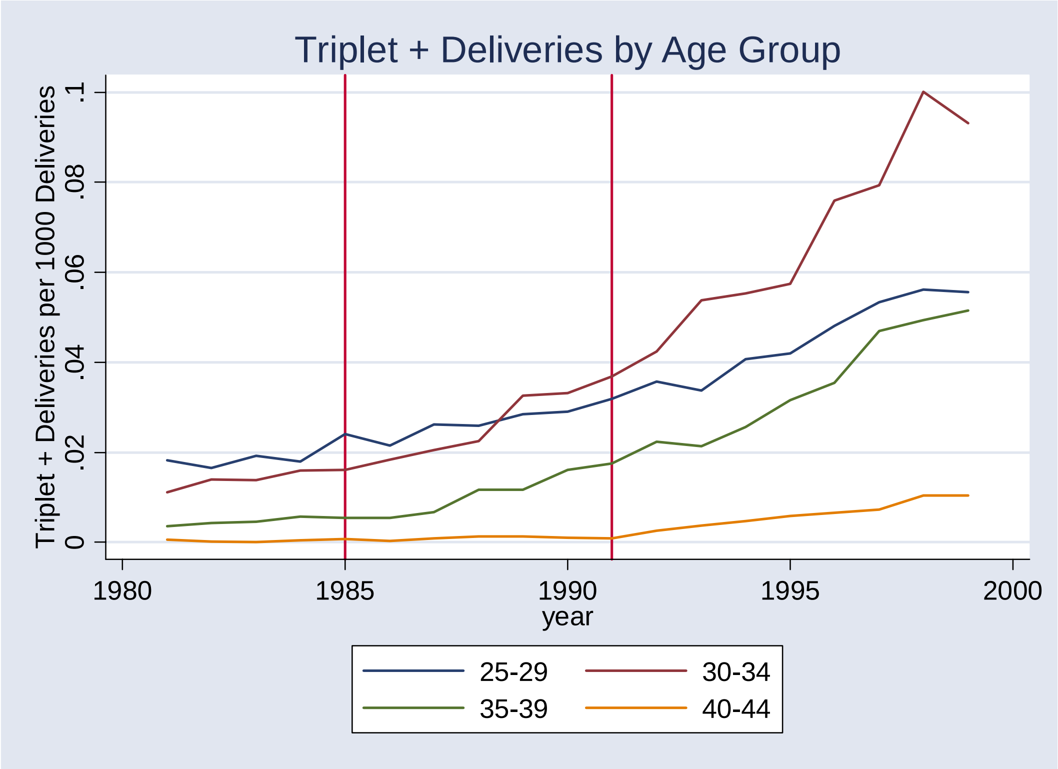

Figure 8: ART Cycles per 1000 Women 25-44 by State Mandate Status

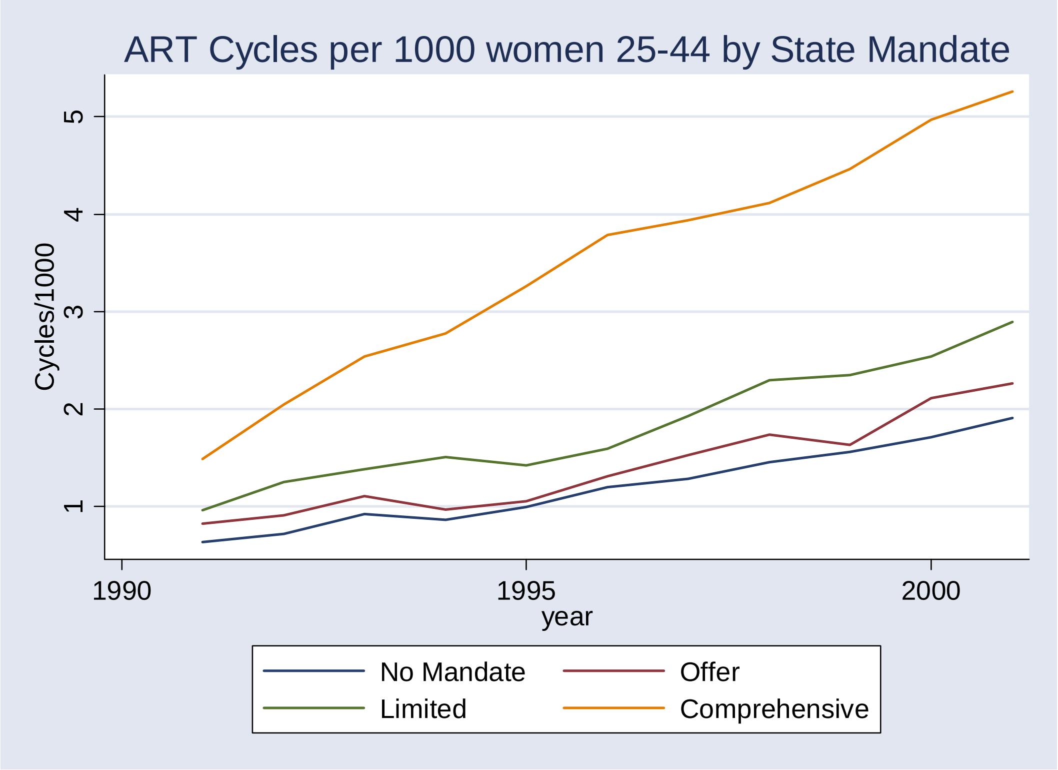

Figure 9: ART Deliveries per Cycle by State Mandate Status

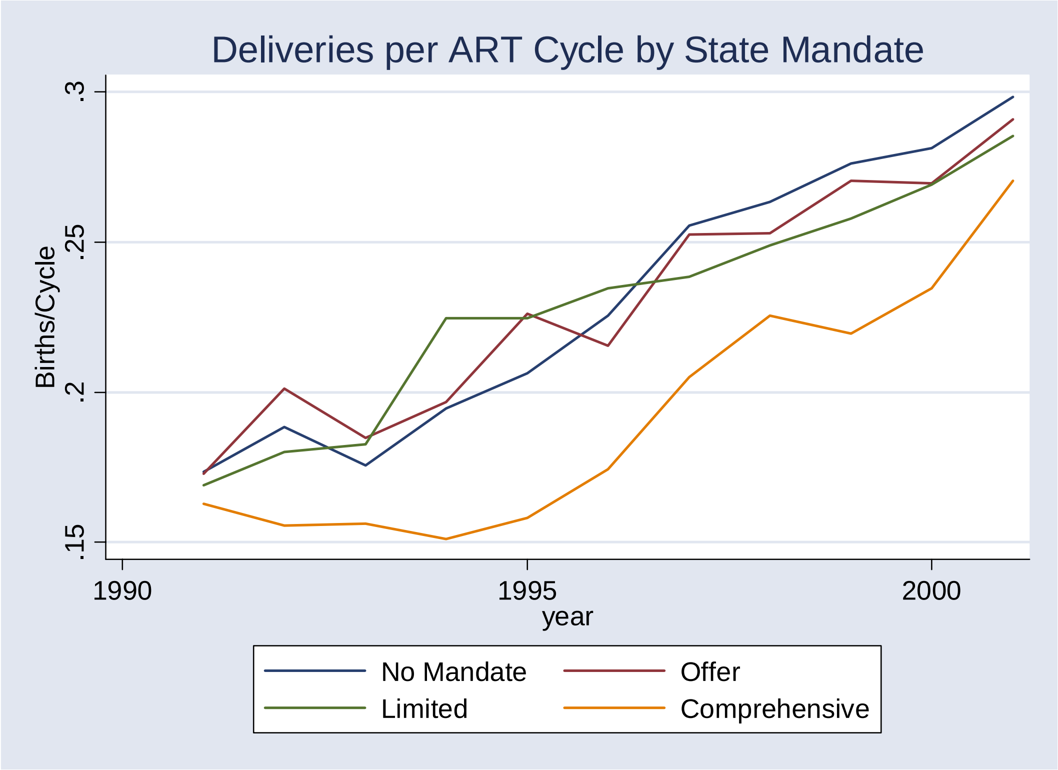

Figure 10: Multiples Deliveries per ART Delivery by State Mandate Status

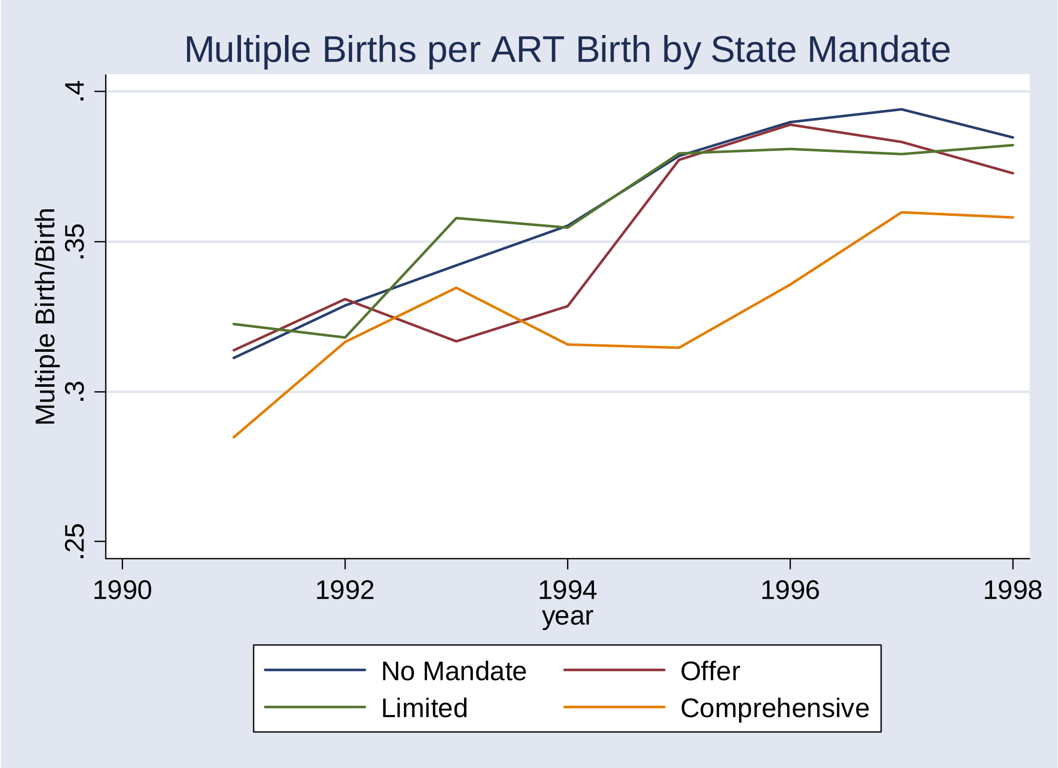

Table 1: Hypothetical Example of the Effects of Treatment on Birth and Multiple Birth Rates for Low and High Fertility Patients

	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
Arkansas	1987	Limited
California	1989	Offer
Connecticut	1989	Offer
Hawaii	1987	Limited
Illinois	1991	Comprehensive
Maryland	1985	Limited
Massachusetts	1987	Comprehensive
Montana	1987	Limited
New York	1990	Limited
Ohio	1991	Limited
Rhode Island	1989	Comprehensive
Texas	1987	Offer
West Virginia	1977	Limited

Table 3: Descriptive Statistics for Control Variables in Delivery Rate Models

N=969

Variable	Mean	Std. Dev.	Min	Max
Offer Only Mandate	0.124	0.329	0.000	1.000
Limited Mandate	0.088	0.284	0.000	1.000
Comprehensive Mandate	0.039	0.194	0.000	1.000
Female Labor Force Participation Rate	0.685	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	0.161	0.035	0.069	0.284
Years of Education Women 20-49: 18+	0.051	0.020	0.006	0.197
Per Capita Income (000s)	24.444	3.727	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	6.422	2.036	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 25-44

Table 4 - The Effects of Benefit Mandates on Delivery Rates

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

	25-29		30-34		35-39		40-44	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Comprehensive Mandate	-2.579*		1.145		1.150*		0.011	
	[1.134]		[0.822]		[0.480]		[0.119]	
Comprehensive (t-2)		1.433		-2.265*		-0.431		0.116
		[1.217]		[1.044]		[0.743]		[0.115]
Comprehensive (t-1)		1.049*		-0.656		0.107		0.005
		[0.492]		[0.516]		[0.701]		[0.076]
Comprehensive (t+1)		0.881		-0.868+		-0.121		-0.194*
		[0.984]		[0.466]		[0.269]		[0.096]
Comprehensive (t+2)		-2.735**		-0.177		1.034		0.058
		[0.922]		[0.857]		[0.810]		[0.116]
Limited Mandate	0.879		-0.496		-0.021		-0.054	
	[1.599]		[1.273]		[0.677]		[0.140]	
Limited (t-2)		0.789		0.277		0.619		0.179
		[0.819]		[0.932]		[0.632]		[0.124]
Limited (t-1)		0.755		0.465		0.141		0.058
		[0.695]		[0.404]		[0.235]		[0.067]
Limited (t+1)		2.536**		0.726		0.322		-0.117
		[0.578]		[0.544]		[0.202]		[0.094]
Limited (t+2)		0.865		-0.643		0.228		0.064
		[2.109]		[1.705]		[0.527]		[0.129]
Offer Only Mandate	0.097		2.379*		0.309		-0.053	
	[1.891]		[0.975]		[0.714]		[0.258]	
Offer Only (t-2)		-5.443*		-3.142+		-0.596		0.313+
		[2.522]		[1.735]		[1.002]		[0.176]
Offer Only (t-1)		-2.813+		-2.307*		-0.429		0.179*
		[1.414]		[0.898]		[0.378]		[0.086]
Offer Only (t+1)		1.03		3.084**		1.070**		0.227
		[1.842]		[0.824]		[0.221]		[0.145]
Offer Only (t+2)		-4.174**		-0.18		-0.624		-0.088
		[1.107]		[0.735]		[0.698]		[0.245]
Observations	969	969	969	969	969	969	969	969
R-squared	0.91	0.92	0.97	0.97	0.98	0.98	0.98	0.98

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.

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

Dependent Variable: Multiple (Twin or higher order) deliveries per 1,000 deliveries among women of indicated age group

	25-29		30-34		35-39		40-44	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Comprehensive Mandate	-0.049		1.072**		1.461*		1.691	
	[0.148]		[0.374]		[0.649]		[1.903]	
Comprehensive (t-2)		-0.069		-0.41		1.647**		0.694
		[0.257]		[0.418]		[0.442]		[0.991]
Comprehensive (t-1)		-0.126		0.233		0.965		-1.043
		[0.362]		[0.321]		[0.578]		[1.466]
Comprehensive (t+1)		-0.069		-0.528+		1.159**		-0.944
		[0.446]		[0.268]		[0.352]		[0.871]
Comprehensive (t+2)		-0.064		1.256*		2.783**		2.394
		[0.173]		[0.503]		[0.910]		[2.774]
Limited Mandate	-0.034		-0.519		-0.357		-1.238	
	[0.200]		[0.451]		[0.501]		[0.981]	
Limited (t-2)		-0.186		0.187		0.348		0.272
		[0.217]		[0.613]		[0.538]		[1.796]
Limited (t-1)		-0.093		0.223		0.726+		-1.174
		[0.239]		[0.210]		[0.410]		[1.049]
Limited (t+1)		-0.019		0.157		-0.26		-1.084
		[0.141]		[0.241]		[0.436]		[0.683]
Limited (t+2)		-0.163		-0.601		0.078		-1.46
		[0.211]		[0.528]		[0.747]		[0.930]
Offer Only Mandate	-0.298		-0.500		-0.746		-0.068	
	[0.223]		[0.473]		[0.512]		[0.989]	
Offer Only (t-2)		-0.315+		-0.061		0.227		1.03
		[0.187]		[0.397]		[0.637]		[1.857]
Offer Only (t-1)		-0.05		0.715+		0.441		-1.166
		[0.148]		[0.361]		[0.634]		[1.166]
Offer Only (t+1)		-0.399		0.128		-0.736		-0.286
		[0.273]		[0.245]		[0.457]		[0.699]
Offer Only (t+2)		-0.304		-0.605		-0.566		0.043
		[0.255]		[0.624]		[0.625]		[1.462]
Observations	969	969		969	969	969	969	969
R-squared	0.91	0.78		0.82	0.81	0.81	0.65	0.66

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.

Table 6: Effects of Comprehensive Mandates on Delivery and Multiple Delivery Rates

	Women 30-34			Women 35-39		
	Deliveries per 1000 women	Multiple Deliveries per 1000 women	Triplet + per 1000 women	Deliveries per 1000 women	Multiple Deliveries per 1000 women	Triplet + per 1000 women
Comprehensive (t-2)	-2.265*	-0.075	-0.0094	-0.431	0.041	0.0018
	[1.044]	[0.060]	[0.0058]	[0.743]	[0.034]	[0.0074]
Comprehensive (t-1)	-0.656	0.006	-0.0084	0.107	0.032	0.0036
	[0.516]	[0.041]	[0.0053]	[0.701]	[0.038]	[0.0062]
Comprehensive (t+1)	-0.868+	-0.052+	0.0089	-0.121	0.037*	0.0101**
	[0.466]	[0.030]	[0.0054]	[0.269]	[0.016]	[0.0032]
Comprehensive (t+2)	-0.177	0.148*	0.0287**	1.034	0.153*	0.0172**
	[0.857]	[0.061]	[0.0075]	[0.810]	[0.059]	[0.0062]
Observations	969	969	969	969	969	969
R-squared	0.97	0.91	0.78	0.98	0.93	0.75

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.

Table 7 - Difference-in-Difference Estimates of the Effects of a Comprehensive Mandate on Rate of Low Birth Weight

Dependent Variable	25-29	30-34	35-39	40-44
Low Birth Weight Births per 1,000 Women	-0.0158	0.151	0.176*	-0.017
	[0.0627]	[0.194]	[0.067]	[0.018]
Very Low Birth Weight Births per 1,000 Women	0.05	0.0562	0.0308*	-0.0053
	[0.123]	[0.0390]	[0.0140]	[0.0044]
Low Birth Weight Births per 100 Births	0.05	0.051	0.173+	-0.231+
	[0.123]	[0.106]	[0.099]	[0.133]
Very Low Birth Weight Births per 100 Births	0.018	0.041*	0.014	-0.084+
	[0.046]	[0.020]	[0.042]	[0.049]

Standard errors 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.

Note: Low birth weight includes infants weighing <2,500 grams and very low birth weight includes infants weighing <1,500 grams.

Table 8: SART Descriptive Statistics

N=408

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 9: Relationship between State Mandate Status and Utilization of ART

	Cycles per 1,000 Women of Indicated Age				
	1999-2001		1995-2001		
	25-44	25-44	25-44	<35	>=35
Comprehensive Mandate	1.236	-0.123	1.463	1.491*	1.486
	[0.838]	[0.561]	[0.963]	[0.721]	[1.173]
Limited Mandate	0.000	-0.329	0.019	0.12	-0.015
	[0.191]	[0.218]	[0.270]	[0.231]	[0.297]
Offer Only Mandate	-0.388	-0.389	-0.471	-0.456	-0.338
	[0.326]	[0.377]	[0.376]	[0.319]	[0.415]
Linear Year		0.079*			
		[0.034]			
Comprehensive*Linear Year		0.222**			
		[0.061]			
Limited*Linear Year		0.051			
		[0.032]			
Offer Only*Linear Year		0.012			
		[0.040]			
Year Fixed Effects	X		X	X	X
Linear Year		X			
Observations	561	561	357	357	357
R-squared	0.75	0.76	0.77	0.75	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.

Table 10: 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 Fresh Birth	Number of Embryos Transferred	Births per Fresh Cycle	Multiples per Birth (Fresh Cycles)	Births per Fresh Cycle	Multiples per Fresh Birth	Number of Embryos Transferred	Births per Fresh Cycle	Multiples per Birth (Fresh Cycles)
Comprehensive Mandate	-0.072**	-0.034**	-0.102	-0.074**	-0.032**	-0.055**	-0.024*	-0.078	-0.055**	-0.024*
	[0.012]	[0.010]	[0.113]	[0.011]	[0.009]	[0.010]	[0.010]	[0.106]	[0.010]	[0.009]
Limited Mandate	0.003	0.004	-0.172*	-0.001	0.007	0.012+	-0.001	-0.061	0.013+	-0.001
	[0.011]	[0.007]	[0.081]	[0.010]	[0.007]	[0.007]	[0.011]	[0.106]	[0.007]	[0.010]
Offer Only Mandate	-0.042*	-0.029	0.072	-0.041*	-0.03	-0.025+	-0.018	0.012	-0.025+	-0.018
	[0.020]	[0.018]	[0.188]	[0.019]	[0.019]	[0.014]	[0.014]	[0.203]	[0.014]	[0.015]
Embryo Transfer Rate				-0.022+	0.018+				0.003	0.009+
				[0.012]	[0.011]				[0.003]	[0.005]
Observations	324	322	324	324	322	324	323	324	324	323
R-squared	0.55	0.24	0.79	0.56	0.24	0.55	0.13	0.41	0.55	0.14

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.