Infertility Treatment, ART and IUI Procedures and Delivery Outcomes: How Important is Selection?[1]

Pooja G. Mookim[2], Randall P. Ellis[3], Ariella Kahn-Lang[4]

May 1, 2008

Abstract: High infertility levels are of great concern in the American society and in the recent years there has been a surge in the use of Assisted Reproductive Technologies (ART) and other infertility procedures such as Intrauterine Insemination (IUI). Women that use these treatments come from a different health risk pool compared to women who are able to conceive naturally, thus creating a selection problem as it is unclear that maternal and neonatal outcomes stem from infertility treatment type or preexisting health condition. Using a database on medical claims made by people working mostly in large firms, we first demonstrate that state mandates that require firms to offer health plans with coverage for infertility treatment are associated with increased ART/IUI usage. Once we use this information to treat the selection problem in ART treatment type, mother's pre-existing health condition is a better predictor of the likelihood of poor outcomes such as complications during pregnancy, miscarriage, abortion and ectopic pregnancy rate and neo-natal health, while infertility treatments such as ART and IUI are at best only weakly associated.

Acknowledgements: The Authors would like to acknowledge DxCG, Inc. for funding and data support and Kevin Lang for the helpful comments and suggestions.

---

[1] Financial support for this chapter and access to this data was provided by Verisk Health, Inc/DxCG of Boston MA. Prof Ellis is a scientist and co-founder at DxCG.
[2] PhD student, Department of Economics, Boston University, Boston MA
[3] Professor of Economics, Boston University, Boston MA
[4] DxCG, Inc. and Washington University in St. Louis

1 Introduction

Medical treatments that help otherwise infertile couples to get pregnant are highly successful and widely used these days. These treatments range from ovulation inducing drugs to invasive procedures such as Intra-Uterine Insemination (IUI) and Assisted Reproduction Technology (ART). While highly successful, there have been concerns regarding the safety of these procedures and their effect on maternal and neonatal outcomes. Specifically, it has been observed that these treatments significantly increase the risk of complications during pregnancy, increase the likelihood of a miscarriage, abortion or ectopic pregnancy and result in babies with poorer health status. A vast literature in the medical field recognizes the fact that while it's true that women undergoing these treatments tend to have poorer delivery outcomes on average compared to those reproducing spontaneously; it's also true that women undergoing these treatments come from a different health risk pool compared to their peers. Thus, some part of the outcomes may partly be an effect of the mother's pre-existing health condition but may still get attributed to these treatments for lack of adequate information. In this paper we suggest one way of separating out these two sources of variation in outcomes. By using an exogenous instrument that changes the probability of these treatments in a very significant way but is uncorrelated with severity, we try to find how much of the poor outcome can be associated with the treatment.

Many couples face the problem of not being able to start a family easily and naturally due to some form of infertility. The 2002 National Survey of Family Growth reported that of the 62 million women of reproductive age in US, nearly 2% or 1.2 million had infertility related medical appointment within the past 12 months. Another 10% had an infertility-related medical visit at some point in the past. Additionally, 7% of married couples in which the woman was of reproductive age (2.1 million couples) reported that they had not used contraception for 12 months and the woman had not become pregnant. Infertility is often defined broadly as not being able to get pregnant after trying for 12 months and infertility treatments includes services like medical tests to diagnose infertility, medical advice and treatments to help a woman become pregnant, and services other than routine prenatal care to prevent miscarriage. Fertility enhancement strategies such as Intrauterine Insemination (IUI) have been available for a long time and are still widely used. Assisted Reproductive Technology (ART) has been used in the United States since 1981 to help women become pregnant, most commonly through the transfer of fertilized human eggs into a woman's uterus (in vitro fertilization or IVF).

Both infertility itself and the use of infertility treatments may increase the risk of adverse obstetrical and perinatal outcomes. Unadjusted analyses suggests a two-fold increased risk of preeclampsia, placental abruption, caesarean section, and vacuum extraction, and a five-fold increased risk of placenta previa in women requiring greater than one year to spontaneously conceive singleton pregnancies, compared with women conceiving without delay (Thomson et.al, 2005). Similarly, annual reports published by Center for Disease Control and Prevention (CDC) on the number of ART cycles performed previous year and their success rates clearly shows that women who have ART are more susceptible to pregnancy complications, miscarriages, ectopic pregnancies and bad neo-natal health outcomes. This indicates that on one hand women who may be faced with a decision whether or not to try assisted methods of reproduction come from a different health risk pool than women who are able to conceive spontaneously and on the other hand using ART may expose women to higher rates of worse outcomes compared to women who conceive spontaneously. A selection problem thus occurs where it is hard to say if a bad pregnancy outcome is due to the utilization of these techniques or because of the poorer health status of the mother. Existing literature often identifies these two sources of variation in outcome rates but is seldom able to offer any way of separating these two effects. Outcome rates may also differ by the severity of the treatment—specifically; treatments like drug therapies and IUI are less invasive and expensive while others like the ART are more invasive and painful.

In deciding whether to undergo these expensive and time-consuming treatments can be difficult if proper information on success rates is lacking. On the one hand, medical indications do not properly guide patients towards the true risks of undergoing such treatments and on the other hand, the high cost associated with these treatments often prevents couples from utilizing these services. Reporting the success rates without properly accounting for the selection problem in ART utilization can introduce a bias in the success rates. In this paper, we propose one way of correcting this problem. Using a database on medical claims made by people working mostly in large firms, we identify all patients with diagnoses, procedure and drug claims that relate to infertility. Using information on states that have mandated insurance coverage of these treatments to control for the selection problem, our state mandate variable provide instruments that enable us to distinguish selection differences from real differences in pregnancy complications, miscarriages, ectopic pregnancies and health of the new-born baby.

2 Related Literatures

Earlier studies in this area have generally been inconclusive about the effects of infertility treatments on perinatal outcomes. This occurs either because they have very little data or because they use inferior estimation strategies to what we propose. Most studies do not use proper controls in the sense that they do not compare ART patients with infertile couples that conceived spontaneously. Matching controls is hard to achieve because so many couples may not even know that they are infertile and may conceive naturally eventually. The Society of Obstetricians and Gynecologists of Canada (SOGC) published a guideline in 2006 based on comprehensive survey of literature relating to ART and perinatal outcomes from 1995 to 2006 (Allen and Wilson, 2006). The first recommendation of this guideline strengthens our point on the selection problem and is worth noting:

"Spontaneous pregnancies in untreated infertile women may be at a higher risk for obstetrical complications and perinatal mortality than spontaneous pregnancies in fertile women. Further research is required to clarify the contribution of infertility itself to adverse obstetrical and perinatal outcomes"

Most studies use logistic regression analysis that does not take into account the endogeneity problem of infertility treatment type variable and therefore they cannot conclusively state if their results are due to the underlying causes of infertility or the actual treatment itself. They also typically use data only on those women who have received some sort of infertility treatment and therefore have the disadvantages of the data being censored to exclude women who can conceive spontaneously. In comparison, we use a much bigger sample of women who do and don't use these treatments and propose to treat the selection problem using two stage regression models where the 'infertility treatment' variable is treated in the first stage for endogeneity with an instrument variable—such as an indicator for whether the state has a mandate governing insurance companies on the coverage of such treatments. Another unique feature of our paper is that we carefully control for the mother's health status by using diagnoses-based risk scores to control for mother's health prior to and in the year of conception.

Several papers from the medical literature discuss maternal and neonatal outcomes following ART. Schieve, et al. (2003) use data on ART patients in US clinics between 1996 and 1998 and show that ART does not pose a risk for spontaneous abortions among pregnancies conceived using ART. The results are based on a logistic regression model that estimates the likelihood of spontaneous abortions for various treatments. In another paper (Schieve, et al. (2004)) perinatal outcomes were studied to conclude that singletons born after ART remain at increased risk for adverse perinatal outcomes like low birth weight Winter, et al. (2002) use a much smaller sample of 1196 clinical pregnancies from three reproductive centers in Australia and perform a multivariate logistic regression analyses to conclude that the risk of early pregnancy loss following ART is the same as that in general population. In another study, Clayton, et al. (2006) conclude that ectopic risk among ART pregnancies varied according to ART procedure type, reproductive health characteristics of the woman carrying the pregnancy, and estimated embryo implantation potential. Hansen, et al. (2004) conducted a systematic review of 25 papers published by 2003 that addressed the association between birth defects among children conceived using ART and ART treatment type. They also did a fixed-effects multivariate analysis to obtain pooled odds ratio for 7 most appropriate studies and found a statistically significant 30-40% increased risk of birth defects associated with ART. However, they themselves note that part of this increased risk could be due to the underlying causes of infertility in couples seeking treatment rather than the treatment itself. Not many papers document the outcome rates following other procedures like the IUI relatively few papers have contrasted outcome rate for ART and IUI procedures, even though they are often used as substitutes. A study in Belgium (Sutter, et al. (2005)) compared 126 IVF patients with 126 IUI patients to conclude that pregnancy outcomes after IUI and IVF are not different from each other. They also conclude that since there is no reason to believe that the IUI technique in itself leads to an increased obstetric or neonatal risk, the worse pregnancy outcome after IVF as compared with spontaneous conceptions is due to the specific patient characteristics rather than to the use of IVF itself.

There has been some research on the effects of the presence of insurance mandates on increased use of infertility treatments. Buckles (2005) use the presence of mandates and multiple birth rates in a state as a proxy for increased ART usage in that state. She concludes that fertility decisions are affected by the availability of infertility treatments. In comparison, we have actual samples of ART utilization rates and can see this association directly in our data. Schmidt (2007) uses a difference-in-differences approach to conclude that the mandates significantly increase first birth rates for women over 35, and these results are robust to a number of specification tests.

3 Motivations

3.1 Background in Assisted Reproductive Technologies (ART)

Many couples who want to start a family face with the disease of infertility. In US, about 12% of women of childbearing age have received an infertility service (ASRM). Depending on the diagnosis, a gamut of treatment options is available to these women to medically assist them in their childbearing endeavor. These treatments range from ovulation inducing drugs to intrauterine insemination to in vitro fertilization. According to the Center for Disease Control (CDC), ART includes all fertility treatments in which both eggs and sperm are handled. In general, ART procedures involve surgically removing eggs from a woman's ovaries, combining them with sperm in the laboratory, and returning them to the woman's body or donating them to another woman. They do not include treatments in which only sperm are handled (i.e., Intrauterine—or artificial—

insemination (IUI)) or procedures in which a woman takes drugs only to stimulate egg production without the intention of having eggs retrieved.

The types of ART include the following:

- IVF (in vitro fertilization) involves extracting a woman's eggs, fertilizing the eggs in the laboratory, and then transferring the resulting embryos in to the woman's uterus through the cervix. For some IVF procedures, fertilization involves a specialized technique known as intracytoplasmic sperm injection (ICSI). In ICSI a single sperm is injected directly into the woman's egg.
- GIFT (gamete intrafallopian transfer) involves using a fiber-optic instrument called a laparoscope to guide the transfer of unfertilized eggs and sperm into the woman's fallopian tubes through small incisions in her abdomen
- ZIFT (zygote intrafallopian transfer) involves fertilizing a woman's eggs in the laboratory and then using a laparoscope to guide the transfer of the fertilize eggs into her fallopian tubes.

In addition, ART often is categorized according to whether the procedure used a woman's own eggs or eggs from a donor and according to whether the embryos used were newly fertilized or previously fertilized, frozen and then thawed. Because an ART procedure includes several steps, it is typically referred to as a cycle of treatment.

By law, any facility in the US that performs ART treatments is required to report its success rates based on live births which then CDC is required to compile into a national report. All ART pregnancies are followed up to determine whether a birth occurred. Thus this data is collected around 9 months after the year-end.

3.2 FERT: Fertility treatments other than ART

Typically a patient diagnosed with infertility will first go through a series of lesser invasive treatments such as drug therapy and/or Intra-Uterine Insemination before moving on to ART. For example a woman may begin by ovulation induction procedure in which she takes hormonal drugs along with timed intercourse in order to achieve a pregnancy. If this treatment fails, or if she is identified with other medical problems or if there is a presence of male factor infertility the couple may be given an IUI treatment. In this treatment washed sperm is injected directly into the woman's cervix in order to give the sperm a better chance to mate with the eggs. Sometimes a tubal surgery may be required on the woman to help facilitate natural conception. Since these procedures do not involve egg retrievals, they fall outside the purview of assisted reproductive treatments. However, it's very important to study the effectiveness and outcome rates following these treatments as success rates in these treatments affect the utilization rates of ART. Cost considerations introduce some ambiguity in the sequencing of these treatments. While couples paying out of pocket may bypass these treatments and move straight on to IVF in order to increase their chances of success in a single cycle, other couples who are covered under insurance may be forced by health plans to take a minimum number of such treatments before moving on to IVF. Although one cannot be sure of the sequence in which these treatments are administered to a patient, the most natural way to think about these treatments is from less invasive and costly to more invasive and costly. Henceforth we refer to all the fertility procedure or drugs associated with diagnoses that do not fall in the category of ART as 'FERT'. We explain later on the creation of these variables in our sample and conduct our analyses while jointly controlling for both FERT and ART.

3.3 National Summary of ART success rates for 2004

Tables 3 and 4 give some comparisons of the national data to our sample. About 74% of ART cycles carried out in 2004 used fresh non-donor eggs or embryos. The average age of women using ART services was 36. 41% of women using ART were less than 35 years and about 20% were 40 years or more. Majority of the women younger than 35 used their own eggs while almost 50% of women 40 and older used donor eggs. About 34% of ART cycles resulted in a clinical pregnancy. 20.3% of ART cycles resulted in a single-fetus pregnancy, 11.3% resulted in a multiple-fetus pregnancy and 2.8% resulted in a miscarriage or ectopic pregnancy. Of all the women who achieved a clinical pregnancy, 82% had a live birth. Of these above mentioned live births, 60% were singleton pregnancies, 29% were twins and about 5% were triplets or more. A woman's age is the most important factor affecting the chances of a live birth when her own eggs are used. Success rates declined steadily after the mid 30s as fertility declined with age. The rates of miscarriages also increased steadily with age.

3.4 Mandate variable as an instrument for ART

A key problem with implementing a simple regression studying the association of ART with each of the above mentioned variables is that ART is an endogenous variable. For example, women who have ART have higher rates of complications but it is also true that the women going for ART come from a different health risk pool than women getting pregnant unassisted. In general, national data shows that women going for ART are older in age and may have some other pre-existing health conditions that prevent them from natural conception. Thus ART affects the complications rate but existing complications lead to greater ART utilization. This circularity makes it important for us to look for instruments that are correlated with ART but not with any of our dependent variables. In this paper we use information on state mandates to treat this selection problem.

3.5 State Laws regarding Insurance coverage of infertility treatments including ART[5]

Fifteen state legislatures in US have passed laws mandating coverage for infertility treatments. However, there is a wide variation in coverage across these states. Table 1 explains this variation in greater detail. Some states require coverage of only select treatments while others offer more comprehensive coverage. There are two broad forms of state mandates—one is a 'mandate to offer' which requires insurers to offer coverage of infertility treatments; however employers are not required to include such coverage in their benefit plans. The other is a 'mandate to cover' which requires health insurance companies to include coverage of infertility treatments as a benefit in every policy. Currently, only Texas and California have a mandate to offer, while the remaining

[5] The following framework is adopted from Hawkins (2006)

thirteen states have mandate to cover laws. These states can be further divided into four categories based on the differences in coverage as follows:

***IVF Coverage only*:** a number of states' mandates address only IVF. For example, Arkansas requires group health insurance companies—exempting HMOs—to cover the cost of IVF. However, there is a marital restriction on cost coverage and the woman is required to use her spouse's sperm. Legislation in Maryland also covers only IVF with a limit of three IVF attempts per live birth and a lifetime maximum level of $100,000 while Hawaii covers only one IVF treatment with a number of preconditions that need to be satisfied. Insurance companies in Texas that offer infertility treatment benefits must provide the same amount of coverage for IVF as for any other pregnancy-related procedures. Furthermore, the patient should have a five year history of infertility, have tried other treatments and must use her spouse's sperm.

Exemption of IVF: In other states, legislation specifically excludes IVF. California, for example, requires insurance companies to offer coverage of infertility treatments, including diagnostic testing and medication. Even though California's mandate does not include IVF, they cover gamete intrafallopian transfer (GIFT), a treatment similar to IVF. Patients are eligible for treatment after one year of infertility or if their infertility is caused by a medically recognized condition. California's mandate does not include age or marital status restrictions. Similarly, in New York insurance companies do not have to provide IVF, GIFT or ZIFT and the age restriction is from 21-44. They do need to cover diagnostic tests and infertility procedures such as a tubal surgery and infertility drugs. Hence mandates in these states would be expected to increase the rate of FERT but not ART.

Comprehensive Coverage**:** the states of Massachusetts, Rhode Island and Illinois provide more comprehensive coverage. Illinois requires any policy that covers more than 25 people to provide coverage for diagnosis and treatment of infertility after one year of infertility. IVF, GIFT and ZIFT are covered only if less expensive treatments have failed. The same is true for New Jersey and Massachusetts. Rhode Island allows co-payments which do not exceed 20% but coverage is limited to married individuals. Connecticut also offers comprehensive coverage but with age and enrollment restrictions. Additionally, Connecticut is the only state to limit the number of embryos transferred per cycle.

Preventative Services only**:** Montana, Ohio and West Virginia have laws that require HMOs to cover infertility services as part of a plan's "preventative health care services". The laws are jotted down in fairly broad terms and thus their scope is unclear. It is unlikely that ARTs are covered since they do not prevent infertility, but, rather, are designed to remedy the problem of infertility.

ERISA Preemption

The Employee Retirement Income Security Act (ERISA) regulates employee benefit plans, such as employer-sponsored health benefits. ERISA distinguishes between insured and self-insured private health care plans and preempts the self-insured plans. These plans are not subject to state laws regulating insurance and are therefore attractive to many employers. In 2005, the majority of covered workers were in a plan that was completely or partially self-insured. This means that the majority of covered workers are unable to benefit from state laws mandating insurance coverage for infertility treatments. However, most firms may offer these benefits anyway as a means of competing with other peers. Some firms for instance will have one high premium plan that covers infertility treatment but is less generous otherwise. This will ensure that only those employees who really need these services will opt for such plan. Since the cost of these services is high enough, one can imagine that a firm can compete with other firms for employees on the basis of such benefits.

3.6 Why mandates should be good instruments

The use and success of infertility treatments are steadily increasing. Between 1996 and 2003, the number of infertility treatments performed each year doubled. The costs of infertility treatments vary depending on the complexity of problem and therapy used. One month of prescription medications to stimulate ovulation can cost anywhere from $40 to $3000, excluding the costs of monitoring the effects of these drugs and other medications often taken in conjunction with these drugs. A typical IUI cycle would cost about $5000 for medication, patient monitoring and the insemination procedure. ARTs are even more expensive with a typical cycle costing upwards of $10,000. In addition, patients who choose these treatments typically undergo numerous cycles before becoming pregnant—if at all—posing a substantial financial burden on such couples.

Thus, insurance coverage for IUI and other less invasive and the more complex ART treatment is bound to affect decision making regarding utilization of these treatments. The presence of an infertility insurance mandate in a state should increase the utilization of these procedures in those states as more couples can internalize the cost by purchasing insurance. At the same time, the presence of a mandate does not affect health outcomes during pregnancy and after delivery. Thus, the presence or absence of a mandate seems to be a good instrument as it affects the rate at which infertility treatments are used but not the rate of their success. It must also be noted that whether or not a mandate will be effective would depend on how generous it is in specifying the minimum coverage. For example, even though states like Montana, Ohio, West Virginia, and Louisiana have mandates, it is unlikely that they will provide enough incentives for insurance companies to offer these treatments and hence are unlikely to affect FERT and ART utilization.

4 Data Description

4.1 MarketScan Database Overview[6]

Our analysis uses the commercial claims and encounters database on privately insured patients in Thompson's MEDSTAT MarketScan database. This represents the inpatient and outpatient healthcare service use of individuals nationwide who are covered by the

[6] Adamson et.al, (2006)

benefit plans of large employers, health plans, and government and public organizations. The MarketScan database links paid claims and encounter data to detailed patient information across sites and types of providers, and over time. The annual medical database includes private sector health data from approximately 100 payers. These data represent the medical experience of insured employees and their dependents for active employees, early retirees, COBRA continuees and Medicare-eligible retirees with employer-provided Medicare Supplemental plans. No Medicaid or Workers Compensation data are included. The medical and surgical claims consist of inpatient admissions and both inpatient and outpatient services. Outpatient prescription drug claims are also covered.

4.2 Sample Selection and Variable Creation

Using the commercial claims and encounters database, for each year, we first selected all women aged 21-54. After creating the state variable based on enrollee zip codes, we dropped all women who could not reliably be grouped into a particular state. We then made a comprehensive list of all procedure codes, diagnosis codes and drug codes that indicated an infertility treatment. There was some fuzziness in identifying women with infertility claims based on codes. For example, a large number of women had submitted claims for the leading hormone drug used for ovulation induction, gonadotropin, but did not have any infertility diagnosis over the five year period. In such cases, we included only those women in our infertility group who were both taking gonadotropin and also had an infertility related diagnosis/procedure. For any woman who was identified as having an infertility issue, we retained information on all her claims—not just infertility related claims

We then grouped the procedure, diagnosis and drug codes into eight categories:

P0 = no diagnoses, procedure or drug indicating infertility.

P1 = 1 if any diagnosis indicating infertility

P2 = 1 if any drug claims that are specific to infertility i.e., a drug claim associated with an infertility diagnosis.

P3 = 1 if any tests or procedures at all related to infertility treatment (almost all of our sample unless drug only people)

P4 = 1 if any artificial insemination procedure indication

P5 = 1 if any egg removal procedure regardless of any indication of a transfer to the mother (includes most incomplete procedure where egg was removed but not transferred either because the patient was an egg donor or because of cryopreservation)

P6 = 1 of Any IVF/GIFT/ZIFT

P7 = 1 if Any micro technique used.

P8 = 1 if any cryogenic procedure used

With the exception of P1, the Pi variables were not created to be mutually exclusive. Rather, these are just binary flags about whether any of these events occurred during a given year. One approach to further analyze these procedure codes is to try to put them into mutually exclusive categories, even if women do not always do them in a given order. The numbering system we use here is one natural order that generally goes from less expensive and complex to more expensive and complex. It is true that many women may skip one or more of these steps. For example, a woman may start off with an ovulation induction in hope of getting an intrauterine insemination (IUI) but if during ultrasound testing she shows a potential risk of hyper-stimulation, her IUI cycle may be cancelled and she may be bumped up to an IVF cycle. Plus, they may occur in different years so that it looks like one or more have been skipped. Still, having some form of hierarchies may prove useful for further analyses such as finding which stages of infertility related process are most affected by mandates. We thus create the following hierarchical categories using the Pi variables.

D0 = 1 if the women did not have any procedure, diagnosis or drug related to infertility—No ART or FERT women

D1 = 1 if only a diagnosis of infertility with no procedures or drugs

D2 = 1 if fertility related drugs associated with an infertility diagnosis but with no fertility procedures

D3 = 1 if any procedures with no higher level removal or transfers

D4 = 1 if any artificial insemination procedure with no egg removal or transfer

D5 = 1 if egg removal without any transfer into a mother

D6 = 1 if any IVF without any micro technique

D7 = 1 if any IVF plus micro technique used

D8 = 1 if any cryogenic procedure used

Once we set up these variables we regrouped them to form measures of infertility treatment. We grouped D1-D4 into a variable called 'FERT' which indicates whether the woman had infertility related diagnosis or procedure or drug associated with an infertility diagnosis but did not get any ART. We also created a variable called 'ART' which indicated the presence of any D5, D6, D7 or D8. This was a natural division since in general; all procedures, drugs and diagnosis in the FERT variable tend to be less invasive and less costly than anything that shows up in ART. In most cases, they also act as baseline procedures for evaluating whether or not the woman should be treated with ART. The health outcomes for new mothers and neo-natal are also deemed to be different for these two groups.

Mother's Risk Score:

DxCG compiles risk scores for each enrollee in the MarketScan data. These risk scores were calculated using the DxCG risk adjustment classification system version 6.2 which uses diagnoses to characterize the medical conditions of each individual with 184 binary flags, and then generates predictions using other samples of expected costs. The risk score used here was the concurrent risk score, predicting the cost of health services used during the same year as the delivery, except that we omitted the components of the risk score related to maternity conditions. Hence this risk score reflects the expected cost of non maternity related health care spending for each woman.

Dependent variables:

In order to accomplish our goal of studying the affect of FERT and ART on health outcomes, we identify three key questions that we may be interested in answering. (1) We want to know if FERT and ART pregnancies have higher rates of miscarriages and ectopic pregnancies than normal pregnancies. To answer this question we created an indicator variable that flags women who initially had a pregnancy diagnosis but later had a miscarriage, abortion or an ectopic pregnancy diagnosis. (2) Whether FERT and ART are associated with higher rates of complications for mothers. For this purpose we created a variable that indicates whether or not a pregnant woman in our sample had any complications. It is somewhat unclear what gets coded as a complication during pregnancy. With advanced technologies replacing old ones certain complications may still be getting coded as a complication even though the mother really didn't go through much discomfort. Thus, 85% of women in our sample had some sort of complication during pregnancy. We therefore decided to include only major complications in our analysis. (3) Do babies conceived through the use of FERT or ART have higher expected medical care costs after birth? For this purpose we used the babies' risk score variable created by DxCG on the same lines as the mother's risk score.

Mandate variable, age and year dummies:

In addition, since a woman's age is the single most important factor affecting success rates following FERT and ART, we also create age dummies from 21 to 54 years. Since, very few records indicate age of the woman to be 50 and above; we grouped these ages in a single 50+ category. We also create year dummies for each year in our sample. Finally, we create a variable called mandate. Basically this variable flags all women that reside in states where it is mandatory by law for health plans to either offer or cover infertility treatment in at least one of their plans. Connecticut was included as a non-mandate state in this variable because at the time of our survey, Connecticut did not have the mandate in effect. We also constructed fourteen dummies for each of the fourteen states that have some form of mandates to allow full variation in the state level mandates to instrument the ART variable.

Once we create the fourteen dummies, we re-group our dummies into two, separating states that have more comprehensive coverage from those that don't. Eight states were identified for this purpose: Massachusetts, Rhode Island, New Jersey, New York, Illinois, Maryland, Texas and California. As we described in section 1.3.5, the first four of these states have comprehensive coverage. Upon further consideration, the 100,000 dollar cap and 3 IVF per live birth condition in Maryland is really not restrictive given the fact that such a dollar cap would allow up to six IVF cycles for a couple on average and typically a couple would go through 3 IVF cycles per live birth. Thus the cap in the Maryland mandate is not binding and we include it in our set of states with comprehensive coverage. Also, Texas and California are the only two states that have a mandate to 'offer' and such a mandate would provide greater incentive to firms to offer plans that cover the lesser expensive FERT treatments. Thus we feel it was appropriate to include these states as well.

One-year sample:

An observation in our one-year sample is one patient year for which a female aged 21-54 appeared in our sample. Thus if a woman gets two ART cycles in two different years then she appears in our one-year sample twice. Creating the sample in this manner we get a total 7,364,102 one-year observations. We also create a sub-sample of 329,524 women who had a completed pregnancy in any year of our five year window. Hence a woman giving birth to two different children in different calendar years would contribute two observations to the one year completed pregnancy file. Finally, of these 329,524 women with a completed pregnancy, there were some for whom we could find the information on a baby. We merged these women with the baby file to produce a sample of 75263 women for whom we also had information on their baby.

Two-year sample:

Using only a one calendar year window to capture information will miss the FERT and ART information for mothers starting their pregnancy in one year and completing it in the following year. We therefore merged information across two-years to create both a sub-sample of all women who appeared in our data for two consecutive years, and more importantly, the sample of completed pregnancies. The reason for this is that women may come in for this treatment at various points during a year so even though they may appear in the year 2000 in our sample, their treatment may have begun only in November—not giving us enough time to identify her as an ART or FERT patient. Typically, where insurance covers these costs, guidelines are set as to the type and number of therapies that need to be performed on such a patient before she can be treated with ART. For example, many insurance companies would require that a patient has had a minimum of 3 failed IUI cycles before she can be given the IVF treatment.

Thus 12 months may not be enough time for us to identify all ART patients and especially see success rates etc. We therefore look for a sub-sample of women who appear for two consecutive years in our sample. For each of these women, we determined whether or not she had ART/FERT in the previous or current year. We also take an average of the woman's health risk score over these two periods and then estimate our model parameters. Thus the mother's risk score used here was the concurrent risk score, predicting the average cost of health services used during the previous and same year as the delivery, except that we omitted the components of the risk score related to maternity conditions. Even though the samples are much smaller at 147,274 women with completed pregnancy and 75,263 women with babies—such a sample more precisely classifies women by the ART/FERT status and has more power. Therefore, we only concentrate on these results.

5 Descriptive Statistics

In this section we describe the sample of women who appear in our data for two consecutive years. For example, if a woman shows up in each of the years 2000, 2001, 2003 and 2004 then we have three records on this woman in our "all women" sample. If additionally she has deliveries in 2001 and 2004, then she contributes two observations to our "completed pregnancy sample". Not all of these women make decisions related to fertility treatments. Those that don't are our control group which we call the "Non ART/FERT Women". Table 2 gives summary statistics on such patient records in the age group of 21-54 and years 2001 to 2004. The average age in our sample is 41 years, reflecting our choice of the age range of 21-54 years and the average mother's risk score is 1.296. About 5.73% women in the two-year sample had pregnancy related claims of which 3.39% had deliveries. Less than 1% of all women had a miscarriage or ectopic pregnancy. This is partly attributable to the fact that not all women in our 'No ART/FERT' group are taking fertility related decisions but we have no way of separating them out.

Table 3 further breaks down the all women sample by fertility status and also compare the results of ART women in our sample with the 2004 CDC results. We see that the mean age of ART women in our sample (35.66 years) matches the CDC reported mean age (36 years). Our sample is much smaller mainly because our sample is based on only a sample of private insurance plans, representing only 3.6 million women-years. Another reason is that some women may be paying out of pocket for ART and FERT procedures or these services could be that they are covered for FERT but have to pay out of pocket for ART. A clinical pregnancy in our sample, which includes both completed and incomplete pregnancies, is not directly comparable to the CDC rate for clinical pregnancies, since the CDC always tracks pregnancy outcomes until they are complete.

Only 12.41% of FERT women resulted in a clinical pregnancy, versus 56.56% of ART women, indicating a much higher success rate of ART therapies.

While theoretically the complete pregnancy sample is the sum of the number of deliveries and failures, they don't add up to 100% in our sample. This is because there is no good information in our dataset that points out deliveries resulting in still birth. Thus, there could be some overlap in the numbers of deliveries and failures if some still births get coded as both a completed pregnancy and a failed pregnancy. It is also possible that a woman delivered twins and one of her child was a still born resulting in her case showing up as both a delivery and a failure. A third possibility is that a woman had two different deliveries during a single calendar year, with different reported outcomes. We do not have enough information in our sample to separate out these effects. The percent of all women with completed pregnancies in our sample of 39.26% compares relatively closely to the 33.7% pregnancy rate in the CDC sample.

The rate of multiple births for ART women in our sample is unrealistically low, only 1.28% compared to the CDC reported rate of 11.30% for ART women. This undercounting is because we determined multiple births for this paper using only diagnoses coded for the mother, and these diagnoses under record multiple births for reasons that we do not understand. When we look at Baby outcomes, we include each multiple birth child as a separate record. We did not attempt to count numbers of births by using counts of babies per mother in our data in part because we cannot easily distinguish mothers with twins and other multiple births using our data from a mother with two separate deliveries during a single calendar year. We do not attempt to do any detailed analysis using multiple birth information in our sample.

Table 4 describes the sub-sample of women who had a completed pregnancy in any of the years that they appear in the entire sample and the pregnancy resulted in either a live birth or a still birth, ectopic pregnancy, miscarriage or abortion. The average age of women having a completed pregnancy is much lower at 31.74 years and the average risk score is much higher at 2.99. In our sample, 83% of pregnancies are completed with birth of a child while 19% pregnancies result in ectopic, miscarriage or abortion. About 11.4% of these women had a major complication during pregnancy. We find that the number of FERT pregnancies far exceeds the number of ART pregnancies. This indicates that it is easier for couples to go for these procedures as they are less expensive and invasive. We also notice that failure rate and pregnancy complications following ART are higher than those following a FERT therapy. Comparing the ART women in our sample to CDC results, we see that there is a higher rate of failure in our sample compared to CDC results.

Figures 1-26 describe the data a bit more. Figure 1 gives age distribution of pregnancies by fertility status. We see that women who get pregnant in our non ART/FERT cohort are much younger than women who become pregnant using ART or FERT. The number of women using ART peaks at ages 33-37, versus ages 29-33 for FERT women and 27-32 for non ART/FERT women. Figure 2 gives the average health risk score of these women by age and fertility status. We see that women who have ART or FERT have a much higher risk score compared to women who conceive spontaneously and there is very small age gradient to these risk scores. There is a very slight age-based pattern of risk scores among non ART/FERT women, which show a 0.3 rise in risk score for ages 21 to 30, followed by a slight decline, with a steady rise starting at age 37. ART and FERT women show a very similar pattern over age, with a nearly constant increment over the non ART/FERT women. Further, since the risk score does not include pregnancy related complications, this strengthens the hypothesis that women who conceive non-spontaneously come from a different health risk pool than women who are able to conceive spontaneously. Figure 3 shows that within each age category, success rates are higher among women who go for treatments as opposed to women who conceive spontaneously. Part of this result is driven by the fact that not all non ART/FERT women want to get pregnant. Much greater success at ages 40 and above for ART women is particularly notable. Success rates remain close to 40% for ART women even at age 50.

We now look at each of our outcomes by age of the mother. Figure 4 tells us that within each age category, the mean complications rate is higher for ART women, and complications rate goes up with age. From figure 5 we see that for non ART/FERT women, the failure rate goes down with age, bottoms out at around age 29, then goes up from ages 29 to 42 and then goes down again. Also, generally the risk of failure for ART women is higher than non ART/FERT women but lower than FERT women. We will see later that our regression results also tell a similar story. This is plausible—FERT procedures such as IUI are performed in a lesser controlled environment and hence have higher risks of multiple gestation and therefore failure. On the other hand in ART procedures a controlled number of embryos are transferred back in the mother's womb and thus chances of multiple gestations are reduced. Figure 6 gives us the mean risk score of the baby in each of the mother's age category. We see that the baby's risk score is pretty flat across age for mothers who conceive spontaneously. Even for women who had ART and FERT the age gradient is very small but these women generally produce slightly less healthy babies.

In figures 7, 8 and 9 we look at our outcomes by the mother's risk score. First of all, it is important to note that the mother's risk score has a lot of outliers. To some extent this problem was accounted for by grouping all women with a risk score of 15 or more into one category. Still, knowing that the average risk score of a woman in these samples are in the range of 3-4, the first half of the figures are more relevant to our discussion than the rest. We see from figure 7 that up until risk score category 3-4 at least, being an ART or FERT woman does not necessarily mean that the complications rate is higher. In fact women who have spontaneous pregnancies have a higher mean complications rate than women who don't. But for more unhealthy women, complications rate is higher if they have treatments to become pregnant. Thus we see that up to a certain health status, having a fertility treatment really does not worsen your chances of complications. It is only the very unhealthy mothers that are likely to have complications after going for fertility treatments. Figure 8 gives similar results to figure 6. Apart from the rather puzzling reduction in failure rate for healthier mothers, for women in the risk score range of 2-9 chances of failure following an ART treatment are higher than a spontaneous pregnancy but lower than a FERT treatment. This suggests that the controlled treatment environment alone does not improve outcome following ART as compared to FERT. Figure 9 shows that more unhealthy mothers will in general have more unhealthy babies, with mothers that had treatments doing worse than mothers that don't.

Figures 10-12 discuss the age profile and treatment success for states that have a mandate vs. those that don't. Figure 10 shows that ART are a very small proportion of all completed pregnancies, with this proportion rising for women in their late thirties. From figure 11 we see utilization of fertility treatments peak at a much older age in mandate states versus non-mandate states. Most women going for treatments in mandate states are older than 35 while in non-mandate states they are around 31-32. This suggests that women that reside in states that have a mandate may wait longer before going for fertility treatments. In figure 12 we see that success rates in mandate vs. non-mandate states are pretty much the same. Thus, presence or absence of a mandate per say does not affect the success rates, although women tend to wait longer before getting these treatments in a mandate state.

Figures 13-26 describe the data at the state level (where we call DC a state). We first plot the mean ART rates against mean FERT rates at state level. We also plot each of our three outcome variables against both mean ART and mean FERT, where each state contributes a single observation. Each plot is done first using raw means for each variable and then using residuals, where residuals are created from a regression of the variable of interest on age, mother's risk score and year dummies.

Figure 13 reveals that ART and FERT are positively but not perfectly associated, which is as expected given the heterogeneity in state mandates and practice styles. These figures also reveal that DC and Hawaii are huge outliers in their rates of ART and FERT in our sample. Since there are very few observations in either of these states, it is unlikely that they will affect our results in our individual level regressions, but these outliers drive the results at the state level when sample sizes for each state are ignored. To correct for this, we add two trend lines in each graph—one while including all states (including DC and Hawaii) and one while excluding DC and Hawaii. We see that the R-square goes down in all figures after we exclude these two states. When we look at our outcome variables, we find that major complications do not seem to be associated with ARTs or FERTs once we remove DC and Hawaii from the analysis. The same is true for the baby's risk score, although in some cases a weak negative association is found between treatment type and baby's risk score. Only failures seem positively associated with ARTs and FERTs.

6 Methodologies

We use two-stage least squares model to estimate the likelihood of complications in pregnancy, failure in achieving a completed pregnancy due to miscarriages or ectopic pregnancy and health risk score of a child if an ART/FERT is used, while controlling for the mother's age, her health risk score before pregnancy and the year she appeared in our sample. The following models were estimated:

$$I(Complications = 1) = \alpha + \sum_{i=1}^{29} \beta_i Age_i + \sum_{t=2001}^{2004} \delta_t Year_t + \gamma(Mother's RiskScore) + \lambda(ART = 1) + \theta(FERT = 1) + \varepsilon$$

$$I(Faiure = 1) = \alpha + \sum_{i=1}^{29} \beta_i Age_i + \sum_{t=2001}^{2004} \delta_t Year_t + \gamma(Mother's RiskScore) + \lambda(ART = 1) + \theta(FERT = 1) + \varepsilon$$

$$Babies' RiskScore = \alpha + \sum_{i=1}^{29} \beta_i Age_i + \sum_{t=2001}^{2004} \delta_t Year_t + \gamma(Mother's RiskScore) + \lambda(ART = 1) + \theta(FERT = 1) + \varepsilon$$

In each of the above models ART and FERT were instrumented by two sets of instruments. In the first case, we instrumented them using 14 dummies, one for each state to allow for full variation in the states that have mandates. We then run it on a tighter set of 8 dummies—one for each state that has a mandate that is likely to affect ART and FERT utilization. These include Massachusetts, Rhode Island, New York, New Jersey, Maryland, Illinois, Texas and California.

7 Results

7.1 Mandates are indeed good instruments

In this section we discuss the first stage regression results from our two stage least squares model. Tables 5 and 6 give OLS regression results of ART and FERT on two different types of instruments. We first regress ART on a single dummy called ‘mandate’ which is an indicator variable for any state that has an infertility insurance mandate. The age and year dummies and mother’s risk score are included in the regression but not shown in the table. In the next regression we include individual state dummies for all the fourteen states that have a mandate in our sample. Both these regressions are run for the entire sample of all women aged 21-54 who appear for two consecutive years, all women who appear for two consecutive years and had a completed pregnancy and all such women with a completed pregnancy for which we had information on their baby. We then repeat this exercise for the FERT Variable.

The coefficient on mandate dummy turns out to be positive and highly significant in all regression model and samples. The coefficient is small but that’s only because in our entire sample of completed pregnancy only 1.3% women actually had ART in the previous or current year. For example, at the mean age of 35 and year 2003, the predicted proportion of women in completed pregnancy sample going for ART is 0.011 in a non-mandated state. This proportion goes up to 0.016—a whopping 45.45% increase—if the state has a mandate. Thus presence of a mandate significantly affects ART utilization rates.

In regressions where fourteen dummies are included, we see that in states where a comprehensive mandate is present in general has a positive effect on ART utilization rate. For example, for both ART and FERT the coefficient on Massachusetts—the state with the most generous and comprehensive mandates on infertility insurance coverage-- is the highest. New York, New Jersey, and Illinois also are amongst the higher coefficients. States where several restrictions are placed on coverage have a lesser significant coefficient. Also, as we suspected in Texas and California, where there is a mandate to 'offer,' only has positive effects on FERT but not ART.

7.2 Complications in Pregnancy

We now discuss the effect of ART on the rate of complications that the mother faces during pregnancy. We estimated the benchmark OLS and Two-stage least squares model to see if after controlling for the age, year and mother's risk score does complication rate get affected by the presence of ART and FERT or not. Column 1 in tables 7 and 8 give results for two different specifications of the instruments for ART and FERT association with complications during pregnancy. We see that in the benchmark OLS model the coefficients on ART and FERT are positive, less than one and statistically significant. This means that the complications rate is higher in women who go for ART or FERT than those who conceive spontaneously. However in all the two stage models the coefficient on ART/FERT is small and statistically insignificant.

The coefficient on mother's risk score is positive and highly significant. This means that once we treat the endogeneity problems in ART and FERT variables, the mother's risk score explains most of the variation in complications and the fact that a patient went for ART or FERT does not seem to affect her chances of complications in a statistically significant way. We also see that there is some growth in the complications rate over time. Even though the coefficients are small, the effect is statistically significant in all models. We believe this is plausible because as technology becomes more advanced the tolerance for complications goes down as doctors strive to reduce discomfort during childbirth. Thus more and more complications that were considered minor previously may become major overtime. This is also true in our sample statistics where only 11% women had major complications but more than 85% had any type of complications. Moving over to the age profile in Tables 7 and 8 and column 1, we see that the predicted likelihood of complications during pregnancy is fairly stable for ages 21-38 and then declines thereafter. This is plausible since older women would go through these treatments only if they are in very good overall health and hence have fewer risks of complications. The results also get affected by the fact that very few women after 40 go for these treatments and thus there may be some outliers in our data. Including eight separate instruments for each state does not particularly change results compared to if we had 14 instruments. We suspect that most of the variation in state dummies is coming from these 8 states and the other states are pretty ineffective at improving ART and FERT utilization rates.

7.3 Failed Pregnancies: Miscarriages, Ectopic Pregnancies and Abortions

We also estimated the benchmark OLS and two-stage least squares model to see if ART and FERT are associated with failure in pregnancy after controlling for the age, year and mother's risk score. Column 2 in tables 7 and 8 give results for two different specifications of the mandate dummy for ART and FERT association with failure rate in pregnancy. We see that in the benchmark OLS model the coefficients on ART and FERT are positive, less than one and statistically significant. This means that the failure rate is higher in women who go for ART or FERT than those who conceive spontaneously. In the TSLS models the coefficient on ART is positive but statistically insignificant and that on FERT is positive but statistically significant. These results suggest that FERT treatment is associated with failure in pregnancy but ART treatment is not. This is plausible since the number of eggs created and fertilized is less controlled with FERT treatments than with ART, and thus FERT pregnancies plausibly have a higher likelihood of multiple gestations and fetal risks. A closer look at the coefficients though makes us suspicious about our model since these coefficients are unrealistically high. The coefficient of 0.939 on the FERT variable tells us that a woman going for FERT treatment will have a failure with almost certainty. This high coefficient is implausible and it is unclear exactly why it becomes so large. One possible explanation is that women who cannot get pregnant without ART/FERT treatments are fundamentally different from non ART/FERT women, so that failure rates are not comparable. If you cannot get pregnant at all unless you get FERT treatment then how can u compare the complications rate with the complications rate of women who can get pregnant without these treatments? Another possible explanation is that the state dummies also belong directly in the model, so that when entered only as instruments for ART and FERT, they capture the extra impact of the state dummies be exaggerating the ART and FERT coefficients.

The coefficient on mother's risk score in the failure model is negative and highly significant. This is true across all model specifications and estimation strategies and is quite counterintuitive. This means that more unhealthy mothers are more likely to carry a clinical pregnancy to full term and that their delivery results in a live birth. One possible explanation for this is that healthy mothers (with low risk scores) are more likely to have failures due to abortions than high risk score mothers. This shows up as a failure in our model, which does not distinguish among various possible reasons for failure. Another possible explanation is that unhealthy mothers (with high risk scores) are put straight on more complex and invasive procedures and they possibly bypass having to go through lesser intense procedures which are not monitored as frequently through ultrasounds and hence have more chances of failures. We suspect something is going on in the health profile of these mothers that's giving us such strong results that are unanimous in all our models.

We also see that there is decline in the failure rate over time. Even though the coefficients are small, the effect is statistically significant in all models. We believe this is plausible because as technology becomes more advanced doctors would strive to reduce the number of miscarriages through better monitoring and clinical evaluation strategies. Moving over to the age profile, we see that the predicted likelihood of failed pregnancies goes down slightly from the ages 21 and 30, then steadily increases from 30 to 43 and then declines again. Again this pattern for older women is plausible since older women would go through these treatments only if they are in very good overall health and hence have fewer risks of complications to begin with. It also seems reasonable for failure rate to go up in ages 30 to 43 since most women in the reproductive age group would start with reproductive therapies in their 30s and medical literature also suggests the concept of "biological clock" of a woman that suggests that women become less fecund once they hit 30. It is somewhat unclear why the failure rate would go down for younger women. However, the absolute decrease is less than .02 percentage points.

7.4 Baby's Health Risk Score:

Finally we estimated a benchmark OLS and TSLS model to estimate the ART and FERT association with the baby's risk score as calculated by the DxCG software. We see that in the benchmark OLS model both ART/FERT and mother's risk score significantly and positively affect the baby's risk-score. However, once we instrument the ART and FERT variables we find that just as in the complications model, most variation in the babies' risk score is explained by the mother's health status before she got pregnant and the fact that the mother had a treatment affects the baby's score in a statistically insignificant way. However, the small Z scores compared with huge coefficient estimates indicate that the coefficients are very imprecisely estimated. The time profile tells us that as time passes more healthy babies are being born which is fairly consistent with our hypothesis of technology advances over time. Moving on to the age profile of the mother we see that the predicted risk score of the baby is fairly stable for age groups 21-40 and then becomes erratic, perhaps due to small samples sizes above that age. Using the age coefficients from the OLS model, there is no obvious relationship between the baby's health status and the mother's age, although once the TSLS coefficients are used, there appears to be an increase in the health risk score of the babies for older mothers.

7.5 Discussion

Our OLS and TSLS results are somewhat puzzling. The primary concern about OLS is that even after including mother's age and risk score there may be unobserved characteristics of ART and FERT women so that OLS is picking up not only the consequences of ART and FERT, but also selection differences due to this unobserved variation. For example women living in a mandate state may have a taste preference for childbearing at a later age in life. If mandate states are uncorrelated with these unobserved errors then they are valid instruments for this selection problem. We have shown that presence of a mandate in a state is associated with higher ART and FERT rates compared to states that do not have mandates. However, if mandates themselves are associated with ART and FERT (such as higher intensity of treatment in a mandate state versus a non-mandate state) then they will no longer be valid instruments and the TSLS results will be biased.

8 Conclusions

We have used medical claims data to analyze the effects of advanced reproductive therapies such as fertility drugs, Intra-Uterine Insemination, and In-Vitro Fertilization on three outcome variables of considerable interest: major complications during pregnancy, pregnancy failures (miscarriages, abortions, and ectopic pregnancies), and expected health spending on the baby, a metric of the newborn's health status. In contrast with the existing literature which has been inconclusive about whether ART and FERT techniques worsen outcomes for mothers and child because they do not control for the mother's health status and unobservable selection variables, we both control for the mothers health status explicitly and use instrumental variable techniques to correct the endogeneity problem with respect to the decision to choose ART and FERT. We use as instruments dummy variables for states that have a mandate regarding insurance coverage of infertility treatments. Our first stage results show state mandate dummies are jointly very significant, although states seem to vary in the impact of their mandates.

While OLS finds significant effects of ART and FERT treatments on our three outcome variables, the TSLS results find much weaker effects. Most outcomes are more strongly associated with the mother's age and health condition. ART and FERT have no significant effect in the major complications, and the effect on babies' risk scores is also insignificant. Only the probability of pregnancy failures appears to be statistically significantly associated with ART and FERT.

Table and Figures

Table 1 Mandate Variations by State

State	Year	Cover/ Offer	Includes IVF	Max IVF?	Age Cap?	Small Employees Exempt?	Dollar Cap	Marriage Required	Facility Restriction	Cyropreservation allowed?	Male Infertility Covered?	HMOs Exempt
Massachusetts	1987	Cover	Yes	No	No	No	No	No	No	Yes	Yes	No
Rhode Island	1989, 2006	Cover	Yes	No	25-40	No	100000	Yes	No	DK	DK	No
New Jersey	2001	Cover	Yes	No	46	<50	No	No	Yes	No	Yes	No
New York	1990, 2002	Cover	No	0	21-44	No	No	No	No	DK	DK	No
Illinois	1991, 1997	Cover	Yes	6	No	<25	No	No	Yes	Dk	Yes	No
Maryland	2000	Cover	Yes	min 3	No	<50	100000	Yes	Yes	DK	Yes	No
Texas	1987	**Offer**	only IVF	No	No	No	No	Yes	Yes	DK	Yes	No
California	1989	**Offer**	No	No	No	No	No	No	No	DK	DK	No
Connecticut	2005	Cover	Yes	2	40	No	No	No	Yes	DK	DK	No
Hawaii	1989, 2003	Cover	Yes	1	No	No	No	Yes	Yes	DK	Yes	No
Arkansas	1987, 1991	Cover	Yes	No	No	No	15000	Yes	Yes	Yes	Yes	Yes
Louisiana	2001	Cover	No	DK	DK	DK	DK	DK	DK	DK	DK	DK
Ohio	1991	Cover	not req	No	No	No	No	No	No	DK	Yes	No
Montana	1987	Cover	DK	DK	DK	DK	DK	DK	DK	DK	DK	No
West Virginia	1995	Cover	DK	DK	DK	DK	DK	DK	DK	DK	DK	No

Table 2: Descriptive Statistics, All Women in Two-Year Sample

All Women:	Means	
Average Age	41.47	
Average Risk Score	1.296	
	N	Percent
Patients	3617579	100.00%
Pregnancies	207326	5.73%
Delivery	122509	3.39%
Failure	28915	0.80%
Major Complications	16805	0.46%
Multiple Birth	1046	0.03%
No ART/FERT Women	3462418	95.71%
FERT Women	150306	4.15%
ART Women	4855	0.13%

Table 3: Descriptive Statistics, All Completed Pregnancies, Two-Year Sample

Infertility Claims	ART		FERT		No ART/FERT Women		2004 CDC Results
	Means		Means		Means		N=Non-donor Cycles & not patients
Average Age	35.66		42.12		41.46		36
Average Risk Score	2.498		1.910		1.227		NA
	N	Percent	N	Percent	N	Percent	Percent
Patients	4855	100%	150306	100%	3462418	100%	94,242=100%
Clinical Pregnancy	2746	56.56%	18660	12.41%	185919	5.37%	
Completed Pregnancy	1906	39.26%	12815	8.53%	132553	3.83%	31,758=33.70%
Delivery	1481	30.50%	9491	6.31%	111537	3.22%	26,059=27.65%
Failure	528	10.88%	3934	2.62%	145767	4.21%	5699=6.05%
Major Complications	406	8.36%	1821	1.21%	24453	0.71%	NA
Multiple Birth	62	1.28%	299	0.20%	6845	0.20%	3589=3.81%

Table 4: Descriptive Statistics, All Women in Two-Year Sample, By Fertility Status

Completed Pregnancy Sample	All women in Completed Pregnancy sample		All ART women in Completed Pregnancy sample		All FERT women in Completed Pregnancy sample		2004 CDC Results
Average Age	31.744		35.268		33.168		36
Average Risk Score	2.987		3.145		2.684		
	N	Percent	N	Percent	N	Percent	
Completed Pregnancies	147274	100.00%	1906	100.00%	12815	100.00%	100.00%
Delivery	122509	83.18%	1481	77.70%	9491	74.06%	82.10%
Failure	28915	19.63%	528	27.70%	3934	30.70%	17.00%
Major Complications	16805	11.41%	406	21.30%	1821	14.21%	
Multiple Birth	1046	0.71%	62	3.25%	299	2.33%	26.70%

Table 5: OLS regression of ART on selected instruments in different datasets *

	All Women Aged 21-54		All Completed Pregnancies		All Mothers with Babies	
	Coefficients	T Ratio	Coefficients	T Ratio	Coefficients	T Ratio
mandate	0.00093	23.42	0.0052	8.61	0.0014	1.86
Massachusettes	0.00224	24.25	0.0148	10.66	0.0063	3.38
Rhode Island	-0.00020	-0.32	-0.0043	-0.47	0.0015	0.13
New Jersey	0.00503	41.59	0.0301	19.18	0.0165	7.83
New York	0.00166	12.37	0.0073	3.88	0.0054	2.26
Illinois	0.00016	1.56	-0.0031	-2.18	-0.0038	-2.12
Ohio	-0.00007	-0.62	-0.0009	-0.48	-0.0014	-0.61
Maryland	0.00152	7.76	0.0105	3.46	-0.0007	-0.18
West Virginia	0.00011	0.36	-0.0010	-0.18	-0.0020	-0.28
Arkansas	0.00036	1.41	0.0022	0.53	-0.0017	-0.34
Texas	0.00103	11.26	0.0103	7.68	0.0082	4.92
California	-0.00010	-1.72	-0.0039	-4.19	-0.0047	-4.01
Hawaii	-0.00030	-0.13	-0.0058	-0.12	-0.0046	-0.10
Louisiana	-0.00009	-0.40	-0.0044	-1.14	-0.0031	-0.63
Montana	0.00015	0.24	-0.0025	-0.26	0.0067	0.54
Observations	3617579		147274		75263	
Dep Var Mean	0.13		1.29		1.11	

*(Included but not shown are Age, Year dummies and risk score--separate regressions with mandate and state dummies)

Table 6: OLS regression of FERT on selected instruments in different datasets *

	All Women Aged 21-54		All Completed Pregnancies		All Mothers with Babies	
	Coefficients	T Ratio	Coefficients	T Ratio	Coefficients	T Ratio
mandate	0.00839	39.02	0.01373	9.13	0.00873	4.50
Massachusettes	0.02839	56.96	0.04795	13.91	0.04134	8.94
Rhode Island	0.01184	3.59	0.05360	2.36	0.03368	1.18
New Jersey	0.00376	5.76	0.02704	6.93	0.02617	4.97
New York	0.00461	6.36	0.01525	3.27	0.01659	2.77
Illinois	-0.00309	-5.60	0.01598	4.59	0.00747	1.69
Ohio	-0.00525	-9.18	0.00168	0.37	-0.00454	-0.79
Maryland	0.00966	9.14	0.01575	2.08	0.03039	3.11
West Virginia	-0.00950	-5.56	-0.00224	-0.16	0.00162	0.09
Arkansas	-0.00382	-2.80	-0.02469	-2.45	-0.02093	-1.69
Texas	0.01234	24.90	0.00786	2.35	0.00446	1.07
California	0.00929	28.48	0.00172	0.74	-0.00391	-1.35
Hawaii	0.02013	1.61	-0.07009	-0.56	-0.05170	-0.45
Louisiana	0.00074	0.61	-0.00104	-0.11	0.01199	0.97
Montana	-0.01009	-2.93	0.02316	0.98	-0.00272	-0.09
Observations	3617579		147274		75263	
Dep Var Mean	4.15		8.7		7.45	

*(Included but not shown are Age, Year dummies and risk score--separate regressions with mandate and state dummies)

Table 7: OLS and TSLS results ART/FERT Association with Various Dependent Variables, 14 State Dummies as Instruments, 2 year sample, state cluster

	Major Complications				*Failure in Pregnancy*				*Babies' Risk Score*			
	OLS		*TSLS*		*OLS*		*TSLS*		*OLS*		*TSLS*	
Variable	*Coeff*	*Z*	*Coeff*	*Z*	*Coeff*	*Z*	*Coeff*	*Z*	*Coeff*	*Z*	*Coeff*	*Z*
ART	*0.061*	*4.97*	*0.076*	*0.25*	*0.097*	*10.26*	*0.652*	*1.73*	*3.324*	*5.47*	*-15.220*	*-0.47*
FERT	*0.009*	*2.02*	*0.018*	*0.13*	*0.131*	*22.39*	*0.939*	*3.60*	*0.940*	*5.42*	*4.548*	*0.67*
Risk Score	*0.042*	*17.11*	*0.041*	*12.42*	*-0.040*	*-12.98*	*-0.058*	*-8.70*	*0.972*	*16.12*	*1.004*	*4.87*
2002	*0.010*	*2.59*	*0.010*	*2.47*	*-0.001*	*-0.11*	*-0.008*	*-1.17*	*0.422*	*3.06*	*0.481*	*2.78*
2003	*0.014*	*5.41*	*0.014*	*4.92*	*-0.006*	*-0.98*	*-0.013*	*-1.53*	*0.211*	*1.36*	*0.193*	*0.90*
2004	*0.017*	*5.67*	*0.017*	*3.72*	*-0.008*	*-1.13*	*-0.006*	*-0.87*	*0.168*	*1.57*	*0.267*	*1.07*
age22	*0.033*	*1.92*	*0.032*	*1.92*	*-0.082*	*-2.59*	*-0.102*	*-2.55*	*-0.383*	*-0.18*	*-0.523*	*-0.24*
age23	*0.042*	*2.66*	*0.042*	*2.58*	*-0.097*	*-2.87*	*-0.118*	*-2.82*	*-0.682*	*-0.33*	*-0.768*	*-0.37*
age24	*0.036*	*1.96*	*0.035*	*1.92*	*-0.111*	*-3.36*	*-0.139*	*-3.21*	*-1.087*	*-0.53*	*-1.183*	*-0.57*
age25	*0.037*	*2.09*	*0.037*	*1.98*	*-0.116*	*-3.16*	*-0.150*	*-3.00*	*-0.917*	*-0.45*	*-1.055*	*-0.51*
age26	*0.033*	*1.84*	*0.033*	*1.75*	*-0.125*	*-3.48*	*-0.168*	*-3.31*	*-0.915*	*-0.45*	*-1.050*	*-0.51*
age27	*0.041*	*2.29*	*0.041*	*2.14*	*-0.129*	*-3.34*	*-0.175*	*-3.22*	*-1.013*	*-0.50*	*-1.181*	*-0.57*
age28	*0.042*	*2.26*	*0.042*	*2.12*	*-0.129*	*-3.41*	*-0.179*	*-3.36*	*-1.088*	*-0.54*	*-1.274*	*-0.61*
age29	*0.045*	*2.64*	*0.044*	*2.33*	*-0.125*	*-3.31*	*-0.184*	*-3.33*	*-0.765*	*-0.37*	*-0.905*	*-0.44*
age30	*0.044*	*2.35*	*0.043*	*2.23*	*-0.130*	*-3.47*	*-0.185*	*-3.54*	*-0.779*	*-0.38*	*-0.839*	*-0.41*
age31	*0.044*	*2.23*	*0.043*	*2.17*	*-0.116*	*-2.90*	*-0.175*	*-3.12*	*-1.108*	*-0.54*	*-1.179*	*-0.57*
age32	*0.044*	*2.37*	*0.043*	*2.19*	*-0.124*	*-3.34*	*-0.188*	*-3.44*	*-1.217*	*-0.60*	*-1.266*	*-0.62*
age33	*0.046*	*2.48*	*0.046*	*2.27*	*-0.107*	*-2.81*	*-0.179*	*-3.15*	*-1.057*	*-0.52*	*-1.103*	*-0.54*
age34	*0.043*	*2.28*	*0.042*	*2.12*	*-0.097*	*-2.50*	*-0.171*	*-3.06*	*-1.058*	*-0.52*	*-1.068*	*-0.51*
age35	*0.037*	*1.73*	*0.036*	*1.62*	*-0.085*	*-2.14*	*-0.168*	*-2.80*	*-1.274*	*-0.62*	*-1.277*	*-0.62*
age36	*0.041*	*2.17*	*0.040*	*1.96*	*-0.067*	*-1.80*	*-0.151*	*-2.65*	*-0.921*	*-0.45*	*-0.941*	*-0.45*
age37	*0.038*	*1.92*	*0.036*	*1.71*	*-0.041*	*-1.08*	*-0.142*	*-2.40*	*-0.762*	*-0.37*	*-0.641*	*-0.31*
age38	*0.032*	*1.51*	*0.031*	*1.42*	*-0.023*	*-0.58*	*-0.126*	*-1.95*	*-1.033*	*-0.51*	*-0.950*	*-0.45*
age39	*0.031*	*1.42*	*0.029*	*1.27*	*0.020*	*0.56*	*-0.094*	*-1.61*	*-0.759*	*-0.38*	*-0.522*	*-0.25*
age40	*0.020*	*0.98*	*0.018*	*0.79*	*0.052*	*1.33*	*-0.071*	*-1.02*	*-1.593*	*-0.78*	*-1.417*	*-0.67*
age41	*0.003*	*0.17*	*0.001*	*0.06*	*0.079*	*2.00*	*-0.068*	*-0.96*	*-0.108*	*-0.05*	*-0.079*	*-0.03*
age42	*0.006*	*0.32*	*0.004*	*0.20*	*0.099*	*2.63*	*-0.032*	*-0.48*	*-1.207*	*-0.58*	*-0.788*	*-0.33*
age43	*-0.010*	*-0.40*	*-0.012*	*-0.45*	*0.120*	*3.03*	*0.002*	*0.04*	*-0.756*	*-0.35*	*0.377*	*0.12*
age44	*-0.027*	*-1.35*	*-0.029*	*-1.19*	*0.067*	*2.17*	*-0.035*	*-0.58*	*-0.590*	*-0.28*	*-0.086*	*-0.03*
age45	*-0.048*	*-1.86*	*-0.050*	*-1.66*	*0.065*	*1.55*	*-0.043*	*-0.77*	*-0.215*	*-0.08*	*2.220*	*0.38*
age46	*-0.081*	*-3.70*	*-0.082*	*-3.47*	*-0.004*	*-0.11*	*-0.082*	*-1.73*	*6.046*	*1.53*	*6.310*	*1.61*
age47	*-0.087*	*-4.30*	*-0.088*	*-4.16*	*-0.031*	*-0.88*	*-0.086*	*-1.69*	*-1.794*	*-0.74*	*-0.867*	*-0.25*
age48	*-0.137*	*-5.66*	*-0.138*	*-5.70*	*-0.090*	*-2.22*	*-0.127*	*-2.32*	*2.443*	*0.53*	*4.595*	*0.60*
age49	*-0.112*	*-5.81*	*-0.112*	*-5.41*	*-0.101*	*-2.61*	*-0.154*	*-2.76*	*-4.841*	*-1.60*	*-1.462*	*-0.21*
age50+	*-0.155*	*-8.14*	*-0.154*	*-7.71*	*-0.070*	*-1.99*	*-0.073*	*-1.83*	*5.940*	*0.88*	*7.384*	*1.39*
Constant	*-0.025*	*-1.32*	*-0.024*	*-1.30*	*0.366*	*8.68*	*0.395*	*7.88*	*1.898*	*0.92*	*1.769*	*0.85*
OBS	*147274*		*147274*		*147274*		*147274*		*75263*		*75263*	
RMSE	*0.3107*		*0.3107*		*0.3875*		*0.451*		*8.961*		*9.223*	
R Square	*0.0455*				*0.0484*				*0.023*			
Dep. Var. Mean	*0.114*		*0.114*		*0.196*		*0.196*		*3.323*		*3.323*	

Table 8: OLS and TSLS results ART/FERT Association with Various Dependent Variables, 8 State Dummies as Instruments for Mandates, 2 year sample, state cluster

	Major Complications				*Failure in Pregnancy*				*Babies' Risk Score*			
	OLS		*TSLS*		*OLS*		*TSLS*		*OLS*		*TSLS*	
Variable	*Coeff*	*Z*	*Coeff*	*Z*	*Coeff*	*Z*	*Coeff*	*Z*	*Coeff*	*Z*	*Coeff*	*Z*
ART	*0.061*	*4.97*	*0.080*	*0.26*	*0.097*	*10.26*	*0.661*	*1.69*	*3.324*	*5.47*	*-13.554*	*-0.41*
FERT	*0.009*	*2.02*	*0.028*	*0.18*	*0.131*	*22.39*	*0.930*	*3.36*	*0.940*	*5.42*	*3.298*	*0.44*
Risk Score	*0.042*	*17.11*	*0.041*	*12.11*	*-0.040*	*-12.98*	*-0.057*	*-8.47*	*0.972*	*16.12*	*1.028*	*4.84*
2002	*0.010*	*2.59*	*0.010*	*2.43*	*-0.001*	*-0.11*	*-0.008*	*-1.17*	*0.422*	*3.06*	*0.483*	*2.75*
2003	*0.014*	*5.41*	*0.014*	*4.76*	*-0.006*	*-0.98*	*-0.013*	*-1.52*	*0.211*	*1.36*	*0.211*	*0.95*
2004	*0.017*	*5.67*	*0.017*	*3.65*	*-0.008*	*-1.13*	*-0.006*	*-0.89*	*0.168*	*1.57*	*0.267*	*1.07*
age22	*0.033*	*1.92*	*0.032*	*1.92*	*-0.082*	*-2.59*	*-0.102*	*-2.53*	*-0.383*	*-0.18*	*-0.486*	*-0.23*
age23	*0.042*	*2.66*	*0.042*	*2.56*	*-0.097*	*-2.87*	*-0.118*	*-2.81*	*-0.682*	*-0.33*	*-0.747*	*-0.36*
age24	*0.036*	*1.96*	*0.035*	*1.89*	*-0.111*	*-3.36*	*-0.139*	*-3.18*	*-1.087*	*-0.53*	*-1.153*	*-0.56*
age25	*0.037*	*2.09*	*0.036*	*1.95*	*-0.116*	*-3.16*	*-0.150*	*-2.96*	*-0.917*	*-0.45*	*-1.012*	*-0.49*
age26	*0.033*	*1.84*	*0.032*	*1.72*	*-0.125*	*-3.48*	*-0.167*	*-3.27*	*-0.915*	*-0.45*	*-0.999*	*-0.48*
age27	*0.041*	*2.29*	*0.040*	*2.09*	*-0.129*	*-3.34*	*-0.174*	*-3.19*	*-1.013*	*-0.5*	*-1.125*	*-0.54*
age28	*0.042*	*2.26*	*0.041*	*2.06*	*-0.129*	*-3.41*	*-0.178*	*-3.31*	*-1.088*	*-0.54*	*-1.213*	*-0.58*
age29	*0.045*	*2.64*	*0.044*	*2.25*	*-0.125*	*-3.31*	*-0.184*	*-3.27*	*-0.765*	*-0.37*	*-0.842*	*-0.41*
age30	*0.044*	*2.35*	*0.043*	*2.17*	*-0.130*	*-3.47*	*-0.185*	*-3.50*	*-0.779*	*-0.38*	*-0.793*	*-0.39*
age31	*0.044*	*2.23*	*0.043*	*2.12*	*-0.116*	*-2.9*	*-0.174*	*-3.08*	*-1.108*	*-0.54*	*-1.125*	*-0.54*
age32	*0.044*	*2.37*	*0.043*	*2.10*	*-0.124*	*-3.34*	*-0.188*	*-3.39*	*-1.217*	*-0.6*	*-1.213*	*-0.59*
age33	*0.046*	*2.48*	*0.045*	*2.19*	*-0.107*	*-2.81*	*-0.178*	*-3.10*	*-1.057*	*-0.52*	*-1.043*	*-0.51*
age34	*0.043*	*2.28*	*0.041*	*2.04*	*-0.097*	*-2.5*	*-0.171*	*-3.02*	*-1.058*	*-0.52*	*-1.012*	*-0.49*
age35	*0.037*	*1.73*	*0.035*	*1.56*	*-0.085*	*-2.14*	*-0.167*	*-2.76*	*-1.274*	*-0.62*	*-1.214*	*-0.59*
age36	*0.041*	*2.17*	*0.039*	*1.87*	*-0.067*	*-1.8*	*-0.150*	*-2.60*	*-0.921*	*-0.45*	*-0.873*	*-0.42*
age37	*0.038*	*1.92*	*0.035*	*1.61*	*-0.041*	*-1.08*	*-0.142*	*-2.35*	*-0.762*	*-0.37*	*-0.577*	*-0.28*
age38	*0.032*	*1.51*	*0.030*	*1.35*	*-0.023*	*-0.58*	*-0.125*	*-1.91*	*-1.033*	*-0.51*	*-0.873*	*-0.42*
age39	*0.031*	*1.42*	*0.028*	*1.20*	*0.020*	*0.56*	*-0.093*	*-1.56*	*-0.759*	*-0.38*	*-0.462*	*-0.22*
age40	*0.020*	*0.98*	*0.016*	*0.71*	*0.052*	*1.33*	*-0.071*	*-0.99*	*-1.593*	*-0.78*	*-1.328*	*-0.62*
age41	*0.003*	*0.17*	*0.000*	*-0.01*	*0.079*	*2*	*-0.067*	*-0.93*	*-0.108*	*-0.05*	*0.044*	*0.02*
age42	*0.006*	*0.32*	*0.003*	*0.13*	*0.099*	*2.63*	*-0.031*	*-0.46*	*-1.207*	*-0.58*	*-0.739*	*-0.31*
age43	*-0.010*	*-0.4*	*-0.013*	*-0.50*	*0.120*	*3.03*	*0.003*	*0.05*	*-0.756*	*-0.35*	*0.359*	*0.11*
age44	*-0.027*	*-1.35*	*-0.030*	*-1.20*	*0.067*	*2.17*	*-0.034*	*-0.56*	*-0.590*	*-0.28*	*0.004*	*0.00*
age45	*-0.048*	*-1.86*	*-0.051*	*-1.68*	*0.065*	*1.55*	*-0.042*	*-0.75*	*-0.215*	*-0.08*	*2.037*	*0.35*
age46	*-0.081*	*-3.7*	*-0.083*	*-3.42*	*-0.004*	*-0.11*	*-0.081*	*-1.69*	*6.046*	*1.53*	*6.402*	*1.65*
age47	*-0.087*	*-4.3*	*-0.088*	*-4.16*	*-0.031*	*-0.88*	*-0.085*	*-1.67*	*-1.794*	*-0.74*	*-0.924*	*-0.28*
age48	*-0.137*	*-5.66*	*-0.138*	*-5.75*	*-0.090*	*-2.22*	*-0.127*	*-2.32*	*2.443*	*0.53*	*4.506*	*0.60*
age49	*-0.112*	*-5.81*	*-0.113*	*-5.33*	*-0.101*	*-2.61*	*-0.154*	*-2.71*	*-4.841*	*-1.6*	*-1.635*	*-0.23*
age50+	*-0.155*	*-8.14*	*-0.155*	*-7.74*	*-0.070*	*-1.99*	*-0.073*	*-1.82*	*5.940*	*0.88*	*7.300*	*1.38*
Constant	*-0.025*	*-1.32*	*-0.024*	*-1.28*	*0.366*	*8.68*	*0.395*	*7.86*	*1.898*	*0.92*	*1.730*	*0.83*
OBS	*147274*		*147274*		*147274*		*147274*		*75263*		*75263*	
RMSE	*0.3107*		*0.3107*		*0.3875*		*0.4499*		*8.9610*		*9.1575*	
R Square	*0.0455*				*0.0484*				*0.0228*			
Dep. Var. Mean	*0.114*		*0.114*		*0.196*		*0.196*		*3.323*		*3.323*	

Note: Mandate States Include MA, RI, NJ, NY, MD, IL, TX and CA

Figures

Figure 1: Among All Women in Two-Year Sample, Age Distribution of Completed Pregnancies by Fertility Status

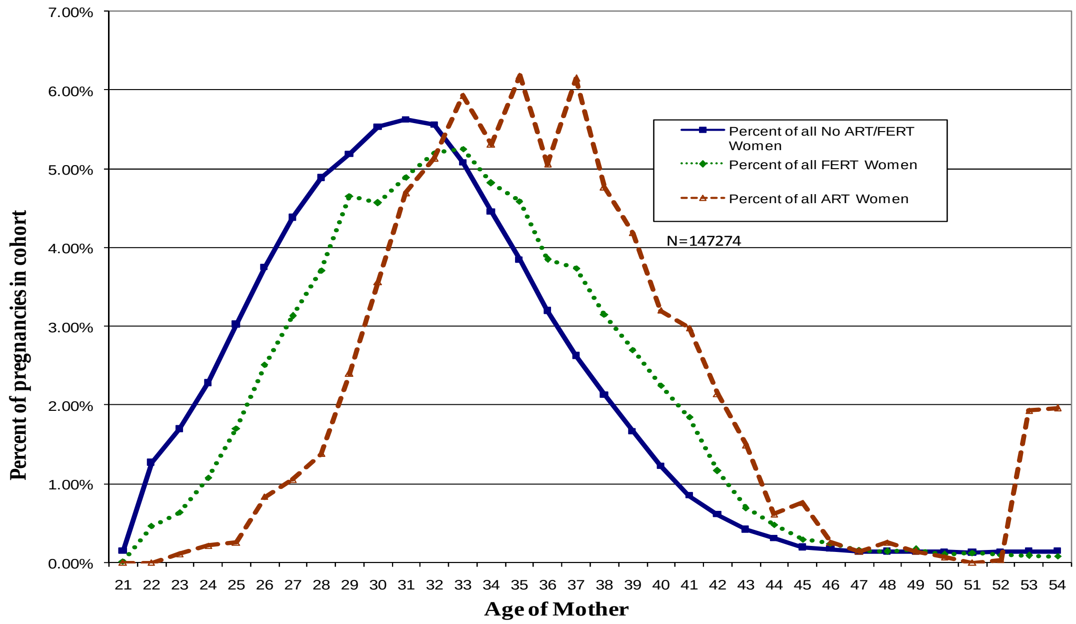


Figure 2: Among All Women in Two-Year Sample, Average Health Risk Score, by Fertility Status

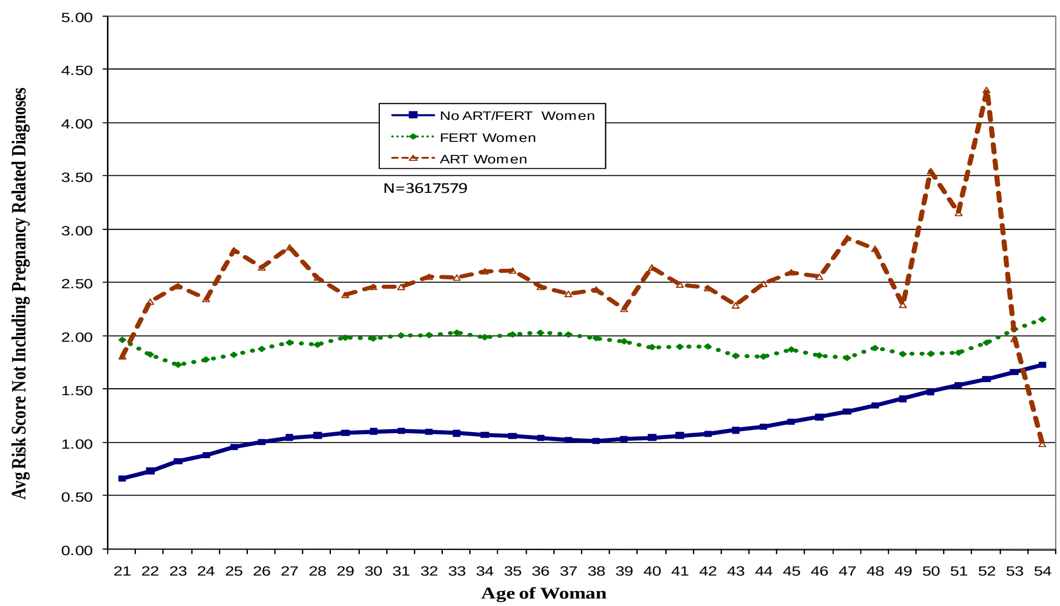


Figure 3: Among All Women in Two Year Sample, Success Rate By Fertility Status and Age of Mother

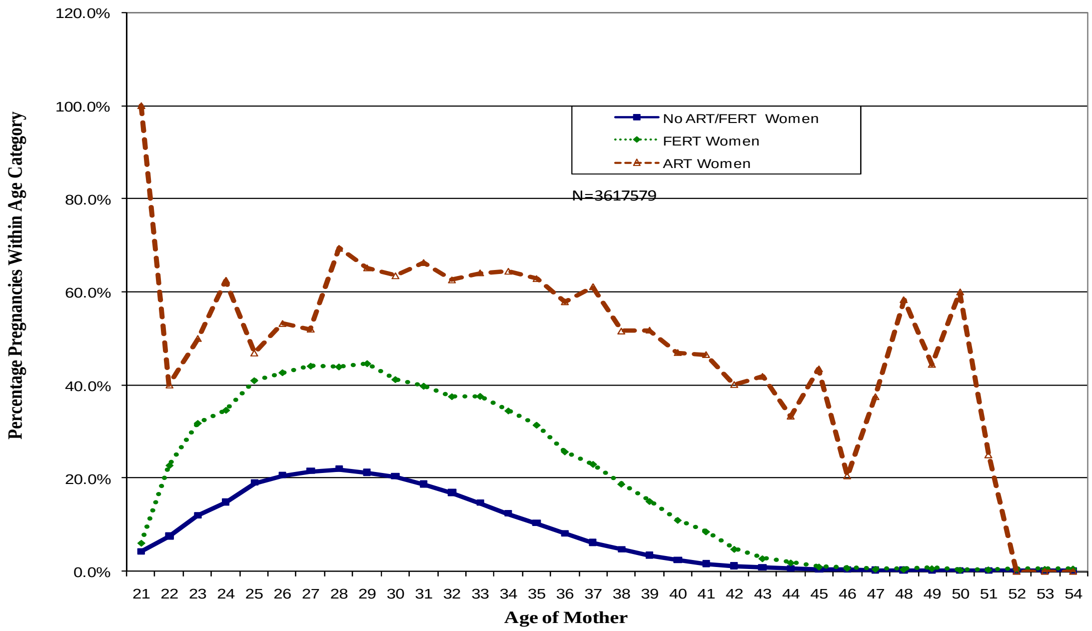


Figure 4: Among Women in Two-Year Sample, With Completed Pregnancy, Complication Rate By Age and Fertility Status

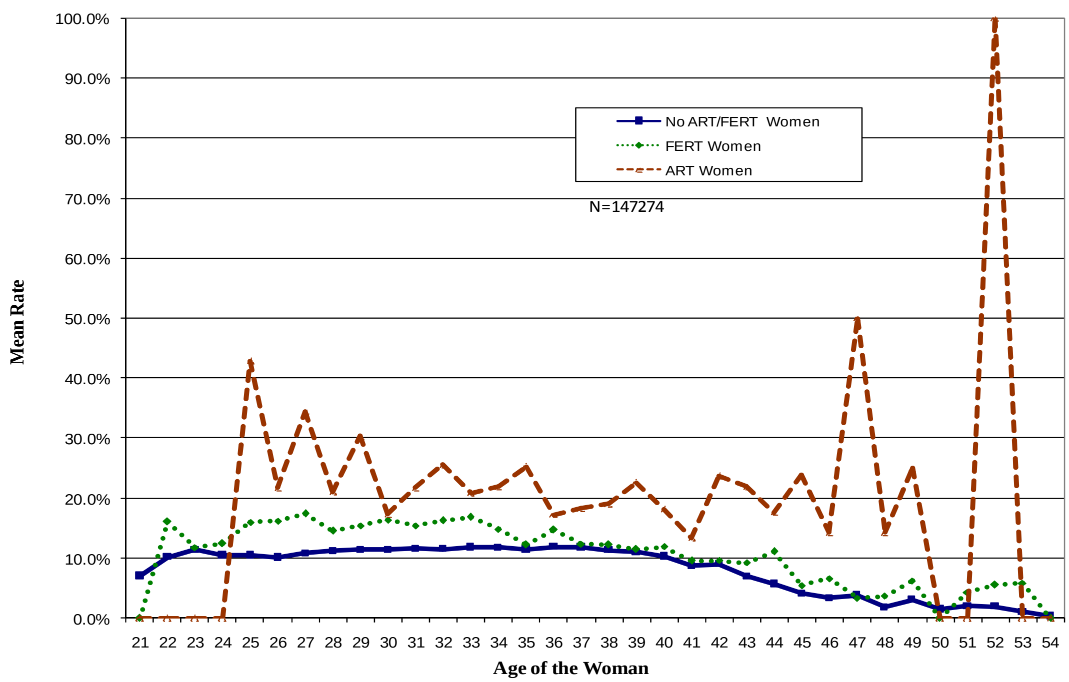


Figure 5: Among Women in Two-Year Sample, With Completed Pregnancy, Failure Rate By Age and Fertility Status

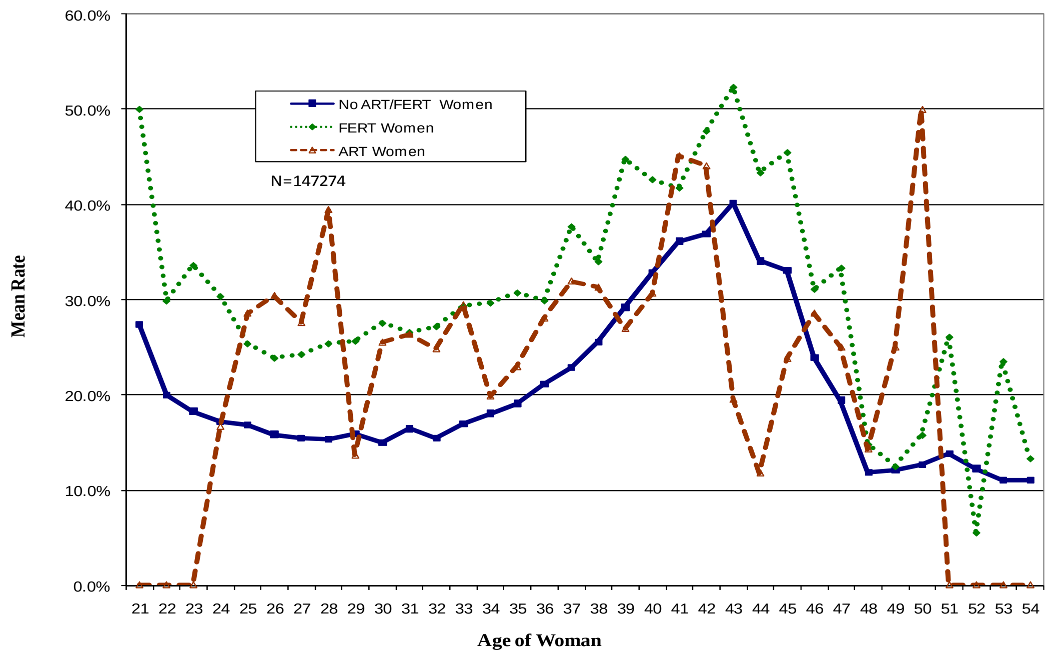


Figure 6: Among All Women in Two Year Sample, With Completed Pregnancies and Baby Information, Mean Risk Score of the Baby By Age of Mother and Fertility Status

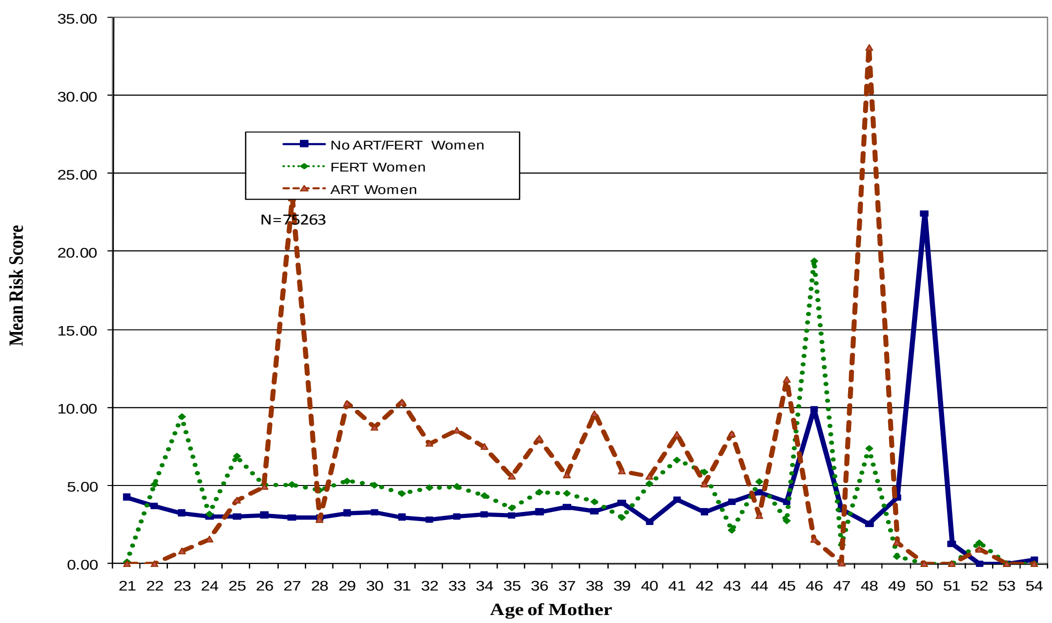


Figure 7: Among All Women in Two Year Sample, With Completed Pregnancy, Mean Complications Rate by Mother's Risk Score Category

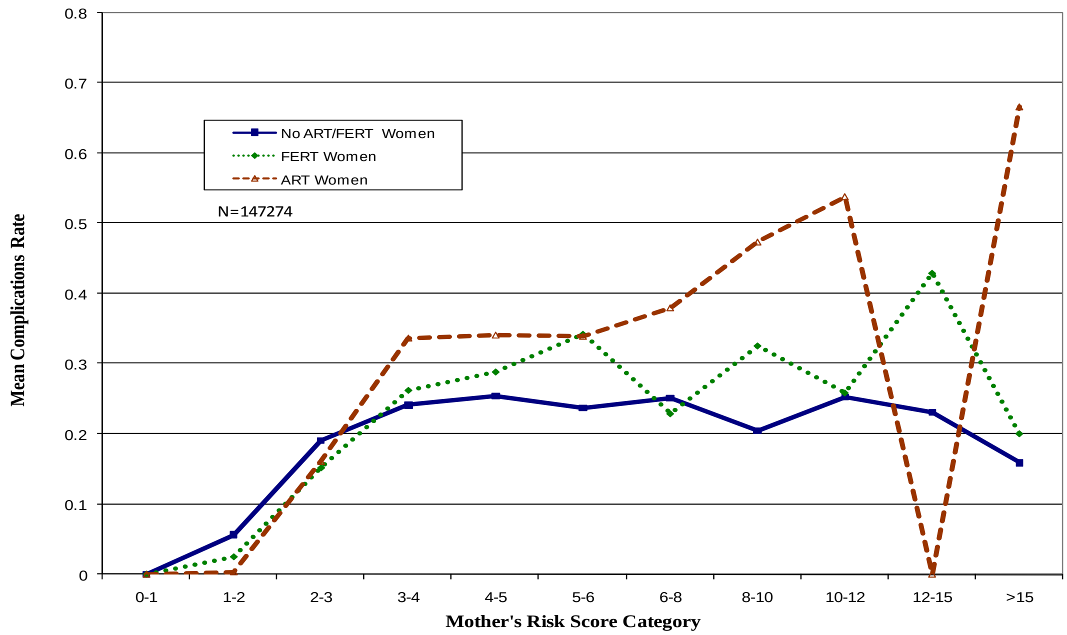


Figure 8: Among All Women in Two-Year Samples, With Completed Pregnancy, Mean Failure Rate by Mother's Risk Score Category

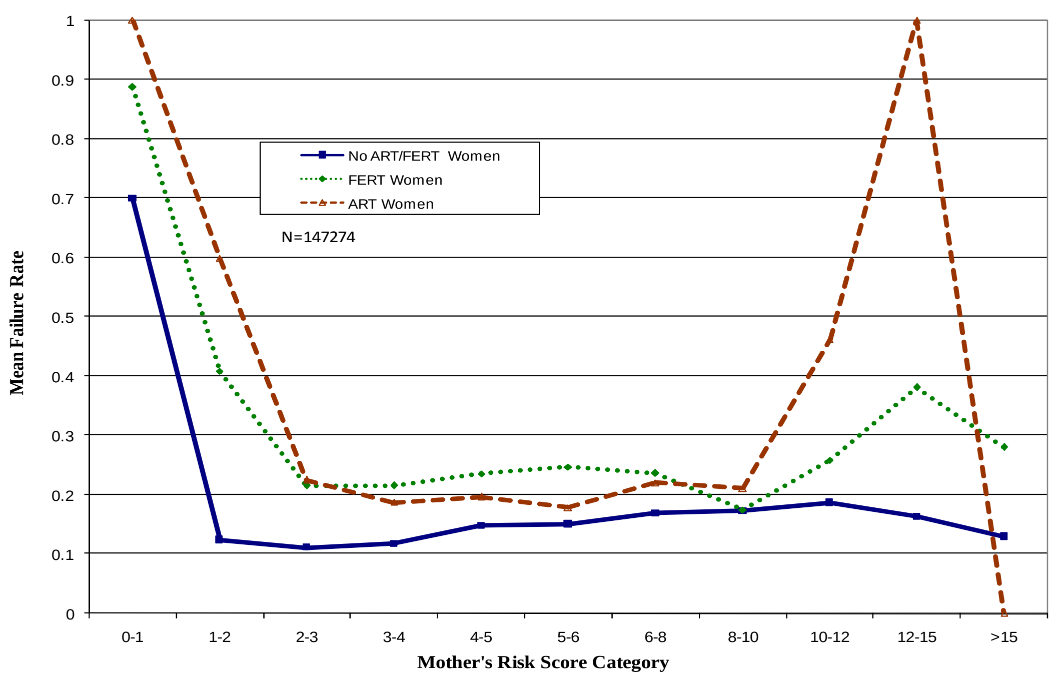


Figure 9: Among All Women in Two-Year Samples, With Completed Pregnancy and Baby Information, Mean Baby's Risk Score For Each of Mother's Risk Category, By Mother's Fertility Status

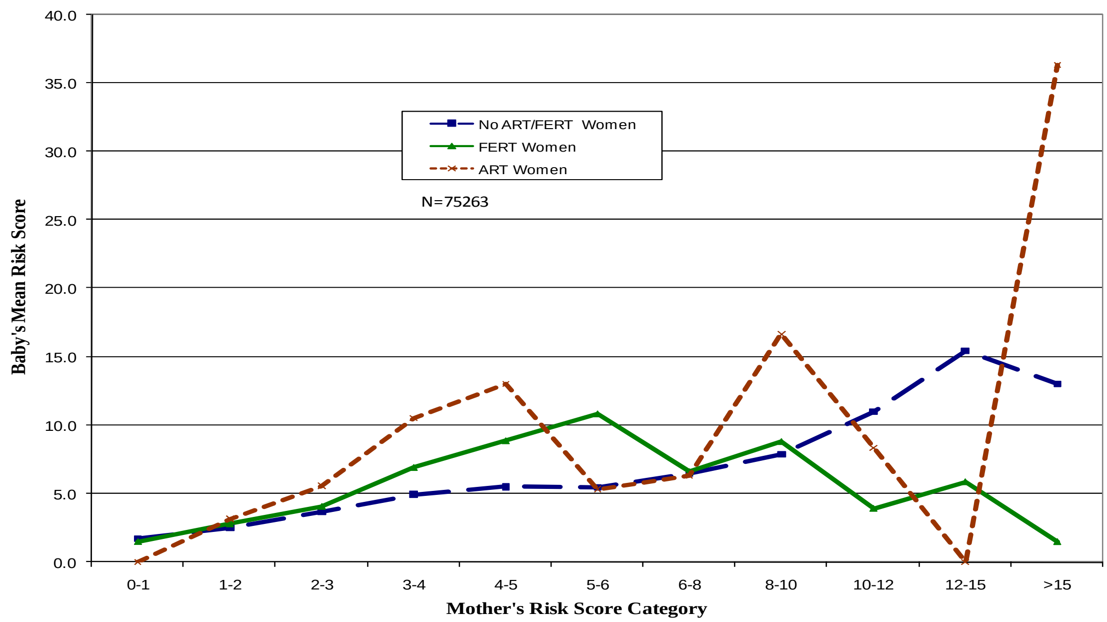


Figure 10: Among All Women in Two-Year Sample, Proportion of All Pregnancies by Age and fertility category

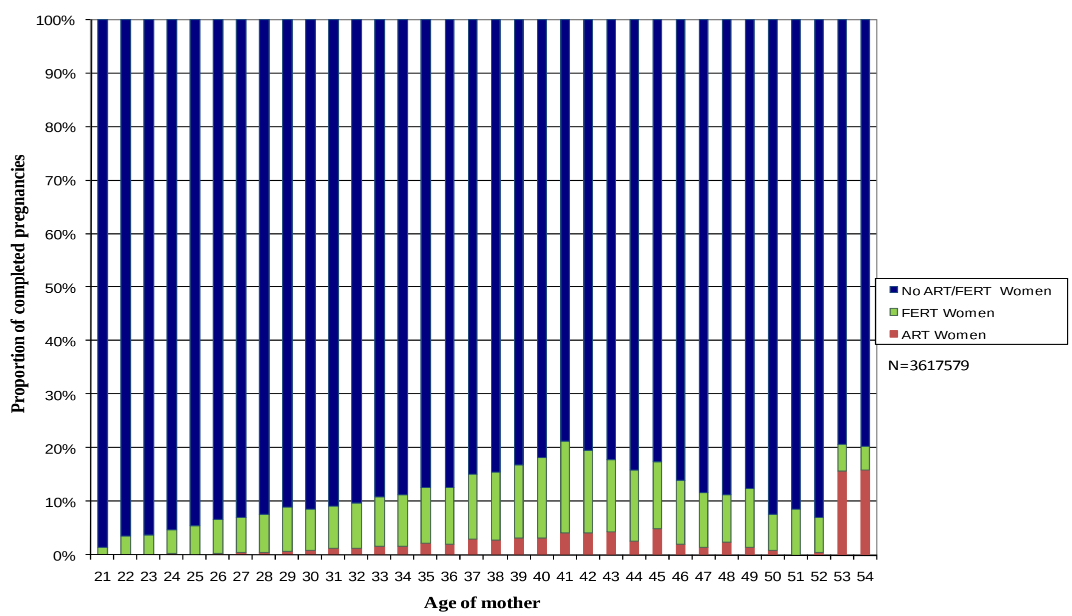


Figure 11: Among All Women in Two-Year Sample, Age distribution of ART women in mandate and no mandate states

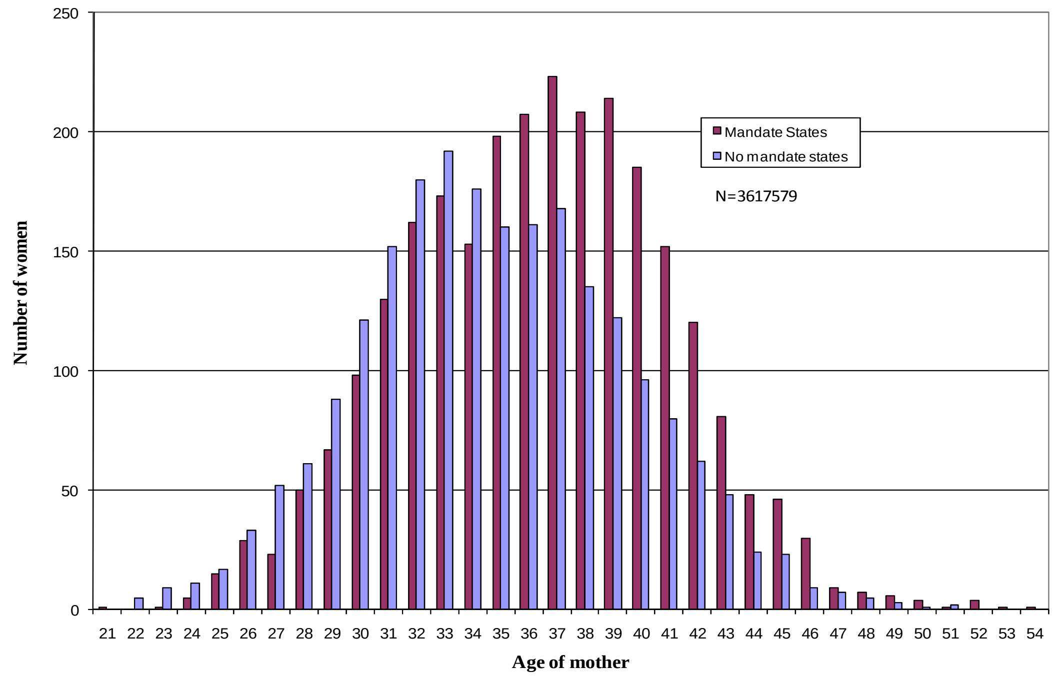


Figure 12: Among All Women in Two-Year Sample, ART Success rate by age of mother, mandate versus non-mandate states

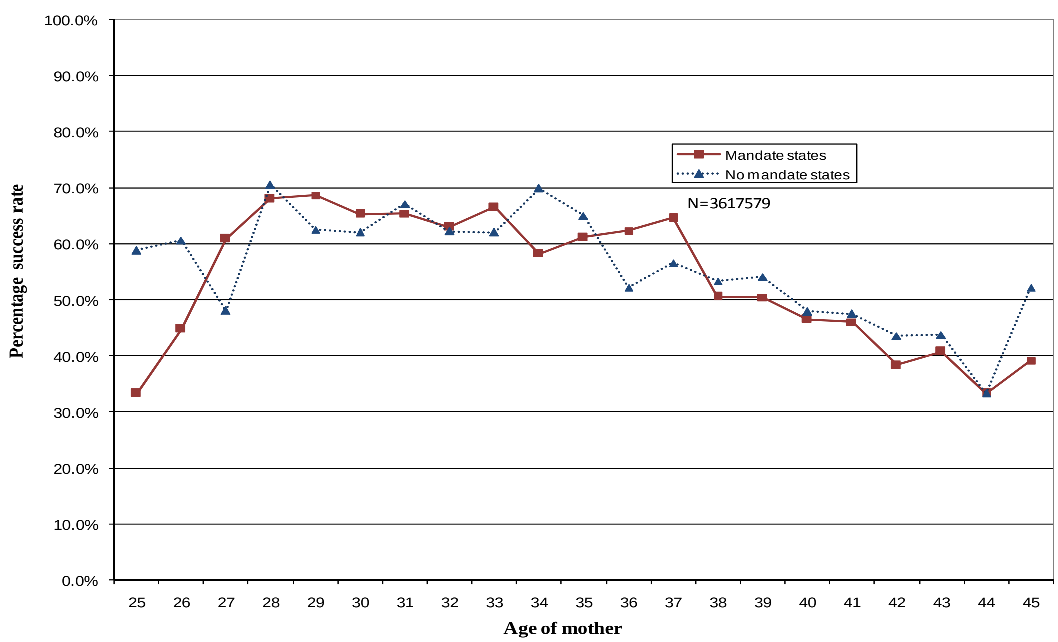


Figure 13: Raw Means, Mean FERT Rate vs. Mean ART Rate, By State

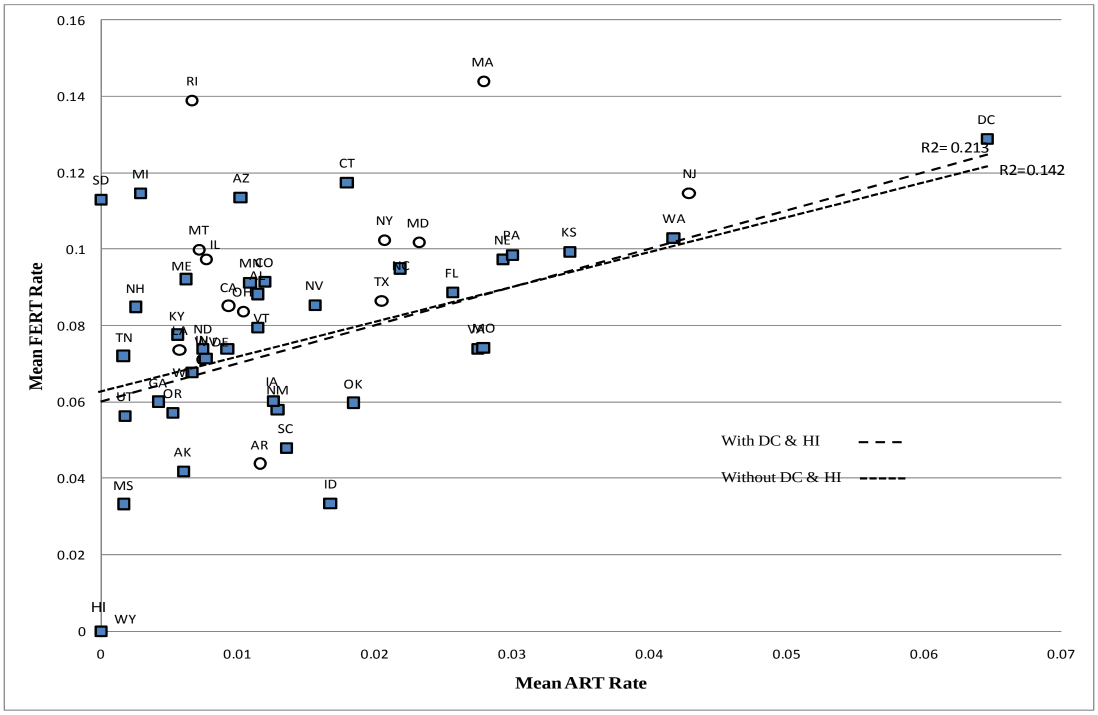


Figure 14: Mean Residual FERT Rate vs. Mean Residual ART Rate, By State

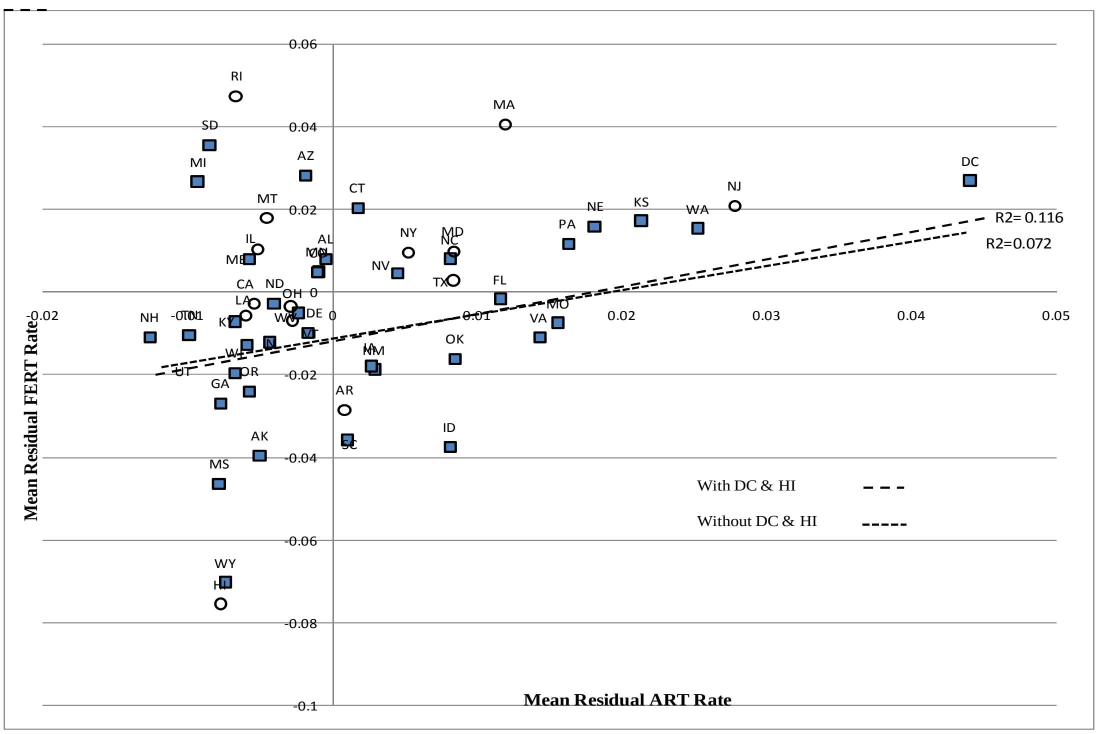


Figure 15: Raw Means, Mean Rate of Major Complications vs. Mean ART Rate, By State

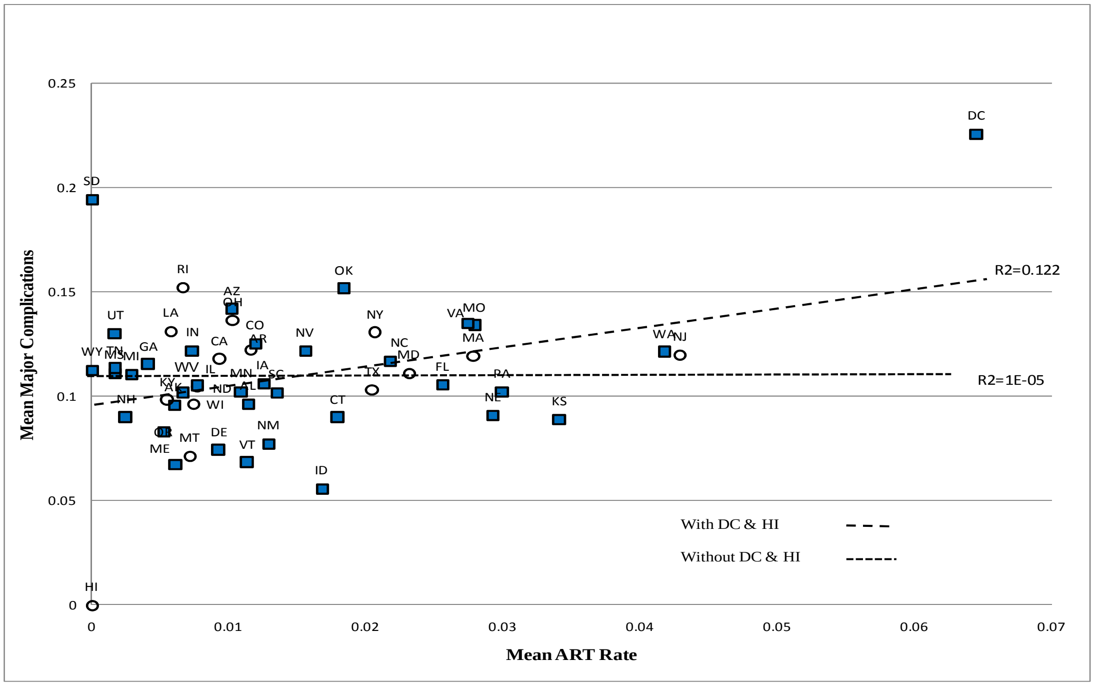


Figure 16: Mean Residual Complications vs. Mean Residual ART Rate, By State

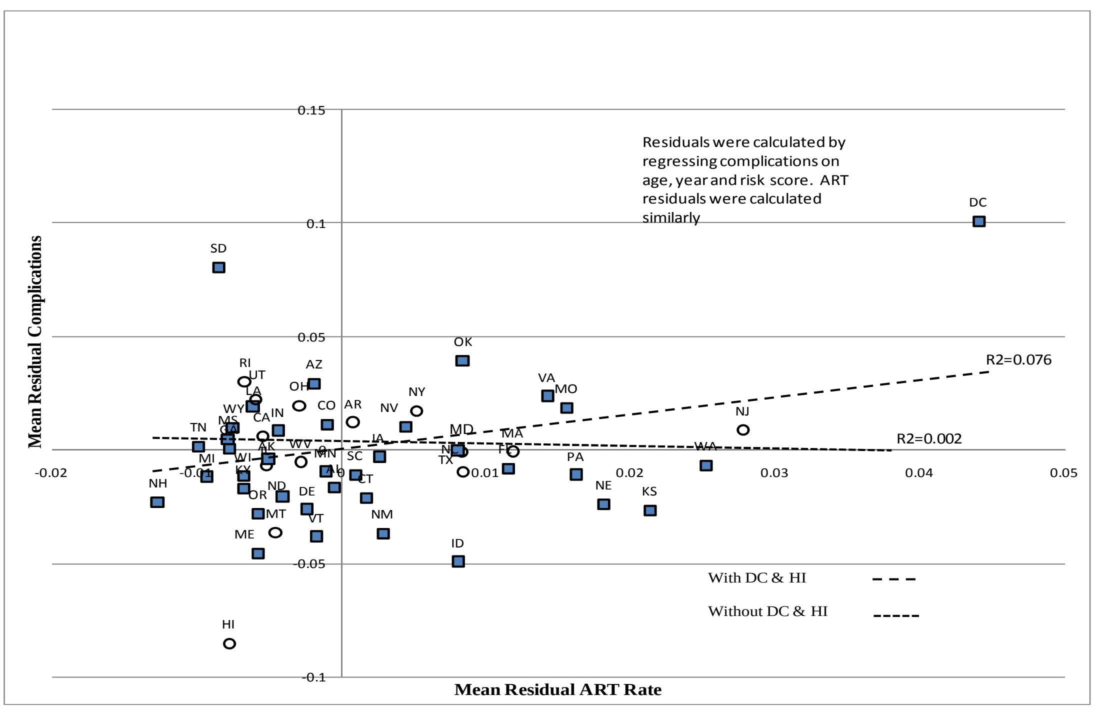


Figure 17: Raw Means, Mean Rate of Major Complications vs. Mean FERT Rate, By State

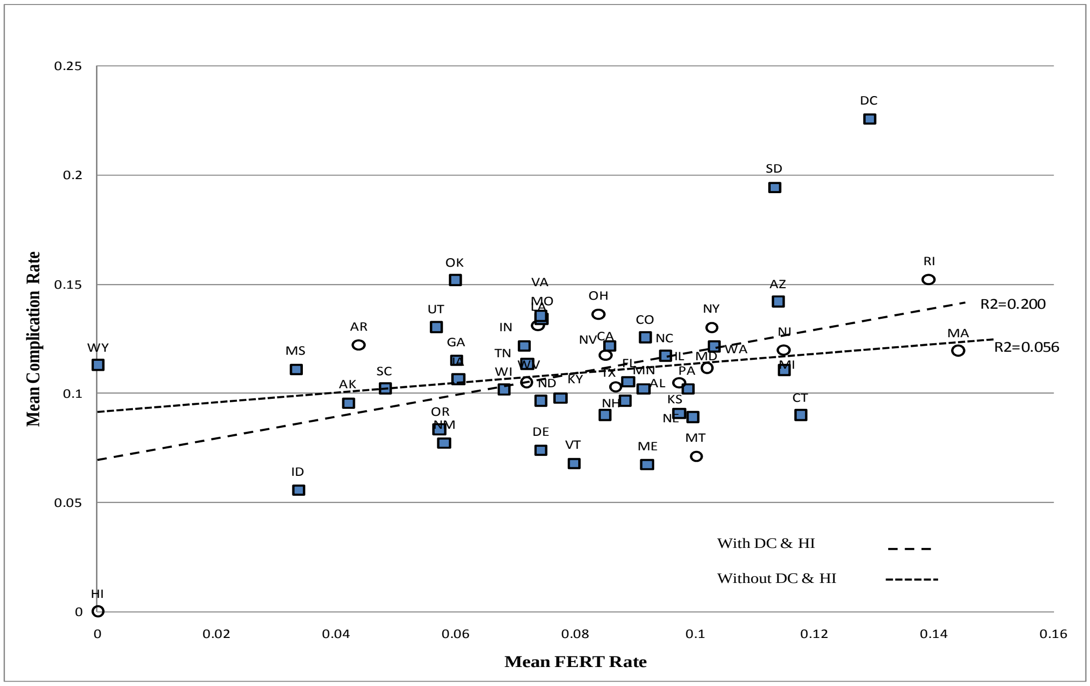


Figure 18: Mean Residual Complications Rate vs. Mean Residual FERT Rate, By State

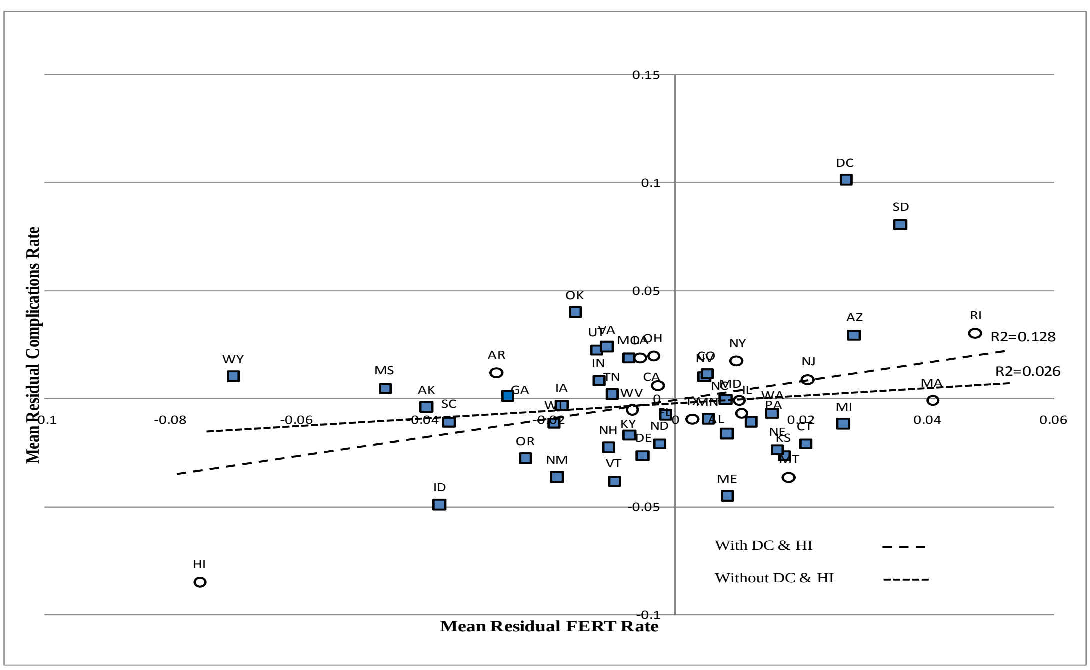


Figure 19: Raw Means, Mean Failure Rate vs. Mean ART Rate, By State

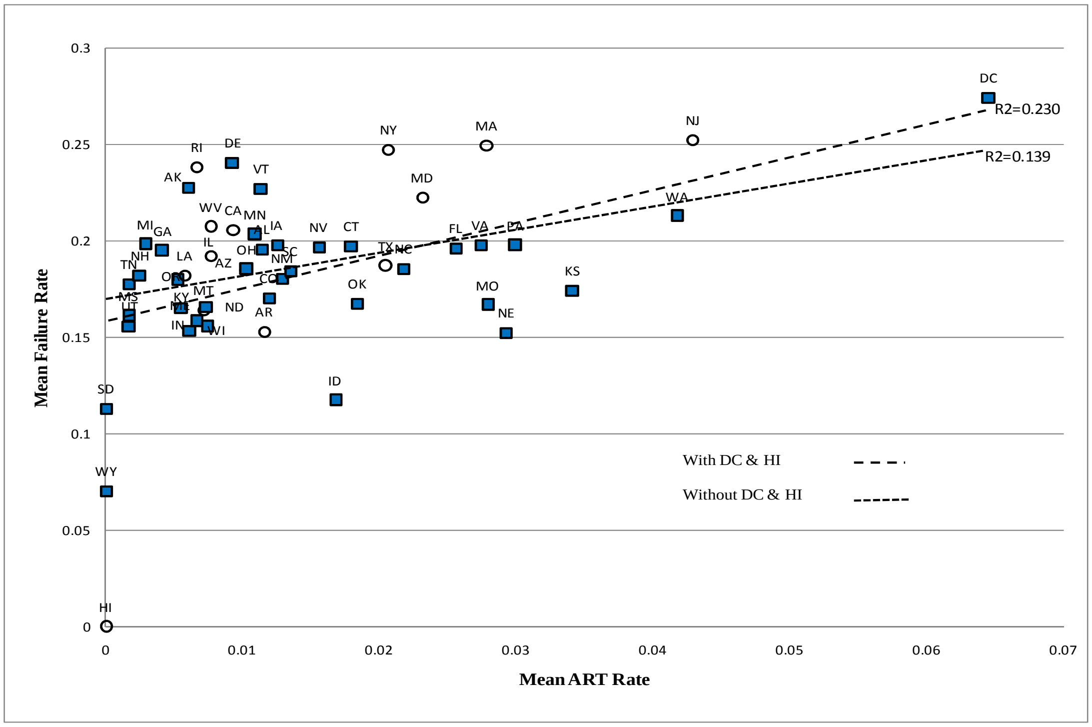


Figure 20: Mean Residual failure Rate vs. Mean Residual ART Rate, By State

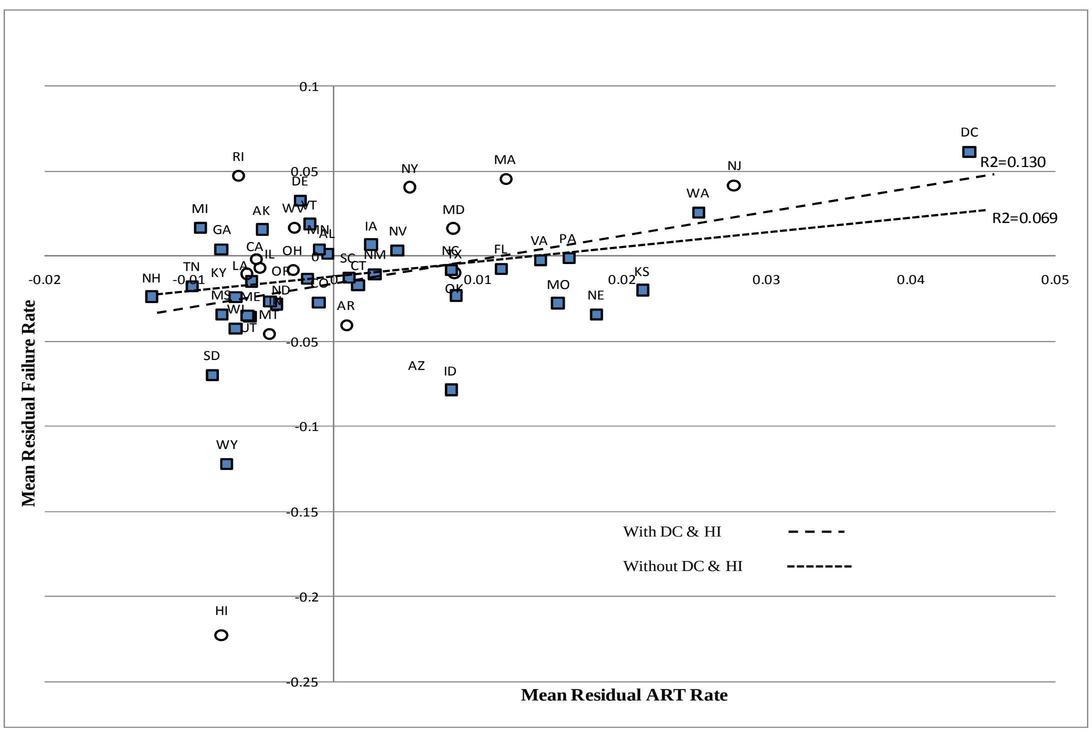


Figure 21: Raw Means, Mean Failure Rate vs. Mean FERT Rate, By State

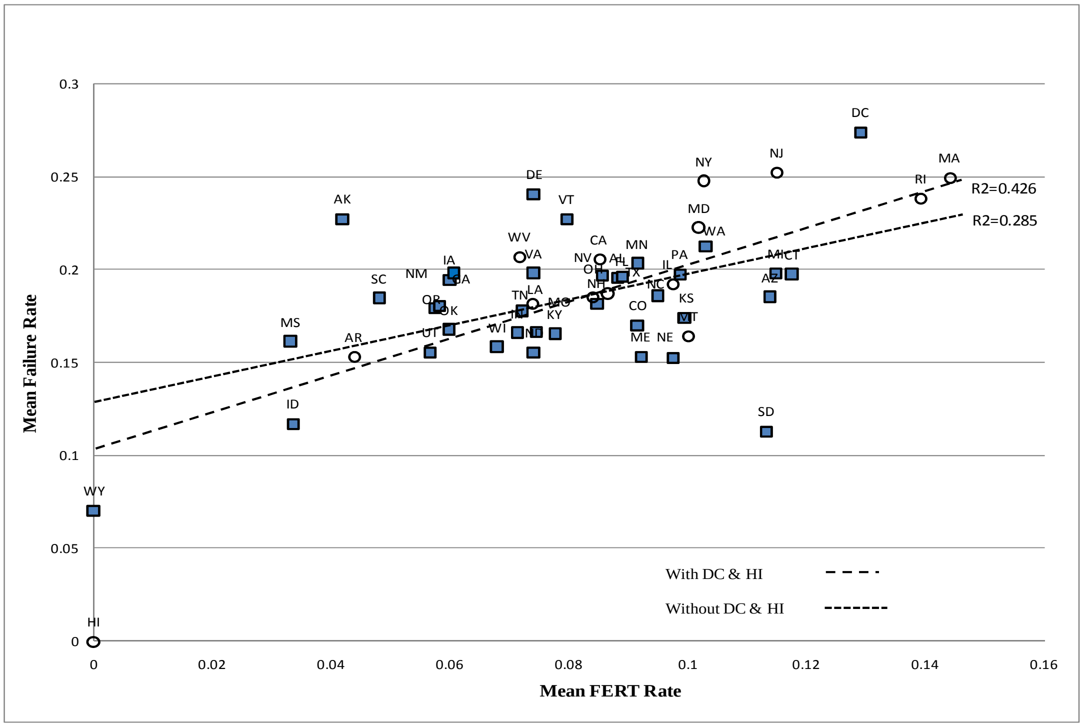


Figure 22: Mean Residual failure Rate vs. Mean Residual FERT Rate, By State

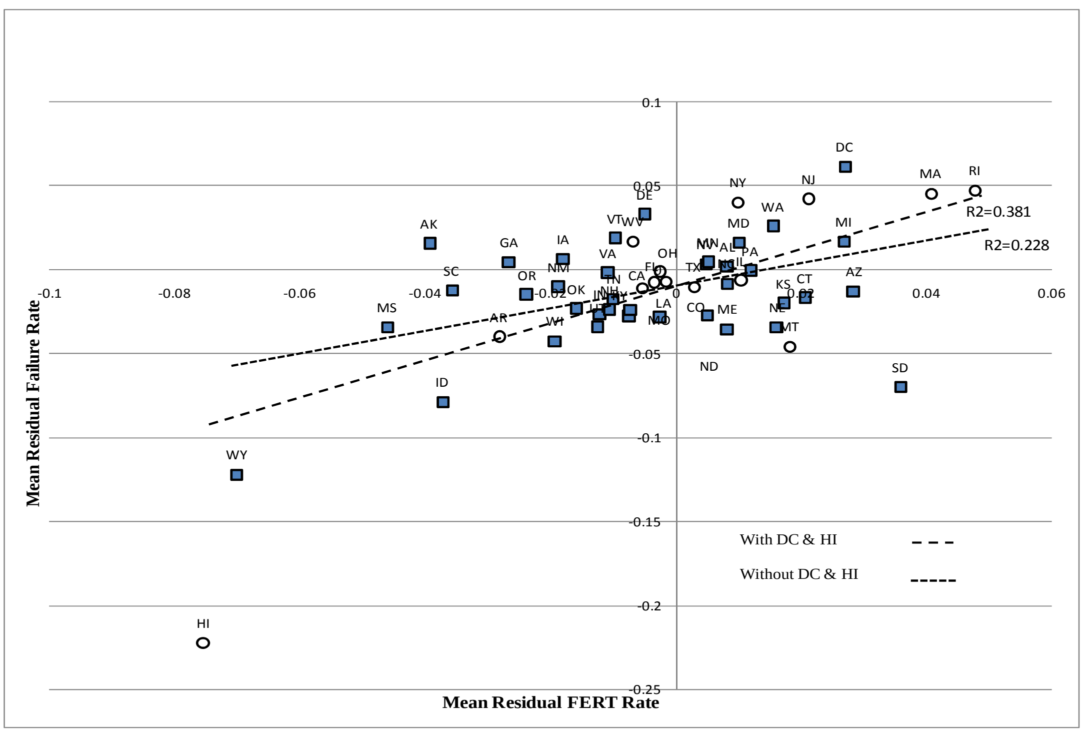


Figure 23: Raw Means Mean Baby Risk Score vs. Mean ART Rate, By State

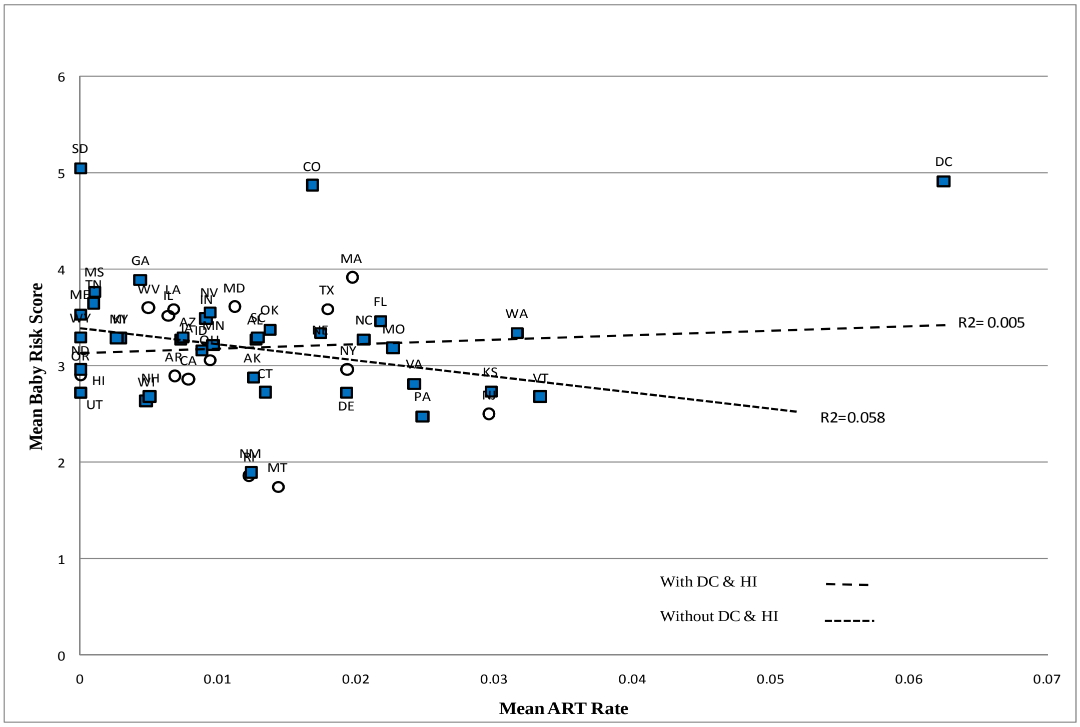


Figure 24: Mean Residual Baby Score vs. Mean Residual ART, By State

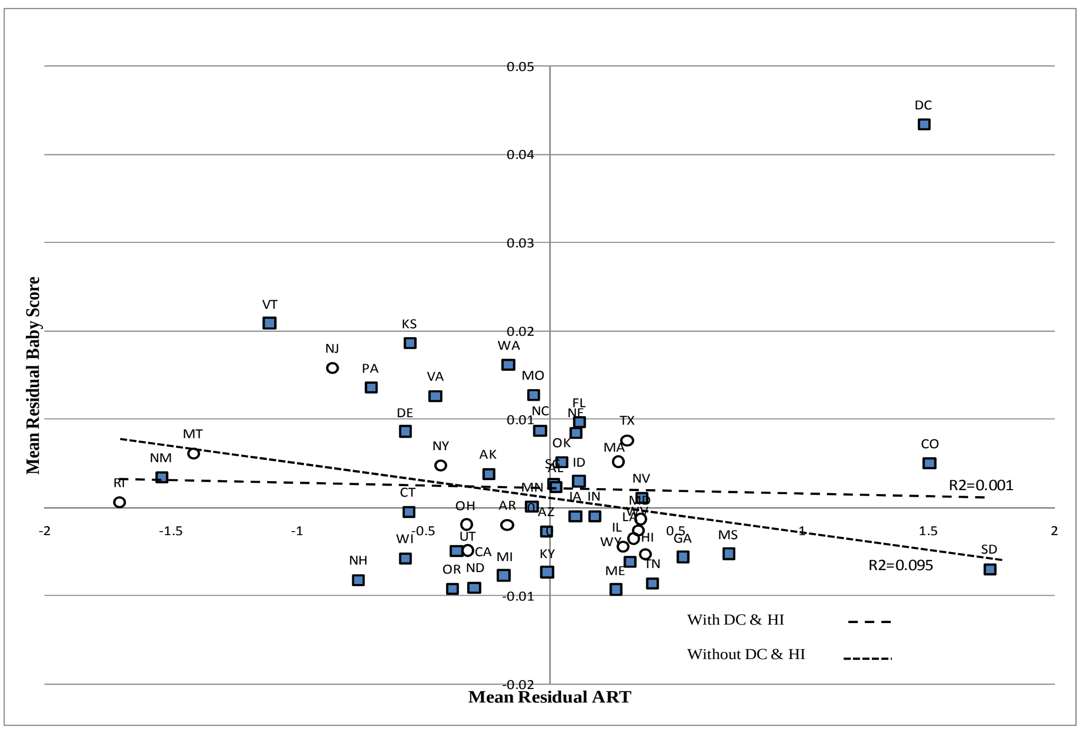


Figure 25: Raw Means Mean Baby Risk Score vs. Mean FERT Rate, By State

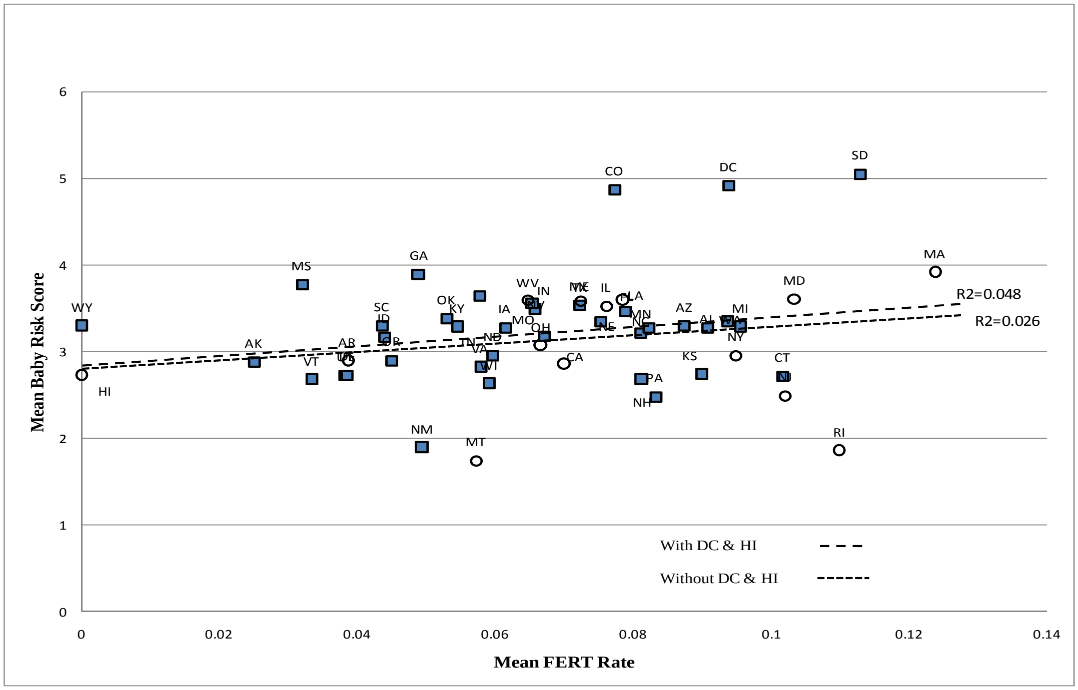
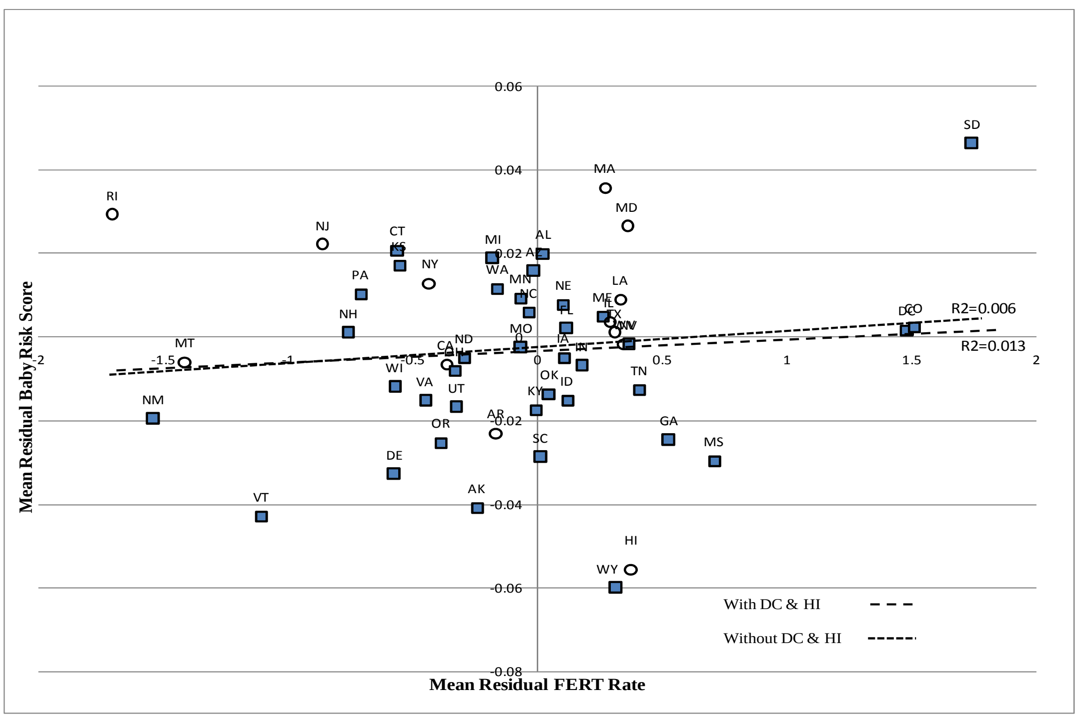


Figure 26: Mean Residual Baby Risk Score vs. Mean Residual FERT Rate, By State



REFERENCES

1. Adamson DM, Chang S, Hansen LG "Health Research Data for the Real World: The Marketscan Databases" Thomson Medstat White Paper, January 2006.

2. Allen VM and Wilson RD "Pregnancy Outcomes After Assisted Reproductive Technology" Journal of Obstetrics and Gynecology Canada 2006; 28(3):220-233.

3. American Society for Reproductive Medicine http://www.asrm.org , 2007.

4. "Assisted Reproductive Technology Success Rates" U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, December 2006.

5. Buckles, Kasey "Stopping the Biological Clock: Fertility Therapies and the Career/Family Tradeoff." working paper.

6. Clayton HB, Schieve LA, Peterson HB, Jamieson DJ, Reynolds MA, Wright VC "Ectopic Pregnancy Risk With Assisted Reproductive Technology Procedures" Obstetrics & Gynecology 2006; Vol.107, No 3 ©2004 by The American College of Obstetricians and Gynecologists.

7. Hansen M, Bower C, Milne E, Klerk ND and Kurinczuk JJ. “Assisted Reproductive Technologies and the Risk of Birth Defects—A Systematic Review” Human Reproduction 2004 Vol. 20, No.2, pp 328-338 © European Society of Human Reproduction and Embryology 2004.

8. Hawkins Jessica L “Separating Fact from Fiction: Mandated Insurance Coverage of Infertility Treatments” Washington University Journal of Law and Policy 2006, Vol. 23:203.

9. Schieve LA, Meikle SF, Ferre C, Peterson HB, Jeng G and Wilcox LS “Low and Very Low Birth Weight In Infants Conceived With Use of Assisted Reproductive Technology” New England Journal of Medicine, 2002, Vol. 346, No. 10.

10. Schieve LA, Rasmussen SA, Buck GM, Schendel DE, Reynolds MA, Wright VC. “Are Children Born After Assisted Reproductive Technology at Increased Risk For Adverse health Outcomes?” Obstetrics & Gynecology 2004; 103: 1154-1163 ©2004 by The American College of Obstetricians and Gynecologists.

11. Schieve LA, Tatham L, Peterson HB, Toner J and Jeng G. “Spontaneous Abortion Among Pregnancies Conceived Using Assisted Reproduction Technology in the United States”. The American College of Obstetrics and Gynecology 2003; Vol. 101, No. 5, Part 1.

12. Schmidt, Lucie “Effects of Infertility Insurance Mandates on Fertility,” *Journal of Health Economics 2007;* 26(3): 431-446.

13. Sutter PD, Veldman L, Kok P, Szymczak N, Elst JVD and Dhont M “Comparison of Outcome of Pregnancy after Intra-Uterine Insemination (IUI) and IVF” Human Reproduction 2005 Vol. 20, No.6, pp. 1642-1646.

14. Thomson F, Shanbhag S, Templeton A, Bhattacharya S “Obstetric Outcome in Women with Subfertility” Br. Journal of Obstetrics and Gynecology 2005; Vol. 112.