EXPANDING GEOGRAPHIC ACCESS TO FERTILITY CARE IN THE UNITED STATES: A STATISTICAL AND MULTIOBJECTIVE OPTIMIZATION APPROACH

by
Micajah Ziegler McGarity

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
March 2022

Abstract

An estimated 25% of potential demand is met in the US for Assisted Reproductive Technology (ART). ART includes *in vitro* fertilization, which is one of the most effective ways to treat infertility and is used generally after alternate treatment options have failed. Of the many barriers to accessing ART, geographic barriers can be addressed using operations research methods to develop cost-effective strategies for expanding access. First, a systematic web-search was performed to collect the locations of all main and satellite ART clinics. 13M reproductive-age women were found to lack geographic access to ART. Geographic access to ART was calculated based on race and ethnicity, and 17% fewer American Indians and Alaska Natives were found to have geographic access to ART compared with all US races. Next, explanatory models of geographic access to care were developed, in which the presence of ART clinics in an area was statistically related to female reproductive-age population and median income. Additionally, infertility treatment demand, as measured by the annual number of ART cycles performed in a county, was explored statistically as a function of socioeconomic and environmental health factors. This statistical model predicted values for ART demand for every US county and provided insights into the importance of county characteristics. Population aside, the two most important factors in predicting ART demand were the positively correlated Primary Care Physician Rate and the negatively correlated Severe Housing Cost Burden. Lastly, outputs from the models predicting demand were embedded in a maximal covering location problem (MCLP) mixed integer programming framework to optimize county-level placement of new ART clinics to improve geographic access to ART, as measured by the reproductive-age population covered and predicted demand covered. A novel multiobjective optimization approach was employed in which Pareto-optimal solutions for new clinic locations were compared with existing clinic locations to identify new clinic locations that maximize unmet demand covered vs. maximize unserved population covered. Finally, an ensemble of recommended locations for new clinics was produced by spatially combining the recommendations from statistical and optimization methods, and the concurrence of model recommendations identified Northwest Arkansas/Southwest Missouri to locate a new ART clinic.

Academic Advisor and Reader:

Benjamin F. Hobbs, Theodore K. and Kay W. Schad Professor of Environmental Management

Department of Environmental Health and Engineering

The Johns Hopkins University

Reader:

Scot M. Miller, Assistant Professor

Department of Environmental Health and Engineering

The Johns Hopkins University

Reader:

Frank C. Curriero, Professor

Department of Epidemiology

The Johns Hopkins University

Dedication

This thesis is dedicated to my loving family, my dear children, and my incredible wife. To my parents and family, thank you for your constant love that has formed me and given me the foundation for a wonderful life. I would not have been able to complete this work were it not for your steadfast support. To my children, I cherish everything that you are and love each of you just the way you are. Your mother and I are so lucky that you even exist and are amazed by each of you every day. To my wife Sara, thank you for making this dissertation possible through believing in me, supporting me, and the countless hours caring for Forrest, Flora, and Fern. Thank you, Sara, for being my person, my best friend, and the love of my life.

Acknowledgements

I would like to thank the many people who helped me along the way. I will be forever grateful for all of the support I received from my family, friends, and committee to help me complete this dissertation.

First of all, thank you to my advisor, Benjamin Hobbs, for providing the perfect combination of encouragement, guidance, and support that was needed to guide me through this process and reach the finish line. It was a difficult but fun process that I'm glad to have been through. Thank you also to my committee members Scot Miller and Frank Curriero for the great questions and insights during the research phase. Thank you for the suggestions on predictive models and expert guidance on spatial statistics. I was so fortunate to receive your world-class advice. Thank you also for serving on my qualifying exam committee.

I would also like to thank the other members of my qualifying exam committee. Many thanks to Edward Bouwer, Takeru Igusa, Meghan Latshaw, and Charles Young. Ed Bouwer asked me what advice I had for the President of the United States, which showed me the high regard he held for his students. Takeru Igusa made the useful suggestion to model the presence of fertility clinics that was incorporated into Chapters 2 and 3. Meghan Latshaw introduced me to the broader discipline of environmental health and suggested possible exposure pathways to environmental toxicants linked to infertility. Charles Young helped me navigate the requirements of the Whiting School of Engineering – Applied Physics Lab Research Assistantship and served as my APL Staff Co-Advisor. Thank you also to my colleagues at APL who encouraged this effort: Alyson, Anthony, Brad, Brendan, Brenna, Bryan, Dave, Ed, Ehab, Ellen, Elizabeth, Eric, Erik, George, Hayden, Jim, Jody, Joe, Jon, Keith, Kevin, Lisa, Mike, Pam, Sheri, and Thomas, and thank you to APL for its tuition support and paid leave to complete this dissertation.

The Department of Environmental Health and Engineering has been a great place to learn and grow. Thank you, Meghan Davis, for directing the doctoral program, and thank you, Marsha Wills-Karp, for leading the department. I will always remember your encouragement to the department's Ph.D. students to be leaders. Thanks also to my colleagues over the years: Brent, Cynthia, Elina, Evangelia, Fengwei, Hongtai, Kathryn, Liang, Pengcheng, Qingyu, Robin, Roshank, Rui, Stephanie, Sara, Sarah, Shen, Stephanie, Yijiao, and Yinong. From the wider Johns Hopkins community, thank you, Valerie Baker and Richard Church, for your helpful feedback.

Thank you to my collaborators outside JHU, Christopher Herndon and John Harris, for asking great questions and for your valuable input on the ART satellite clinic study that formed the basis of Chapter 2. Christopher Herndon helped me understand the need to investigate ART satellite clinics, and John Harris helped me better communicate my findings to a medical audience.

Lastly, I am eternally grateful to Santiago Padilla, Katherine Bass, and Joseph Osheroff who taught me so much about ART, but more significantly helped my wife and me overcome infertility and grow our family. Thank you!

Finally, thank you to my family for your love and support. To my mother and father in-law Carol and Jim, thank you for caring for the kids while I completed the web-search of every fertility clinic in the US. To my siblings, Kate, Owen, and Sam, thank you for the constant backing. To my mother, Jane, thank you so much for proofreading this dissertation and for coming to help countless times over the years. To my father, Arthur, thank you for reading my manuscripts and for your continuous encouragement and inspiration. To my wife Sara, thank you for everything that you did over the years to make this possible. I could not have completed this dissertation were it not for your unwavering support.

Contents

List of Tables

List of Figures

Glossary

Acronyms

ABOG – American Board of Obstetrics and Gynecology

ACOG – American College of Obstetricians and Gynecologists

ACS – American Community Survey

AIC – Akaike information criterion

ART – Assisted reproductive technology

ASRM – American Society for Reproductive Medicine

ATC – Access to care

BMI – Body mass index

CART – Classification and regression trees

CBSA – US Census Bureau core-based statistical areas

CDC – US Centers for Disease Control and Prevention

GBM – Gradient boosting machine

GC – Gestational carrier

GLM – Generalized linear model

HIV – Human immunodeficiency virus

IUI – Intrauterine insemination

IVF – *In vitro* fertilization

kNN – k-Nearest neighbors

LIFE – Longitudinal Investigation of Fertility and Environment Study

MAE – Mean absolute error

MCLP – Maximal covering location problem

MILP – Mixed integer linear program

ML – Machine learning

MLE – Maximum likelihood estimation

NB – Negative binomial regression

NSFG – National Survey of Family Growth

OB/GYN – Obstetrics and gynecology

PBDEs – Polybrominated diphenyl ethers

PCBs – Polychlorinated biphenyls

PCP – Primary care physician

PG – Prince George County, Maryland

PGT – Preimplantation genetic testing

$PM_{2.5}$ – Particulate matter 2.5, particles or droplets in the air that are 2.5 microns or less in width

R^2 – R-squared, the proportion of the dependent variable's variation that is predictable from one or more independent variables

REI – Reproductive Endocrinology and Infertility, a subspecialty of American Board of Obstetrics and Gynecology

RMSE – Root mean squared error

STD – Sexually transmitted disease

SVM – Support-vector machine

WHO – World Health Organization

Definitions

Assisted reproductive technology (ART)

Clinical and laboratory procedures to aid in fertilization in which eggs or embryos are handled, including *in vitro* fertilization (IVF) (Wilcox, 2010). IVF is the most common ART treatment, so the terms IVF and ART are used interchangeably in this dissertation.

Biological infertility

Infertility related to advanced maternal age (Menken et al., 1986) (e.g., infertility for women over 38 years of age (American Society for Reproductive Medicine, 2014a)).

Core-based statistical areas (CBSAs)

US Census Bureau delineated metropolitan and micropolitan core areas plus adjacent territory with a “high degree of social and economic integration with the core as measured by commuting ties” (U.S. Census Bureau, 2021).

Ecological study

Ecological studies are observational studies performed at the population or group level vs. the individual level, and they are performed to “monitor population health so that public health strategies may be developed and directed” (Levin, 2006).

Fecundity

A person’s (or couple’s) capacity to conceive and deliver a baby (Wilcox, 2010).

Fertility

The number of babies actually born from a woman (Wilcox, 2010).

***In vitro* fertilization (IVF)**

Fertilization by medical procedures outside the body (Wilcox, 2010). IVF is the most common ART treatment, so the terms IVF and ART are used interchangeably in this dissertation.

Infertility

Term to describe couples who are having trouble conceiving a child but who may still be fecund. Generally fecund and infecund couples are deemed infertile if they do not become pregnant after more than a year of trying (Wilcox, 2010). ACOG and WHO defines infertility as the inability of a woman or a couple to achieve a clinically recognized pregnancy after attempting for more than a year.

IVF mandate

State laws in the US requiring health insurance plans to cover IVF/ART (RESOLVE, 2021c).

Pareto-optimal solutions / Pareto frontier

Pareto-optimal solutions along the Pareto frontier are solutions in which an increase in benefit to one objective results in a reduction in benefit to another objective (Cohon, 2004).

Reproductive-age women

Term used to describe the age range for women who are generally fecund and are also likely to seek IVF/ART treatment. The age range defining reproductive-age women varies in the literature. In this thesis, the age range used was 20-49 years old.

Chapter 1 Introduction

1.1 Problem overview

In the United States, about 13% of women ages 15 to 49 have impaired fecundity making it difficult or impossible for them to get pregnant and carry a baby to term (National Survey of Family Growth, 2021). While infertility can be medically treated using several methods, there exist a variety of barriers in the US to accessing the latest treatment methods of Assisted Reproductive Technology (ART), which includes *in vitro* fertilization (IVF), resulting in substantial unmet need. An estimated 25% of demand for ART is met in the US (Chambers et al., 2009) based on the standard demand estimate of 1,500 cycles per million population per year (Collins et al., 2001). Only 0.5% of women in the US ages 15 to 49 have ever received ART treatments (National Survey of Family Growth, 2021). In response, access to infertility care is an important topic to the community as demonstrated by the American Society of Reproductive Medicine's (ASRM) recent creation in 2015 of an infertility Access to Care (ATC) Initiative, ATC Special Interest Group, and ATC Task Force (American Society for Reproductive Medicine, 2015b).

In general, access to healthcare is an important societal priority throughout the world. Access to health care services has many dimensions, but fairness is the ultimate goal. Various barriers to accessing care exist throughout world, and particular public health questions arise in setting government policy to reduce barriers, which include the role of the private and public sector, cost allocation, location of facilities, educational outreach, and prevention of disease. Access to fertility care carries profound questions related to fairness in procreation. Is it fair for only the wealthy to have access to the best fertility treatments? Is it morally acceptable for certain racial or ethnic groups in a country to have better access than others? Should access to care also prioritize older patients who intentionally delayed childbearing and are naturally less able to reproduce? To correct disparities in access to care, where are the best places to locate new clinics? These types of questions have been addressed in a wide range of health care contexts, which I review later in the thesis using quantitative analysis in Chapter 2, statistical models in Chapter 3, and optimization methods in Chapter 4.

Access to fertility care is primarily a concern in developed countries with advanced infrastructure and healthcare systems where access to advanced fertility care has grown since the first human was born in the United Kingdom through IVF in 1978 (Edwards et al., 1980). More recently, the World Health Organization (WHO) first declared infertility a disease in 2009 (Zegers-Hochschild et al., 2009), and in 2012 the Inter-American Court of Human Rights determined the banning of IVF in Costa Rica to be a human rights violation (Inter-American Court of Human Rights, 2012; Valerio et al., 2017). With declining fertility rates in recent decades in the US, Americans are having fewer children than are needed to maintain a stable population (Kearney et al., 2021). This is likely due to societal trends of having fewer children and the delaying of childbearing to later in life when fecundity (the ability to have children) naturally declines rapidly (Klein & Sauer, 2001). However, there are other contributors to increases in infertility, including increased cannabis use, the rise in sexually transmitted diseases, obesity, and environmental exposures to air pollution and endocrine disrupting chemicals (EDCs), which have all been associated with reduced male and female reproductive function (Carlsen et al., 1992; Ehrlich et al., 2012; Gaskins et al., 2018; Meeker et al., 2010; Mumford et al., 2021; Peretz et al., 2014; Swan et al., 2003).

Increased access to ART could potentially improve US fertility rates while treating the disease of infertility, but there are many barriers to accessing ART including cultural/religious, educational, governmental, economic, and geographic. An example of a cultural/religious barrier is the Catholic Church's ban on ART (U.S. Conference of Catholic Bishops, 2018). Educationally, ART is a highly complex intervention with much uncertainty involved requiring extensive education of patients prior to starting treatment. There are few governmental barriers to ART in the US with the industry being largely unregulated; however, regulation of ART exists around the world. For example, Germany severely regulates ART (Jasanoff & Metzler, 2018), and Costa Rica banned it altogether until it was lifted in 2016 (Valerio et al., 2017). Regulation of ART may come to certain US states in the future as evidenced by the recent passage of laws that increasingly restrict reproductive rights to progressively earlier in the pregnancy (Mallampati et al., 2017), so it is conceivable that if the trend on restrictions continues, states will eventually pass laws regulating the human embryos created through ART.

Economic barriers to ART impact a large portion of the US reproductive-age population with most insurance not covering ART (SHRM, 2015) and a single treatment costing roughly half of US median household income (Katz et al., 2011). As a result, in the US there are huge disparities in the utilization of IVF among economic and racial/ethnic groups with wealthier, highly educated white women being treated at much higher rates (Jain, 2006; Jain & Hornstein, 2005). Lastly, geographic access is a crucial dimension of access to ART because ART treatments take place over an extended period of time, on the order of months to years, and involve numerous office visits, on the order of 10-20 visits, for a single IVF cycle. Monetary, physical, and time-costs associated with traveling long distances for care increases the burden of infertility as well as renders treatment impossible to patients without the means of paying these associated costs.

The focus of this original work is the geographic component of accessing fertility care in the US. Several important questions can be asked about the present and potential future configurations of the US fertility service delivery system. In this thesis, I focus on three specific questions:

1. How many reproductive-age women have geographic access to fertility care in the US?
2. What local factors contribute to more demand for fertility services and a greater likelihood of a clinic operating in an area?
3. Where should new fertility clinics be located to best expand geographic access to care?

This dissertation is comprised of three essays. In Chapter 2, results of a systematic web-search of fertility clinics are presented and current geographic access to care is characterized with respect to main and satellite fertility clinics. In Chapter 3, a geospatial statistical model of fertility treatment demand is shown that provides insights into socioeconomic and environmental health drivers of ART demand. This statistical model also uses recent IVF treatment data in conjunction with socioeconomic and environmental health data to generate values for predicted ART demand, in units of number of ART cycles, for every county nationwide. The dissertation culminates with Chapter 4, which uses the clinic locations from Chapter 2 to determine *access* combined with treatment *demand* from Chapter 3 as the inputs to location analysis models in a multiobjective optimization framework. The location analysis generates location assignments to optimally site new clinics to increase access through maximizing coverage of population or demand.

To succeed at producing useful recommendations, optimization models must include representations of how proposed strategies will affect the desired outcomes. This is preferably achieved through validated models created by analysis of empirical data. Therefore, Chapter 3 of the dissertation consists of the assembly and statistical analysis of databases of empirical data to create models that describe the relationships between socioeconomic and environmental health factors and infertility treatment demand. In addition to predicting values for demand, this geospatial modeling improves the understanding of the spatially variable socioeconomic and environmental health drivers of ART demand. Datasets used include the self-created "Fertility Clinic Locations and Details in the United States" dataset for clinic locations (McGarity & Hobbs, 2021), the CDC's "ART Success Rates from 2018" dataset for ART demand (CDC, 2020a), and the Robert Wood Johnson Foundation's "County Health Rankings" (County Health Rankings & Roadmaps, 2021a) for environmental and socio-economic factors. Additionally, Chapter 3 uses these same predictors along with the clinic location data from Chapter 2 to statistically model clinic presence and applies the clinic presence model to predict the potential for an ART clinic in every US county.

Chapter 4 uses output from the models in Chapter 2 and Chapter 3 along with location analysis optimization methods to identify strategies to increase access. Predicted infertility treatment demand and reproductive-age population at the county level are used as input to location analysis models to optimally locate ART clinics to expand geographic access to care. Finally, optimization models for clinic placement are used to identify locations of highest uncovered population and unmet demand to investigate racial equity of current geographic access and identify new locations to optimally improve coverage of underserved racial groups.

The explanatory geographic access models, demand models, and location analysis models help bring new understanding into what socioeconomic and environmental factors lead to increased access and demand for fertility treatments. Because there are equity trade-offs between locating clinics to meet demand and locating clinics to cover populations, a multi-objective optimization approach which contrasts Pareto-optimal solutions was taken.

This dissertation is organized as follows. Prior to the introduction, a glossary provides definitions for acronyms, key terminology, and concepts. The three essays following the introduction describe the completed research, and are separated into Chapter 2 - US Fertility Clinics Placement Patterns: Analyses of The Impact of Satellite Clinics on Extending Geographic Access and Racial Equity, Chapter 3 - Predictive Modeling of Assisted Reproductive Technology Demand and Fertility Clinic Presence, and Chapter 4 - Expanding Geographic Access to Fertility Care: A Multiobjective Optimization Approach. Finally, in Chapter 5, conclusions are made from the three essays and potentially useful directions for future research are described.

In the remainder of this introductory chapter, I provide a broad context for the chapters in this dissertation. Section 1.2 covers how geographic access to fertility care fits into the broader context of access to fertility care and healthcare in general. Section 1.3 describes the value of ecological studies in general and how ecological studies of ART demand and clinic presence add value through hypothesis generation and prediction. Section 1.4 summarizes how generic facilities are optimally located, how healthcare facilities are optimally located, and the benefit to taking a multiobjective approach in siting fertility clinics. Finally, Section 1.5 summarizes the research scope and contributions.

1.2 Characterizing geographic access to fertility care

In order to characterize current geographic access to fertility care, it is important to understand how access to healthcare is characterized in general and when applied to fertility care. Chapter 2 explores geographic access to fertility care in the wider context of general access to healthcare. In this section, I discuss how access to healthcare is characterized in general and specific to fertility care, and I cover the studies found in the literature which have explored geographic access to fertility care.

Access to healthcare is a function of many factors in addition a clinic being located in an area. For example, Donabedian (1973) summarized access as follows:

> The proof of access is use of service, not simply the presence of a facility . . . Access can, accordingly, be measured by the level of use in relation to 'need.' . . . barriers to access are not only financial but also psychological, informational, social, organizational, spatial, temporal, and so on.

Penchansky & Thomas (1981) defined a construct to define access to care known as the five A's: affordability, availability, accessibility, accommodation, and acceptability. In the US, the National Academy of Medicine formed a committee for the Access Monitoring Project that defined access as "the timely use of personal health services to achieve the best possible health outcomes" (Institute of Medicine, 1993). In the case of access to fertility care, conceivable factors influencing access include: economic factors (income, time availability, subsidies, insurance coverage); private and public transportation efficiency and reliability; travel distance; time availability; trust of the medical establishment; willingness of providers to educate patients about treatment options; and formal rules and informal cultural norms that provide preferential access or discriminate against particular subpopulations, and finally geographic among others. In this thesis, I emphasize the impact of geographic access to fertility clinics on general access to fertility care, recognizing that it is important but not the only influential factor.

Characterizing geographic access to fertility care facilities is very similar to characterizing geographic access to generic healthcare facilities. Geographic access to healthcare facilities is generally a function of the distance, time, and out-of-pocket expense involved in travelling from where people live to facilities to receive care. This earliest recorded observation of this relationship, known as Jarvis' Law, was made by a medical doctor from Massachusetts named

Edward Jarvis who found that “the people in the vicinity of lunatic hospitals send more patients to them than those at a greater distance” (Jarvis, 1852). A century later in 1952, a study of utilization of medical services in rural Missouri found patients that lived farther from a physician had a lower physician utilization rate and more days of bed illness compared with patients living nearby. Two years later in 1954, a study was published by the US Public Health Service which found a statistically significant inverse relationship between utilization and distance traveled to a specialist physician, but not to a general practitioner. Summarizing these studies, a 1980 review paper published again by the US Public Health Service captured the role of distance in healthcare access as: “preventing the conversion of health need into demand for services” (Government Studies & Systems, 1980).

Based the belief that physicians were geographically maldistributed in the US, in 1972 the National Health Service Corps was created (Redman, 1973) to address the lower physician per population ratios in non-metropolitan areas than in metropolitan areas (Newhouse, 1990). In a 2005 follow-on study, M. B. Rosenthal et al. (2005) found improvements in geographic access to healthcare from a doubling in the number of physicians between 1979 and 1999, which resulted in diffusion of physicians to more rural areas. In particular, their study found the number of obstetrics and gynecology (OB/GYN) physicians grew from about 4k in 1979 to 7k in 1999, but it did not analyze the number of OB/GYN physicians certified in the subspecialty of reproductive endocrinology and infertility (REI), which is required to perform ART/IVF. Since the creation of the REI subspecialty in 1972, success rates of IVF have increased steadily every year, but the number of physicians completing their REI fellowships has remained constant in the past decade at around 45 per year (Gambone et al., 2015). This constant rate of REI certification is potentially a factor limiting growth and affordability of the ART industry as well as geographic access to ART care.

Geographic access to fertility care has only been examined recently in the literature. In the last decade, geographic access to US fertility clinics was first characterized by Nangia et al. (2010a) who tallied the number of reproductive-age women within a 60-minute drive of an IVF clinic. They found a higher percentage (90%) of reproductive-age women with geographic access to ART clinics in IVF insurance mandated states than in nonmandated states (71%). Later, Harris et al. (2017) took a different approach, defining geographic access as living in a US Census Core-Based Statistical Area (CBSAs) that has an ART clinic. That study found about 29% of the female reproductive-age population lived in an area with no fertility clinics, and 40% lived in an area with a single or zero fertility clinics. However, Nangia et al. (2010a) and Harris et al. (2017) only studied access to main ART clinics, which have embryology labs, and did not include satellite clinics in their analyses. Most satellite clinics play a critical role in ART cycles by supporting the regular monitoring required for proper ovarian stimulation. Satellite clinics could be clustered in areas that already have geographic access to fertility care to compete for patients as described in Hotelling's Law in economics (Hotelling, 1929) (defined in Section 2.1.1) or they could be located to expand geographic access to new areas. Therefore, Chapter 2 of this dissertation studies the role of satellite fertility clinics on geographic access to care, investigates if state IVF insurance mandates influence the rate of satellite clinics per population, and explores their relationship to the annual number of IVF cycles a practice performs.

1.3 Ecological studies predicting assisted reproductive technology demand and fertility clinic presence

In order to produce predictions of where demand for ART is unserved and where promising locations exist for siting new ART clinics, ecological studies are utilized in Chapter 3. These ecological studies use predictive models using population-level socioeconomic and environmental health factors as inputs to predict ART demand and clinic presence for every US county. In this section, I discuss the role of ecological studies in the context of ART, and the potential drivers of ART demand and clinic presence.

Ecological studies of demand and utilization of ART treatments and the presence of fertility clinic is limited in the literature. Ecological studies are observational studies performed at the population or group level vs. the individual level, and they are performed to “monitor population health so that public health strategies may be developed and directed” (Levin, 2006). Although there are many ecological studies for many diseases and chronic conditions (Pickett & Wilkinson, 2007; Tu & Ko, 2008), only one study was found in the literature that ecologically examined the drivers of ART utilization (Hammoud et al., 2009), and ecological studies on ART clinic presence were not found in the literature. This is likely due to the aversion among epidemiologists to perform ecological studies due to concerns over ecological fallacies (Idrovo, 2011) with ecological studies considered appropriate for hypothesis generation and not appropriate for hypothesis testing (Ben-Shlomo, 2005). However, hypothesis generation in the context of ART utilization and clinic presence would be a valuable addition to literature on ART because it could direct future studies exploring potential causal relationships identified in ecological studies. Additionally, the output predictions from ecological studies of ART demand and clinic presence are the overarching purpose of the studies, and their identified drivers are helpful for generating trust in the models’ predictions. With the predictive purpose in mind, it is helpful to understand what population-level factors have been found important in the literature to inform selection of similar predictors utilized in predictive models of ART demand and clinic presence.

Several studies have investigated the factors potentially driving ART demand and clinic presence. With IVF costing roughly half of US median household income (Katz et al., 2011), plus the situation that Catholic religious doctrine forbids IVF, economic factors and religious views have been identified as the most important barriers to IVF (Lunenfeld et al., 2004). Economically, most patients pay out-of-pocket for IVF. There is not a federal mandate for health insurance coverage of IVF, but as of April 2021, 13 states have passed laws requiring IVF insurance coverage (RESOLVE, 2021c). These “IVF mandates” have been shown to increase utilization rates (Crawford et al., 2016), but these mandates do not apply to all insurance plans held in each state. For example, in 2020, about half of Americans obtained their health insurance through their employer (Katerine Keisler-Starkey & Lisa Bunch, 2021), and about 60% of employees have insurance from a company that self-insures (Fronstin, 2012) and are exempt from state IVF mandates. Therefore, employer coverage of IVF is largely optional, and a 2015 survey of 462 mainly large, private employers found only 27% included coverage for IVF (SHRM, 2015). A 2002 health economics study of IVF estimated a 10% decrease in treatment cost would result in a 30% increase in utilization (Collins, 2002). In summary, the economics of ART indicate that its high cost due to it generally not being covered by health insurance is a primary barrier to accessing care.

The impact of religion is often cited as a barrier to access and utilization of ART, but evidence does not support this claim. The Catholic church prohibits its members from using ART even to treat infertility, but in practice, it still allows, for example, Penn State Health to operate an ART center in a Catholic health facility with the required disclaimer that it “abides by the U.S. Bishops’ Ethical and Religious Directives for Catholic Health Services” (Penn State Health, 2022). This disclaimer may have turned away some Catholics, but the center still performed 2,950 ART cycles in 2018 alone (CDC, 2020a). Surveys on religion cannot be performed by the US Government because it is unlawful for it to collect information on religious affiliation (Public Law 94-521, 1976), but a 2013 Pew Research survey on a representative sample of 4,006 adults nationwide found only 13% of Catholics, 13% of Protestants, and 9% of those religiously unaffiliated believe IVF is morally wrong (Pew Research Center, 2013). In conclusion, religion is likely not a primary barrier to ART access, and it probably only reduces utilization slightly.

Whereas religious and economic factors act as limiters of utilization of ART, environmental exposures increase the risk of infertility and therefore increase the need for ART treatments. The CDC lists the following factors as being known to reduce female fecundity: age, smoking, excessive alcohol use, extreme weight gain or loss, and excessive physical or emotional stress that results in amenorrhea (absent periods) (CDC, 2021). The existing literature has established a strong relationship between environmental exposures and reproductive health (Schettler et al., 2000), and some studies suggest that the environment plays a role in fertility treatment effectiveness (Carré et al., 2017b; Choe et al., 2018). At the national scale, a few studies have investigated spatial and temporal infertility trends (Odisho et al., 2014; Swan et al., 2003). Hammoud et al. (2009) was the only study found in the literature that ecologically modeled IVF utilization as a function of demographic, social, and economic factors, but it did not explore environmental factors. Of the literature reviewed, ART utilization or the presence of fertility clinics has never been studied in relation to environmental exposures despite the strong evidence indicating exposures to certain toxicants reduce fecundity, creating the need for ART treatment. Large-scale ecological studies are a promising approach for uncovering potential unknown regional patterns relating infertility treatment demand and clinic presence to environmental health factors that may warrant further study and investigation. I present an initial investigation of this type in Chapter 3.

1.4 Application of optimization to siting new fertility clinics

In order to produce optimal recommendations for where to site new fertility clinics, it is important to understand how generic facilities and specifically healthcare facilities are optimally sited. Chapter 4 uses the maximal covering location problem (MCLP) optimization model to site new fertility clinics that is commonly used to locate healthcare facilities to maximize the number of people covered by placed facilities. In this section, I describe where the MCLP model fits into the wider literature of optimal facility placement and how applying the MCLP model to place new fertility clinics is a valuable contribution to the literature.

Historically, the problem of optimal siting of facilities began with Alfred Weber who considered the problem of locating a factory on a plane between two resources and a single market such that the sum of weighted Euclidian distances is minimized (Weber, 1962). Weber's factory location problem applies to the private sector with its overarching goal of profit maximization, as opposed to public sector goals, which might include maximizing societal utility subject to a budgetary constraint on the number of facilities (Revelle et al., 1970). ReVelle & Swain (1970) formulated a mixed integer program to locate facilities along a road network to minimize the population weighted travel costs with the *p*-median problem. Subsequently, Church & ReVelle (1974) formulated a location set covering problem known as the maximal covering location problem (MCLP) which maximizes the population covered within a distance or travel time.

Optimal siting of healthcare facilities has a rich literature. Schilling (1976) applied the MCLP to optimally locate fire stations in Baltimore, MD in a multi-objective framework with competing objectives of property value covered vs population covered. ReVelle et al. (1977) applied MCLP to optimally place emergency medical services (EMS) to maximize population covered. The MCLP methodology for siting healthcare facilities has also been applied in contexts both urban (Eaton et al., 1985) and rural (Bennett et al., 1982). More recently, capacitated MCLP was applied to locate preventative healthcare facilities (Verter & Lapierre, 2002), and a modular capacitated MCLP was applied to locate emergency vehicles (Yin & Mu, 2012).

Studies evaluating the optimal siting of fertility clinics were not found in the literature. Geographic access to fertility clinic facilities has been examined by a few studies (Gharani et al., 2015; Harris et al., 2017; Nangia et al., 2010a), but there are no applications of location analysis to improve fertility clinic access. The literature has focused on understanding where geographic access exists from existing clinics and what factors are contributing to better or worse access, but no studies reviewed have taken the next step to ask where *new* clinics should be placed to best increase geographic access to care. This step is the focus of Chapter 4 of this thesis.

1.5 Research scope and contributions

The primary contributions of this thesis are as follows. First, a geospatial analysis of fertility clinics in the US is performed to improve the understanding of geographic access to care, including satellite clinics, which have not been studied previously (Chapter 2). Then, ecological studies of ART utilization and clinic presence are performed at the national scale using county-level socioeconomic and health metrics as predictors (Chapter 3). These studies create estimates of ART demand and clinic presence for each county in the US using county socioeconomic and environmental factors in addition to population. This approach has not been taken previously in the literature and potentially provides an estimate of demand at a high enough spatial resolution to allow for the ART industry to identify new areas of unmet demand and locations likely to support new fertility clinics. Finally, location analysis is applied to identify locations for new fertility clinics via the MCLP framework to optimally expand geographic access to fertility care (Chapter 4) to address the gap in the existing literature, which has only characterized access from existing clinics and has never studied where new clinics should be placed.

This thesis takes a novel approach in the context of healthcare facility placement by adopting a multi-objective optimization framework including competing objectives of demand and population. This thesis argues that for facility location problems in which there are several factors contributing to demand, solely using population to estimate demand is insufficient. The rationale for this argument is that for-profit businesses select facility locations to maximize profit by covering the most demand. The alternative strategy of covering the most population arguably has the greatest social utility, but the identified optimal locations could be unviable economically because of insufficient demand. But by taking a multi-objective approach, balanced solutions can be found that are hopefully good for both business and society.

Chapter 2 US Fertility Clinics Placement Patterns: Analyses of the Impact of Satellite Clinics on Extending Geographic Access and Racial Equity

2.1 Introduction

2.1.1 Background and literature review

Geographic access is a major barrier to utilizing assisted reproductive technology (ART) for millions of infertile women and men throughout the United States. Only an estimated 25% of demand for ART is met in the US (Chambers et al., 2009) based on the standard demand estimate of 1,500 cycles per million population per year (Collins et al., 2001), and as shown in the following studies, geographic access is an important contributor to this failure to meet demand. Using 60-minute driving time to define access, a 2010 study examined regional differences in geographic access to care and found disparities in geographic access to ART when comparing US geographic regions and when comparing states with and without IVF mandates (Nangia et al., 2010a). A 2017 study using US Census Core-Based Statistical Areas (CBSAs) to define geographic access estimated that 28.8% of reproductive-age women in the US do not live in an area with a fertility clinic (Harris et al., 2017). However, these studies used only ART clinics that have on-site embryology labs ("main clinics") and did not examine the role of clinic sites without embryology labs on-site that do not perform *in vitro* fertilization (IVF) procedures. This is a crucial omission, because such clinics ("satellite clinics") provide non-ART fertility treatments and also play a critical role in supporting ART cycles through consultations, basic infertility diagnostic evaluation, bloodwork, and ultrasound monitoring.

In this thesis, I correct this shortcoming of the current literature by creating a publicly available dataset of all fertility clinics in the US that includes the location of both main clinics and satellite clinics. This comprehensive dataset, including both fertility clinic types for all ART practices that report to the CDC, is the first available to the community. Analysis on this dataset in this thesis allows for several insights to be gained with respect to geographic access to fertility care. An improved understanding is gained into how many people in the US have geographic access, whether state IVF insurance mandates influence the number of main or satellite fertility clinics per capita in a state, and if satellite clinics are generally used to reach underserved geographic areas or compete for patients in areas with existing access.

Satellite clinics could be opened by an ART practice to extend the geographic reach of its main clinics, or they could be opened to compete for patients in an area with an existing clinic. The tendency for competing businesses to locate near each other is described by Hotelling's Law. This law of economics states that competing sellers will tend to locate close to each other to maximize their individual market share, and "as more and more sellers of the same commodity arise, the tendency is not to become distributed in the socially optimum manner but to cluster unduly" (Hotelling, 1929).

The role satellite clinics play in extending geographic access to ART services has not been studied. To assess the role of such clinics in providing access requires data on their location vis-à-vis populations of potential users. But even knowing where all fertility clinics in the US are located is non-trivial. The locations and use of satellite clinics in the US are not reported to or tracked by the Centers for Disease Control and Prevention (CDC). ART programs are only required to report clinic locations with embryology labs to the CDC (CDC, 2020a), as stipulated by the "Fertility Clinic Success Rate and Certification Act of 1992" (Fertility Clinic Success Rate and Certification Act of 1992, 1992). The Society for Assisted Reproductive Technology (SART) provides a "Find an IVF Clinic" tool that contains satellite clinic locations for SART member practices (Society for Assisted Reproductive Technology, 2021), but it does not differentiate whether a given clinic is a satellite clinic or a main clinic. Also, SART's tool does not include non-member clinics, and 20% of all clinics reporting to the CDC in 2018 were not SART members (CDC, 2020a).

To my knowledge, satellite clinic locations have never been analyzed or even been recorded in a comprehensive list. Because this information is critical to understanding the role of satellite clinics, I conduct a systematic search for satellite clinics and a geospatial analysis in this thesis in order to better understand how many satellite clinics exist, where they are located, and how they impact geographic access to ART.

Finally, no studies found in the literature have investigated the relationship between race/ethnicity and geographic access to ART. It is unknown whether disparities exist based on race or ethnicity in geographic access to ART. Investigating racial access to ART based on access to main or satellite clinics could shed light on the specifics of racial disparities in geographic access to care.

2.1.2 Contribution

In this thesis, I correct this shortcoming of the literature first by creating a comprehensive dataset of all US fertility clinics that includes both main and satellite clinics. This dataset is used in the geospatial analysis of clinic placement patterns of this chapter to improve the understanding of where clinics of both types have been sited. I hypothesize (H1) that satellite clinics will follow clinic placement patterns of main clinics found in the literature and be placed primarily in large cities concentrated in the Northeast because it was found in the literature to be "overserved" (Nangia et al., 2010a), and I hypothesize (H2) that there will be significantly more fertility clinics per capita in IVF mandate states. Additionally, I hypothesize (H3) that areas with one or more main clinics have more reproductive-age women than areas with one or more satellite clinics but no main clinics, and areas with no clinic of either type will have the fewest reproductive-age women.

Second, I then analyze driving times between satellite clinics and main clinics to better understand if satellite clinics are generally used to improve geographic access to care or instead to compete for patients in areas with existing geographic access. I consider several hypotheses. One is that I hypothesize (H4) that satellite clinic placement patterns will follow Hotelling's Law (Hotelling, 1929) and "cluster unduly" around other ART practices' main clinics. I further hypothesize (H5) that the more satellite clinics an ART practice has, the more ART cycles it will perform because it is able to compete for more patients and it will be easier for patients to complete ART cycles with greater convenience to satellite clinics. I also hypothesize (H6) that fertility clinics are more likely to locate in areas with larger populations and higher incomes because they are likely drivers of higher demand for ART. Finally, since Asian/Pacific Islander women have the highest utilization rates of ART and American Indian/Alaska Native women have the lowest (Dieke et al., 2017), I hypothesize (H7) that these racial groups will respectively have the best and worst geographic access to ART.

2.1.3 Chapter scope

This essay (Chapter 2) is organized as follows. First, all of the methods used in the essay are detailed (Section 2.2). Then, results are reported including the findings of the systematic web-search, tests of hypotheses, and a case study in identifying promising areas to locate new clinics (Section 2.3). The discussion in Section 2.4 explores the larger implications of this essay's results. Finally, a summary is given in Section 2.5.

2.2 Methods

This essay's methods are detailed in this section, starting with how the locations of all fertility clinics in the US are collected in Section 2.2.1 and how reproductive-age population is estimated in Section 2.2.2. Next, statistical methods are described to model clinic presence in Section 2.2.3, clinic counts in Section 2.2.4, and the impact of state IVF insurance mandates in Section 2.2.5. Finally, the approach to explore racial and ethnic geographic access to care is explained in Section 2.2.6. Specific software used and study approval details are covered in Sections 2.2.7 and 2.2.8, respectively.

2.2.1 Systematic web-search: Data collection of main and satellite clinic locations

The approach taken to collect the locations of all main and satellite ART clinics across the US was to perform a systematic web-search of each clinic listed in the CDC's ART Success Rates dataset from the most recent year available to establish the number and location of all satellite clinics. The search was performed in December 2020 and January 2021. The CDC's 2018 Success Rates dataset (CDC, 2020a) was used to identify practices to find online. This systematic search consisted of using the Google search engine to search for the first and second listed names of the practice, the listed city/state, and the name of the medical director. This search method generally resolved the issue that arises when two clinics have similar practice names but different medical directors.

The data were collected using a web form and stored in a database using the ArcGIS Survey123 platform. Fields collected include: practice name, clinic location's name, medical director, clinic's website, address, phone number, fax number, days of the week clinic operates, geographic location, type of clinic, if an embryology lab is onsite, if clinical services are provided at the clinic, and if the clinic is open to the general public. A screenshot of the web form used is shown in Appendix A.

2.2.2 CBSA population data

To assess the impact of satellite clinics on geographic access to ART, core-based statistical areas (CBSAs) published by the United States Census Bureau were used to define geographic access, similar to Harris et al. (2017). The 938 CBSAs are defined by the US Census Bureau as metropolitan and micropolitan core areas plus adjacent territory with a “high degree of social and economic integration with the core as measured by commuting ties” (U.S. Census Bureau, 2021). In 2019, the American Community Survey (ACS) 5-year estimate of the total US population was 324,697,795 (U.S. Census Bureau, 2019b), and the total US population living in CBSAs was estimated at 309,804,779 (U.S. Census Bureau, 2019a) or 95% of the total US population. With respect to US female reproductive-age population, of the 63,961,819 total women age 20-49 (U.S. Census Bureau, 2019b), 96% are estimated to live in CBSAs (U.S. Census Bureau, 2019a).

2.2.3 Statistical modeling of clinic presence in a CBSA

To better understand the factors potentially driving the number of main or satellite clinics, the presence of clinics in a CBSA was modeled as a function of CBSA female reproductive-age population and CBSA median income using logistic regression (shown in Equation 2-1) with standardized covariates, also known as *z*-scores (Bock et al., 2007). The z-scores for reproductive-age population and median income were calculated by first “centering” by subtracting the mean and then “scaling” by dividing by the standard deviation. Upon fitting the regression model, variable importance was calculated using the absolute value of the *t*-statistic, computed from the covariate’s regression coefficient divided by its standard error.

Logistic regression is a form of generalized linear models (GLM) commonly used in epidemiological research when modeling binary response variables that are true or false. When using logistic regression to model presence of ART clinics, an assumption is made that the presence of ART clinics is Bernoulli distributed, defined by a single parameter p representing the probability. The *logit* function of p is assumed to be a linear combination of fit parameters (Waller & Gotway, 2004). These relationships are shown in Equation 2-1.

Equation 2-1. The mathematical relationships of logistic regression with probability p to model clinic presence as a function of CBSA median income and female reproductive-age population.

$$\mathrm{Y} = \begin{cases} 1 \; if \; ART \; clinic \; present \\ 0 \; if \; ART \; clinic \; is \; not \; present \end{cases}$$

$$\mathrm{Y} \sim \mathrm{Bernoulli}(p)$$

$$logit(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 Median\ Income_{z-score} + \beta_2 Female\ Repr.Age\ Pop_{z-score}$$

2.2.4 Statistical relationship between clinic counts and practice number of ART cycles

The number of main and satellite clinics was also investigated at the practice level. Using the CDC’s 2018 ART-Success Rates dataset, the clinics found in the systematic web-search were associated with their corresponding ART practice using the name of the medical director for the ART practice. For clinics whose names and addresses were matched to the 2018 Success Rates dataset from the CDC, it was assumed that the medical director was the medical director listed in the CDC dataset. Otherwise, the clinic’s medical director was inferred based on information on the clinic’s website or the medical practice’s closest nearby clinic with a matching medical director. The number of ART cycles a practice performs was modeled as a function of the number of main and satellite clinics using a negative binomial generalized linear model (GLM) because the mean and variance of the number of cycles were found to be unequal and over-dispersed. The negative binomial model was selected for its ability to model over-dispersed count response variables (Cameron & Trivedi, 2013). Negative binomial regression is a Poisson-gamma distribution mixture model with a mean of μ and a variance of $\mu + \mu^2/\theta$. This two-stage model with an unobserved random variable E distributed $gamma(\theta)/\theta$ and the Practice's Number of ART Cycles is Poisson distributed with mean μE (Venables & Ripley, 2002). These mathematical relationships are shown in Equation 2-2.

Equation 2-2. The mathematical relationships of the Negative Binomial GLM with mean μ and fit θ, modeling a practice's number of ART cycles as a function of the number of main or satellite clinics it has.

$$\text{Practice's Number of ART Cycles} \mid \text{E} \sim \text{Poisson}(\mu E), \qquad E \sim gamma(\theta)/\theta$$

$$\log(\mu) = \beta_0 + \beta_1 * (number\ of\ main\ or\ statellite\ clinics)$$

2.2.5 Role of state IVF insurance mandate

Additionally, the number of clinics per million women aged 20-49 was calculated for each state, and each state was grouped based on if it had an insurance IVF mandate implemented before 2018. The states with insurance mandates for IVF included: AR, CT, HI, IL, MA, MD, NJ, and RI (RESOLVE, 2021c). For main and satellite clinics, each state's average number of clinics per capita was calculated, and Welch's independent one tailed *t*-test (assuming unequal variances) (Bock et al., 2007) was performed, comparing the group means of clinics per million women between states with and without IVF mandates.

2.2.6 Racial and ethnic geographic access to ART

To investigate if geographic access to ART is equitable based on race and ethnicity, the 2019 American Community Survey (ACS) 5-year estimate for population by race (U.S. Census Bureau, 2019e) and ethnicity (U.S. Census Bureau, 2019c) were used along with the fertility clinic locations from the web-search spatially joined to the corresponding CBSA. The racial categories the US Census Bureau collects include: White alone, Black or African American alone, American Indian and Alaska Native alone, Asian alone, Native Hawaiian and Other Pacific Islander alone, Some Other Race alone, and Two or More Races. The ethnic categories the US Census Bureau collects include: Not Hispanic or Latino, and Hispanic or Latino. The percentage of all US racial and ethnic groups in CBSAs with access to either a main or satellite clinic was calculated, and the percentages for each racial and ethnic category with access to either a main or satellite clinic were calculated. Pearson's chi-squared test (Bock et al., 2007) was used to statistically test for differences among the racial and ethnic groups.

2.2.7 Software used in analysis

ArcGIS Survey123 was used to collect and store, in a cloud-based database, the locations and details of fertility clinics found during the systematic web-search. R 4.0.5 was used for statistical analyses. ArcGIS Pro 2.7.3 (ESRI Corp) was used for GIS analysis and for calculating driving times (regardless of time of day) between main and satellite clinics using a network routing layer from ArcGIS 2019 Business Analyst.

2.2.8 Study approval

The Johns Hopkins Institutional Review Board determined that the study did not require the IRB's review and approval because only publicly available data were used.

2.3 Results

The results section first covers the findings from the systematic web-search on the locations of all main and satellite clinics in the US (Section 2.3.1). Next, each hypothesis posited in the introduction is tested (Section 2.3.2). Finally, a case study in identifying areas for new clinics is presented in Section 2.3.3 that identifies new areas to locate economically viable clinics compared with new areas to address racial disparities in utilization of ART.

2.3.1 Systematic web-search findings for main and satellite clinics

The systematic-search criteria identified 441 unique practices based in the United States. Together, these practices had a total of 1,052 clinic locations, comprising 469 main clinics and 583 satellite clinics in the US. The locations of all the clinics in the US were compiled in a user-friendly Fertility Clinic Finder web application (McGarity, 2021) that allows users to enter their address and find their closest main or satellite clinic and includes information gathered in the web-search. A screenshot of the web app is shown in Appendix B. Additionally, the dataset was archived with the Johns Hopkins University Data Archive and is publicly available for download (McGarity & Hobbs, 2021).

2.3.1.1 Main clinics found

Of the 441 practices found, 1% (n = 5) did not have a main clinic and used another practice's embryology lab, and in effect operated a satellite clinic at the address listed in the CDC dataset; 93% (n = 411) had one main clinic; 5% (n = 20) had two main clinics; and 1% (n = 5) had three or more main clinics. Most ART programs report their embryology labs separately and were treated as separate practices, but some programs with multiple embryology labs only have a single record in the CDC Success Rates dataset and were treated as a single practice.

There were also a few inconsistencies found between the CDC's Success Rates dataset and the data from the web-search. For example, there were several practices that were listed in the Success Rates dataset that do not actually have an embryology lab at the city and state listed for the practice. For example, a practice located in Illinois has its administrative office's city/state listed in the Success Rates dataset, not the location of its embryology lab which is nearly a 40-minute drive away. Two other clinics located in New Jersey reported results to the CDC but do not operate embryology labs at their listed city/state. Instead, they partner with a separate practice and use its embryology lab about a 20-minute drive away. The physicians at these clinics direct IVF cycles and perform the egg retrievals, so their reporting to the CDC is separate from the practice that operates the embryology lab, but their listed locations in the Success Rates dataset are not the location of the embryology lab but the location of their separate clinical offices which are essentially satellite clinics. These discrepancies found between the listed locations and the actual locations support the need for the systematic web-search to determine accurate locations of main clinics with embryology labs and satellite clinics versus using the CDC's Success Rates dataset alone.

Additionally, several practices that operate multiple embryology labs report only the location of one embryology lab. For example, one practice operates embryology labs in New York, San Francisco, and Los Angeles but it is only listed in New York, NY in the 2018 Success Rates dataset. In some cases, this was due to new embryology labs recently starting up and a reporting lag between when a clinic was opened and when its data are released by the CDC. But in other cases, the practice of underreporting embryology labs was caused by CDC guidance to ART practices that instructs them to group reporting of multiple embryology labs if there is the perception to the public that they are essentially the same practice. This guidance from the CDC does not appear to be consistent with the 1992 statute which required the annual reporting of "identity of each embryo laboratory" (Fertility Clinic Success Rate and Certification Act of 1992, 1992).

The final benefit of using the web-search dataset vs. the Success Rates dataset alone is that it does not suffer from the two-year lag in reporting, and as such gives the most up-to-date picture of the locations of fertility clinics across the country. For example, a practice in Maryland that closed in July of 2020 will no longer be listed in the Success Rates dataset starting in 2021, which will not be released until 2023. Similarly, a new practice in Florida opened in late 2018, but it was not listed in 2018 Success Rates dataset. It will likely be listed in the 2019 Success Rates dataset, but this will not be released until 2021. This two-year lag in Success Rates reporting could present a challenge to the industry in locating new clinics from an inaccurate picture of where clinics are operating in the present.

2.3.1.2 Satellite clinics found

Of the 441 practices found, 50% (n = 221) had zero satellite clinics, 18% (n = 81) had one satellite clinic, 12% (n = 54) had two satellite clinics, 8% (n = 36) had three satellite clinics, 4% (n = 17) had four satellite clinics, and 7% (n = 32) had five or more satellite clinics.

By and large the satellite clinics that I found appear from their websites to be owned and operated by an ART center as part of the outreach of a nearby main clinic, but for a few ART centers, the satellite clinics found were owned and operated by partnering OB/GYNs. For example, the only ART clinic that serves the State of Alaska is a partner OB/GYN practice for an ART practice in Minnesota.

2.3.1.3 Non-US clinics found

Five clinics associated with practices based in the US were found located abroad. Three satellite clinics were found outside the US, one each in Bermuda, Canada, and Mexico. Two main clinics were located outside the US, one each in Mexico and China. Aside from reporting these findings, ART facilities located outside the US were excluded from this study.

2.3.2 Tests of hypotheses

In Section 2.1.2, above, seven hypotheses were offered. In this section, the hypotheses and the results of the analysis are summarized as follows:

H1: Satellite clinics will be placed primarily in large cities concentrated in the Northeast: *Confirmed* (see Section 2.3.2.1, below)

H2: There will be significantly more fertility clinics per capita in states with IVF insurance mandates: *Confirmed for satellite clinics, not for main clinics* (see Section 2.3.2.2, below)

H3: Areas with one or more main clinics have more reproductive-age women than areas with one or more satellite clinics but no main clinics, and areas with no clinic of either type will have the fewest reproductive-age women: *Confirmed* (see Section 2.3.2.3, below)

H4: Satellite clinic placement patterns will follow Hotelling's Law and cluster unduly near main clinics: *Confirmed* (see Section 2.3.2.4, below)

H5: The more satellite clinics an ART practice has, the more ART cycles it will perform*:* *Confirmed* (see Section 2.3.2.5, below)

H6: Fertility clinics are more likely to locate in areas with larger populations and higher incomes: *Confirmed* (see Section 2.3.2.6, below)

H7: Racial groups with the most and least utilization of ART will respectively have the best and worst geographic access to ART: *Confirmed* (see Section 2.3.2.7, below)

2.3.2.1 H1 - Geographic coverage of main and satellite clinics

Of the 938 CBSAs, 137 CBSAs contained at least one main clinic, and in total cover 46M (72%) US reproductive-age women; 91 CBSAs did not contain a main clinic but contained at least one satellite clinic covering 5.1M (8%) reproductive-age women; and 710 CBSAs did not contain either clinic type, leaving 13M (20%) reproductive-age women uncovered. Further, 78% (n = 451) of satellite clinics were found in CBSAs that also have a main clinic, and 22% (n = 129) of satellite clinics were found in CBSAs without a main clinic. All main clinics were located within a CBSA, but there were three satellite clinics not located within a CBSA. A map depicting these findings is shown in Figure 2-1.

Figure 2-2 illustrates the concentration and geographic distribution of clinics by CBSA for main clinics and satellite clinics, respectively. Both main clinics and satellite clinics are located across the US. However, the CBSAs with the highest concentration of clinics are generally urban centers.

The number of clinics with respect to US regions was also studied. Regionally, the number of main clinics per practice was 1.1 in the North, South, and Midwest, and 1.0 in the West and Puerto Rico. There was greater variation in the number of satellite clinics per practice regionally. The rate of satellite clinics per practice (ordered from most to least) was 1.8 in the Northeast, 1.5 in the Midwest, 1.3 in the South, 1.0 in the West, and 0.3 in Puerto Rico.

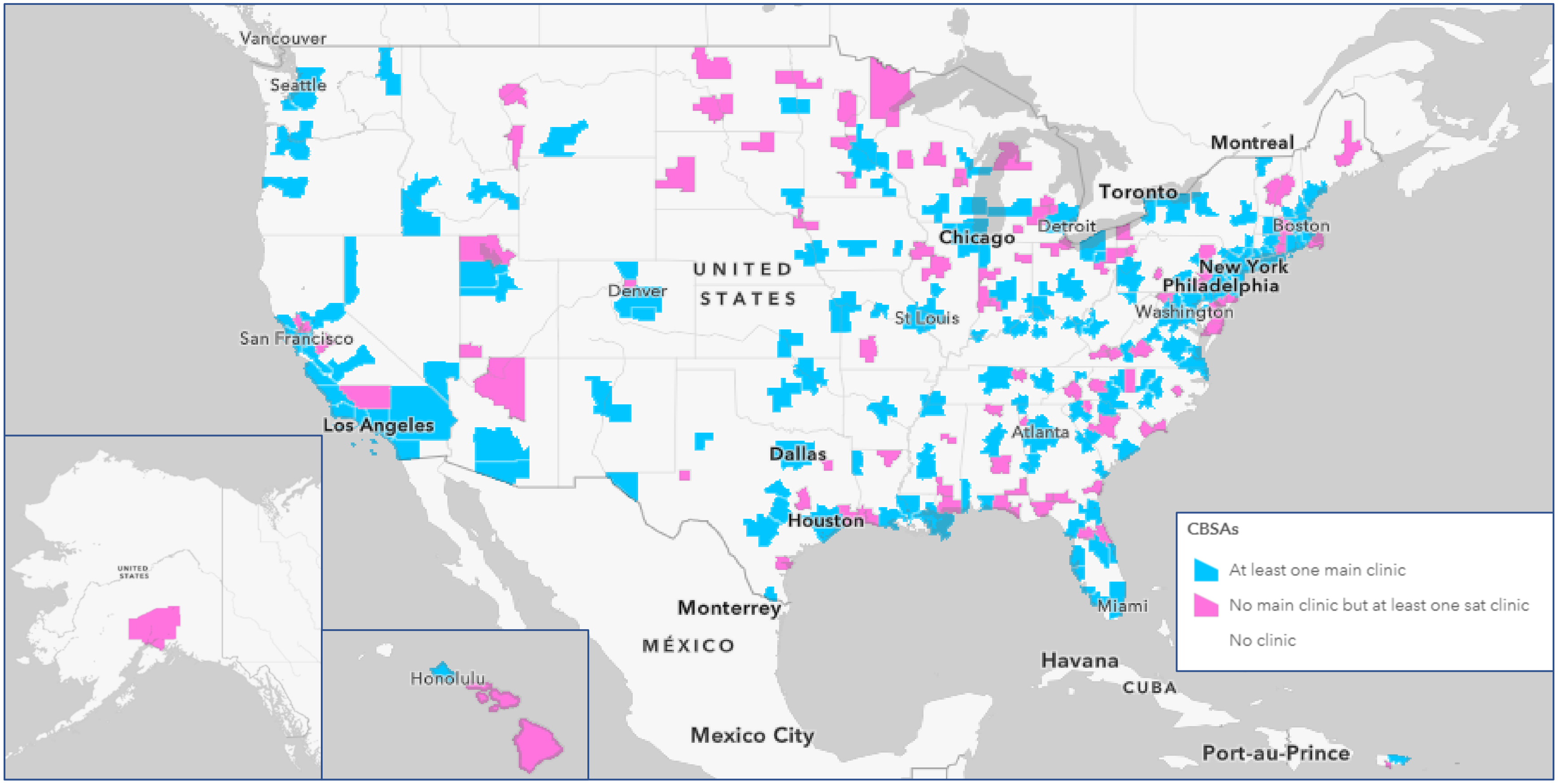


Figure 2-1. The geographic distribution of main and satellite clinics in the US. CBSAs with at least one main clinic are colored in blue, the CBSAs without a main clinic but with at least one satellite clinic in pink, and CBSAs without any main or satellite ART clinics in grey. Note: all main clinics were located in a CBSA, but three satellite clinics were not located in a CBSA.

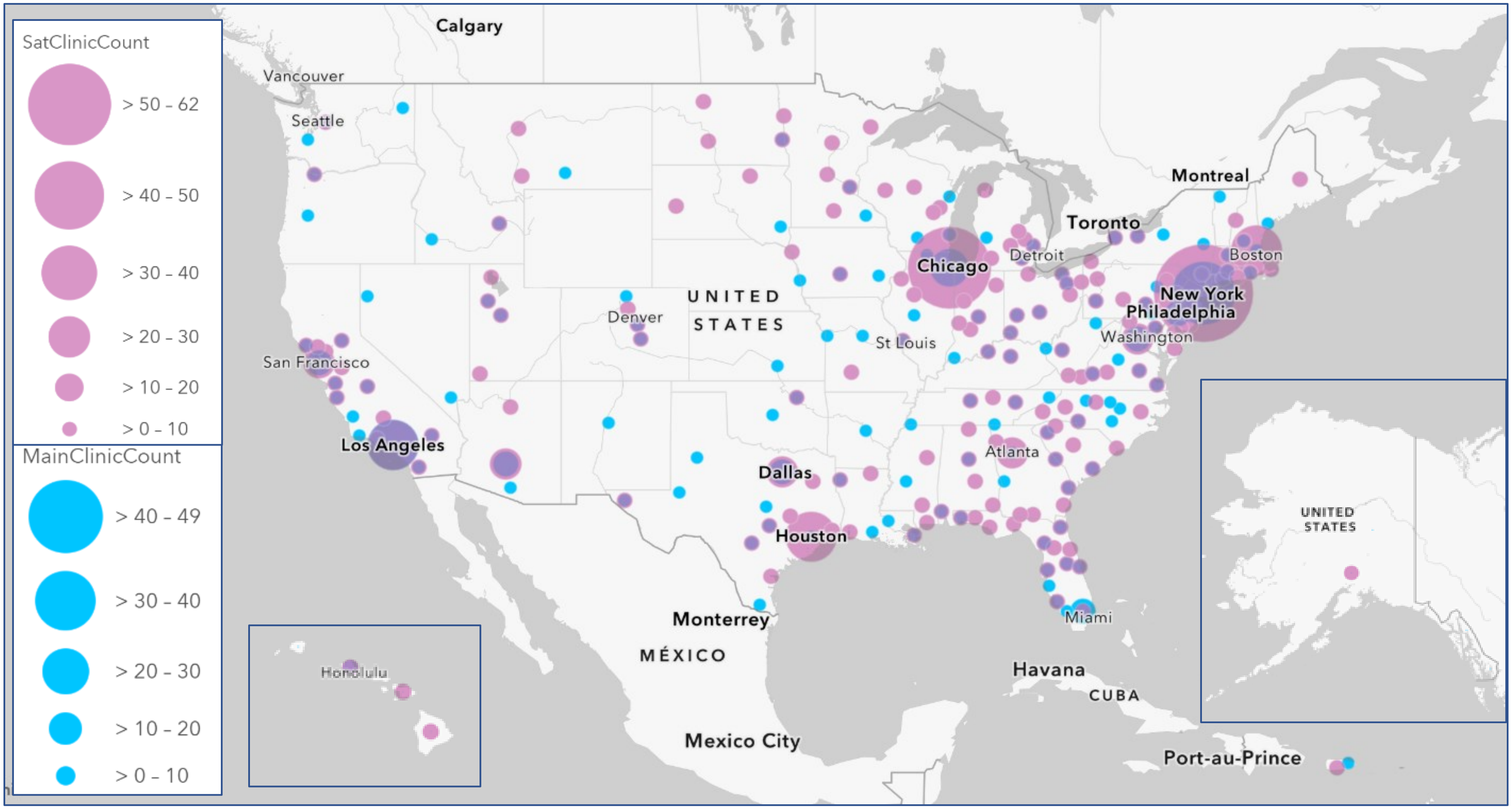


Figure 2-2. Geographic distribution of fertility clinics sized by the number of main clinics in blue and satellite clinics in pink. Areas with both main and satellite clinics appear purple due to transparency of the satellite clinic layer.

2.3.2.2 H2 - Impact of state IVF insurance mandates on the number of clinics per capita

The impact of a state's IVF insurance mandate on the number of main and satellite clinics per million reproductive-age women was also examined. A significant difference was found for satellite clinics, with a difference of means of six more satellite clinics per million women ($p = 0.04$) in insurance mandate states. For main clinics, however, this comparison found a non-significant difference in means of two more main clinics per million women ($p = 0.18$).

2.3.2.3 H3 - Coverage efficiency of CBSAs with fertility clinics

CBSAs with no main clinics and at least one satellite clinic were found to have an average of 55k women apiece. In contrast, CBSAs with at least one main clinic were found to have an average population of 343k reproductive-age women. Coverage efficiency of CBSAs with at least one main clinic ranged from 16k to 4M, and coverage efficiency of CBSAs without a main clinic but at least one satellite clinic ranged from 8k to 182k. Plots comparing the efficiency of CBSAs with at least one main clinic to CBSAs without a main clinic but at least one satellite clinic are shown in Figure 2-3 and Figure 2-4 respectively.

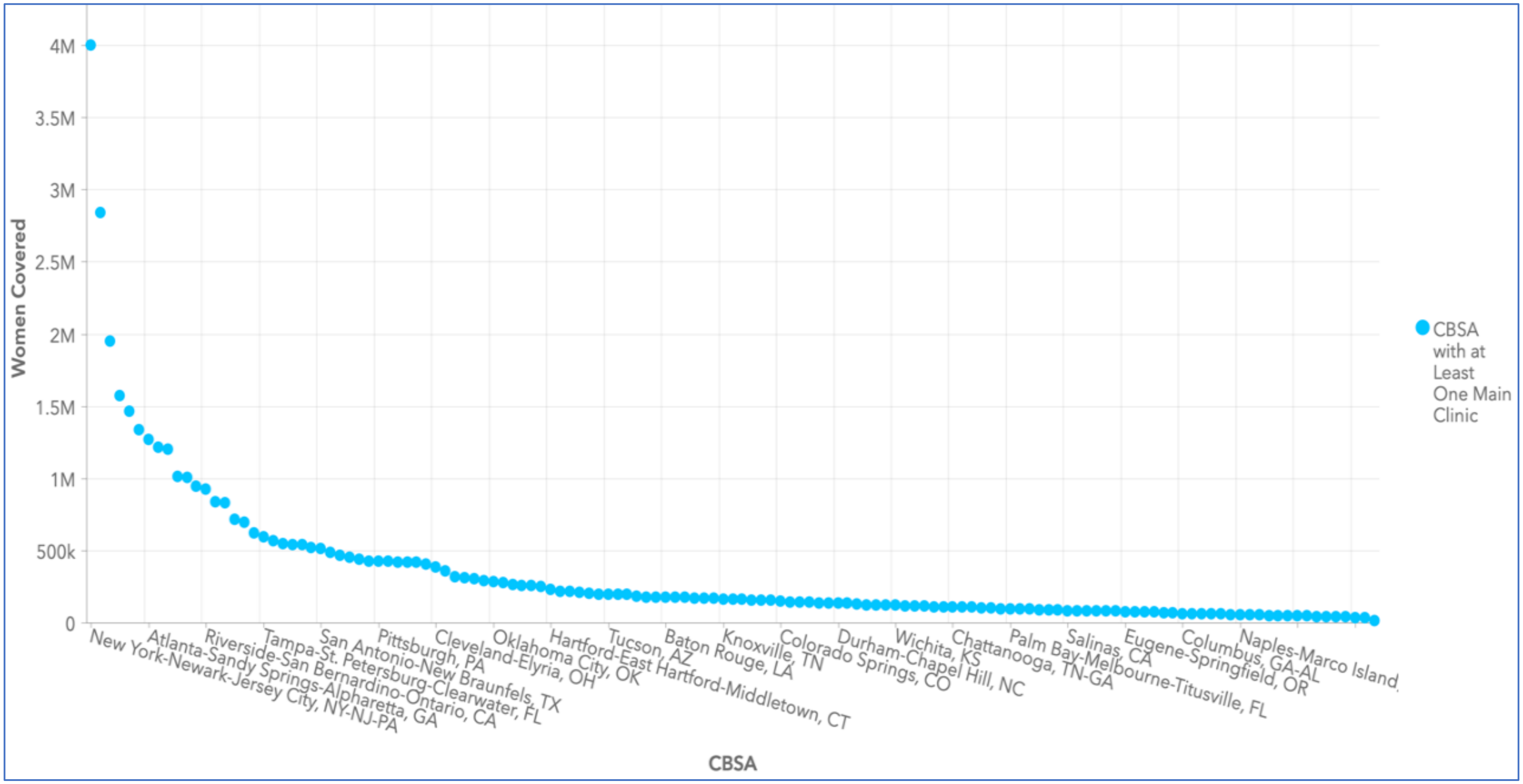


Figure 2-3. Coverage efficiency of CBSAs with at least one main clinic. Note: CBSAs listed on horizontal axis are CBSAs that fall on vertical lines and are a subset of the CBSAs with one or more main clinic.

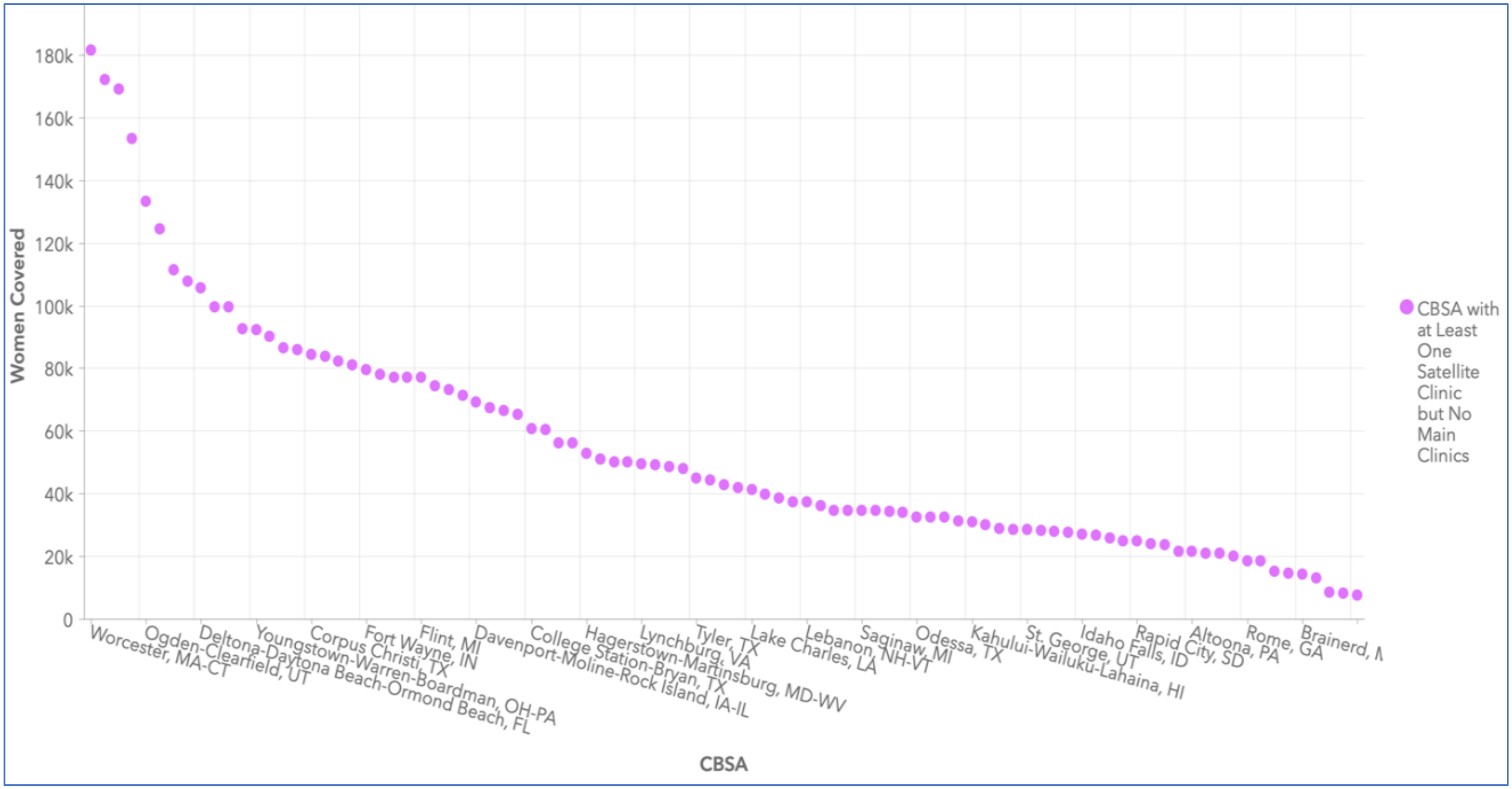


Figure 2-4. Coverage efficiency of CBSAs without a main clinic but with at least one satellite clinic. Note: CBSAs listed on horizontal axis are CBSAs that fall on vertical lines and are a subset of the CBSAs with one or more satellite clinic but no main clinic.

To better understand cumulative CBSA coverage efficiency, I have created a stratified cumulative coverage plot as follows. CBSAs with clinics were separated into two groups. One group consisted of all CBSAs that have at least one main clinic. The second group consisted of all CBSAs without a main clinic but with at least one satellite clinic. Then, the CBSAs within the groups were ordered based on female reproductive-age population and the cumulative coverage was calculated. The starting cumulative coverage for the group without a main clinic but at least one satellite clinic was the total coverage of CBSAs with at least one main clinic. This stratified cumulative coverage of reproductive-age women by ART clinics based on clinic presence in CBSA is shown in Figure 2-5.

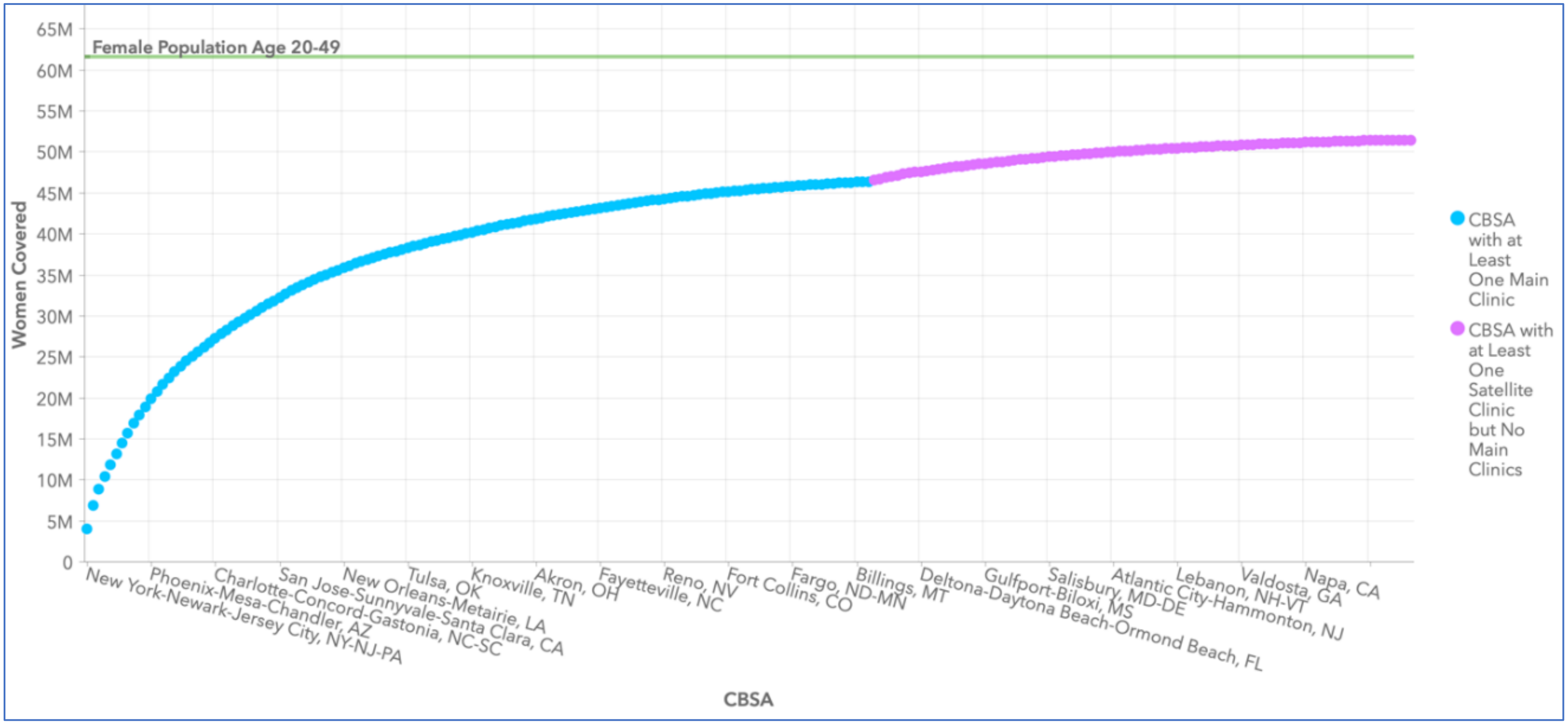


Figure 2-5. The cumulative coverage of reproductive-age women by ART clinics based on CBSA. Note: CBSAs listed on horizontal axis are CBSAs that fall on vertical lines and are a subset of the CBSAs with at least one main clinic or at least one satellite clinic but no main clinics.

There are areas with only satellite clinics that contain more reproductive-age women than some areas with main clinics. The populations in the former areas might therefore justify upgrading a satellite clinic to a main clinic to provide more service options to such a large population. It appears from Figure 2-5 that the coverage efficiency has plateaued, and all the advantageous locations for a clinic have already been "taken," but there do remain locations with no clinics that contain more women than areas with either a main clinic or a satellite clinic. This overlap is illustrated in a cross section of CBSAs with a population between 50k and 150k reproductive-age women shown in Figure 2-6.

Finally, there were 25 CBSAs found that did not have either a main or satellite clinic and have the potential to extend access to ART to 50-147k women per clinic or 1.7M women in total. These areas would seem from this initial analysis to be promising areas to locate a new main or satellite clinic. A plot of these CBSAs vs. potential increase in geographic ART access is shown in Figure 2-7.

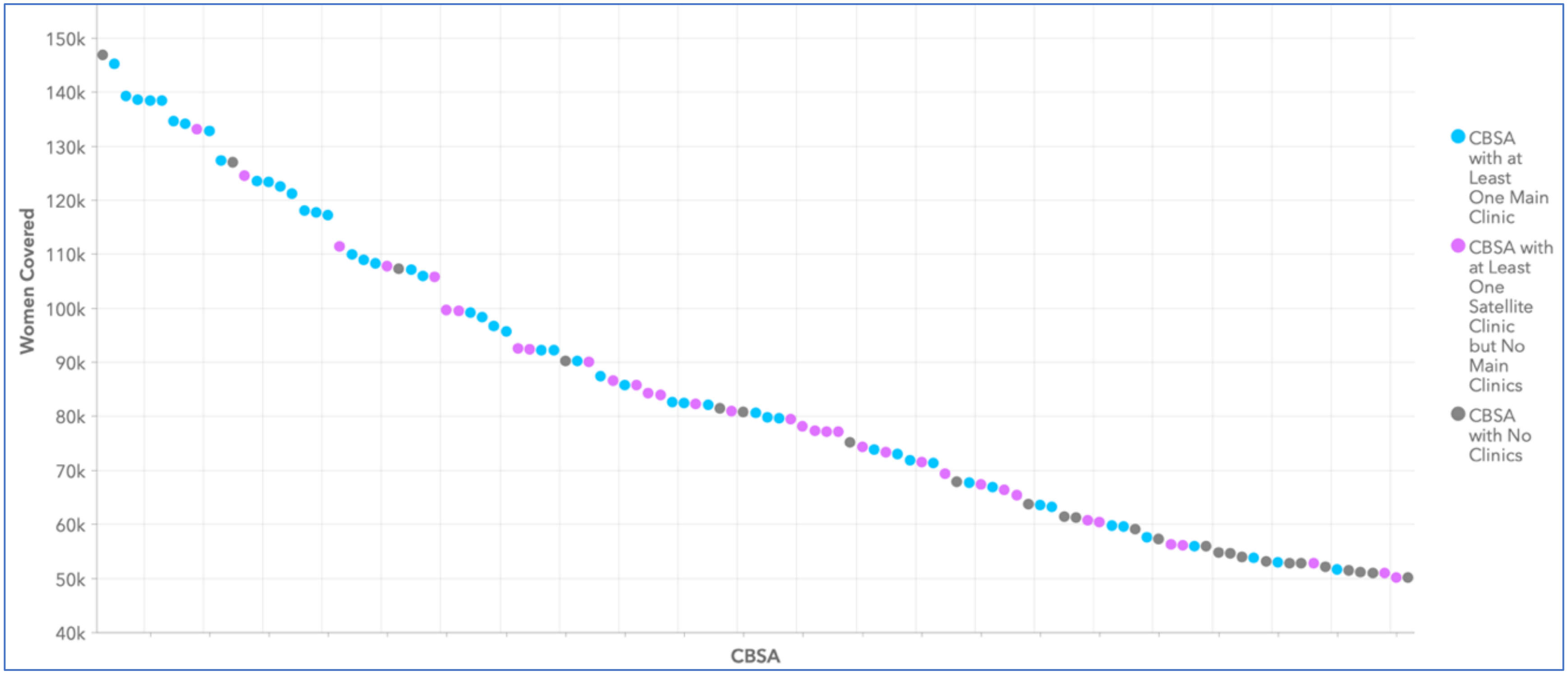


Figure 2-6. The efficiency of locating a clinic in CBSAs covering between 50k and 150k reproductive-age women (age 20-49).

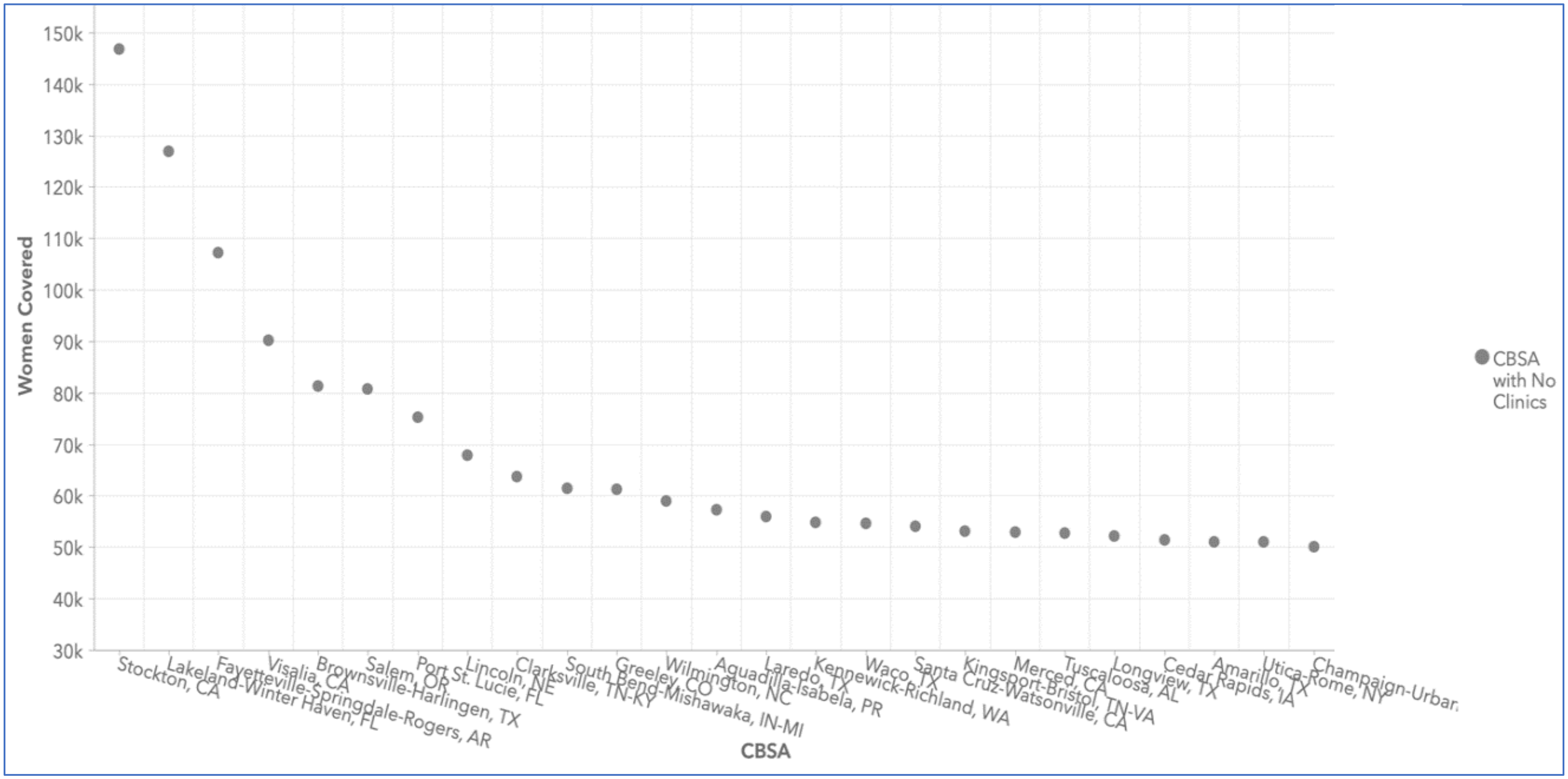


Figure 2-7. Top 25 CBSAs without an ART main or satellite clinic based on the number of reproductive-age women (age 20-49) in each CBSA (vertical axis).

2.3.2.4 H4 - Satellite clinic proximity to main clinics

Most satellite clinics were found located close to a main clinic. For clinics in the contiguous US, the driving time was calculated between each satellite clinic and its practice's closest main clinic. The median driving time found was 42 minutes, while the mean driving time found was 66 minutes with a 95% confidence interval of 56 to 76 minutes. Practices appear to be placing satellite clinics a good distance away from their main clinics, giving the appearance of expanding geographic access, but most of these satellite clinics are located close to other practices' main clinics. 85% of satellite clinics are located closer to the main clinic of other practices than to their own main clinic. When calculating the driving time between each satellite clinic and any practice's closest main clinic, the median driving time found was 16 minutes and the mean driving time found was 31 minutes with a 95% confidence interval of 28 to 34 minutes.

2.3.2.5 H5 - Relationship between an ART practice's number of clinics and the number of ART cycles it performs

At the practice level, the negative binomial regression modeling found a statistically significant ($p < 0.001$) relationship with a positive correlation between a practice's number of satellite clinics and the number of ART cycles it performs. The mathematical relationship found from the negative binomial regression is shown in Equation 2-3 below.

Equation 2-3. Mathematical relationship between an ART practice's number of satellite clinics and the number of ART cycles it performs.

$$log(CountSatelliteClinics) = -0.208 + 0.052 * \frac{Practice\ ART\ Cycles}{100}.$$

For main clinics, however, there was a positive but statistically nonsignificant ($p = 0.26$) relationship found between a practice's number of main clinics and the number of cycles performed.

2.3.2.6 H6 - Statistical modeling of both main clinic and satellite clinic presence

Logistic regression with standardized covariates was used to model the presence of both satellite and main clinics in a CBSA. This modeling found both female reproductive-age population and median income to have coefficients with significance values $p < 0.001$. For both clinic types, population was found to be a more important covariate. For main clinics, population was about 3 times more important than median income (based on covariate t-statistics), and for satellite clinics, population was about 2 times more important than median income. The results of this modeling have been publicly archived (McGarity et al., 2021) and are shown in Table 2-1 through Table 2-3. An increase in one standard deviation in female reproductive-age population increases the odds of a clinic being located in a CBSA by 27,000 for main clinics and 66 for satellite clinics. An increase in one standard deviation in median income increases the odds of a clinic being in a CBSA by 1.8 for main clinics and 1.6 for satellite clinics.

Table 2-1. Logistic regression odds ratios exploring the relationship between population and median income on the likelihood of a main or satellite clinic's presence in a CBSA.

	Main Clinic OR [95% CI]	Sat Clinic OR [95%CI]
Intercept	0.19 [0.13 – 0.28]	0.26 [0.21 – 0.33]
Female Pop 20-49 Z-Score	27,000 [5,000 – 185,000]	66 [27 – 173]
Median Income Z-Score	1.8 [1.3 – 2.5]	1.6 [1.3 – 2.0]

Table 2-2. Mean and standard deviation of population and median income used to calculate z-scores.

	Mean	Standard deviation
Female Pop 20-49	65736.04	227646.8
Median Income	53380.88	12140.34

Table 2-3. Logistic regression coefficients exploring the relationship between population and median income on the likelihood of a main or satellite clinic's presence in a CBSA.

	Main Clinic Presence	Sat Clinic Presence
Intercept (p-value)	-1.66 (< 2e-16)	-1.34 (< 2e-16)
$\beta_{\text{Female Pop 20-49 Z-Score}}$ (p-value)	10.3 (< 2e-16)	4.18 (< 2e-16)
$\beta_{\text{Median Income Z-Score}}$ (p-value)	0.595 (0.000525)	0.467 (3.19e-05)

2.3.2.7 H7 - Findings on geographic access to ART by US racial and ethnic groups

Racial and ethnic geographic access to ART was explored, and geographic access varied across racial and ethnic groups. Outlined in Table 2-4 and shown in Figure 2-8, 82% of people in CBSAs of all racial and ethnic groups were found to have access to either a main or satellite clinic, with 73% having access to a main clinic and 9% having access to only a satellite clinic. White alone and American Indian and Alaska Native alone were the two racial categories with poorer geographic access to either clinic type compared with all groups. White alone had the same access to satellite clinics (9%), but 3% less access to main clinics (70%). American Indian and Alaska Native alone had similar access to satellite clinics (8%), but 17% lower access to main clinics (56%), likely because American Indians and Alaska Natives have the smallest portion of their population living in urban areas (U.S. Census Bureau, 2019f). All other racial categories had better geographic access than all US racial groups likely because they have higher rates of living in urban areas. Asian alone had the best access to main clinics (90%) but lower access to satellite clinics (5%). Black or African American alone had better access than all groups to main clinics (81%) and only slightly lower access to satellite clinics (8%). Finally, Pearson's chi-squared statistical test was applied to these differences in access compared with the mean access of all US races, and the observed differences were found to deviate significantly from the mean for both race ($p < 2.2e-16$) and ethnicity ($p < 2.2e-16$).

Geographic access to ART by ethnicity is summarized in Table 2-5. Compared with all US ethnicities, Not Hispanic or Latino had 1% poorer geographic access to main clinics (72%) and the same level of geographic access to satellite clinics (9%). Hispanic or Latino had better access to main clinics (80%) but 4% poorer access to satellite clinics (5%). The better geographic access to ART for Hispanics and Latinos is likely due to 98% of Hispanics and Latinos living in urban areas (U.S. Census Bureau, 2019d).

Table 2-4. Geographic access to main or satellite fertility clinics based on race.

Access to ART	All Racial Groups	White alone	Black or African American alone	American Indian and Alaska Native alone	Asian alone	Native Hawaiian and Other Pacific Islander alone	Some Other Race alone	Two or More Races
At least one main clinic	73%	70%	81%	56%	90%	74%	84%	76%
No main clinics but at least one satellite clinic	9%	9%	8%	9%	5%	13%	4%	9%
At least one main or satellite clinic	82%	79%	88%	65%	94%	87%	89%	85%
Difference from all races in access to at least one main or satellite clinic	0%	-3%	6%	-17%	13%	6%	7%	3%
Percentage of population living in urban areas	82%	78%	92%	63%	96%	92%	98%	89%

Table 2-5. Geographic access to main or satellite fertility clinics based on ethnicity.

Access to ART	Both Ethnicities	Not Hispanic or Latino	Hispanic or Latino
At least one main clinic	73%	72%	80%
No main clinics but at least one satellite clinic	9%	9%	5%
At least one main or satellite clinic	82%	81%	86%
Difference from both ethnicities in access to at least one main or satellite clinic	0%	-1%	4%
Percentage of population living in urban areas	82%	78%	98%

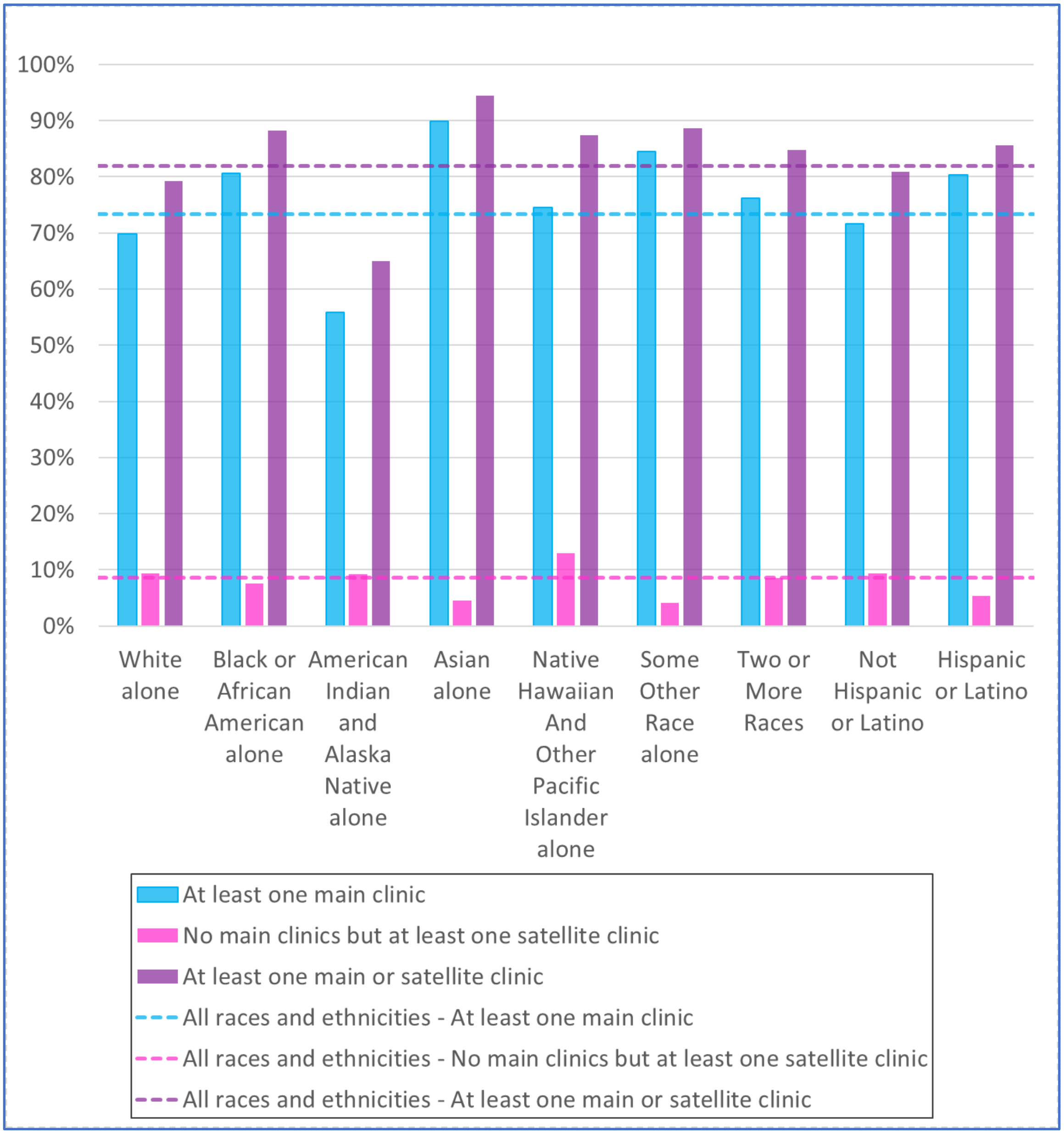


Figure 2-8. Geographic access to main or satellite ART clinics based on race and ethnicity. Blue bars represent geographic access to at least one main clinic, pink bars represent geographic access to no main clinics but at least one satellite clinic, and purple bars represent geographic access to at least one main or satellite clinic. Dashed lines show the geographic access to ART for all races and ethnicities based on geographic access to different clinic types.

2.3.3 Case study in identifying areas for new clinics

A case study in identifying areas for new clinics is presented in this section that extends the utility of the models and analysis of this chapter to identify specific solutions to expand access to ART. The models of clinic presence are applied in Section 2.3.3.1 to calculate the likelihood of a main or satellite clinic being present in every CBSA in the US. When compared with the actual presence of clinics of either type, CBSAs are identified that I conjecture could likely support economically viable fertility clinics (Section 2.3.3.1). This allows for the models to be used to inform industry decision making on new investments.

Additionally, the analysis on racial equity in Section 2.3.2.7 above discovered a relative lack of geographic access for American Indians and Alaska Natives. This analysis is extended in Section 2.3.3.2 to identify both the top-five CBSAs for new ART clinics to cover the most American Indians and Alaska Natives, and the specific CBSAs where new ART clinics could be opened to bring the percentage of American Indians and Alaska Natives with geographic access to ART on par with the geographic access level of all US racial groups.

The case study results are presented in the next sections on new areas for economically viable clinics in Section 2.3.3.1, and the new areas to improve racial equity in utilization of ART in Section 2.3.3.2.

2.3.3.1 Identifying new areas for economically viable clinics

Economic access to ART treatments is a primary barrier to care in the United States, and the majority of insurance plans do not include coverage for fertility treatments. As of April 2021, only 13 states have laws that require insurance coverage of ART (RESOLVE, 2021c). These “IVF mandates” have been shown to increase utilization rates (Crawford et al., 2016), but they do not apply to all insurance plans held in each state. In 2020, of the 91.4% of Americans with health insurance, 52.9% obtained it through their employer (Katerine Keisler-Starkey & Lisa Bunch, 2021). Because of federal law (Employee Retirement Income Security Act of 1974, 1974), state IVF mandates do not apply to the estimated 58.5% of workers with health plans self-insured by their employer (Fronstin, 2012). Therefore, it is unclear how many employers provide IVF coverage, but a 2015 survey of 462 mainly large, private employers found 27% included coverage for IVF (SHRM, 2015). The federal government does not offer insurance plans with ART coverage to its employees (U.S. Office of Personnel Management, 2021), and Americans who obtain health insurance through the federal government do not receive coverage for ART, except for specific wounded veterans (U.S. Dept. of Veteran Affairs, 2019) and Medicaid recipients in Utah with genetic disorders (ASRM, 2020).

Consequently, most patients in the US are required to pay out-of-pocket for ART treatments. Practices that desire to select an economically viable location to place a new clinic should consider the potential patient base both with respect to population as well as its ability to afford services based on income. For example, if a new clinic is opened in an area with a large population but only a small percentage of that population can afford treatment, the clinic will likely fail to generate the revenue needed to cover costs and the clinic will ultimately close. From a financial point of view (if not necessarily from the point of view of equity or other social welfare metrics), new clinics should be placed where they can expand access to care to the most people while also being economically viable. Assuming that existing practice locations are economically viable, the current approach to siting clinics may be the best guide for locating new clinics with the aim of incrementally increasing access to care while maximizing the chance of new clinics staying in business.

Therefore, with the aim of finding potential new locations for successful ART clinics, the logistic regression model (detailed in Table 2-3) predicting the presence of main clinics was applied to find 5 CBSAs that the model predicts are most likely to have a main clinic, but actually have no clinic of either type at present. For example, one CBSA without a fertility clinic, Stockton, CA Metro Area, has a female population of 146,788 women and a median income of $64,432, so its likelihood of having a main clinic is calculated from Equation 2-4 to be 93%, and the likelihood of it having a satellite clinic is calculated from Equation 2-5 to be 64%. The calculated likelihoods for the top five CBSAs are shown in Table 2-6.

Equation 2-4. Calculation for Stockton, CA Metro Area's likelihood of a main clinic using the logistic regression model for main clinic presence.

$$93\% = logit^{-1}(\beta_0 + \beta_1 Median\ Income_{z-score} + \beta_2 Female\ Repr.Age\ Pop_{z-score})$$

$$93\% = logit^{-1}(-1.6635 + 10.3151 * \frac{146788 - 65736.04}{227646.8} + 0.5953 * \frac{64432 - 53380.88}{12140.34})$$

Equation 2-5. Calculation for Stockton, CA Metro Area's likelihood of a satellite clinic using the logistic regression model for satellite clinic presence.

$$64\% = logit^{-1}(\beta_0 + \beta_1 Median\ Income_{z-score} + \beta_2 Female\ Repr.Age\ Pop_{z-score})$$

$$64\% = logit^{-1}(-1.3394 + 4.177 * \frac{146788 - 65736.04}{227646.8} + 0.4665 * \frac{64432 - 53380.88}{12140.34})$$

Table 2-6. Top five CBSAs without an ART main or satellite clinic based on logistic regression modeled likelihood of presence of a main clinic.

CBSA	Female Pop. Age 20-49	Median Income ($)	Likelihood Main Clinic	Likelihood Sat. Clinic
Stockton, CA Metro Area	146,788	64,432	93%	64%
Lakeland-Winter Haven, FL Metro Area	126,869	50,584	72%	42%
Fayetteville-Springdale-Rogers, AR Metro Area	107,239	57,603	61%	40%
Salem, OR Metro Area	80,727	60,178	35%	31%
Visalia, CA Metro Area	90,163	49,687	32%	26%

The logistic regression models indicate that these unserved areas are representative of current clinic locating practices and are likely promising areas to locate a new main or satellite clinic. These CBSAs would likely support economically viable clinics while extending access to ART to about a half a million reproductive-age women in total.

2.3.3.2 Identifying new locations to improve racial equity in utilization of ART

The analysis on geographic access to ART clinics based on race and ethnicity found underserved racial groups to be American Indian and Alaska Native alone and White alone, and Not Hispanic or Latino to be the underserved ethic group. Compared with geographic access to a main or satellite clinic for all US races and ethnicities, the White only group was found to be underserved by 3%, the American Indian and Alaska Native alone racial group was found to be underserved by 17%, and the Not Hispanic or Latino ethnic group was found to be underserved by 1%. All other racial and ethnic groups were found to have better geographic access to main or satellite clinics compared with all US races and ethnicities, likely in part because more of their population lives in urban areas.

Strategies for reducing geographic disparities among races and ethnicities could target certain CBSAs based on population of underserved groups, but strategies should ultimately target increasing actual utilization of ART, so utilization rates among racial and ethnic groups should also be considered in addition to geographic access. A recent study found that American Indian and Alaska Native, Black, and Hispanic women utilized medical services to achieve pregnancy significantly less than white women (Janitz et al., 2019). Another study investigating utilization of ART as a function of race and ethnicity found higher utilization by Asian/Pacific Islander and white populations and lower utilization by black, Hispanic, and American Indian/Alaska Native populations (Dieke et al., 2017).

Because the only racial/ethnic group that is geographically underserved, at least by my CBSA-based analysis, that also underutilizes ART or medical services to achieve pregnancy is American Indian and Alaska Native alone, it follows that a potentially equitable strategy for improving utilization via improved geographic access is to locate new clinics in CBSAs in order to maximize geographic access exclusively for American Indians and Alaska Natives. If this strategy were taken, the CBSAs without a fertility clinic that would extend geographic access to the largest number of American Indians and Alaska Natives are: Gallup, NM, Lumberton, NC, Farmington, NM, Show Low, AZ, and Tahlequah, OK. These locations along with logistic regression (Table 2-3) modeled likelihood of each clinic type are shown in Table 2-7.

Table 2-7. Top five CBSAs maximizing coverage of geographically unserved American Indian and Alaska Native populations.

CBSA	American Indian/ Alaska Native Population	Female Pop. Age 20-49	Median Income	Likelihood Main Clinic	Likelihood Sat. Clinic
Gallup, NM Micro Area	54,977	14,224	33,834	1%	5%
Lumberton, NC Micro Area	52,863	25,860	34,976	1%	6%
Farmington, NM Metro Area	50,133	23,761	50,518	2%	10%
Show Low, AZ Micro Area	48,795	17,902	40,067	1%	6%
Tahlequah, OK Micro Area	16,307	9,735	42,774	1%	6%

All of the top five CBSAs prioritized for expanding geographic access to American Indians and Alaska Natives have low modeled likelihood for having a main or satellite clinic. This makes sense because the statistical modeling was based on industry practices in locating clinics that have ultimately led to lower geographic access for American Indians and Alaska Natives. Though these locations appear to not be promising locations for clinics with below average female reproductive-age population and median income, just five clinics are estimated to increase geographic access to the American Indian and Alaska Native population by about a quarter of a million people (223k), which is approximately 10% of the total US American Indian and Alaska Native population of 2.3 million living in CBSAs or 8% of the total American Indian/Alaska Native population of 2.8 million living anywhere in the US (U.S. Census Bureau, 2019e). The estimated populations covered by these five clinics are listed in Table 2-8 by racial group and Table 2-9 by ethnicity.

Table 2-8. Racial coverage of the top five CBSAs selected to maximize coverage of unserved American Indian and Alaska Native populations.

CBSA	White Alone	Black Or African American Alone	American Indian and Alaska Native Alone	Asia Alone	Native Hawaiian and Other Pacific Island Alone	Some Other Race Alone	Two or More Races
Gallup, NM Micro Area	10,426	474	54,977	829	30	3,010	2,692
Lumberton, NC Micro Area	37,096	31,585	52,863	798	34	7,451	2,769
Farmington, NM Metro Area	66,160	1,041	50,133	984	76	4,270	3,851
Show Low, AZ Micro Area	51,496	1,005	48,795	652	141	4,033	3,148
Tahlequah, OK Micro Area	24,777	661	16,307	359	139	1,076	5,345
Total	**189,955**	**34,766**	**223,075**	**3,622**	**420**	**19,840**	**17,805**

Table 2-9. Ethnic coverage of the top five CBSAs selected to maximize coverage of unserved American Indian and Alaska Native populations.

CBSA	Not Hispanic or Latino	Hispanic or Latino
Gallup, NM Micro Area	62,118	10,320
Lumberton, NC Micro Area	120,822	11,774
Farmington, NM Metro Area	100,847	25,668
Show Low, AZ Micro Area	96,791	12,479
Tahlequah, OK Micro Area	45,182	3,482
Total	**425,760**	**63,723**

The five new clinics selected to maximize coverage of American Indians and Alaska Natives raise the percentage of that group that is covered nationally from 65% to 75%, which is still short of the CBSA coverage level across all US racial groups of 82%. An additional 385k American Indians and Alaska Natives need to be covered by new fertility clinics to have geographic access at parity to all US racial groups. For this to be achieved, a total of 25 new clinics in unserved CBSAs (shown in Table 2-10) are needed to bring geographic access for American Indians and Alaska Natives up to the current level of access across all US races and ethnicities. Again, these locations have a low predicted likelihood of a clinic being sited there with the notable exception of Fayetteville-Springdale-Rogers, AR, the modeled third best location to site a new main clinic (see Table 2-6) while also being able to reach nearly 6k American Indians and Alaska Natives.

Table 2-10. CBSAs to locate 25 new fertility clinics to bring the percentage of American Indian and Alaska Native population with geographic access to ART up to 82% (the level of geographic access of all US racial and ethnic groups).

CBSA	Female Pop. Age 20-49	Median Income	Likelihood Main Clinic	Likelihood Sat Clinic	American Indian and Alaska Native Alone
Gallup, NM Micro Area	14,224	33,834	1%	5%	54,977
Lumberton, NC Micro Area	25,860	34,976	1%	6%	52,863
Farmington, NM Metro Area	23,761	50,518	2%	10%	50,133
Show Low, AZ Micro Area	17,902	40,067	1%	6%	48,795
Tahlequah, OK Micro Area	9,735	42,774	1%	6%	16,307
Muskogee, OK Micro Area	12,946	43,078	1%	6%	12,381
Fort Smith, AR-OK Metro Area	46,573	44,750	5%	12%	11,471
Grants, NM Micro Area	4,736	39,413	1%	5%	11,172
Shawnee, OK Micro Area	14,403	49,250	2%	8%	10,595
Yakima, WA Metro Area	46,453	51,637	7%	15%	10,004
Bemidji, MN Micro Area	8,665	49,160	1%	7%	9,395
Payson, AZ Micro Area	7,535	43,524	1%	6%	8,733
Houma-Thibodaux, LA Metro Area	41,093	51,837	6%	13%	8,495
Riverton, WY Micro Area	6,736	55,896	2%	9%	8,283
Fairbanks, AK Metro Area	20,175	76,992	8%	22%	7,849

Table 2-10, Continued. CBSAs to locate 25 new fertility clinics to bring the percentage of American Indian and Alaska Native population with geographic access to ART up to 82% (the level of geographic access of all US racial and ethnic groups).

CBSA	Female Pop. Age 20-49	Median Income	Likelihood Main Clinic	Likelihood Sat Clinic	American Indian and Alaska Native Alone
Cullowhee, NC Micro Area	10,884	45,288	1%	7%	7,711
Shawano, WI Micro Area	7,230	55,152	2%	9%	6,969
Lawton, OK Metro Area	24,979	52,076	3%	10%	6,879
Eureka-Arcata, CA Micro Area	27,087	48,041	3%	9%	6,846
Durant, OK Micro Area	8,876	44,212	1%	6%	6,717
Pueblo, CO Metro Area	30,254	46,783	3%	10%	6,688
Bellingham, WA Metro Area	45,476	62,984	11%	20%	6,437
Española, NM Micro Area	6,741	39,952	1%	5%	6,380
Visalia, CA Metro Area	90,163	49,687	32%	26%	6,118
Fayetteville-Springdale-Rogers, AR Metro Area	107,239	57,603	61%	40%	5,972

2.4 Discussion

To understand the problem of geographic access to fertility clinics, it is important to understand how fertility care fits into the broader picture of general access to healthcare. Access to healthcare services in the US is generally focused on routine care provided by a primary care physician (PCP) via increasing two factors: the availability of PCPs and the percentage of insured Americans (Office of Disease Prevention and Health Promotion, 2022). Improving these two factors are also helpful in a public health context of fighting infertility. Access to PCPs and having insurance helps reduce the need for fertility treatments by screening for and treating sexually transmitted diseases (STDs) (CDC, 2022), and infertile couples benefit from access to PCPs who can diagnose infertility and refer patients to an OB/GYN who specializes in Reproductive Endocrinology and Infertility (REI) who perform IVF/ART.

The number of REI fellows certified by the American Board of Obstetrics and Gynecology (ABOG) is limited to about 45 annually (Gambone et al., 2015). Unless ABOG plans to increase the number of REI fellows certified annually in the future, this certification rate will limit access to ART. Additionally, most Americans do not have health insurance that includes coverage for ART, and many do not live in an area with either a main or satellite ART clinic. Satellite ART clinics have been cited as a tool to potentially improve geographic access to underserved areas (American Society for Reproductive Medicine, 2015b), but the impact of satellite clinics on geographic access to ART was not previously explored in the literature, so the focus of this essay was to better understand the contribution of satellite clinics in providing geographic access to ART services.

2.4.1 Trends in satellite clinic utilization

Several insights have been obtained in this essay on how the fertility care industry tends to utilize satellite clinics. The first is that higher volume ART practices are more likely to establish satellite clinics. Secondly, and more importantly, 78% (n = 451) of satellite clinics do not increase geographic access to care (as measured by a satellite clinic being located in a CBSA without a main clinic). However, there are 129 satellite clinics (22%) located in CBSAs without a main clinic, thereby extending geographic access to care to 5.1 million (8%) reproductive-age women.

2.4.2 Driving times between main and satellite ART clinics

The analysis of driving times between satellite and main clinics suggests that most satellite clinics are sited by practices in order to compete for patients, rather than to increase geographic access to underserved populations. Satellite clinics are located on average 66 minutes away from *their* practice's main clinic but only 31 minutes away from *any* practice's main clinic. These driving times suggest that the primary roles of satellite clinics are to improve patient convenience and increase market share by encroaching on market areas served by other practices' main clinic.

In general, the tendency for competing businesses to locate near each other is described in economics with Hotelling's Law, which states that competing sellers will tend to locate close to each other to obtain the largest market share, and "as more and more sellers of the same commodity arise, the tendency is not to become distributed in the socially optimum manner but to cluster unduly" (Hotelling, 1929). Considering that travel time was found to be the second largest contributor to the substantial time-cost of fertility care (Wu et al., 2013) and that patients are willing to accept a clinic with a lower pregnancy rate in exchange for reductions in travel time (van Empel et al., 2011), the satellite clinic placement strategy described in this essay is likely effective in competing for patients with long travel times within their CBSA, but ineffective at reaching patients in unserved CBSAs who could potentially gain access to ART through a local satellite clinic.

2.4.3 Relationship between practice ART cycles and number of satellite clinics

At the practice level, the analysis supports hypothesis H5 that satellite clinics allow practices to reach more patients. Looking at the roughly half of all practices that operate satellite clinics, those practices tend to perform more cycles than practices lacking satellite clinics. For example, in 2018, only one practice with zero satellite clinics performed more than 2,000 cycles, whereas 25 practices with at least one satellite clinic performed more than 2,000 cycles. Exploring these trends further, the negative binomial regression modeling in Section 2.3.2.5 estimates one satellite clinic for practices performing 400 cycles, two for 1,700 cycles, three for 2,500 cycles, four for 3,000 cycles, five for 3,500 cycles, etc.[1] Based on this trend,

---

[1] Using Equation 2-3:

$$log(Count\ Satellite\ Clinics) = -0.208 + 0.052 * \frac{Practice\ ART\ Cycles}{100}$$

practices operating fewer satellite clinics than their number of cycles indicate should consider opening a new satellite clinic. Conversely, practices operating more satellite clinics than their number of cycles would indicate should consider if there are better locations to operate their satellite clinics that could reach more patients.

2.4.4 Ad-hoc monitoring arrangements

It is possible for individual patients to arrange for ad-hoc ART monitoring of themselves with a local OB/GYN and clinical laboratory and then travel long distances for egg retrievals and embryo transfers. However, the option to use ad-hoc monitoring is seldom described on ART practice websites, so many prospective patients likely do not know it is potentially an option and do not pursue these ad-hoc monitoring arrangements. Such arrangements are possible to expand geographic access to ART to anywhere within a reasonable distance to a receptive OB/GYN and a clinical laboratory. Moreover, networks of partnering OB/GYNs could be established that partner with any ART practice in the US. SART is a professional society for ART that seeks to ensure patients receive high quality care, but it could also facilitate monitoring agreements between ART practices and other ART practices or OB/GYNs. Conceivably, any SART practice could perform monitoring for any other SART practice, and SART could establish an OB/GYN partner network that any of its member practices could leverage for monitoring support of patients living far from SART ART clinics.

2.4.5 Economically viable clinics and the influence of state IVF insurance mandates

State IVF insurance mandates likely have an impact in certain CBSAs and influence the establishment and continued operation of satellite clinics. This study found significantly more satellite clinics per million women in IVF mandate states, and of the five CBSA's identified as likely to economically support a new clinic shown in Table 2-6, only Fayetteville, AR is in an IVF insurance mandate state. Either the areas identified are simply undiscovered markets or the lack of state IVF mandates in California, Oregon, and Florida tipped the scales away from economic viability of clinics in these areas. If these states adopt IVF mandates in the future, the industry should be poised to open new clinics in these promising but unserved areas. Finally, future studies should explore the effect of insurance mandates on geographic access to care, particularly considering the recent increase in number of states adopting mandates for IVF insurance coverage.

2.4.6 Policy implications of findings on equitable geographic access to ART based on race and ethnicity

This study's findings on equitable geographic access to ART based on race and ethnicity carry policy implications in favor of subsidizing ART clinics that increase geographic access to American Indians and Alaska Natives. Compared with the coverage percentage among all US racial groups, American Indians and Alaska Natives were found to be geographically underserved by 17%. This large disparity in geographic access likely contributes to their underutilization of ART medical care to achieve pregnancy found in the literature (Dieke et al., 2017; Janitz et al., 2019). Poor geographic access is likely a primary barrier to care for American Indians and Alaska Natives, warranting strategic subsidizing of new clinics to better reach these underserved people.

The policy implications of this essay's findings are particularly notable considering historically that the US government was directly responsible for the oppression of indigenous peoples with the passage of the Indian Removal Act of 1830 (Indian Removal Act, 1830), the resultant population losses from the Trail of Tears (Thornton, 1984), and the forced sterilization of Native American women in the 1970s without their informed consent (Comptroller General of the United States, 1976; Lawrence, 2000).

As a result of treaty obligations and the Indian Health Care Improvement Act (Indian Health Care Improvement Act, 1976), the US government provides free healthcare at Indian Health Service (IHS) facilities to Native Americans. IHS's goal for Native Americans is to "raise their health status to the highest possible level" through a "comprehensive health service delivery system" (Indian Health Service, 2021a). The IHS acknowledges infertility as a disease based on its website (Indian Health Service, 2021b), but unfortunately does not offer treatment for infertility using ART at its facilities.

The IHS works within yearly budgets approved by Congress and only meets around half of the healthcare services needed by American Indians and Alaska Natives (U.S. Centers for Medicare & Medicaid Services, 2016). If a medical service is not available at an IHS facility for an eligible person, an IHS Purchase Referred Care (PRC) authorization may be obtained, but the service must be deemed medically necessary, and there must be sufficient PRC funds available (Indian Health Service, 2018).

I submitted a Freedom of Information Act (FOIA) request to determine if IHS would pay for ART via PRC, and I learned that "PRC considers [ART] an Excluded Service" (Mitchell, 2022). So, for ART treatments to be eligible for IHS payment under the PRC program, IHS would need to deem ART treatments to be medically necessary and ART would have to be removed from the list of excluded services. Then, ART treatments could be covered under PRC, but this would only reduce the economic barrier to care. Geographic access would still be deficient based on where American Indians and Alaska Natives live relative to existing ART clinics. For example, from the center of the Hopi Reservation, the closest satellite fertility clinic is about two hours away in

Flagstaff, AZ and the closest main fertility clinic is about four hours away in either Phoenix, AZ or Albuquerque, NM.

The lack of ART access for American Indians and Alaska Natives could also have human rights implications due to the US Government's relationship to Native Americans and its failure to provide ART. The US Government's relationship to Native American tribes is one of "guardian" to "ward" (Green, 1975). Native American tribes do not have the autonomy of independent nation states, and tribal matters are micromanaged by the Bureau of Indian Affairs (Hanke, 2019). The United Nation's 1948 Universal Declaration of Human Rights declared that "Men and women of full age, without any limitation due to race, nationality or religion, have the right to marry and to found a family" (United Nations, 1948). Additionally, the World Health Organization (WHO) first defined infertility as a disease in 2009 (Zegers-Hochschild et al., 2009), and in 2012 the Inter-American Court of Human Rights declared in 2012 the banning IVF in Costa Rica a human rights violation. Though the US Government has not banned IVF on tribal lands, the guardian-ward arrangement effectively prevents ART on Indian Reservations because IHS does not provide it directly or indirectly via PRC, and the arrangement has stunted tribal economic development making privately operated ART practices economically unviable.

One possible solution is for the IHS to deem IVF treatments to be medically necessary for infertile American Indians and Alaska Natives such that treatment could be paid for via the PRC system. IHS could then provide ART monitoring support from its clinics. Then, patients would only need to travel several hours to a main clinic occasionally for weekend monitoring, egg retrievals, and embryo transfers. A more long-term solution is for Congress to specifically allocate funds for establishing ART embryology labs at IHS facilities to bring full-service treatments directly to these unserved populations. There are no laws prohibiting the federal government from operating ART clinics, and the federal government currently operates ART clinics at six military medical centers (RESOLVE, 2022). Congressional funding of new ART clinics through the IHS would likely be only a budgetary matter, passable under current US Senate rules via the budget reconciliation process, and not subject to the 60 votes in the Senate required to overcome a filibuster.

2.4.7 Limitations

The conclusions of this chapter are limited by the methods and data used. Several such limitations are discussed here, along with possible avenues for overcoming them in future research. The first limitation is the web-search relied on information publicly available on practice websites. This may not reflect the true state of clinics due to out-of-date or incorrect website information. Additionally, several ART practices have partnerships with non-ART OB/GYN practices to provide ultrasound and monitoring support for an IVF cycle. If the partnership was detailed on an ART practice's website, then the non-ART OB/GYN practice was considered a satellite clinic of the ART practice. If the partnership was omitted from the ART practice's website, it was not captured in this study. Practices should include these monitoring partnerships if they exist, and formal surveys from the CDC and SART should include the ability for these *de facto* satellite clinics to be captured.

This study did not investigate the level of care provided by each satellite clinic and treated all satellite clinics the same, although services provided at satellite clinics in fact vary widely. For instance, some satellite facilities are open one day per week for consultation only, while others provide monitoring for ART and non-ART fertility treatments 7-days per week. Formal surveys of fertility clinic locations performed by CDC and SART should collect information on what specific services are available at each practice's locations versus their current approach to only track services offered by the practice as a whole.

The use of US Census CBSAs to define geographic access to care is an approximation of actual geographic access. The geographic area defining a CBSA is selected by the Census Bureau based on high levels of social and economic integration measured by commuting times (U.S. Census Bureau, 2021). However, long travel times within large CBSAs may inhibit access to care for some patients. On the other hand, patients living outside of a CBSA with a fertility clinic but within a reasonable travel time to a clinic would not be counted as having access. Similarly, some patients may be willing to travel outside of their home CBSA for care. The approach taken in Chapter 4 overcomes these issues by using optimization models that define coverage in terms of driving times between population centers.

Further, some patients do not own cars and rely on public transportation that varies widely across the US in its reliability and travel times in some areas are likely much longer via public transportation vs. driving. For example, a patient without a car living near the West Baltimore MARC (Maryland's commuter rail) station can access a satellite fertility clinic in Baltimore's inner harbor by taking a bus then walking a short distance in 25 minutes. However, on weekends and for egg retrievals/embryo transfers, that patient would need to travel 69 minutes (via two busses) to the closest main clinic in Towson, MD when driving would only take 23 minutes. This hypothetical carless patient is assumed to have geographic access based on living in a CBSA with a clinic, but the assumption that all patients living in a CBSA with a fertility clinic have geographic access to that clinic may not hold for patients with low socioeconomic status or racial and ethnic groups with low rates of car ownership, considering that 5.4% of white households, 24% of black households, and 12% of Latino household do not own a car (Raphael & Stoll, 2001). Future studies should use travel times on public transportation along with local reliability estimates to better estimate geographic access to ART.

Finally in the analysis of racial/ethnic access to care, there are two plausible explanations for why ART clinics are underserving American Indians and Alaska Natives. This essay took the position that there are not economically viable locations for ART clinics to operate while serving large numbers of American Indians and Alaska Natives, which has led to low utilization rates. Alternatively, there could be a cultural aversion to using ART among American Indians and Alaska Natives that contributes to low American Indian/Alaska Native utilization of ART and the decision to not locate ART clinics near American Indians and Alaska Natives. Future studies should attempt to estimate the receptivity of American Indians and Alaska Natives towards using ART to clarify if it is commensurate with their lack of utilization and geographic access to ART.

2.5 Summary

In January 2021, I completed a systematic web-search of ART clinics in the United States. I found 469 main clinics with embryology labs and 583 satellite clinics for a total of 1,052 clinic locations in the United States performing or facilitating ART. Most satellite clinics were found located close to a main clinic, and only 22% of satellite clinics extended geographic access to 8% of US reproductive-age women. Satellite clinics were found to be located on average 65 minutes from their practice's main clinic and 31 minutes from any main clinic.

US Census Bureau's Core-Based Statistical Areas (CBSAs) were used to approximate geographic coverage. 78% of satellite clinics were found in CBSAs that also have a main clinic, and 22% of satellite clinics were found in CBSAs without a main clinic. 72% of US reproductive-age (ages 20-49) women live in a CBSA with at least one main clinic, and 8% of US reproductive-age women live in a CBSA without a main clinic but with at least one satellite clinic. 20% of US reproductive-age women do not live in an area with a main clinic or satellite clinic.

In my statistical analysis, I found that CBSA reproductive-age population and median income were statistically significant when modeling presence of both clinic types, with population more important than median income. ART practices with satellite clinics tend to perform more IVF cycles. States with IVF insurance mandates have significantly more satellite clinics per million women, but not significantly more main clinics per million women. Satellite clinics are located throughout the country primarily in areas that also have main ART clinics to compete for patients via increased convenience.

In summary, about half of all practices operate satellite clinics, and they tend to perform more cycles than practices that do not operate satellite clinics. Satellite clinics are predominantly located near main clinics, and only a fraction of satellite clinics are located in areas without a main clinic. Just 22% of all satellite clinics extend geographic access to areas without a main clinic, accounting for only 8% of reproductive-age women. Satellite clinics clearly play an important role in extending geographic access to millions of women, but they do so inefficiently. There are still many promising areas in the US with large numbers of unserved infertile couples for whom a nearby clinic would make all the difference in accessing care. New and growing practices should consider these areas when determining where to locate new clinics such that geographic access to care will no longer be a barrier to millions of couples seeking care.

Finally, increasing racial and ethnic equity in geographic access to care should be considered when selecting locations for new clinics. This study's findings suggest that geographic access is a barrier for certain racial and ethnic groups but not others. American Indians/Alaska Natives were found to be geographically underserved in the US compared to all US racial groups by 17%. Considering the long history of repression of the indigenous population in the United States by the United States Government itself from the Indian Removal Act of 1830 (Indian Removal Act,

1830) to the forced sterilization of Native American women in the 1970s without their informed consent (Comptroller General of the United States, 1976; Lawrence, 2000), this large gap in geographic access to fertility care is alarming. The US Government does not offer ART at its Indian Health Services (IHS) facilities, so Federal government funding of ART for American Indians and Alaska Natives through the IHS to correct this deficiency is warranted.

Chapter 3 Predictive Modeling of Assisted Reproductive Technology Demand and Fertility Clinic Presence

3.1 Introduction

3.1.1 Purpose

This essay investigates ways to estimate ART demand in every county of the US. The simplest way to estimate ART demand is to multiply the expected number of ART cycles per capita by the reproductive-age population of a county. But could a more accurate estimate of ART demand be derived using location-specific predictors in addition to population? To answer this question, this essay creates predictive models of ART demand using a range of socioeconomic and environmental health factors to hopefully better predict ART demand compared with models using exclusively population. The predictive models employed include a range of machine learning methods ranging from simpler statistical models to complex tree-based ensemble models.

The predictive models are tested for their ability to predict ART demand using the county aggregate number of ART cycles performed in a year. The testing of models is done via repeated cross-validation that holds out a fraction of data and uses the remaining dataset to fit predictive models that are tested for predictive accuracy of the held-out data's response variable. This testing is repeated for a total of 30 trials, and the predictive performance is estimated for each model tested. A final model is selected based on comparing predictive performance metrics and relative complexity with a preference for models with better predictive performance that are simpler than their peers.

Ultimately, the selected model is utilized to extrapolate trends in proven ART demand from the counties with ART main clinics to predict hypothetical demand for all US counties.

These estimates of hypothetical ART demand are likely overestimates because current data on ART demand are available only for counties with ART clinics that likely also serve patients from nearby counties without ART clinics. But recognizing this limitation, these predictions still add value to the dissertation as a whole by capturing macro trends in ART demand to be used in Chapter 4's multiobjective optimization between covering population and covering demand that allows for placement of new clinics that balance the needs of for-profit ART practices with the social benefit of providing geographic access to the most people.

3.1.2 Background and literature review

The ultimate purpose of statistically modeling ART demand and clinic presence is to produce predictions useful in making decisions concerning the locations to place new ART clinics and what populations to serve. There are two types of ART business models currently operating in the US: for-profit enterprises and academically affiliated not-for-profit centers. The vast majority of ART clinics in the US operate as for-profit businesses, with only a small fraction (less than 15%) of ART clinics that are not-for-profit academically affiliated (Patrizio et al., 2022). Each business model has differing inherent motivations, but they both require sufficient patient volumes to operate.

Not-for-profit academically affiliated clinics operate to support research efforts of faculty members and to train the next generation of reproductive endocrinologist through REI fellowships, which require an academic institution (Patrizio et al., 2022). Though not-for-profit ART centers do not have an inherent purpose of generating profit, they still need to generate adequate revenue to justify operating costs, and they need an adequate number of patients to facilitate their educational and research activities. Therefore, they often locate in areas that will have higher demand for ART. For example, Johns Hopkins University's fertility center is located in the more affluent Baltimore County rather than in less affluent East Baltimore where the main campus of Johns Hopkins Medicine is located. Therefore, not-for-profit ART centers consider likely patient volume when deciding where to open/operate clinics and are drawn to areas with at least adequate expected ART demand.

ART healthcare practices are primarily for-profit enterprises (Patrizio et al., 2022) and as such are incentivized to locate clinics in areas with high enough demand to make a high enough rate of return to attract investors. ART practices have increasingly attracted investors with the industry expanding from an academic research activity focused on treating infertility into a commercialized industry focused on preserving future fertility and genetic screening of embryos. This expanding role of ART has grown the industry in the US from about $2B per year in 1992 to $8B in in 2022 (Patrizio et al., 2022). For-profit ventures have the inherent purpose to maximize returns to the owner in the form of profit (revenue minus cost) (although for-profits can also have a complex range of social missions as well (Peredo & McLean, 2006)). Therefore, in locating for-profit healthcare facilities anticipated operating costs and expected revenue need to be considered to identify the locations with the greatest potential profit.

A large investment is required to establish an ART clinic, but these investment costs are not in the public domain. I could not find estimates for the cost of establishing a new ART clinic in the academic literature or from wider internet searches, so I asked members of the industry at the 2021 ASRM Scientific Congress. A vice president of a new ART practice that is opening new clinics across the US provided me a rough estimate of costs. She estimated the cost of opening of a new ART clinic to be about $5M for a main clinic with an embryology lab and between $1-2M to establish a satellite clinic with ART monitoring capabilities dependent on its capabilities. With these large outlays to setup new clinics, it is critical to have a good estimate of demand for ART in a prospective area.

Estimates for demand for health services are typically calculated as a fraction of the population likely to seek care, but five factors make it difficult to accurately predict ART demand and clinic presence using population alone: racial/ethnic, religious, economic, social, and environmental factors. A single study was found in the literature that ecologically modeled ART utilization as a function of demographic, social, and economic factors (Hammoud et al., 2009). It found ART utilization to be higher in states with higher availability of IVF clinics, a greater degree of urbanization, and a greater prevalence of people with college degrees. This would indicate that, in general, highly educated urban areas are accessing treatment, but the several factors influencing the level of demand more specifically need to be better understood.

The purpose of this section is to provide information on the various factors that influence ART demand and clinic presence as a background for quantitative analyses later in this chapter that use various socioeconomic and environmental metrics in predictive models. This background section includes the role of race and ethnicity on ART utilization in Section 3.1.2.1, the role of religion in promoting or limiting ART demand in Section 3.1.2.2, the economic factors influencing ART demand and clinic presence in Section 3.1.2.3, the elective and non-elective reasons for using ART in Section 3.1.2.4, and the environmental factors contributing to infertility in Section 3.1.2.5.

3.1.2.1 Role of race and ethnicity on ART utilization

Utilization of ART by race and ethnicity has been studied in the literature. Dieke et al. (2017) found American Indian/Alaska Native non-Hispanic women to have the lowest rate of ART utilization at 807 per million women. The National Survey of Family Growth (NSFG) sampled 15,303 married women in the United States aged 15-44 and found a decline in infertility from 1982 to 2002 from 8.5% to 7.4%, and a multivariate analysis found that infertility was more common for older, nulliparous (women with no children), non-Hispanic black or Hispanic women without a college degree (Stephen & Chandra, 2006). Jain (2006) conducted a cross-sectional survey of 1,500 women in Illinois and found that African Americans significantly delay seeking care longer than Caucasian woman (4.3 vs. 3.3 years respectively), despite Illinois being a state that mandates insurance coverage for IVF. Any analysis of race or ethnicity is confounded by numerous contributing factors such as social, economic, geographic, environmental factors, so it is important to highlight that race and ethnicity are correlated with lower ART utilization but are not inherently causes of lower ART utilization.

3.1.2.2 Role of religion in promoting or limiting ART demand

Some religious groups and leadership hold generally a "cautiously positive" view towards ART with its potential to treat infertility but express reservations with the intent to protect both marriages and human embryos. Judaism has generally embraced ART based on the interpretation of Jewish scripture that the human soul does not enter an embryo until 40 days after conception, and many Jewish leaders actively encourage use of ART to follow God's instruction to "be fruitful and multiply" (Cha, 2018). Of all the major religions, only Roman Catholic Christianity opposes ART in all forms, and Islam opposes only ART in which donated eggs or sperm are used (Dutney, 2007).

The issue of embryo destruction is a concern for some religions' leadership and membership. The standard treatment regimen for IVF likely destroys several embryos. The standard IVF treatment regimen is to stimulate the production of many eggs with the expectation that only a fraction will fertilize and develop adequately (over 3-6 days *in vitro*) to be considered viable. Embryos that stop developing or are considered unlikely to be viable are discarded. The extra embryos not transferred to the woman are frozen for future frozen embryo transfers (FET), donated to another couple, donated to research, or discarded (Embryo Options, 2022). Several Christian denominations, Hinduism, and Buddhism officially oppose ART when embryos are destroyed (Biggers, 2012). Therefore, strict adherents of those faiths may believe the standard IVF treatment regimen to be morally wrong for that reason.

However, there are often big differences between the official positions of an organized religion, and the actual attitudes of its members. In a 2013 survey on a representative sample of 4,006 adults nationwide found only 13% of Catholics, 13% of Protestants, and 9% of those religiously unaffiliated believe IVF is morally wrong (Pew Research Center, 2013). Therefore, most religious people in the US likely do not agree with their religion's doctrine on ART, and religion is likely not a primary barrier to accessing ART.

The increasing secularization of the US could also play a role in ART utilization. An increasing number of Americans are identifying as non-religious. In a 2021 survey of a representative sample of 3,937 US adults, 29% identified as non-religious, 63% of US adults identified as Christian and 6% identified with other religions, and of the Christians, only 21% identified as Catholic (Smith, 2021). Even in societies that nominally follow a particular religion, there are large numbers of non-religious/secular people who do not have religious reservations or inclinations to utilize ART. For example, nominally Christian societies in Western and Northern Europe have higher rates of ART utilization when compared to the US (Nachtigall, 2006). Theses societies are also more secular than the US, and secularization is also associated with lower country-level fertility rates (Schnabel, 2021). Additionally, highly religious women in the US have both higher intended fertility and actual fertility rates (Hayford & Morgan, 2008), so the increasing secularization of the United States could potentially result in higher or lower utilization of ART dependent on the reduced desire for children caused by secularization vs. the increased receptivity to use ART from the lack of religious objections.

3.1.2.3 Economic factors influencing ART demand and clinic presence

Economic factors, such as personal income, pricing, and time- and travel-related costs, are highly impactful on demand for ART. A 2011 study found the median cost of one cycle of IVF to be about $25k, and the median cost of a successful outcome (delivery or pregnancy at 18 months) was found to be about $61k (Katz et al., 2011). Considering the median household income in 2011 was about $51k (U.S. Census Bureau, 2019d), the cost of ART is a considerable barrier to access care unless heavily subsidized.

Insurance coverage for ART treatments is not standard in the United States. Diagnostic services are typically covered, but ART treatments are typically not covered. In 2021, zero health insurance plans offered by the federal government to its employees covered ART, and only Care First and Kaiser Permanente offered plans that covered intrauterine insemination (IUI) (U.S. Office of Personnel Management, 2021). IUI is performed by ART clinics in the US, but it is not as capable as ART with lower success rates and certain causes of infertility not being treatable with IUI.

Prior to 2018, only 16 states had infertility insurance coverage laws, 8 had an IVF insurance mandate, and 5 had fertility preservation laws addressing iatrogenic (medically induced) infertility (RESOLVE, 2018). Iatrogenic infertility is generally caused by the treatment of cancer. Fertility preservation treatments are performed prior to cancer treatments so that patients are hopefully still able to have biological children post cancer treatment using frozen embryos, sperm, and oocytes.

In the past four years, several states have adopted new laws to mandate insurance coverage to treat infertility, and as of April 2021, 19 states had fertility insurance coverage laws, 13 of those states mandated IVF coverage, and 11 states had fertility preservation laws to cover

IVF in the event of medically induced infertility (RESOLVE, 2021c). The differences between 2018 and 2021 in prevalence of state fertility insurance laws is summarized in Table 3-1.

Table 3-1. Recent changes in number of states with fertility insurance laws and IVF insurance mandates (RESOLVE, 2018, 2021c).

	States with fertility insurance laws	States with IVF insurance mandate	States with fertility preservation laws
Prior to 2018	16	8	5
April 2021	19	13	11

State insurance mandates are limited in their reach because the federal government and companies that self-insure are not subject to IVF insurance mandate laws of the states. Only insurance plans where an insurance company itself takes on the healthcare claim liabilities are subject to these state IVF mandate laws (RESOLVE, 2021b).

The time-cost of infertility treatments is also considerable. One study found the average time spent on fertility care by 319 couples to be 125 hours with the majority of the time spent by couples on provider visits (73 hours) (Wu et al., 2013). Many of these provider visits are often scheduled early in the morning to allow for patients to balance fertility treatments with work commitments, but several procedures are scheduled according to strict fertility treatment schedules that likely require taking time off from work unless the procedures fall on weekends or holidays. For example, after 2-3 weeks of regular monitoring of ovarian follicular development, egg retrievals for IVF are performed 36 hours after an ovulation trigger shot is given. Paid time off or a flexible work schedule are often necessary to allow patients to be able to receive care without losing income — income that can be a prerequisite to affording treatment for those paying out-of-pocket.

3.1.2.4 Elective and non-elective reasons for using ART

To better understand if ART is utilized electively or non-electively, the reasons for undergoing ART must be understood. According to the CDC's 2018 ART Success Rates dataset, the reasons for undergoing an ART cycle include male factor, endometriosis, tubal factor, ovulatory dysfunction, uterine factor, preimplantation genetic testing (PGT), gestational carrier, diminished ovarian reserve, egg or embryo banking, recurrent pregnancy loss, other infertility, other non-infertility, and unexplained (CDC, 2020a). Table 3-2 shows the reasons for performing an ART cycle for the 306k cycles performed in the US in 2018.

Table 3-2. Reason for using ART in 2018 of 306k cycles reported to the CDC. The percentage of cycles sum to 190% because there can be more than one reason for performing an ART cycle, so a normalized percentage of cycles was calculated by dividing each percentage of cycles by 190%.

Reason for Using ART	Percentage of Cycles	Normalized Percentage of Cycles
Male factor	28%	14.7%
Endometriosis	7%	3.7%
Tubal factor	11%	5.8%
Ovulatory dysfunction	14%	7.4%
Uterine factor	6%	3.2%
Preimplantation genetic testing	13%	6.8%
Gestational carrier	2%	1.1%
Diminished ovarian reserve	30%	15.8%
Egg or embryo banking	34%	17.9%
Recurrent pregnancy loss	5%	2.6%
Other, infertility	24%	12.6%
Other, non-infertility	5%	2.6%
Unexplained [infertility]	11%	5.8%

Of the reasons for performing an ART cycle, most are clearly to treat infertility, but there are also non-infertility reasons to perform ART. Infertility treating reasons include: male factor, endometriosis, tubal factor, uterine factor, recurrent pregnancy loss, other non-infertility, diminished ovarian reserve, and unexplained [infertility]. One likely non-infertility treating reason is "Other, non-infertility," but reasons (described below) including preimplantation genetic testing (PGT), gestational carrier, and egg or embryo banking could be either to treat infertility or electively unrelated to infertility.

PGT is selected as a reason when the "primary reason for using ART was for conducting preimplantation genetic testing, which includes diagnosis or screening to detect chromosomal or genetic abnormalities and prevent an inherited disease. This includes cycles performed for aneuploidy screening" (CDC, 2020b). PGT is also used to perform gender selection, which is offered at some clinics throughout the US, but PGT is typically performed when the man or woman are carriers for a genetic disorder that could be screened via PGT or when the female is at a higher risk for aneuploidy that increases with maternal age.

According to ASRM, a gestational carrier (GC) could be used "when an intended parent wants to have a child and either does not have a uterus or has a medical condition that would prevent carrying a pregnancy safely. Also, a GC may be considered for women who have a history that suggests a problem with her uterus, such as recurrent miscarriage or IVF failure or when a female partner is absent (single male or gay couple)" (American Society for Reproductive Medicine, 2012). Therefore, a gestational carrier could be used in the case of a medical condition that causes infertility or for a non-medical reason such as when a female partner is absent.

According to the CDC's Success Rates Report, causes of diminished ovarian reserve could be "congenital, medical, or surgical causes or advanced age," but it does not specify what age constitutes "advanced age" (CDC, 2020b). Menken et al. (1986) concluded that "the evidence, therefore, is persuasive that biological infertility rises, but only moderately, with age, at least until the late thirties and early forties" (Menken et al., 1986). Moreover, ASRM's guide to elective egg freezing notes that "most experts agree that elective egg freezing is most successful for woman younger than 38 years of age" (American Society for Reproductive Medicine, 2014a). The CDC's 2018 ART Success Rates Report provides a breakdown of all cycles by patient age (CDC, 2020b) shown in Table 3-3, which could help in estimating the portion of diminished ovarian reserve cycles caused by advanced maternal age. For example, 40% of cycles were performed for women 38 years and older, suggesting that about half of diminished ovarian reserve ART cycles are performed to address biological infertility.

Table 3-3. The number and percentage of cycles performed in 2018 stratified by patient age (CDC, 2020b).

Patient Age	<35	35–37	38–40	41–42	≥43	Total
Total number of cycles	113,845	69,614	60,038	28,836	33,864	306,197
Percentage of cycles	37%	23%	20%	9%	11%	100%

Lastly, the egg or embryo banking reason for ART is for the "purpose of freezing eggs or embryos for future use." The reasons for freezing eggs or embryos vary from elective postponing of children (American Society for Reproductive Medicine, 2014a), for military personnel prior to deployments (NYU Langone Health, 2021), and preserving fertility of patients prior to oncology treatments (described in Section 3.1.2.3) (American Society for Reproductive Medicine, 2015a).

3.1.2.5 Environmental factors contributing to infertility

Several studies have investigated the environmental drivers of infertility. There are a number of factors that were found to consistently reduce fecundity, the ability to reproduce (Louis & Platt, 2011). These variables include: age (Dunson et al., 2002; Menken et al., 1986; Schwartz et al., 1983), body mass index (BMI) (Bolúmar et al., 2000; Ramlau-Hansen et al., 2007), occupational exposures, environmental chemicals, and cigarette smoking (Augood et al., 1998). There are also a number of factors that are suggestive of reduced fecundity: alcohol consumption, caffeine consumption, stress, anabolic steroids, biking (for men), and diet (Chavarro et al., 2007).

Environmental chemicals associated with reduced male and female reduced fecundity include: mercury (Cole et al., 2006), polybrominated diphenyl ether (PBDE) (Harley et al., 2010), polychlorinated biphenyls (PCBs) (Buck Louis et al., 2009; Gesink Law et al., 2005; Hauser, 2006), 1,1-dichloro-2,2-bis(p-chlorophenyl) ethylene (DDE) (Gesink Law et al., 2005), dioxin (Eskenazi et al., 2010), and phthalates (Hauser, 2006; Hauser & Sokol, 2008). Possible exposure pathways to these chemicals include: occupational exposures (mercury, dioxin), dermal exposures (phthalates), or inadvertent ingestion of dust (PBDE), and ingestion of food (mercury, PCBs, DDE, phthalates).

In a 2017 systematic review of air pollution's effect on infertility, it was concluded from a qualitative review of 61 studies that "both animal and human epidemiological studies support the idea that air pollutants cause defects during gametogenesis leading to a drop in reproductive capacities in exposed populations. Air quality has an impact on overall health as well as on the reproductive function" (Carré et al., 2017a). A 2018 study found an increase in fertility after the closure of coal and oil power plants in California as a function of distance to the powerplants (Casey et al., 2018), and in an animal study on rats, $PM_{2.5}$ exposure resulted in reduced sperm quality (Cao et al., 2017).

A population-based, prospective cohort study named the "Longitudinal Investigation of Fertility and Environment" (LIFE) study (2005– 2009) explored the role of the environment on fertility. From this study, Mendola et al. (2017) linked proximity to major roadways with prospectively measured time-to-pregnancy and infertility where the "likelihood of pregnancy was increased 3% for every 200 meters further away the couple's residence was from a major roadway," suggesting "major roadways may be related to reductions in fecundity," though this effect estimate was reduced when controlling for race/ethnicity, education, and income. Similarly in animal studies, Ogliari et al. (2013) observed a reduction in ovarian reserve (the reproductive potential remaining in a woman's ovaries) from diesel exhaust in intrauterine-exposed mice.

In the US nationally, two studies have investigated the geographic relationship of infertility. Swan et al. (2003) explored the geographic differences in semen quality of fecund US males by analyzing sperm from four prenatal clinics across the country. The study found that sperm quality "may be reduced in semirural and agricultural areas relative to more urban and less agriculturally exposed areas." Odisho et al. (2014) studied temporal and geospatial trends in male factor infertility with ART at a national scale using data from ART clinics and highlighted that "there are many factors known to be associated with male infertility," but "none of these associations has been explored spatially at the national level."

While there is an understanding in the literature of environmental factors that impair individual fecundity, there is a gap in understanding what environmental factors potentially reduce regional trends in fecundity and drive spatial differences in demand for ART services. Fertility clinic presence was modeled with a limited set of predictors in Chapter 2, and no ecological studies were found in the literature that investigated high-level drivers of the presence of fertility clinics. Only three studies were found that mapped US fertility clinics at the national level (Harris et al., 2017; Odisho et al., 2014; Pham et al., 2014), but no studies have explored the macro/population-level trends potentially driving the presence of fertility clinics in certain geographic areas over others. A single ecological study was found in the literature that modeled IVF utilization as a function of demographic, social, and economic factors (Hammoud et al., 2009). This study helped improve the understanding of the drivers of utilization, but environmental factors were crucially omitted, and the study was performed at the state level, which is too coarse to inform decision making on placement of new clinics.

3.1.3 Contribution

In this essay, I address the lack of ecological studies in the literature on ART utilization and ART clinic presence by conducting a large-scale ecological study to estimate the demand for ART treatment and presence of ART clinics as a function of social, demographic, economic, and environmental factors on the county scale for the entire US. I hypothesize (H8) that ART demand will be driven by socioeconomic trends on education and income and less on environmental factors because economic access to ART likely continues to be a barrier to accessing treatment, including for most people suffering from infertility due to environmental factors. However, these factors interact, and environmental factors are hypothesized to act to magnify the impact of socioeconomic factors. Additionally, I hypothesize (H9) that the presence of ART clinics will be best modeled using the rates of primary care physicians (PCPs) since the economics driving higher rates of PCPs likely also contribute to higher rates of ART clinics.

3.1.4 Chapter scope

This essay (Chapter 3) is organized as follows. First, in the methods section, data sources are summarized, and various machine learning (ML) models are described to be tested for predictive performance of ART demand and clinic presence. Then, results of applying the selected methods to the data are reported, including the predictive performance of the ML models and the selected final model's predictions, and hypotheses are tested. The discussion explores the significance of the essay's results. Finally, a summary is given focusing on the pertinent details of the essay.

3.2 Methods

This essay's methods are detailed in this section, beginning with where the ART demand data are sourced in Section 3.2.1, how clinic presence in counties is determined in Section 3.2.2

and what socioeconomic and environmental health data are compiled in Section 3.2.3. Next, machine learning methods tested for predictive performance are described in Section 3.2.4, including methods to model both counts and binary response variables to model ART demand and clinic presence respectively.

3.2.1 ART demand data

Data for ART treatment demand were drawn from the CDC's 2018 ART Success Rates dataset (CDC, 2020a). This dataset is publicly available and provides data on all ART clinics in the United States from 1995 to the current year with a two-year reporting lag to determine birth outcomes. This dataset contains annual statistics from each fertility clinic, detailing the numbers of ART cycles that were attempted for various disease diagnoses and age ranges. Cycle counts from each clinic were aggregated to the county level to produce a derived "demand" statistic for each county. A given clinic's demand was aggregated to the county level to calculate county demand, and Equation 3-1 defines mathematically how demand was calculated.

Equation 3-1. Demand for each county was calculated as the sum of all ART cycles performed by each ART practice j located within the county i.

$$Demand_{County\ i} = \sum_{Clinic\ j\ in\ County\ i} Total\ cycles_j$$

3.2.2 ART clinic presence data

Clinic presence data were sourced from the systematic web-search described in Chapter 2. This web-search, completed in December 2020 and January 2021, found 469 main clinics with embryology labs and 583 satellite clinics for a total of 1,052 clinic locations in the United States performing or facilitating ART. These clinic locations were spatially assigned to their corresponding county and the number of main and satellite clinics were calculated for each county.

3.2.3 Environmental, social, and economic data

A range of environmental, social, and economic datasets were included in the analysis of ART demand and the presence of ART clinics. Female reproductive-age population by county was calculated using five-year estimates from the 2019 American Community Survey (U.S. Census Bureau, 2019a). Because ART demand data were from the 2018 reporting to the CDC, state IVF insurance mandates implemented before 2018 were used and obtained from RESOLVE: The National Infertility Association's website on insurance coverage by state. States with IVF insurance mandates implemented before 2018 include: AR, CT, HI, IL, MA, MD, NJ, and RI (RESOLVE, 2021c).

The primary source of data on environmental, social, and economic factors came from the County Health Rankings & Roadmaps dataset from the University of Wisconsin Population Health Institute supported by the Robert Wood Johnson Foundation (County Health Rankings & Roadmaps, 2021a). These data are at the county-level nationwide and provide numerous measures of county health, compiled for each county in the US with data sourced between 2010 and 2019. Based on the literature summarized in this chapter above, I hypothesize that several measurers are likely to be good predictors of ART demand, including tobacco use, diet and exercise, alcohol and drug use, sexual activity, access to care, quality of care, education, employment, income, family and social support, community safety, air and water quality, and housing and transit. A subset of all the variables in the County Health Rankings & Roadmaps dataset were selected as potential predictors of ART demand based on applicability and potential for redundancy with other similar variables. The variables included in the statistical modeling are listed in Table 3-4.

Table 3-4. Subset of county health rankings variables used in statistical modeling.

Variable	Details
State IVF insurance mandate	1 if state has IVF insurance mandate before 2018, 0 otherwise
Female total population ages 20 to 49	Reproductive ages with higher likelihood of seeking fertility treatments
Poor physical health days	Average number of physically unhealthy days reported in past 30 days (age-adjusted)
Poor mental health days	Average number of mentally unhealthy days reported in past 30 days (age-adjusted)
Low birthweight	Percentage of live births with low birthweight (< 2,500 grams)
Adult smoking	Percentage of adults who are current smokers
Adult obesity	Percentage of the adult population (age 20 and older) that reports a body mass index (BMI) greater than or equal to 30 kg/m^2
Food environment index	Index of factors that contribute to a healthy food environment, from 0 (worst) to 10 (best)
Physical inactivity	Percentage of adults age 20 and over reporting no leisure-time physical activity
Access to exercise opportunities	Percentage of population with adequate access to locations for physical activity
Excessive drinking	Percentage of adults reporting binge or heavy drinking
Sexually transmitted infections	Number of newly diagnosed chlamydia cases per 100,000 population
Uninsured	Percentage of population under age 65 without health insurance
Primary care physician rate	Rate of primary care physicians to 100k population
Mental health provider rate	Rate of mental health providers to 100k population
Flu vaccinations	Percentage of fee-for-service (FFS) Medicare enrollees that had an annual flu vaccination
Some college	Percentage of adults ages 25-44 with some post-secondary education
Unemployment	Percentage of population ages 16 and older unemployed but seeking work
Income inequality	Ratio of household income at the 80th percentile to income at the 20th percentile
Social associations	Number of membership associations per 10,000 population
Violent crime rate	Number of reported violent crime offenses per 100,000 population

Table 3-4, Continued. Subset of county health rankings variables used in statistical modeling.

Variable	Details
Air pollution-particulate matter	Average daily density of fine particulate matter ($PM_{2.5}$) in micrograms per cubic meter
Drinking water violations	Indicator of the presence of health-related drinking water violations.
Long commute and drives alone	Among workers who commute in their car alone, the percentage that commute more than 30 minutes
Infant mortality	Number of all infant deaths (within 1 year), per 1,000 live births
Diabetes	Percentage of adults aged 20 and above with diagnosed diabetes
HIV prevalence	Number of people aged 13 years and older living with a diagnosis of human immunodeficiency virus (HIV) infection per 100,000 population
Drug overdose deaths	Number of drug poisoning deaths per 100,000 population
Insufficient sleep	Percentage of adults who report fewer than 7 hours of sleep on average
Median household income	The income where half of households in a county earn more and half of households earn less
Traffic volume	Average traffic volume per meter of major roadways in the county
Severe housing cost burden	Percentage of households that spend 50% or more of their household income on housing
Broadband access	Percentage of households with broadband internet connection

Each of these covariates was normalized to facilitate better predictive performance of the machine learning algorithms. This was done by first applying the Box-Cox power transform, which transforms the predictors to resemble the normal distribution more closely. Then, the predictors were centered and scaled by first subtracting their means and then dividing by their standard deviations. This process was streamlined in R using the preProcess function in the caret package (RDocumentation, 2022b).

3.2.4 Machine learning models tested

Accurate prediction of demand (treated as a continuous dependent variable) and presence/absence of fertility clinics (treated as a classification dependent variable) was the overarching modeling goal, and the secondary goal of this analysis was to gain insights into the drivers of fertility clinic demand and presence. With these goals in mind, a range of statistical and machine learning (ML) models (described in the subsections 3.2.4.4 through 3.2.4.6 below) were compared with respect to their accuracy in prediction of demand and presence of clinics. Prior to testing predictive performance, it is impossible to know which ML model will have the best predictive performance of a particular response variable using a set of predictors. These predictive models vary in their approaches with more traditional statistical methods assuming linear relationships between the independent variables and the response variable, and newer machine learning models allowing for non-linear relationships to be captured.

Though traditional statistical models are believed generally to be worse at prediction compared with more complex ML models (Obermeyer & Emanuel, 2016), newer ML models are able to yield better predictions through increasing complexity. Highly complex ML models may be able to predict better than traditional statistical models or simpler ML models, but they do so by giving up explainability. For example, predictions from traditional statistical methods are relatively easily explained with a single mathematical equation. The drivers of prediction of the relatively simple ML model classification and regression trees (CART) are also relatively easy to explain with a graphical tree showing the training data split based on predictors. Conversely, the highly complex random forest algorithm usually contains hundreds of CART trees, and it is usually able to yield better predictions than a single CART model. But it is difficult to explain specifically how a prediction was reached when hundreds of CART models contributed to it.

In the case of placing new ART clinics, I argue that the ability to explain how model predictions are made is critical to gaining trust of decision makers whose use of those predictions hinges on that trust. This argument is supported by findings in the literature that explainability generates trust in predictive models (Shin, 2021) and that physician understanding, explainability, and trust in ML models are related (Diprose et al., 2020). Therefore, the statistical and ML models used to predict ART demand and clinic presence were selected with a preference for models with greater ability to explain their predictions. Now, returning to the topic of how predictive performance was evaluated for this essay, in Section 3.2.4.1 that follows I provide some further information about the cross-validation procedures used to evaluate predictive performance of each model tested.

3.2.4.1 Predictive model selection procedure

The approach to evaluate the various ML models is depicted in Figure 3-1 and is detailed as follows. For the demand modeling, only the counties with reported demand to the CDC were used to fit models, but for the presence of clinics modeling, all US counties were used to fit the models. The approach used to test the predictive performance of various models considered was to break the full dataset into two parts: the training dataset and the test dataset. The training dataset comprised of a random selection of 80% of the original dataset, and the remaining 20% was saved for final testing after model selection.

The training dataset was then used in 10-fold cross-validation, repeated three times facilitated by R's caret package (V6.0-88). 10-fold cross-validation consists of randomly assigning the training dataset into 10 bins. Then, 10 cross-validation "runs" are performed whereby data from 9 of the 10 bins are used to fit each of the models to be tested, and each fit model is used to predict the held-out data from 1/10th of the training dataset. Each of the 10 cross-validation runs uses a different set of held-out data. This process is repeated three times with different random assignments each time of the training dataset into 10 bins, so at the end of the process, 30 estimates of predictive performance are generated. This cross-validation process allows the models with the best predictive performance to be found and guards against the dangers of overfitting by estimating performance by predicting out-of-sample performance metrics.

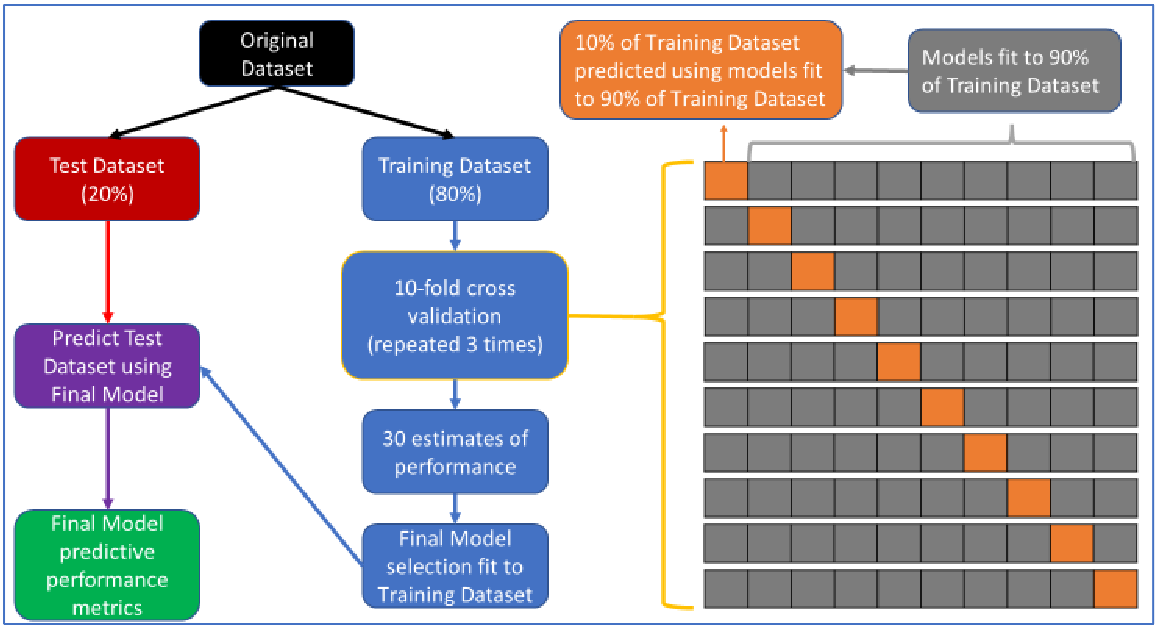


Figure 3-1. A depiction of the model selection process. The original dataset (show in black) is split into the training dataset (in blue) and the test dataset (in red). 10-fold cross-validation repeated three times is performed on the training dataset to generate 30 estimates of predictive performance for each ML model considered to allow for a Final ML Model to be selected. This Final Model is fit to the Training Dataset and is used to predict the never-seen Test Dataset to create unbiased estimates of predictive performance of the Final Model.

During the training phase, each model may be tuned to improve its performance but only using the training dataset. The test dataset is never considered in model tuning and selection, and the final selected model is fit to the entire training dataset to be tested against the never-seen 20% of the original dataset that was held out initially. This allows for an estimate of predictive performance to be generated that is unbiased (Kuhn & Johnson, 2013).

Predictive performance of models performing regression of ART demand was measured using mean absolute error (MAE), root mean squared error (RMSE), and goodness of fit (R^2). In each cross-validation iteration, MAE was calculated by averaging the absolute value of the differences between the predicted values and the actual values, RMSE was calculated by taking the square root of the average of the differences between the predicted values and the actual values squared, and R^2 was calculated by squaring the correlation between the observed and predicted values.

Predictive performance of models performing classification of ART clinic presence/absence was measured using accuracy and Cohen's kappa (Cohen, 1960), known simply as "kappa." Accuracy was calculated as the number of correctly predicted counties divided by the total number of counties. The kappa statistic accounts for the likelihood that a correct prediction was made by chance. The kappa statistic was calculated according to Equation 3-2. Kappa is particularly useful in the case of an imbalance of classes. Since the percentage of counties without a main clinic was 94% and the percentage of counties without a satellite clinic was 91%, a predictive model would be expected to be more accurate than the model that simply predicts no clinic in every county.

Equation 3-2. The calculation of Cohen's kappa statistic. p_o is the proportion of observed agreements and p_e is the proportion of agreements expected by chance.

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

3.2.4.2 Machine learning models to predict ART demand

To predict ART demand, a range of machine learning models were selected as a good representation of various supervised learning methods. Supervised learning algorithms use training data with a known label or result to fit a model to achieve a desired level of accuracy on the training data. ML algorithms were selected from a full range of different types of ML algorithms for performing regression, including linear regression, instance-based methods, decision tree algorithms, and ensemble methods. Artificial neural network algorithms were not tested because they are generally unexplainable “black box” models. This decision was made because I believe the ability to explain the mechanisms of prediction is critical to engendering trust in model predictions. Therefore, the following popular (Brownlee, 2019) ML models were chosen as a representative set of the full range of different types of machine learning algorithms used for regression: linear regression, log-linear regression, Poisson regression, negative binomial regression, classification and regression trees (CART), support-vector machines, k-nearest neighbors, random forest, generalized boosted regression modeling (GBM), and cubist. Each of these models is conceptually described in Section 3.2.4.4 or Section 3.2.4.5 below, and their relative predictive performance was compared using the metrics MAE, RMSE, and R^2, which are detailed in Section 3.2.4.1 above. Some of these models treat the count data as a continuous variable (e.g., linear regression), and others explicitly represent the dependent variable as integer non-negative counts (e.g., Poisson regression). In contrast, the models of the next section are classification models in which the dependent variable being predicted is membership in a class.

3.2.4.3 Presence ML models

To predict ART clinic presence/absence, a range of machine learning models were selected as a good representation of various supervised learning methods as in the previous section. ML algorithms were selected from a full range of different types of ML algorithms for performing classification, including linear regression, instance-based methods, decision tree algorithms, and ensemble methods. Again, artificial neural network algorithms were excluded from the analysis because they are “black boxes.” Therefore, the following popular (Brownlee, 2019) ML models were chosen as a representative set of the full range of different types of machine learning algorithms used for classification: logistic regression, classification and regression trees (CART), support-vector machines (SVM), k-nearest neighbors, random forest, generalized boosted regression modeling (GBM), and C5.0. Each of these methods modeled the presence or absence of a clinic as a categorical variable, and each model is conceptually described in Section 3.2.4.5 or Section 3.2.4.6 below. Relative predictive performance of each model was compared using the metrics accuracy and kappa, which are detailed in Section 3.2.4.1 above.

3.2.4.4 Regression-only capable machine learning algorithms tested

3.2.4.4.1 Linear regression

Linear regression is the first of five methods that I describe that are capable of performing regression only. It assumes the relationship between the response variable and its predictors is linear. When modeling ART demand, linear regression takes the form shown in Equation 3-3, and maximum likelihood estimation (MLE) was used to determine the best values of β_0 through β_p to model ART demand.

Equation 3-3. The mathematical relationships of linear regression for parameters 1 through p with residual error distributed normally with a mean of 0 and variance σ^2.

$$\text{Demand} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon$$

$$\epsilon \sim N(0, \sigma^2)$$

Linear regression models were then fit sequentially using backward stepwise Akaike information criterion (AIC) feature selection. The AIC metric estimates predictive error by balancing the benefit of R^2 (goodness of fit) with the number of parameters in the model. This discourages the overfitting that occurs when increasing the number of model parameters increases R^2. Backward stepwise AIC feature selection sequentially removes variables until the minimum AIC is achieved (Venables & Ripley, 2002).

3.2.4.4.2 Log-linear regression

Log-linear regression is almost identical to linear regression, but it guarantees that the predictions are positive, which is appropriate for count data, which is never negative. Log-linear regression simply log transforms the response variable, such that it assumes the relationship between the natural log of the response variable and its predictors is linear. When modeling ART demand, log-linear regression takes the form shown in Equation 3-4, and maximum likelihood estimation (MLE) was used to determine the best values of β_0 through β_p to model ART demand. Stepwise backward AIC feature selection was then performed iteratively to produce a simpler log-linear regression model to reduce overfitting and hopefully improve predictive performance.

Equation 3-4. The mathematical relationships of log-linear regression for parameters 1 through p with residual error distributed normally with a mean of 0 and variance σ^2.

$$\log\,(\text{Demand}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon$$

$$\epsilon \sim N(0, \sigma^2)$$

3.2.4.4.3 Poisson regression

Poisson regression is a form of generalized linear models (GLM) commonly used in epidemiological research when modeling counts. When using Poisson regression to model demand, an assumption is made that demand is Poisson distributed defined by a single parameter μ representing the mean and the variance, which are assumed to be equal. The logarithm of μ is assumed to be a linear combination of fit parameters (Waller & Gotway, 2004). These relationships are shown in Equation 3-5 with N representing female reproductive-age population as an offset. Stepwise AIC feature selection was then performed iteratively to produce a simpler Poisson regression model to reduce overfitting and hopefully improve predictive performance.

Equation 3-5. The mathematical relationships of the Poisson GLM with mean μ for parameters 1 through p.

$$\text{Demand} \sim \text{Poisson}(\mu)$$

$$\log(\mu) = \log(N) + \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$$

3.2.4.4.4 Negative binomial regression

Negative binomial (NB) regression is a variant of Poisson regression also used to model counts that allows for the mean and variance to be unequal. Negative binomial regression is a Poisson-gamma distribution mixture model with a mean of μ and a variance of $\mu + \mu^2/\theta$. This two-stage model with an unobserved random variable E distributed $gamma(\theta)/\theta$ and Demand given E is Poisson distributed with mean μE (Venables & Ripley, 2002). These relationships are shown in Equation 3-6 with N representing the population as an offset. Stepwise AIC feature selection was then performed iteratively to produce a simpler negative binomial regression model to reduce overfitting and hopefully improve predictive performance. Upon fitting the regression model, variable importance was calculated using the absolute value of the t-statistic, computed from the covariate's regression coefficient divided by its standard error.

Equation 3-6. The mathematical relationships of the Negative Binomial GLM with mean μ and fit θ for parameters 1 through p.

$$\text{Demand} \mid \text{E} \sim \text{Poisson}(\mu E), \qquad E \sim gamma(\theta)/\theta$$

$$\log(\mu) = \log(N) + \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$$

3.2.4.4.5 Cubist

Cubist is a regression algorithm that combines a CART model with linear regression. This hybrid approach generally results in better predictions than linear regression alone, while allowing for easier interpretation of linear regression. The Cubist algorithm binarily partitions the training dataset, then fits linear regression models to each of the partitioned datasets (Quinlan, 1992). Cubist contains the tuning parameters "committees" and "neighbors." The committees parameter allows for a boosting-like methodology to be applied in which subsequently fits cubist models are fit based on the residual errors from the previously fit model. The neighbors parameter allows for combining the predictions from cubist with the nearest neighbor algorithm (Kuhn et al., 2012). 15 randomly selected combinations of these parameters were used to tune the model for the best predictive performance.

3.2.4.5 Both regression and classification capable machine learning algorithms tested

3.2.4.5.1 Classification and regression tree (CART)

Classification and regression tree (CART) models iteratively binarily partition the dataset used to train the model based on the value of the predictor that minimizes the sum of the squared errors between the fit values and the observed values. CART models, as their name suggests, can be used for either classification or regression. In the case of regression, the fit value of a given partition is simply the mean of the observed values of the response variable for the observations that were partitioned into that side of the split, and in the case of classification, each partitioned observation votes using its observed classification to determine the fit (Breiman et al., 1984). This process is repeated until no more than 30 levels are fit, and, to reduce overfitting, the tree is pruned (Therneau & Atkinson, 1997).

3.2.4.5.2 Random forest

Random forest (RF) models are an ensemble method of classification and regression in which numerous CART models are fit to the training dataset. Each CART model is fit independently using a randomly chosen subset of the full set of predictors. In the case of classification, each tree in the forest votes, and the most popular class is selected as the prediction of the forest. In the case of regression, the forest prediction is the average prediction of each tree of the forest (Breiman, 2001). The number of variables randomly sampled as candidates at each binary partition is set via the parameter *mtry*, and the best *mtry* parameter was found from a set of 15 randomly generated values.

3.2.4.5.3 Gradient boosting machine

Gradient boosting machine (GBM) is a classification or regression algorithm that is very similar to the random forest model except that the trees are fit sequentially rather than independently. Each subsequent tree fit attempts to better predict the cases where the prior model performed poorly (Friedman, 2001). Stochastic Gradient Boosting is a slight variant of GBM used for this essay that randomly selects a subset of the training dataset without replacement and has been observed to produce better predictions (Friedman, 2002). A few tuning parameters exist for the GBM model (number of boosting iterations, maximum tree depth, shrinkage, and minimum terminal node size). 15 randomly selected combinations of these parameters were used to tune the model for the best estimated predictive performance.

3.2.4.5.4 Support-vector machine

Support-vector machines (SVMs) are machine learning algorithms that find a hyperplane that optimally separates predictors into different classes. SVMs can be used for classification or regression using a linear or non-linear kernel function regardless of the type of classification. In this essay, a non-linear or radial kernel function was used to model demand and presence of fertility clinics. The best hyperplane is one that separates the two classes with the largest margin between the hyperplane and each class. SVMs allow for certain samples to be within a "soft margin" via a cost parameter that penalizes samples inside the margin (Noble, 2006). This cost parameter C was tuned by selecting the best value of C from a set of 15 randomly generated values.

3.2.4.5.5 k-Nearest neighbors

k-Nearest neighbors (KNN) algorithm can perform both classification and regression. It is a non-parametric method that finds the *k* closest (in Euclidian distance) training set vectors. In KNN classification, each of the *k* training set vector votes its classification, which is decided by plurality vote, and ties are broken randomly. In KNN regression, the prediction is determined by averaging the value of the response variable of the *k* nearest training set vectors (Fix & Hodges, 1989). The parameter *k* was tuned by selecting the best value of *k* from a set of 15 randomly generated values.

3.2.4.6 Classification-only capable machine learning algorithms tested

3.2.4.6.1 C5.0

C5.0 is a classification algorithm very similar to the Cubist algorithm detailed in Section 3.2.4.4.5. It also combines tree-based learning with linear regression, but it is used for classification rather than regression (Quinlan, 1993). Similar to Cubist, C5.0 has a boosting capability, but it doesn't utilize the nearest neighbor algorithm. The two tuning parameters available are the number of boosting iterations and whether to perform feature selection to "winnow" the model (Kuhn & Johnson, 2013). 15 randomly selected combinations of these parameters were used to tune the model for the best predictive performance.

3.2.4.6.2 Logistic regression

Logistic regression is a form of generalized linear models (GLM) used for classification when modeling binary response variables that are true or false. When using logistic regression to model presence or absence of ART clinics in US counties, an assumption is made that the presence of an ART clinic is Bernoulli distributed defined by a single parameter p representing the probability. The *logit* function of p is assumed to be a linear combination of fit parameters (Waller & Gotway, 2004). These relationships are shown in Equation 3-7.

Equation 3-7. The mathematical relationships of the Logistic GLM with probability p for parameters 1 through p.

$$\mathrm{Y} = \begin{cases} 1\ if\ ART\ clinic\ present \\ 0\ if\ ART\ clinic\ is\ not\ present \end{cases}$$

$$\mathrm{Y} \sim \mathrm{Bernoulli}(p)$$

$$logit(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$$

Logistic regression models were fit sequentially using stepwise Akaike information criterion (AIC) regression. The AIC metric estimates predictive error by balancing the benefit of R^2 (goodness of fit) with the number of parameters in the model. This discourages the overfitting that occurs when increasing the number of model parameters increases R^2. The stepwise AIC feature selection sequentially removes variables until the minimum AIC is achieved (Venables & Ripley, 2002).

3.3 Results

The results section first covers the findings on the predictive performance of the machine learning models tested to predict ART demand in Section 3.3.1.1 and fertility clinic presence in Section 3.3.1.2. Next, the final models selected for their predictive strength and relative simplicity were used to predict ART demand in Section 3.3.2 and clinic presence in Section 3.3.3 for every county in the US. Then, each hypothesis posited in the introduction is tested in Section 3.3.4. Finally, a case study in identifying areas for new clinics is presented in Section 3.3.5 that identifies new areas to locate economically viable clinics across the US.

3.3.1 Predictive performance of models

3.3.1.1 ART demand modeling results

The quality of predictions from different models was compared using predictive performance measures calculated using the 30 cross-validation runs. These predictive performance metrics include mean absolute error (MAE), root mean squared error (RMSE), and goodness of fit (R^2). These measures were calculated for each machine learning model tested in the prediction of demand. The out-of-sample predictive MAE performance results are shown in Figure 3-2, with the range of predictive performance for each model with the boxes showing the 25^{th} percentile to the 75^{th} percentile, the whiskers showing the local minimum and maximum, the horizontal lines showing the medians, and the ‘X’ symbols showing the means. The Cubist model performed on average the best and the log-linear model performed on average the worst.

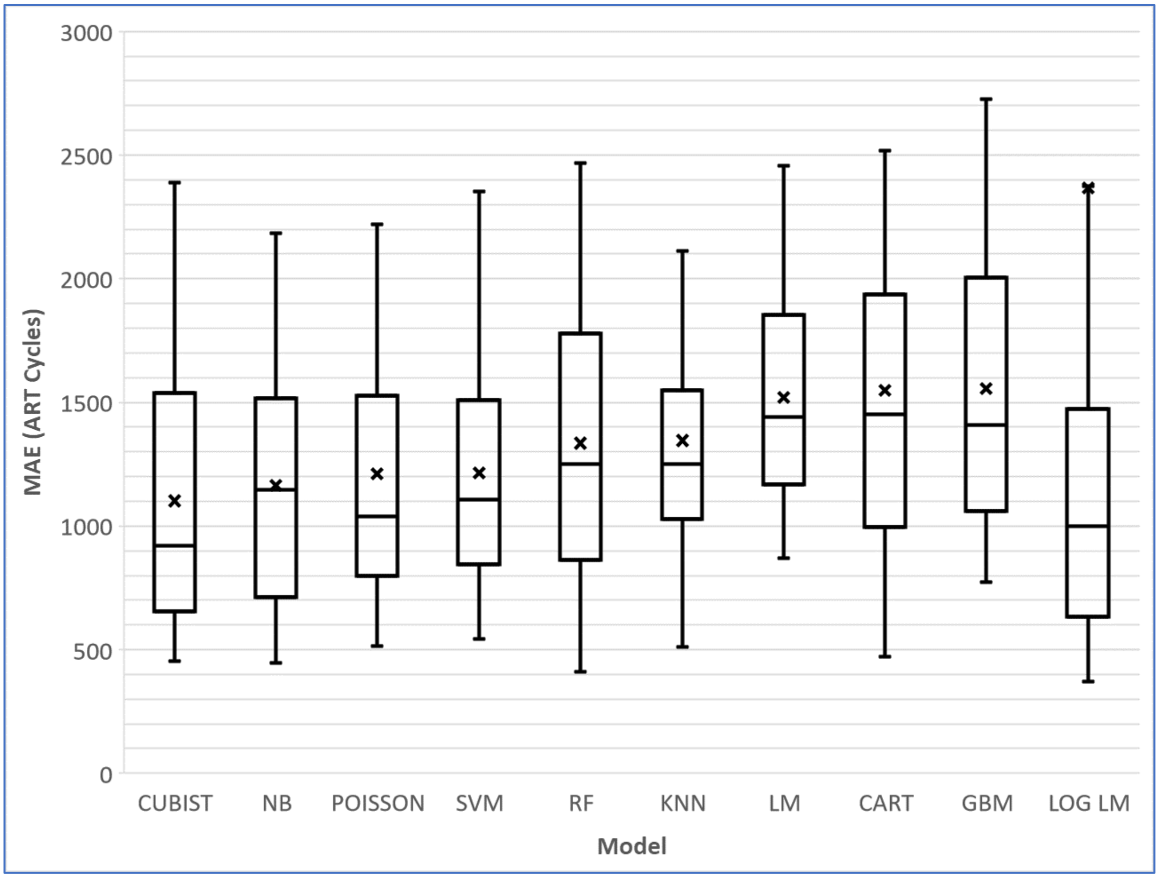


Figure 3-2. Comparison of machine learning algorithms estimated MAE predictive performance (out-of-sample) in predicting demand for ART (ordered from best to worst average MAE). Each box shows the 25th and 75th percentile, whiskers show local minimum and maximum, the horizontal lines are the medians, and the 'X' symbols are the means. Independent variable (y axis) is counts of number of ART cycles within a county.

Legend: CUBIST = cubist, NB = negative binomial regression, POISSON= Poisson regression, SVM = support-vector machines, RF = random forest, KNN = k-nearest neighbors, LM = linear regression, CART = classification and regression tree, GBM = gradient boosting machine, and LOG LM = log-linear regression.

The RMSE results (Figure 3-3) show the range of out-of-sample predictive performance of the various models, with the support-vector machines (SVM) model performing on average the best and the log-linear model performing on average the worst. The R^2 results (Figure 3-4) also show the range of out-of-sample predictive performance, with the log-linear model performing on average the best and the CART model performing on average the worst.

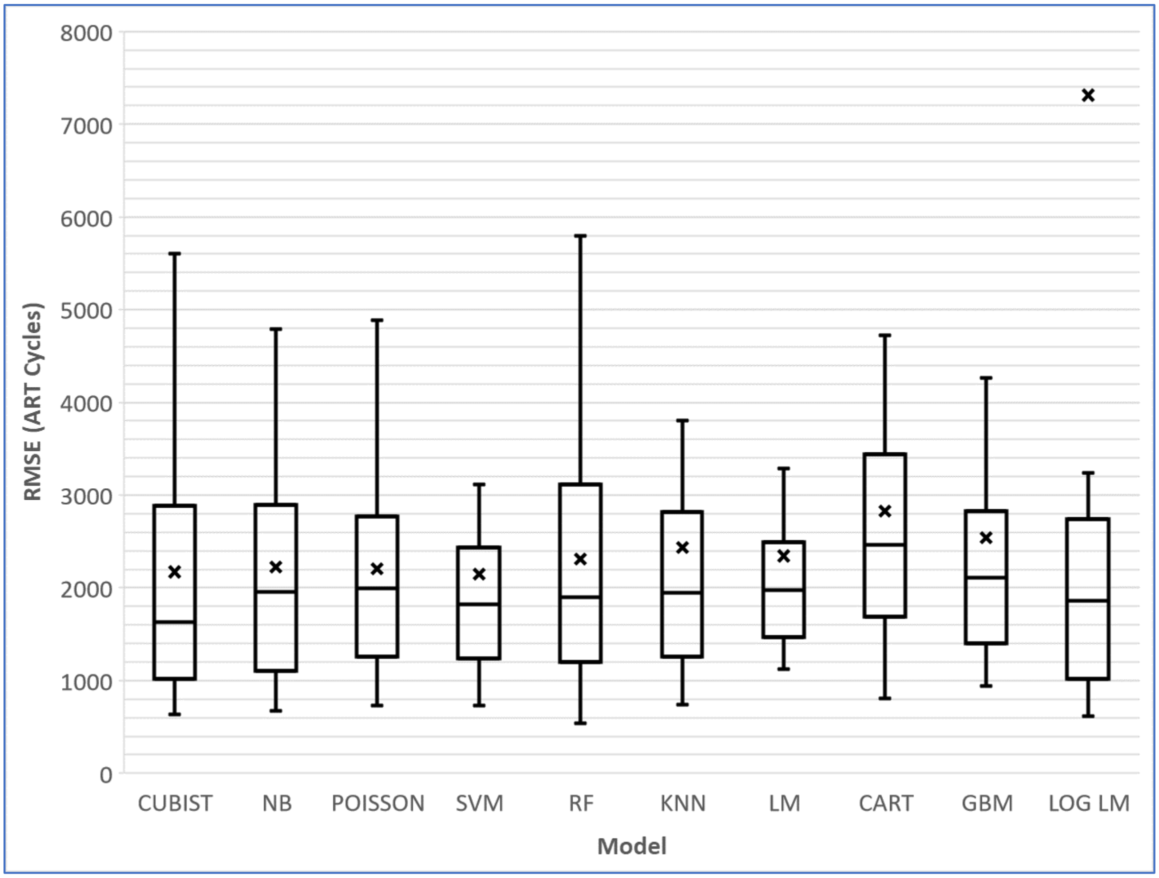


Figure 3-3. Comparison of machine learning algorithms estimated RMSE predictive performance in predicting demand for ART (ordered to be consistent with Figure 3-2). Each box shows the 25th and 75th percentile, whiskers show local minimum and maximum, the horizontal lines are the medians, and the 'X' symbols are the means. Independent variable (y axis) is counts of number of ART cycles within a county.

Legend: CUBIST = cubist, NB = negative binomial regression, POISSON= Poisson regression, SVM = support-vector machines, RF = random forest, KNN = k-nearest neighbors, LM = linear regression, CART = classification and regression tree, GBM = gradient boosting machine, and LOG LM = log-linear regression.

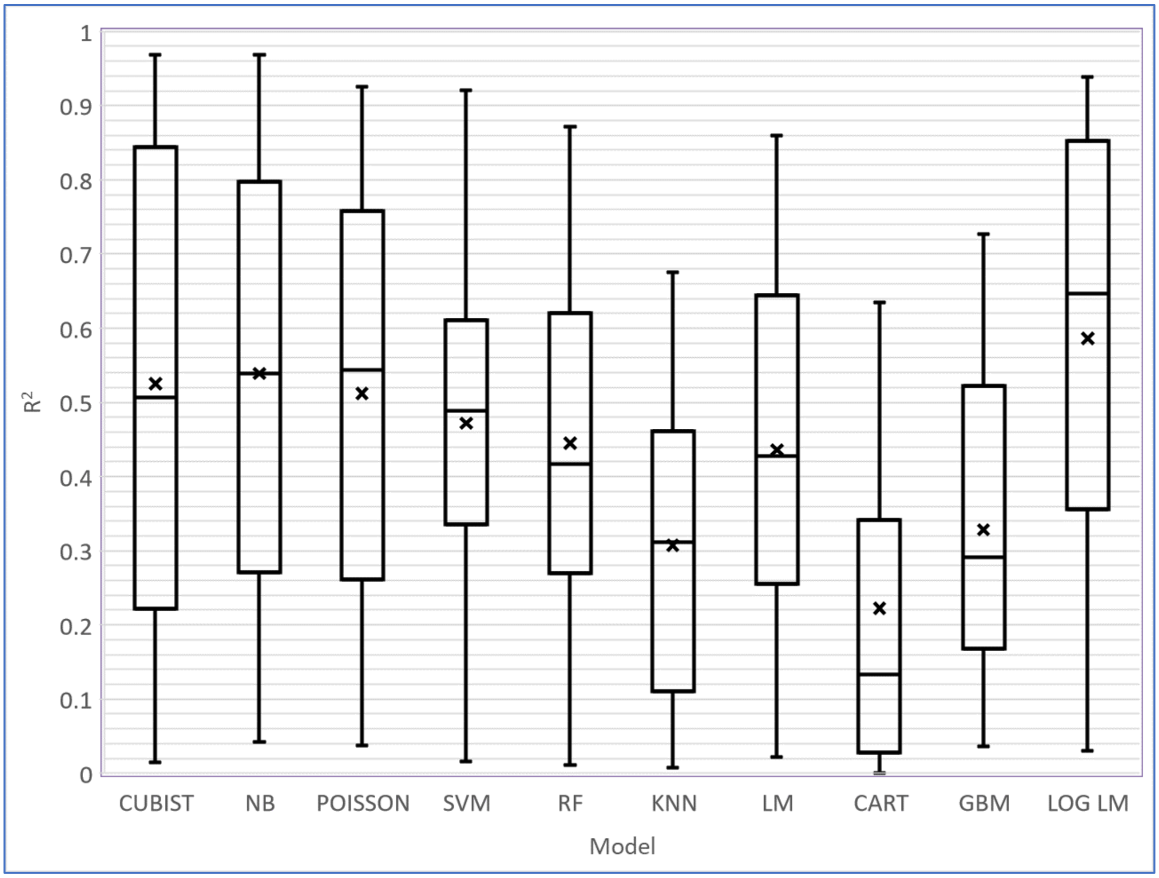


Figure 3-4. Comparison of machine learning algorithms estimated out-of-sample goodness of fit (R^2) in predicting demand for ART (ordered to be consistent with Figure 3-2). Each box shows the 25th and 75th percentile, whiskers show local minimum and maximum, the horizontal lines are the medians, and the 'X' symbols are the means.

Legend: CUBIST = cubist, NB = negative binomial regression, POISSON= Poisson regression, SVM = support-vector machines, RF = random forest, KNN = k-nearest neighbors, LM = linear regression, CART = classification and regression tree, GBM = gradient boosting machine, and LOG LM = log-linear regression.

Interestingly, the log-linear model had the best average and median R^2 (0.65 and 0.59 respectively), but it had the worst average MAE and RMSE (2,368 and 7,314). However, the log-linear model had the second-best median MAE (998) and the third-best RMSE (1,858). The best explanation for the wide swings in predictive performance between the log-linear model's metrics is that the log-linear model generally predicts very well resulting in good R^2 (the square of the correlations between the observed and predicted values). Similarly, the log-linear model's MAE and RMSE metrics were sometimes best for a given cross-validation run, but occasionally it predicted very poorly. For example, on three cross-validation runs, its MAE was above 11,000, but none of the other models ever had predictive error in excess of 4,000.

Top performing models from the cross-validation included cubist, negative binomial regression, support-vector machines (SVM), and log-linear regression. Log-linear regression had the highest average R^2, and its median MAE and RMSE performance were both top ranked, but its mean performance was the worst with 10% of its MAEs and RMSEs being extremely high (see Section 3.3.1.1 for more details). Therefore, log-linear was not selected as the final model. Cubist and SVM models had the best MAE and RMSE respectively, so they were strongly considered to be selected as the final model. However, the much simpler model, negative binomial regression, had the second-best mean and median R^2 and very good MAE and RMSE performance (within 6% of the performance of Cubist for MAE and within 4% of the performance of SVM for RMSE).

With different performance metrics supporting the selection of different models, I applied the heuristic of Occam's razor heuristic, which is often applied when choosing between machine learning models. Occam's razor heuristic can be described "other things being equal, simple theories are preferable to complex ones" (Witten et al., 2016). Therefore, when applying the Occam's razor heuristic, the negative binomial model was selected as the simplest of the top performing models when modeling ART demand.

The negative binomial model also has the added benefit of easier interpretation that I believe is critical to establishing trust in its predictions. This is particularly important because the predictions are used as the inputs to for optimization methods in Chapter 4. If an overly complex model is selected in this chapter that cannot be explained easily and its predictions are not trusted by a hypothetical decision maker, I risk this lack of trust in the predictions from this chapter spreading to the recommendations from the optimization model in Chapter 4.

Upon selecting negative binomial regression, predictive performance was tested against the held-out test dataset to perform an unbiased validation of performance. The negative binomial model that was fit to the training dataset (80% of total) was used to predict the held-out test dataset (20% of total), and performance metrics were calculated. The MAE predicting the test dataset was 646 ART cycles, and the RMSE predicting the test dataset was 990 ART cycles. These validation test results fall within the performance estimates generated from the training phase cross-validation results used in model selection.

3.3.1.2 Clinic presence modeling results

Various machine learning algorithms were tested for their performance in predicting the presence of main and satellite clinics. Logistic regression was found to be the best at predicting the presence of main clinics. The logistic regression model was found to be the best performer with respect to both accuracy and kappa. The estimates of performance for each model tested are shown in Figure 3-5 for accuracy and Figure 3-6 for kappa.

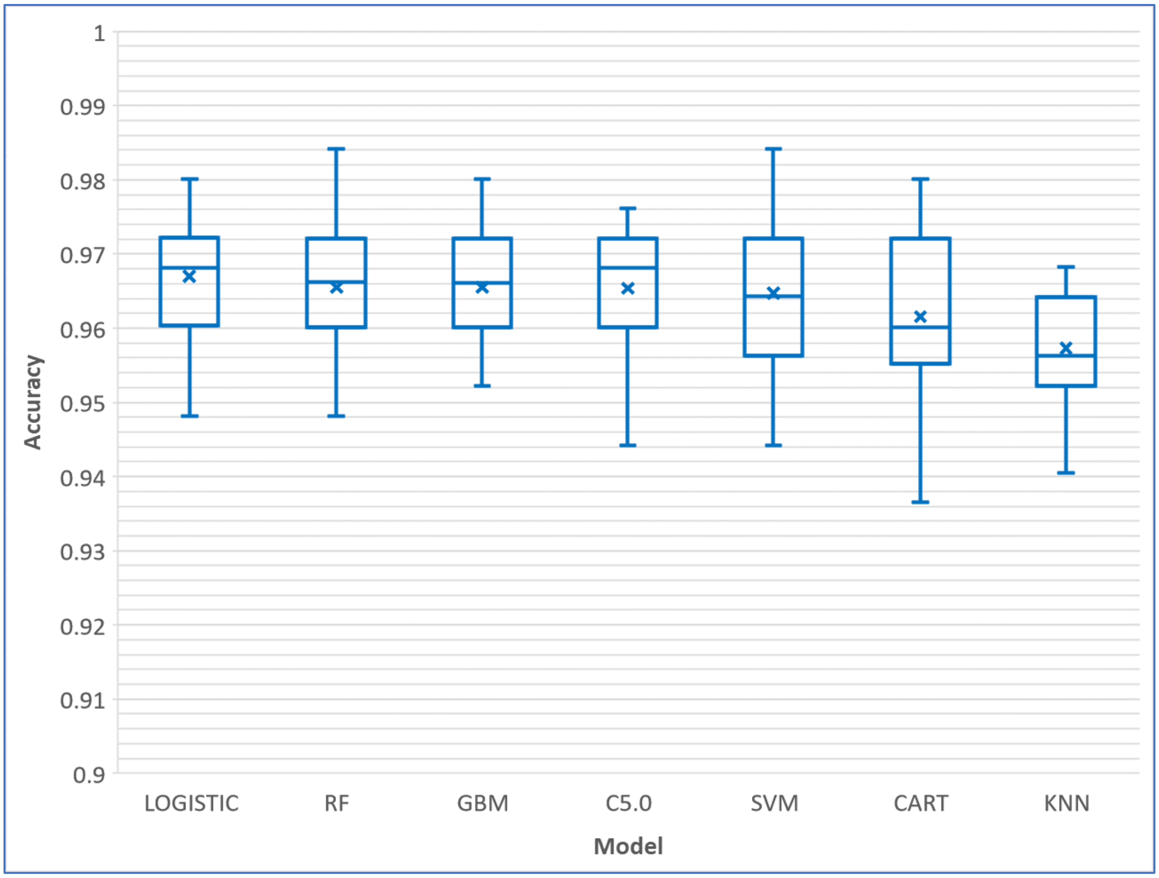


Figure 3-5. Estimated accuracy of machine learning algorithms used to predict the (out-of-sample) presence of main ART clinics (ordered from best to worst average accuracy from left to right). Each box shows the 25th and 75th percentile, whiskers show local minimum and maximum, the horizontal lines are the medians, and the 'X' symbols are the means.

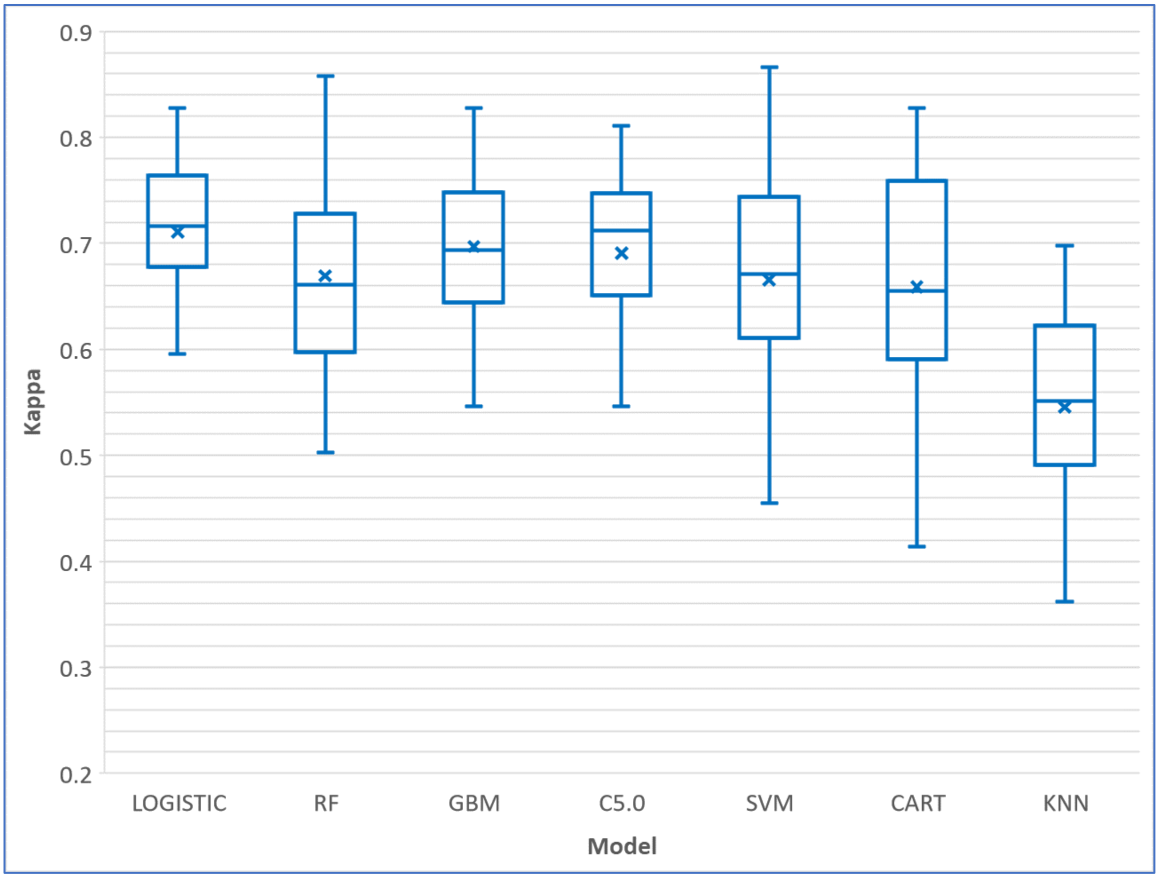


Figure 3-6. Estimated kappa of machine learning algorithms used to predict the out-of-sample presence of main ART clinics (ordered to match Figure 3-5). Each box shows the 25th and 75th percentile, whiskers show local minimum and maximum, the horizontal lines are the medians, and the 'X' symbols are the means.

For satellite clinic presence modeling, the gradient boosting machine (GBM) model was found to have the best estimated kappa but only the third best accuracy, and logistic regression was found to be the best model with respect to accuracy and the second best with respect to kappa. The estimates of performance for each model tested is shown in Figure 3-7 for accuracy and Figure 3-8 for kappa.

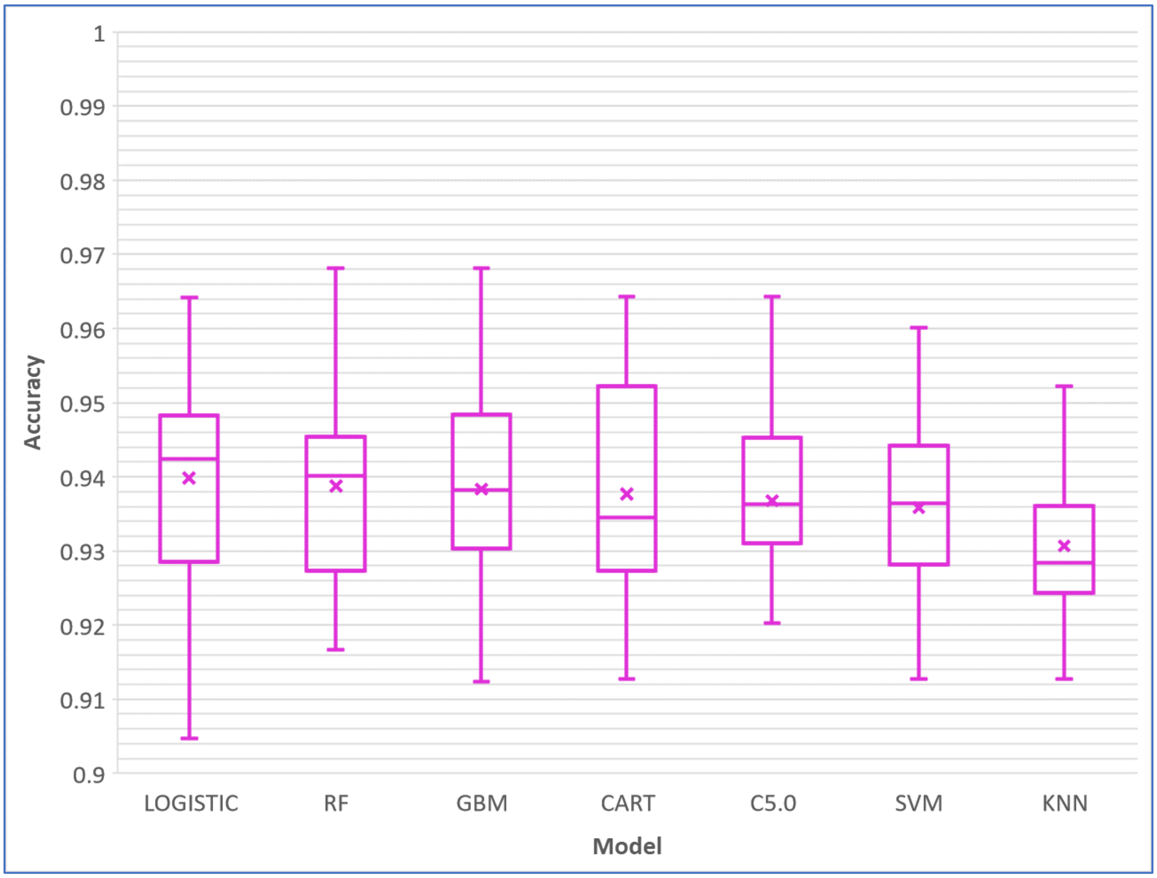


Figure 3-7. Estimated accuracy of machine learning algorithms used to predict the presence of satellite ART clinics (ordered from best to worst average accuracy from left to right). Each box shows the 25th and 75th percentiles, whiskers show local minimum and maximum, the horizontal lines are the medians, and the 'X' symbols are the means.

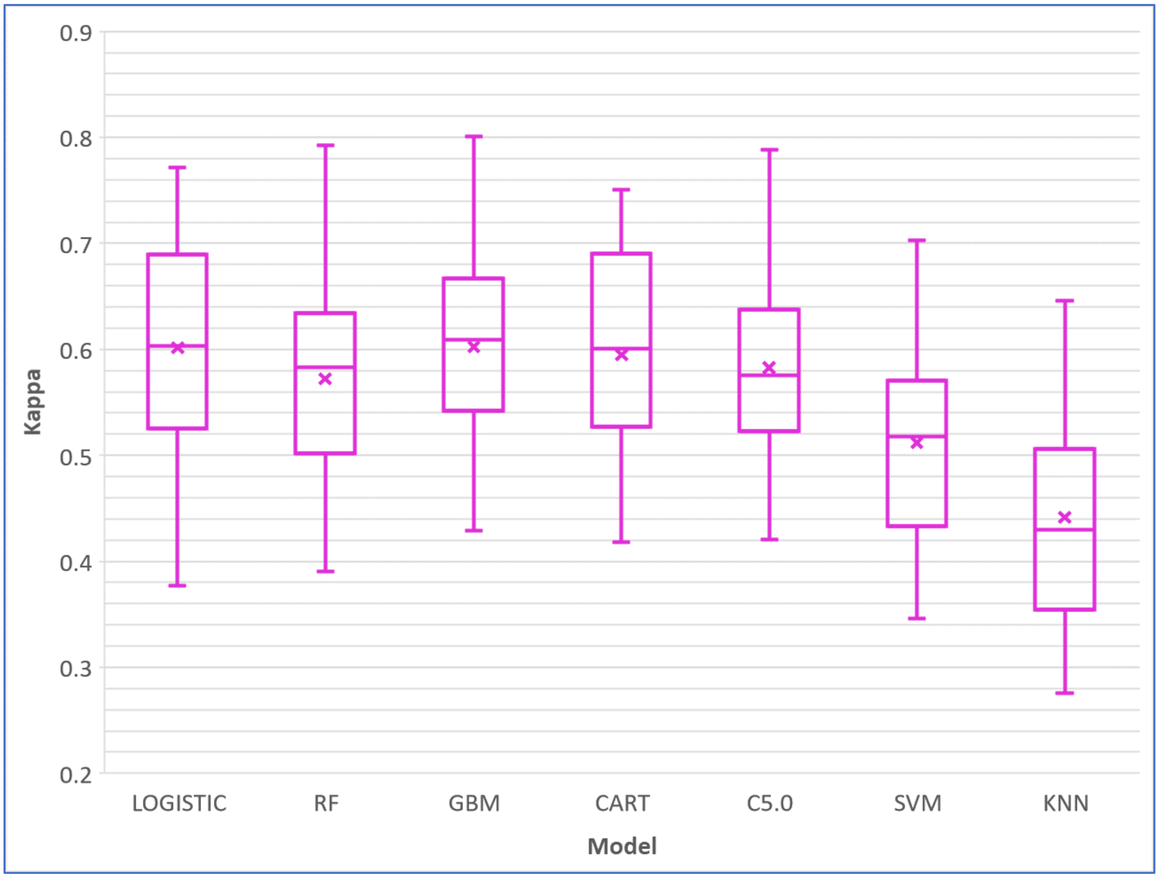


Figure 3-8. Estimated kappa of machine learning algorithms used to predict the presence of satellite ART clinics (ordered to match Figure 3-7). Each box shows the 25th and 75th percentiles, whiskers show local minimum and maximum, the horizontal lines are the medians, and the 'X' symbols are the means.

Based on the relative complexity and performance of the various models at predicting the presence of main or satellite clinic presence, logistic regression was selected as the final model. Logistic regression was one of the top performers for predicting the presence of both main and satellite clinics, and because logistic regression is a relatively simple model compared with the other models tested and it was a top predictor of both clinic types, it was selected as the final model.

Performance of the logistic regression model was then validated against the 20% test dataset that was held out. This was done by fitting a logistic regression model to the training dataset (80% of the original) using the predictors described in Section 3.2.3. This model was then simplified using backward stepwise AIC feature selection to reduce overfitting by penalizing the number of variables in the model. This simplified logistic regression model was then used to predict the never-seen/held-out test dataset and predictive performance metrics were calculated.

The logistic regression model predicted the likelihood of an ART clinic for the 628 counties in the test dataset (20% of the original dataset). If the predicted likelihood of a clinic for a county was greater than 50%, then the county was predicted to have an ART clinic. The predictions for every test dataset county were compared against actual presence or absence of a clinic. The accuracy of prediction of main clinics was 98% with a 95% confidence interval between 96% and 99%, and kappa was 0.78. The confusion matrix of all predictions compared with referenced truth for main clinics is shown in Table 3-5. The accuracy of prediction of satellite clinics was 95% with a 95% confidence interval of 93% to 96%, and kappa was 0.62. The confusion matrix of all predictions compared with referenced truth for satellite clinics is shown in Table 3-6. For both clinic types, there was a slightly higher proportion of false negatives over false positives, but those differences were small in both cases. These results fall within the performance estimates generated from the cross-validation process used in model selection.

Table 3-5. Performance of logistic regression at predicting main clinics presence in a county for the held out test dataset (20% of the orignal dataset).

Prediction \ Reference	FALSE	TRUE
FALSE	584	10
TRUE	5	29

Table 3-6. Performance of logistic regression at predicting satellite clinics presence in a county for the held out test dataset (20% of the orignal dataset).

Prediction \ Reference	FALSE	TRUE
FALSE	563	19
TRUE	15	31

3.3.2 Prediction of ART demand for every county in the US

Lastly, the same demand modeling approach used for the training dataset was replicated for the full dataset of all counties with ART cycles reported to the CDC in 2018. As with the training dataset, prior to model fitting, predictors (independent variables) that were highly correlated (>70%) were removed to reduce redundant variables and reduce multicollinearity. Multicollinearity is not a problem that affects prediction generally, but it is a problem when making model inferences because two highly correlated predictors' presence in a model can cause neither to be significant when either would be significant with the other removed. Therefore, multicollinearity reduces the quality of the interpretation of the significance of the predictors and therefore reduces the hypothesis generation benefit of an ecological study.

To remove highly correlated predictors, the absolute value of the correlation between each independent variable was calculated. The highest pairwise correlation that exceeds 70% was resolved by removing the variable with the largest mean absolute correlation to all other predictors. This process was repeated until no independent variables were correlated with each other more than 70% (RDocumentation, 2022a). For counties with reported ART cycles, the following predictors were removed due to high correlation with other variables: median household income, food environment index, infant mortality, poor physical health days, adult smoking, adult obesity, physical inactivity, and violent crime rate. The remaining predictors were correlated no more than 70% and are shown in Figure 3-9.

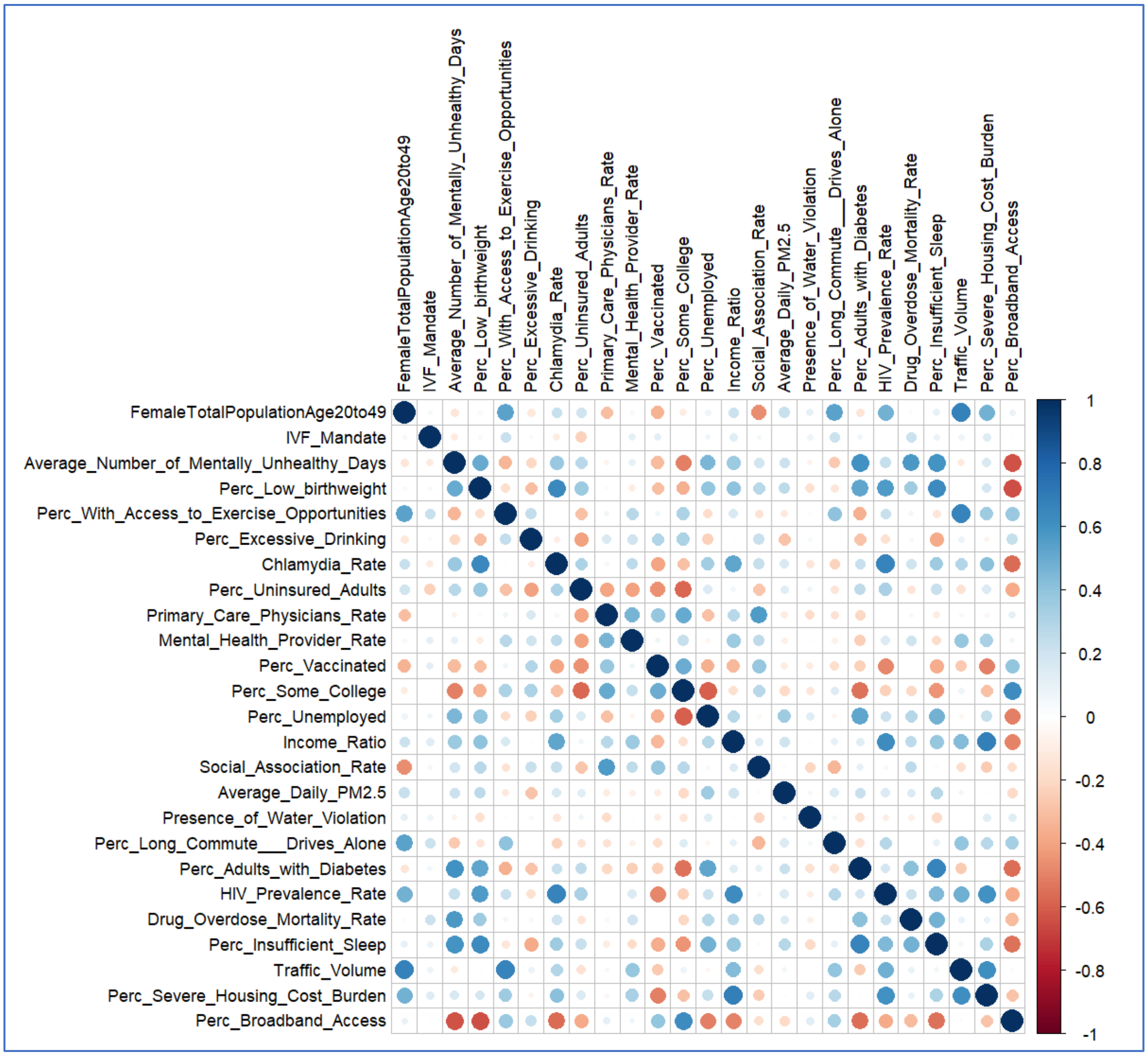


Figure 3-9. Variables correlated less than 70% in counties with reported ART cycles to the CDC in 2018.

Next, the negative binomial regression model was fit to the entire dataset of uncorrelated variables, and stepwise backward AIC feature selection was performed to reduce overfitting. Statistically significant variables ($p < 0.05$) found in order of importance were: Primary Care Physicians Rate, % Severe Housing Cost Burden, % Some College, HIV Prevalence Rate, IVF Mandate (binary), % Uninsured Adults, and Average Number of Mentally Unhealthy Days. Details of the final model found are captured in Table 3-7. Possible reasons why these variables were found to be significant are proposed in this chapter's discussion (Section 3.4.1).

Table 3-7. The negative binomial model of demand for ART. Normalized variables are ordered by importance defined as the absolute value of the coefficient divided by its standard error. The fit coefficient and significance values are also shown. Significant variables at the 0.05 signifcance level are bolded.

Normalized Variable	Importance	Coefficient	Significance	Signif. Codes
Primary Care Physicians Rate	**4.3**	**0.33**	**1.8E-05**	***
% Severe Housing Cost Burden	**3.6**	**-0.29**	**3.3E-04**	***
% Some College	**3.2**	**0.30**	**1.5E-03**	**
HIV Prevalence Rate	**2.9**	**0.23**	**3.2E-03**	**
IVF Mandate (binary)	**2.8**	**0.44**	**5.2E-03**	**
% Uninsured Adults	**2.6**	**-0.19**	**8.7E-03**	**
Avg. Number of Mentally Unhealthy Days	**2.2**	**-0.16**	**3.0E-02**	*
Social Association Rate	1.9	-0.15	6.2E-02	.
% Broadband Access	1.5	-0.13	1.4E-01	
(Intercept)		**-4.90**	**< 2e-16**	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

To determine if possible spatial drivers of demand were being captured in the predictive modeling, the model's residual errors were spatially analyzed. First, standardized Pearson residuals were calculated by dividing the raw residuals by their standard error. Standardized Pearson residuals are, in theory, normally distributed with a mean of 0 and variance of 1. These residuals were then mapped for the US counties with reported ART demand to the CDC in the contiguous 48 states, using the centers of population of each county as the point representation of each county, and a semivariogram plot[2] was created for up to half of the maximum distance between counties (1,499 mi). This semivariogram plot is shown in Figure 3-10. There is a spike in semivariance at 173 miles, but generally the semivariogram is flat indicating the negative binomial model appears to capture the spatial processes of demand, and there is limited residual spatial autocorrelation (Waller & Gotway, 2004). If spatial processes driving demand are

[2] A semivariogram plot visualizes spatial autocorrelation. The semivariogram function is defined as: $\hat{\gamma}(h) = \frac{1}{2|N(h)|}\sum_{N(h)}\left(y(s_i) - y(s_j)\right)^2$, where $N(h)$ contains the pairs of locations (s_i, s_j) that are a distance h apart. (Waller & Gotway, 2004)

not accounted for and there is spatial autocorrelation, the model is misspecified. Additionally, residual spatial autocorrelation indicates that some spatially related variable is missing from the model, and better predictions may be possible with the inclusion of the unknown missing variables.

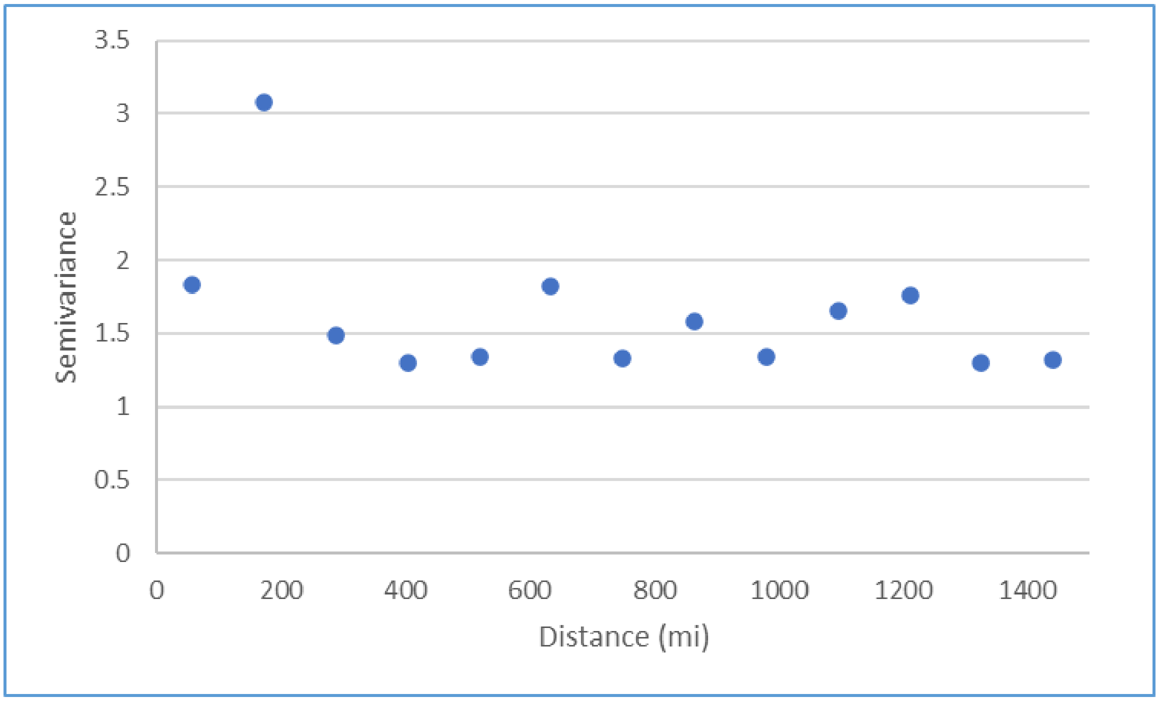


Figure 3-10. ART demand semivariogram for counties in the 48 contiguous US states.

Next, demand was predicted for the entire country with one change to the dataset of all US counties used for prediction. Instead of using the state IVF insurance mandates implemented before 2018 that were used when fitting the demand model, the state IVF insurance mandates passed by April of 2021 were used. This allowed for recently passed IVF mandates to be used in predictions rather than out-of-date mandates implemented before 2018 that were appropriately used to model ART demand using 2018 data. The list of states with IVF insurance mandates laws passed by April of 2021 were: AR, CT, CO, DE, HI, IL, MA, MD, NH, NJ, NY, RI, and UT. States with new mandates include: CO, DE, NH, NY, and UT (RESOLVE, 2021c). Demand estimates were mapped using proportionally sized circles for each county with a predicted demand over 250 cycles. Demand predictions are shown in Figure 3-11.

Demand rate (cycles per capita) was also calculated using the predicted demand divided by the female reproductive-age population for each county. This demand rate was then mapped nationwide with the lowest demand rate counties colored green and counties with a demand rate over 6% of the population colored in red. This mainly allowed for the modeled effect of state IVF mandates to be visualized nationally, with IVF mandate states having predicted higher rates of demand. These results are shown in Figure 3-12.

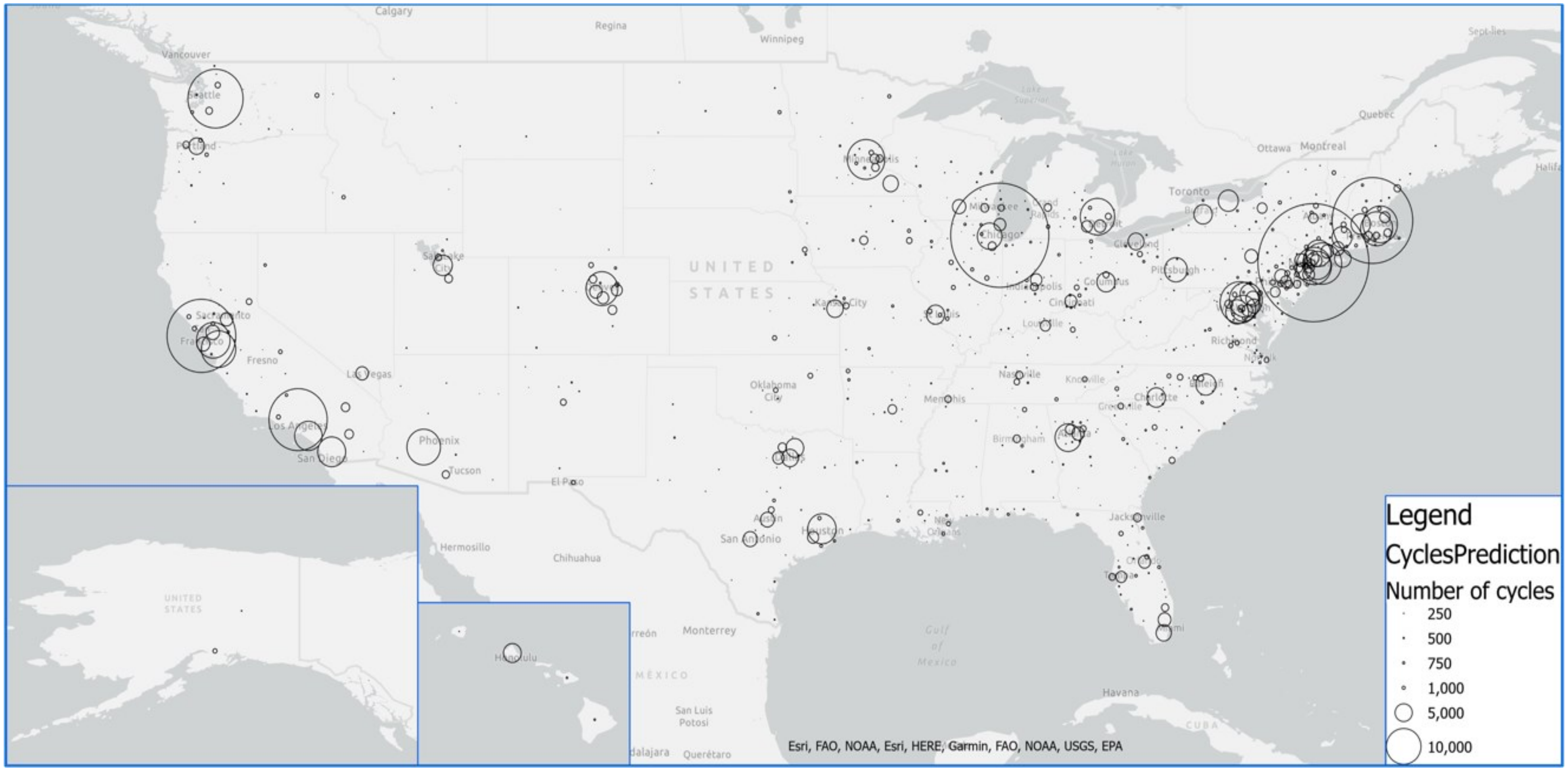


Figure 3-11. Nationwide prediction of ART demand for every county in the US. Predicted number of ART cycles are shown with proportionally sized circles for counties with a predicted number of ART cycles greater than or equal to 250.

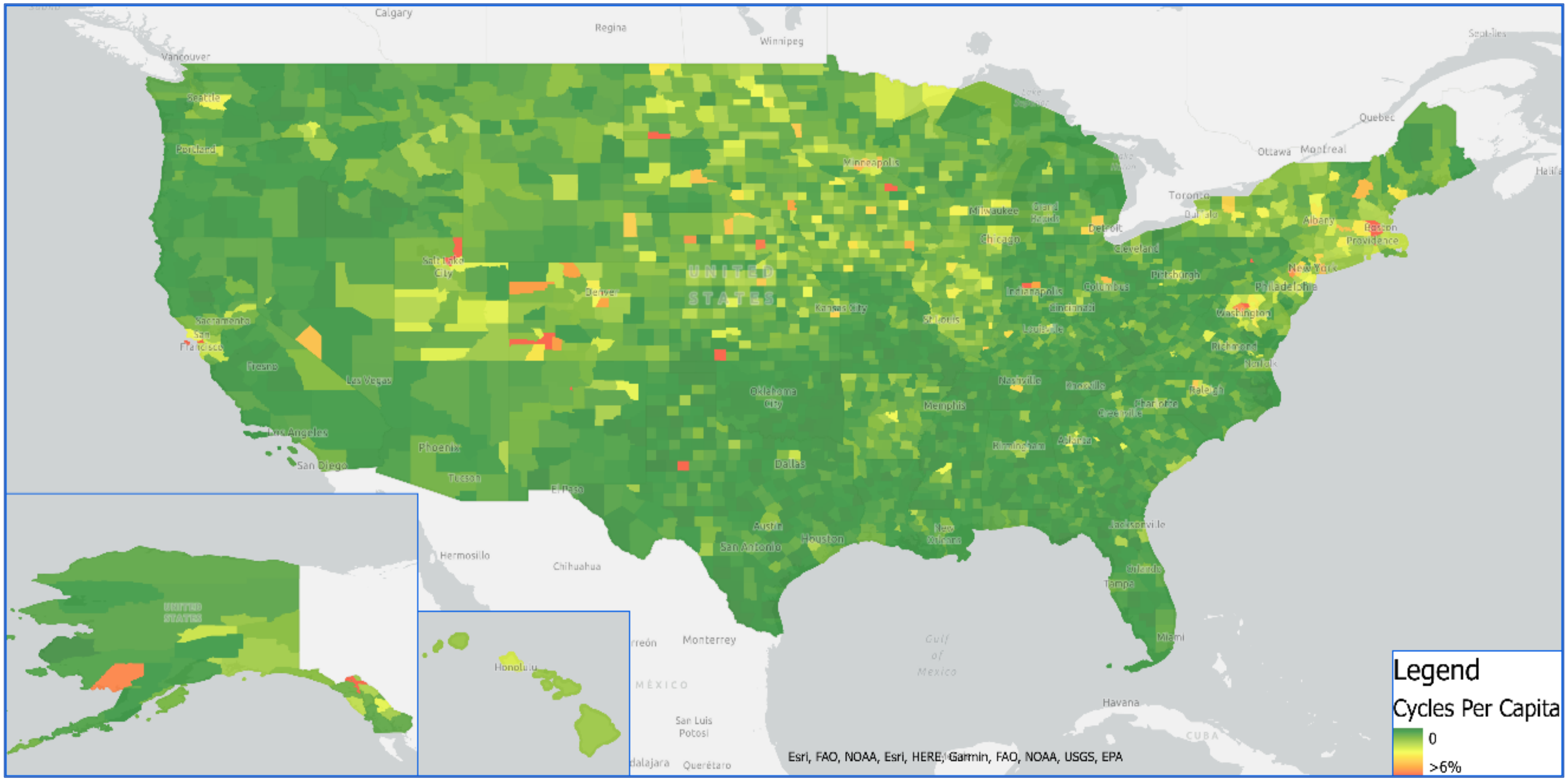


Figure 3-12. The demand rate (cycles per capita) for ART for each county in the US. Predicted cycles per capita was calculated by dividing the predicted number of cycles by the female reproductive-age (20-49) population in each county. The lowest demand rate is shown in green, and demand rates greater than 6% are shown in red with rates in-between shown in yellow.

The modeled effect of IVF mandates was estimated by totaling the predicted demand for all counties under different real or hypothetical IVF mandate scenarios. This approach found the total predicted demand in all US counties with IVF mandates implemented prior to 2018 to be 950k cycles, with IVF mandates passed by April 2021 to be 1.0M cycles, without any IVF mandates to be 870k cycles, and with IVF mandates in every state to be 1.4M cycles. Compared with the actual number of ART cycles in 2018 (306k cycles), the predicted demand numbers are likely overestimates of actual demand because the current number of ART cycles from counties with ART clinics includes patients from other counties without ART clinics. The predictions also appear to be overestimates when compared with the hypothetical total number of ART cycles of about 500k if the US was meeting the international standard of 1,500 ART cycles per million population per year (Collins, 2002). Though the predicted number of ART cycles appear to be overestimates, the ART demand model still captures the relative trends between counties, allowing the estimates to be used for optimizing coverage of ART demand in Chapter 4.

Additionally, the percentage change of predicted ART demand from 2018's prediction under different state IVF mandate scenarios is useful in capturing the aggregate modeled effect of IVF mandates. The IVF mandates passed in the five new states between 2018 and 2021 result in a predicted increase in ART demand by 6% nationally compared with 2018. Conversely, if IVF mandates were removed from every state, the predicted demand for ART would go down by 8% nationally compared with 2018, and if IVF mandate were passed in every US state, national demand for ART is predicted to go up by 56% compared with 2018.

3.3.3 Prediction of ART clinic presence for every county in the US

Lastly, the same clinic presence modeling approach used for the training dataset was replicated for the full dataset. As with the training dataset, prior to model fitting, predictors that were highly correlated (>70%) were removed. As discussed in Section 3.3.2, this was done with the intent of reducing multicollinearity, which reduces the quality of the interpretation of the significance of the predictors and therefore reduces the hypothesis generation benefit of an ecological study. For all counties, the following predictors were removed due to high correlation with other variables: Poor physical health days, median household income, and poor mental health days. The remaining predictors, correlated no more than 70%, are shown in Figure 3-13.

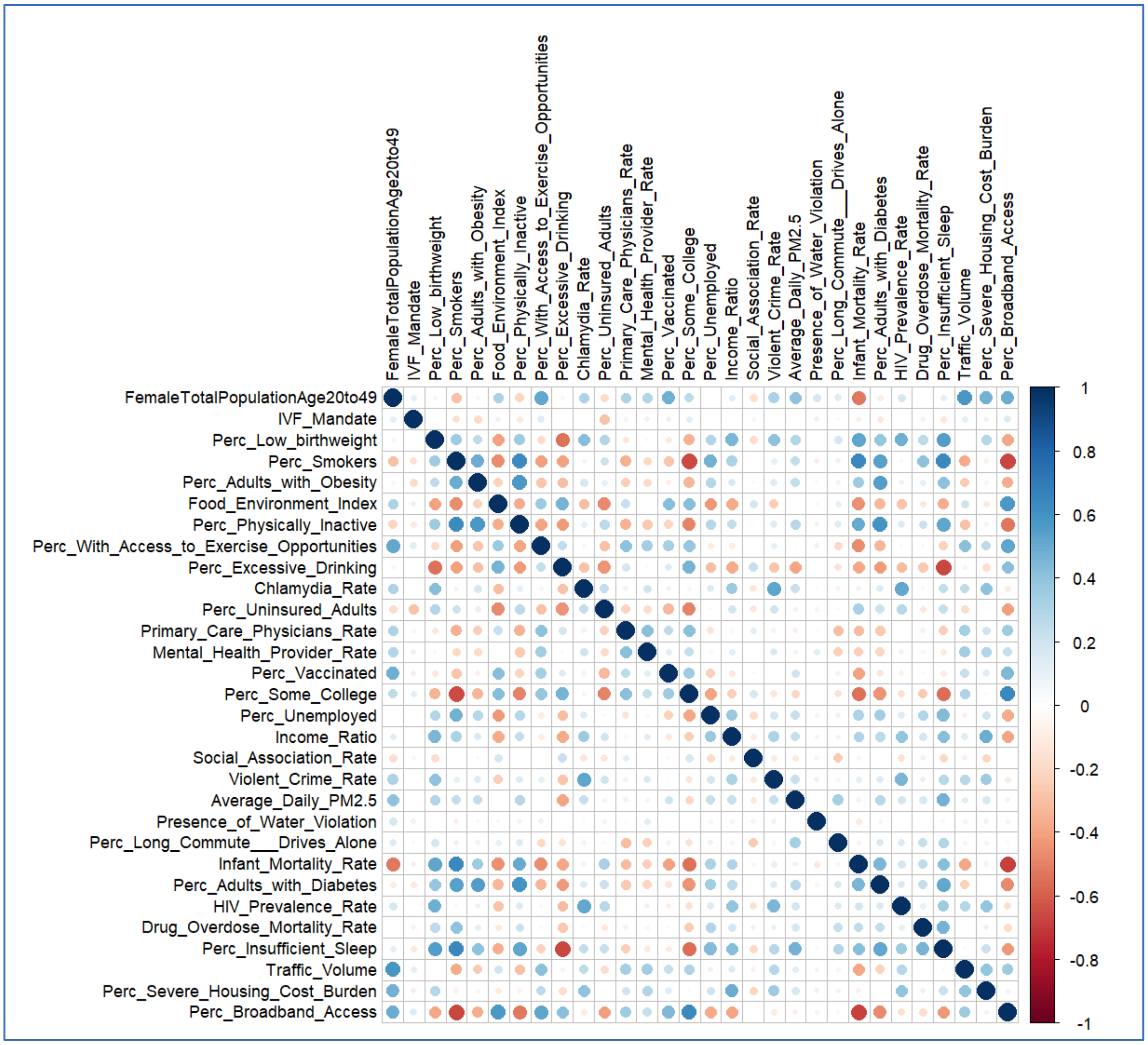


Figure 3-13. Socioeconomic and environmental health factors correlated less than 70% in all counties.

For both main and satellite clinic presence, the full dataset was used to fit a logistic regression model followed by stepwise AIC backwards feature selection using AIC. This produced two final models for each clinic type with several variables with significance values under 0.05. For main clinics, statistically significant variables in order of importance included: Female Total Population Age 20 to 49, Primary Care Physicians Rate, % Long Commute and Drives Alone, % Low birthweight, % Smokers, % Excessive Drinking, Drug Overdose Mortality Rate, HIV Prevalence Rate, and % Broadband Access. For satellite clinics, statistically significant variables in order of importance included: Female Total Population Age 20 to 49, % Physically Inactive, % Low birthweight, % Long Commute & Drives Alone, HIV Prevalence Rate, Food Environment Index, % Severe Housing Cost Burden, % Smokers, Violent Crime Rate, % Insufficient Sleep, and Primary Care Physicians Rate. These findings are shown in Table 3-8 and Table 3-9, and possible reasons why these variables were found to be significant are proposed in this chapter's discussion (Section 3.4.2).

Table 3-8. Logistic regression on main clinic presence in a county. Selected normalized variables shown, sorted by variable importance, defined as the absolute value of the coefficient divided by its standard error.

Normalized Variable	Importance	Coefficient	Significance	Signif. Codes
FemaleTotalPopulationAge20to49	**11.0**	**4.5**	**< 2e-16**	***
Primary Care Physicians Rate	**7.0**	**1.1**	**2.6E-12**	***
% Long Commute and Drives Alone	**4.5**	**-0.9**	**6.4E-06**	***
% Low birthweight	**3.6**	**1.0**	**3.7E-04**	***
% Smokers	**2.6**	**-0.9**	**8.1E-03**	**
% Excessive Drinking	**2.4**	**0.4**	**1.5E-02**	*
Drug Overdose Mortality Rate	**2.1**	**0.3**	**3.3E-02**	*
HIV Prevalence Rate	**2.1**	**-0.4**	**3.8E-02**	*
% Broadband Access	**2.1**	**0.7**	**4.0E-02**	*
IVF_Mandate1	1.9	-0.8	5.7E-02	.
% Physically Inactive	1.8	0.5	6.4E-02	.
(Intercept)		**-9.9**	**< 2e-16**	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 3-9. Logistic regression on satellite clinic presence in a county. Selected normalized variables shown, sorted by variable importance, defined as the absolute value of the coefficient divided by its standard error.

Normalized Variable	Importance	Coefficient	Significance	Signif. Codes
Female Total Population Age 20-49	**12.4**	**3.0**	**< 2e-16**	***
% Physically Inactive	**3.7**	**0.7**	**2.6E-04**	***
% Low birthweight	**3.6**	**0.7**	**3.3E-04**	***
% Long Commute & Drives Alone	**3.2**	**-0.5**	**1.3E-03**	**
HIV Prevalence Rate	**3.0**	**-0.4**	**3.1E-03**	**
Food Environment Index	**2.8**	**0.7**	**5.9E-03**	**
% Severe Housing Cost Burden	**2.7**	**0.4**	**7.8E-03**	**
% Smokers	**2.6**	**-0.6**	**1.1E-02**	*
Violent Crime Rate	**2.2**	**-0.3**	**2.7E-02**	*
% Insufficient Sleep	**2.1**	**0.4**	**3.7E-02**	*
Primary Care Physicians Rate	**2.0**	**0.2**	**4.1E-02**	*
% Vaccinated	1.8	-0.3	7.4E-02	.
% Some College	1.6	0.3	1.1E-01	
% Excessive Drinking	1.6	0.2	1.2E-01	
Drug Overdose Mortality Rate	1.5	0.1	1.3E-01	
Social Association Rate	1.5	0.3	1.3E-01	
% Broadband Access	1.5	0.4	1.4E-01	
(Intercept)		**-5.5**	**< 2e-16**	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Next, the clinic presence models were checked for spatial autocorrelation. This was done by first Pearson standardizing their residual errors. Standardized Pearson residuals are normally distributed with a mean of 0 and variance of 1. These residuals were then mapped for all US counties in the contiguous 48 states, and a semivariogram plot was created for up to half of the maximum distance between counties (1,606 mi). The semivariogram plot for main clinics is shown in Figure 3-14 and satellite clinics is shown in Figure 3-15. The semivariogram of residual error of the main clinic presence model is mostly flat indicating the model adequately captured spatial effects, but the semivariogram of residual error of the satellite clinic presence model is slightly curved upwards indicating a minor spatial process has not been accounted for in the satellite clinic model.

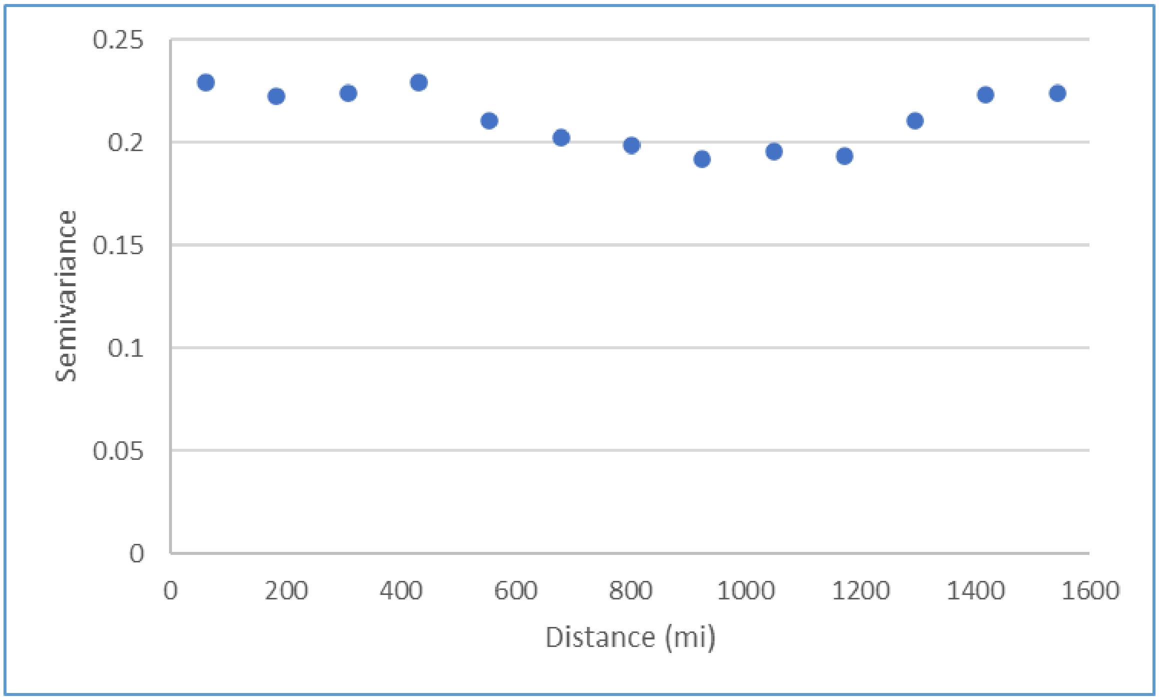


Figure 3-14. Semivariogram of Pearson standardized residuals of logistic regression model of main clinic presence.

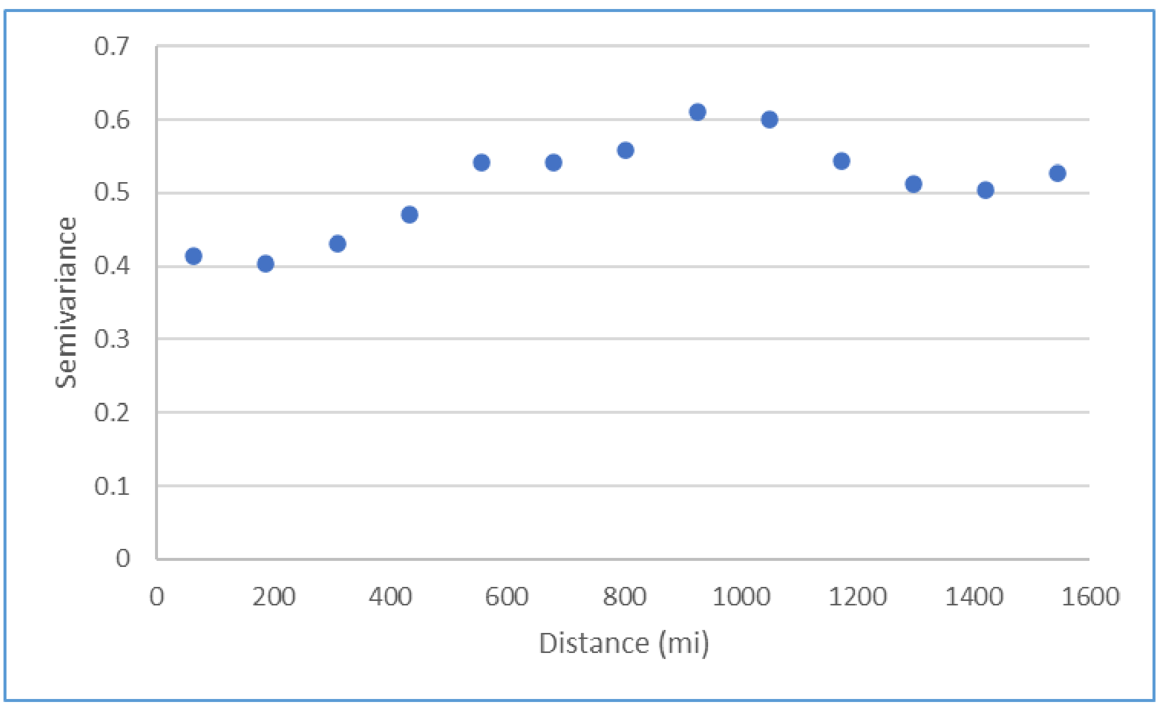


Figure 3-15. Semivariogram of Pearson standardized residuals of logistic regression model of satellite clinic presence.

3.3.4 Tests of hypotheses

In Section 3.1.3, above, two hypotheses were offered. The hypotheses and the results of the analysis are summarized as follows:

H8: ART demand will be driven by socioeconomic trends on education and income and less on environmental factors: *Confirmed* (see Section 3.3.4.1, below)

H9: The presence of ART clinics will be best modeled using the rates of primary care physicians: *Confirmed* (see Section 3.3.4.2, below)

3.3.4.1 H8 - Drivers of county-level ART demand

I hypothesized that ART demand is driven by socioeconomic factors of education and disposable income and less on environmental factors because economic access to ART continues to be a primary barrier to obtaining care whether or not environmental factors modify potential demand. This hypothesis was supported with socioeconomic factors found to be the primary drivers of demand for ART. The following predictors related to the environment were all found to be non-significant: Access to Exercise Opportunities, Excessive Drinking, Sexually Transmitted Infections, Air Pollution-Particulate Matter, Drinking Water Violations, Long Commute and Drives Alone, Insufficient Sleep, and Traffic Volume. The significant variables in the final model included: Primary Care Physician (PCP) Rate, % Severe Housing Cost Burden, % Some College, HIV Prevalence Rate, IVF Mandate, % Uninsured Adults, and Average Number of Unhealthy Days. The only predictors strongly related to the environment that were found to be significant were HIV prevalence Rate and Average Number of Unhealthy Days. Note that female reproductive-age population was included inherently as an exposure offset parameter rather than a conventional covariate. Modeling population as an exposure offset is the standard practice when modeling rates using Poisson and negative binomial regression (Waller & Gotway, 2004).

3.3.4.2 H9 - Drivers of ART clinic presence in counties

I hypothesized that the presence of ART clinics will be best modeled using the rates of primary care physicians (PCPs) since the economics driving higher rates of PCPs likely also contribute to higher rates of ART clinics. This hypothesis was supported by the finding that PCP rate is a significant predictor of both main and satellite clinic presence. PCP rate was found to be the second-most important predictor for main clinics, and the least important predictor for satellite clinics. Though PCPs do play a critical role in making the initial diagnosis of infertility and referring patients to ART clinics, this finding of correlation should not be misunderstood to imply causation. Both PCPs and ART clinics are likely influenced by the same set of geographic push/pull factors, resulting in PCP rate being a valuable predictor in modeling both ART demand and ART clinic presence.

3.3.5 Case study in identifying counties for new ART clinics

Finally, the logistic regression models fit to the entire dataset were used to predict new main and satellite clinics for every county in the US. This was achieved by predicting the presence of a clinic in each county. Though the response variable to be predicted by logistic regression is binary (either a 1 or a 0), predictions from logistic regression are real numbers between 0 and 1 and represent the likelihood of clinic presence in a county. This predicted presence category was true if the prediction was 0.5 or greater and false if less than 0.5. The predicted category was then compared to the actual presence of a clinic in each county. Locations where the logistic regression models predicted the presence of a main or satellite clinic when a clinic *of the type predicted* does not exist in reality were then mapped and are shown in Figure 3-16.

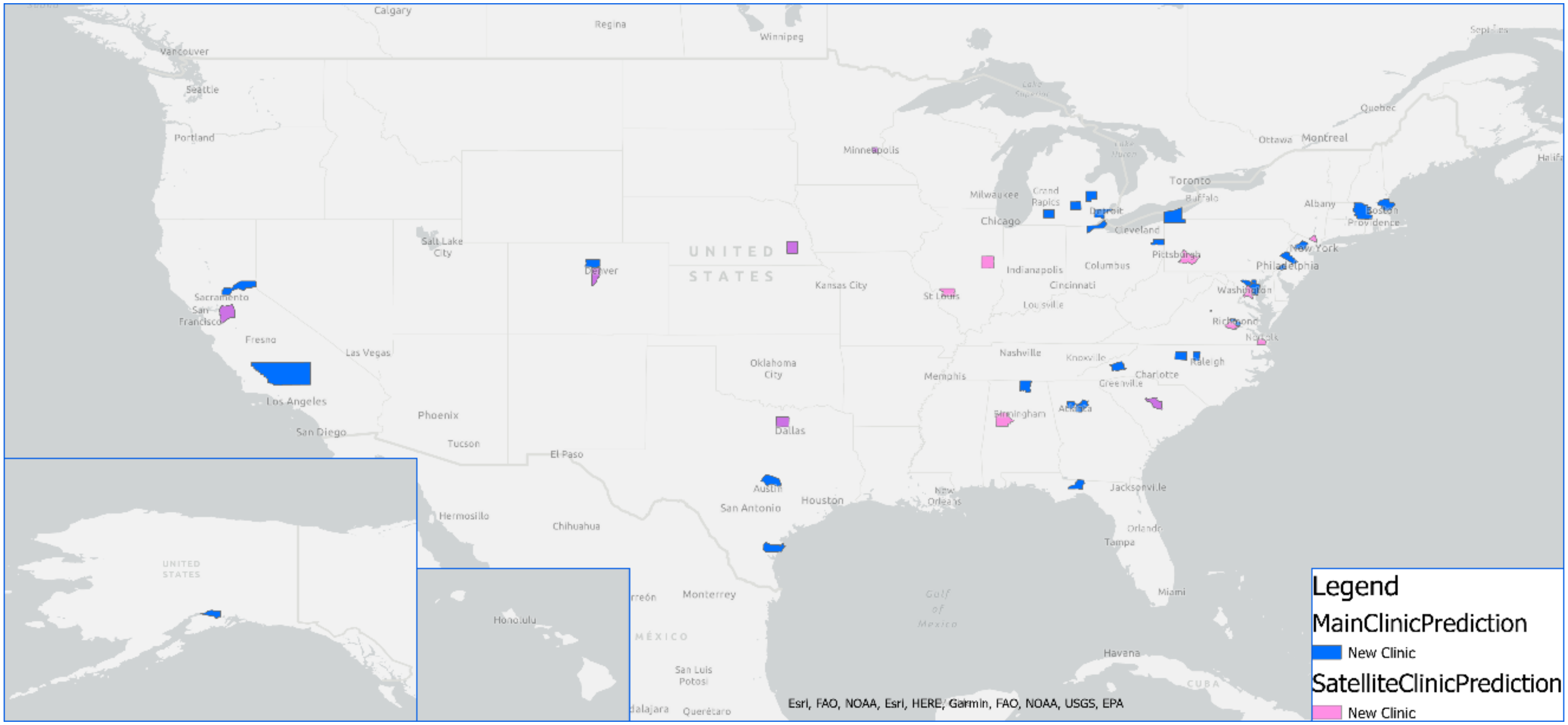


Figure 3-16. The presence of statistically predicted main or satellite clinic in counties where a clinic of the type predicted is not present. These "new" clinic counties are shaded in blue when a main clinc was predicted, pink when a satellite clinic was predicted, and purple when both clinic types were predicted.

Logistic regression was then used to identify locations for clinics where neither clinic type was present. Figure 3-16 took a similar approach, but a county with a predicted satellite clinic that didn't contain a satellite clinic was identified even if that county did contain a main clinic. Similarly, a county with a predicted main clinic that didn't contain a main clinic was identified even if that county did contain a satellite clinic. Though such information may be useful for identifying areas with the potential for a main clinic that only have a satellite clinic or areas where an additional satellite clinic could be supported even where a main clinic operates, it doesn't highlight areas that lack a clinic of either type that were statistically predicted to have one. Therefore, the locations without an existing clinic *of either type* that were predicted to have either a main or satellite clinic were mapped and are shown in Figure 3-17.

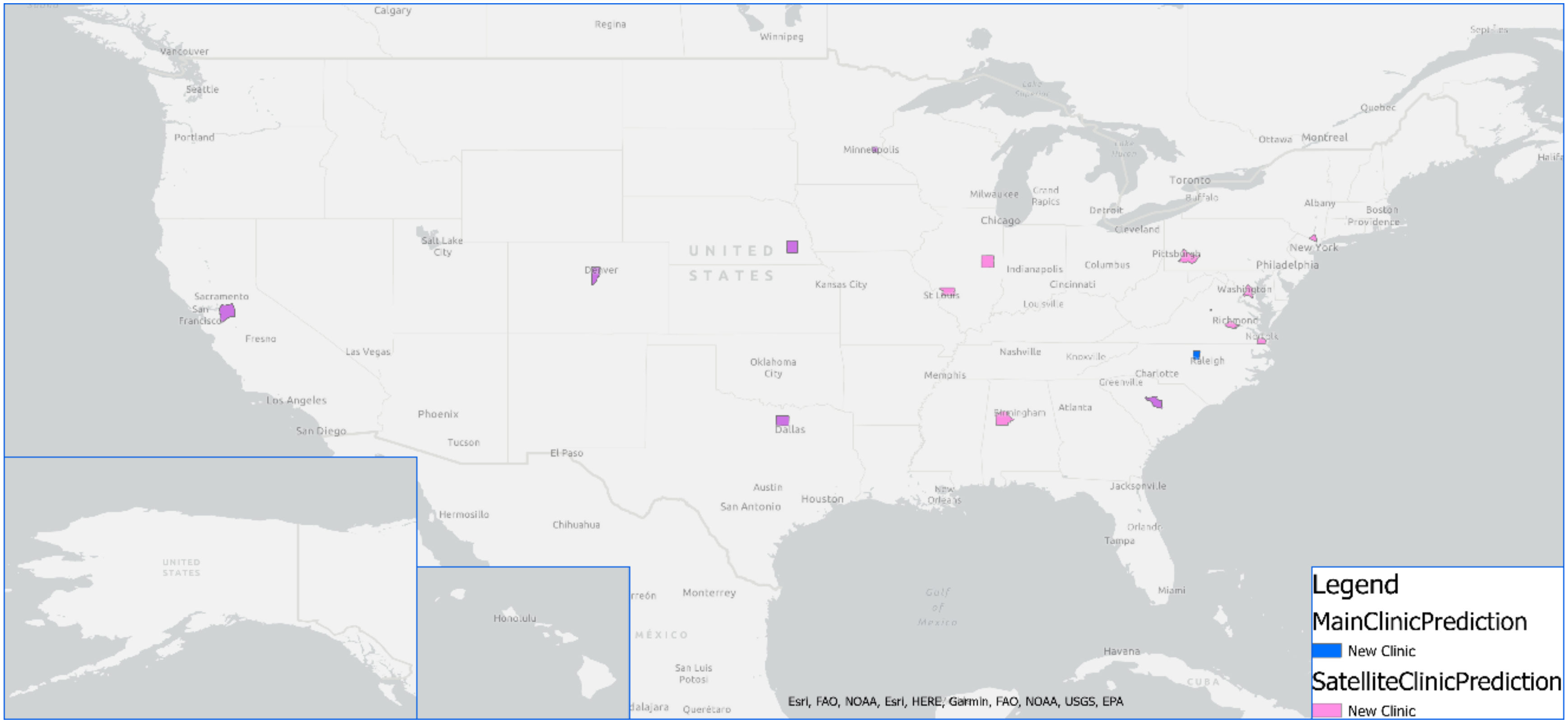


Figure 3-17. Counties that do not have a fertility clinic that were predicted to have a main or satellite clinic. These "new" clinic counties are shaded in blue for main clinics, pink for satellite clinics, and purple for both.

Predicted new locations for fertility clinics where no clinic of either type currently operates were found primarily near large cities (San Francisco, Denver, Dallas, Minneapolis, St. Louis, Pittsburgh, Washington D.C., Richmond, and New York) in nearby counties without clinics. However, new clinics also were found outside of areas with large cities in “college towns” with large state-run universities: Orange Co, NC (UNC Chapel Hill), Tuscaloosa, AL (University of Alabama), Champaign, IL (University of Illinois), Lincoln, NE (University of Nebraska), Richland, SC (University of South Carolina), and Charlottesville, VA (University of Virginia). After manually investigating these locations, UNC Chapel Hill, University of Virginia, and University of South Carolina all have nearby main or satellite ART clinics in adjacent counties. The remaining three college towns without nearby ART clinics (University of Alabama, University of Illinois, and University of Nebraska) appear to be overlooked areas of opportunity to reach underserved populations that the industry traditionally would serve but is not currently reaching. Possible reasons why those locations do not presently have clinics are proposed in this chapter’s discussion (Sections 3.4.3 and 3.4.4) below.

3.4 Discussion

In this section, insights obtained from the above results found are discussed. First, the statistically significant variables found in modeling ART demand are examined in Section 3.4.1, and the significant variables found in modeling clinic presence in counties are explored in Section 3.4.2. Next, the results from the case study identifying promising counties to operate new ART clinics are discussed to try to understand why ART clinics do not operate in less affluent urban counties (Section 3.4.3) and “college town” counties with large state-run universities (Section 3.4.4). Policy implications are covered in Section 3.4.5 based on the findings that PCP rate is impactful on ART demand and clinic presence. Finally, the limitations of this study are covered in Section 3.4.6.

3.4.1 Significance of findings modeling ART demand

PCP rate, the most important ART demand predictor, was found to be positively associated with demand. I now hypothesize some direct and indirect mechanisms for this finding. Indirectly, geographic pull-factors that attract PCPs to an area are likely similar to the geographic pull-factors that draw reproductive endocrinologists and fertility clinics. Additionally, a higher PCP rate likely means more frequent visits to PCPs by the reproductive-age population of a county and earlier diagnoses of infertility by the PCP and referral for treatment to fertility clinics. Infertility among physicians themselves is also a potential direct link between higher PCP rates and higher demand for ART. Nearly 1 in 4 (24.1%) female physicians will suffer from infertility (Stentz et al., 2016), which is about double the rate of infertility in the general population (CDC, 2021). This higher rate of infertility is likely due to biological infertility caused by female physicians delaying childbearing until after their many years of medical training. This likely results in increased demand for ART from the physicians themselves and increased demand for ART in areas with higher rates of physicians.

Percent Severe Housing Cost Burden being negatively associated with demand for ART and Percent Some College being positively associated with demand for ART is in line with findings in the literature. Both wealth and high levels of education were found to be associated with higher rates of IVF utilization even in a state with an IVF insurance mandate (Jain & Hornstein, 2005).

HIV Prevalence Rate was found to be positively associated with ART Demand. This factor could be directly associated with increased demand, or a confounder could be playing a role. Because HIV has been associated with a reduction in female fertility (Chirgwin et al., 1996), and the use of contraception among couples with one HIV positive partner would considerably reduce fecundity, ART may be one of the safest reproductive options for couples with one partner who is HIV positive (Ethics Committee of the American Society for Reproductive Medicine, 2015). This benefit noted, only 3% of ART practices in the United States have the facilities to allow for treating HIV positive couples (American Society for Reproductive Medicine, 2014b). This is due to concern over risk of transmission of HIV and higher cost to provide separate laboratory and equipment to reduce the risk of transmission. Therefore, HIV prevalence rate is likely not directly related to increased demand for ART, and a confounder is likely playing a role.

The finding that State IVF Insurance Mandates are positively associated with more demand for ART makes sense economically considering the high cost of ART. State IVF mandates have been shown in various studies to be associated with increased utilization for ART (Hamilton & McManus, 2012; Hammoud et al., 2009; Henne & Bundorf, 2008), so this finding was expected and supports previous findings in the literature. In Section 3.3.2, the effect on national demand for ART was estimated to be 8% higher as a result of IVF mandates implemented before 2018, and this finding also supports the finding in Chapter 2 of a statistically significant difference in mean satellite clinics per million reproductive-age women between states with and without IVF insurance mandates.

The Percentage of Uninsured Adults in a county was found to be negatively associated with ART demand. There are several reasons why this might be expected. Though many insurance plans do not cover IVF treatments, most cover at least the testing required as part of an infertility diagnosis (Advanced Fertility Center of Chicago, 2021). Without any insurance coverage, diagnoses of infertility would be less likely to occur. Additionally, some insurance plans do cover IVF treatments, especially in the states with IVF insurance mandates. Insurance plans provided by employers who self-insure in these states are exempt from the mandate, but state mandates likely create an expectation in those states for employers even who self-insure to provide ART coverage. Outside of state IVF mandates, employees have been successful in petitioning their employers to add coverage for IVF (RESOLVE, 2021a). So, employers with private/self-insured insurance plans sometimes still do offer IVF coverage even if not required by law.

Finally, the predictor Average Number of Mentally Unhealthy Days was found to be negatively associated with ART demand. There have been several studies examining mental health, infertility, and ART. Infertility has been found to be associated with an increased risk of mental disorders (Klemetti et al., 2010), and infertility treatments are stressful, but psychological interventions have been found to be beneficial to infertility patients (Cousineau & Domar, 2007). This study's finding of a negative association between poor mental health and increased demand for ART could be due to better access and treatment of mental illness in counties that also have better access to ART treatment, or it could suggest a benefit of IVF treatments on average mental health in a county. Counties with fewer IVF cycles per capita also likely have a greater fraction of untreated people with infertility who are more likely to have more mentally unhealthy days.

3.4.2 Significance of findings from modeling ART clinic presence in counties

For both clinic types, Female Population Age 20 to 49 was the most important variable. In the case of main clinics, female reproductive-age population was two times as important as the second most important variable (PCP rate). In the case of satellite clinics, female reproductive-age population was about three times as important as the second most important variable (percent physically inactive). These results were similar to the relative variable importance found in Chapter 2 (Section 2.3.2.6) when modeling clinic presence within CBSAs.

Differing from Chapter 2's CBSA models, median income was not included in the county-level models because it was highly correlated with many of the county health rankings variables and was removed prior to model fitting. However, the finding of female reproductive-age population being the most important variable in predicting clinic presence in both counties and CBSAs is revealing about trends in the industry. The consistent finding that population drives placement and continued operation of clinics indicates that most areas with sufficient number of reproductive-age women can support the operation of either a main or satellite clinic.

PCP rate was also found to be a significant predictor positively associated with the presence of both main and satellite clinics. This finding was consistent with the modeling of demand for ART that found PCP rate as the most important predictor of demand (shown in Table 3-7). It makes intuitive sense that the links between PCP rate and demand would also apply between PCP rate and the presence of fertility clinics.

HIV Prevalence Rate was found to be negatively associated with the presence of both main and satellite clinics. The negative association with fertility clinics and HIV rates is consistent with the lack of ART availability for patients with HIV with only 3% of SART practices offering ART treatment for patients with HIV (American Society for Reproductive Medicine, 2014b). This finding conflicts with the positive association found between ART demand and HIV Prevalence Rate, again suggesting the presence of a confounder.

The predictor Percent with Long Commute and Drives Alone was found to be a significant predictor negatively associated with the presence of both main and satellite ART clinics. This measure is equal to the percentage of solo commuters whose drive is longer than 30 minutes. Though long commutes are associated with poorer health outcomes like obesity (Frank et al., 2004), which is associated with reduced fecundity (Louis & Platt, 2011), this finding is not likely related to health, but rather the trend to locate fertility clinics in urban areas. Counties where workers have long commutes are generally either rural or suburban, which are less frequently areas with ART clinics (Harris et al., 2017).

Percent Low Birthweight was found to be positively associated with the presence of both clinic types. This measure is defined as the percentage of live births with low birthweight (less than 2,500 grams/5.5 lbs.). This is an interesting finding since IVF is associated both with both multiple births and prematurity (Grady et al., 2012), so higher rates of low birthweight may be an effect of ART clinics. The rationale for this supposition is that it has been common in US IVF treatments to transfer more than one embryo with the intention of increasing the likelihood of a pregnancy at the risk of multiple pregnancy (Maheshwari et al., 2011). Guidance has been issued to reduce the incidence of multiple pregnancy and its associated increased health risks, but the current guidance still allows for more than one embryo to be transferred for all age groups unless transferring a later stage embryo known as a blastocyst (Practice Committee of the American Society for Reproductive & Technology, 2013), which has a higher likelihood of pregnancy. This guidance appears to have been effective at reducing the rate of triplets and higher-order multiple birth rates (Martin et al., 2017), but this study's findings linking low birthweight to the presence of fertility clinics could indicate that the practice of transferring more than one embryo was still more common as of 2018.

Percent Smokers was found to be negatively associated with the presence of both main and satellite ART clinics. Smoking has been found to consistently reduce fecundity (Augood et al., 1998; Louis & Platt, 2011) and smoking reduces the success rate of IVF (Lintsen et al., 2005). Because ART clinic success rates are tracked by the CDC, ART practices may be drawn to counties with lower smoking rates and repulsed by counties with higher smoking rates. Locating clinics in counties with higher smoking rates could result in lower success rates for practices that could hurt the reputation of ART practices and disincentivize the locating of clinics in these counties.

3.4.3 Neglected less affluent counties near large cities

The findings in literature suggest that highly educated urban areas have higher utilization of ART (Hammoud et al., 2009), and urban areas have better geographic access to ART (Harris et al., 2017), but this study still identified promising counties in/near large cities throughout the US that do not currently have either a main or satellite ART clinic. These neglected urban counties are generally the less affluent counties in an urban area, with neighboring more affluent counties having ART clinics. State and local subsidies could potentially be employed to encourage ART practices to do a better job of locating clinics in these neglected counties to increase convenient access to ART to less affluent counties. For example, one of the less affluent counties in the Washington D.C. area, Prince George (PG) County in Maryland, does not have an ART clinic, but neighboring more affluent counties all have ART clinics. Many residents of PG County could potentially be more likely to seek care if geographic access were improved, especially considering Maryland has an IVF insurance mandate.

3.4.4 Overlooked college towns

The identification of college towns by the ART clinic presence modeling suggests that the industry has overlooked promising areas to locate ART clinics. The more rural location of these universities has likely resulted in them being overlooked by the industry, but there are compelling reasons suggesting that locating an ART clinic in a college town would be advantageous. Universities typically pay competitive wages to members of their faculty who are required to obtain doctorates. These faculty also likely delay childbearing to later in life until after completing their doctorates and sometimes until achieving tenure. Additionally, egg or sperm donation programs could be highly successful in these areas with large numbers of highly fecund women and men in their 20s with college expenses to pay. Therefore, ART practices in the vicinity of large state-run universities should investigate if the economics would support opening a new clinic in these areas.

3.4.5 Policy implications considering the role of primary care physicians on ART demand and clinic presence

This study's finding that PCP rate is an important predictor of both ART demand, and the presence of both ART clinic types is valuable because it could help the industry identify previously overlooked candidate sites for new ART clinics. PCP locations have historically been maldistributed with PCP practices clustered in urban areas (Kindig & Movassaghi, 1989; M. M. Rosenthal & Frederick, 1984). To correct this maldistribution, the federal government created the National Health Service Corps (NHSC), which pays off medical school student loans in exchange for new doctors practicing in underserved areas. However, the NHSC benefit does not apply to specialty clinics even if they are located in a designated Health Professional Shortage Area (HPSA) unless there was an ART clinic operating as part of a Federally Qualified Health Center or Indian Health Service Facility (NHSC, 2021). The only ART clinics that are run by the federal government are the six ART clinics run out of military treatment facilities (MTFs), and these sites do not qualify for loan repayment under the NHSC program. Therefore, the only way newly graduated REI physicians could qualify for such a subsidy would be if the federal government first establishes an ART clinic in a Federally Qualified Health Center or Indian Health Service Facility.

The ART industry following the patterns of the broader medical industry is unsurprising considering it is a subset, but that also means that it should receive similar scrutiny with respect to clinic placement. The findings from Chapter 2 (Section 2.4.2) suggest that the ART industry follows Hotelling's Law and generally co-locates clinic locations near each other. Instead of simply locating new ART clinics in proven markets near existing ART clinics, the findings of this essay suggest that ART practices could identify unserved areas with high PCP rate to identify new areas with the potential to support an ART clinic.

3.4.6 Limitations

The statistical methods selected to model ART demand and clinic presence largely captured the spatial processes of their response variables as evidenced by the relatively flat semivariograms. However, for satellite clinic presence modeling, the slight rising semivariogram curve with the increase in distance (shown in Figure 3-15) indicates a minor residual spatial dependence that could be investigated in future studies. Adding additional predictive variables to capture these spatial processes could be useful. For example, the number of ART clinics in adjacent counties or spatially intersecting CBSAs would likely be helpful in reducing the "near-sightedness" of the county-level statistical models that did not consider clinic presence or absence in neighboring counties that, in fact, can provide some service in a county under consideration. Alternatively, the spatial processes could be captured explicitly for the demand modeling using the spatial statistics method of universal kriging. This would require the use of linear regression instead of negative binomial regression, but the residuals of the linear regression model of demand appear to be relatively normal (shown in Appendix C), so the value added from universal kriging is likely low. However, the predictive performance of the linear regression model was not much poorer than the negative binomial model (shown in Figure 3-2 through Figure 3-4), so if universal kriging improves predictions even incrementally, the linear model could potentially generate superior predictions.

Estimates for current demand come from counties where clinics currently operate and report their number of ART cycles to the CDC (4% [n=120] of all counties), but predictions were made for every county in the US. If the CDC were to track infertility diagnoses similar to how cancer is tracked in the US, then ART demand could then be better estimated and served if infertility, the primary driver of ART demand, was thereby better tracked.

Underestimates of demand likely exist in counties with very few ART clinics and where intra-county access is limited by the time-cost of travel within the county. Another source of underestimation of demand is from the six ART practices operated by the US military. Patients of these practices need to hold TRICARE insurance and have permission to access the facilities where the clinics are located. The general population living nearby would likely utilize these clinics if given access. On the other hand, overestimates of demand are likely more common because of the willingness for couples to travel outside their county of residence to seek treatment in a different county. This could be due to lack of any clinic in their county of residence, or because some population residing near county borders might actually be closer to practices in other counties, or finally because of actual or perceived better-quality care at a clinic located elsewhere. To mitigate this study limitation, the CDC could consider adding to their collection of data from ART clinics on the distance traveled by each patient and provide summary statistics for each clinic on its service areas.

Finally, demand for ART from international patients from US clinics is another factor unaccounted for in the data reported to the CDC. Though international demand for ART in the US is likely to continue with American success rates being some of the highest in the world (Kushnir et al., 2017), closure of borders due to pandemics or for political reasons could negatively impact medical tourism for ART. The CDC should begin collecting and reporting patient nationality data in its ART surveillance system as an aid to the industry in understanding where international demand for ART in the US exists. Additionally, medical tourism for ART in the US inflates estimates for the percentage of demand met in the US by Americans. International ART patients likely constitute only a small fraction of the ART cycles performed in the US, but it would be valuable for the CDC to capture the *exact* number of ART cycles associated with those living in the US.

3.5 Summary

Demand for ART and presence of ART clinics were modeled using an array of machine learning algorithms. Predictive performance of linear regression, instance-based methods, decision tree algorithms and ensemble methods were compared using a training dataset. Negative binomial regression was selected to model ART demand, and logistic regression was selected to model main and satellite ART clinic presence. These models were selected based on strong predictive performance compared with other models, relatively low model complexity, and ease of interpretation.

Socioeconomic factors were found to be most important predictors of demand compared with environmental factors, supporting H8. State IVF insurance mandates were found to be significant and positively associated with higher ART demand. When modeling clinic presence, female reproductive-age population was 2 to 3 times more important than the next most important variable, and the primary care physician (PCP) rate was found to be a valuable predictor of ART clinic presence, supporting H9.

Statistical models were used to extrapolate demand estimates for every county in the US and to identify promising new counties to locate clinics. ART clinic presence logistic regression models identified underserved counties that were predicted to have either a main or a satellite clinic based on trends of clinic placement in counties home to large state-run universities and less affluent counties near San Francisco, Denver, Dallas, Minneapolis, St. Louis, Pittsburgh, Washington D.C., Richmond, and New York. This work addresses a gap in the literature of ecological studies of ART demand and clinic presence by creating the first model of clinic presence and the first model of demand that uses predictors in addition to female reproductive-age population.

Chapter 4 Expanding Geographic Access to Fertility Care: A Multiobjective Optimization Approach

4.1 Introduction

4.1.1 Purpose

This essay builds on the previous two chapters by applying optimization to also identify locations to site ART clinics. However, this essay takes an approach that is fundamentally different than the approach taken previously in the dissertation. Earlier chapters used the statistical method of logistic regression to predict the presence or absence of an ART clinic at the CBSA-level (Chapter 2) and at the county-level (Chapter 3). The approach taken in earlier chapters is not flawed *per se*, but it is limited by the predictions being made using only predictors specific to each geographic area without considering the influence of surrounding CBSAs or counties. This chapter addresses the previous chapters' limitations by utilizing *optimization* instead of *statistics* to place clinics, and a combined approach incorporating both methods may also have advantages (see Section 4.3.5). In the sections that follow in the introduction, more background is provided along with a literature review in Section 4.1.2, the contribution of this essay is summarized with various hypotheses proposed in Section 4.1.3, and the scope of the chapter is outlined in Section 4.1.4.

4.1.2 Background and literature review

To form recommendations for how to expand geographic access to ART, first an understanding of where people have access must be established. Progress towards the question of who has existing geographic access was made using two geographic constructs to approximate geographic access. CBSAs were used in Chapter 2, and counties were used in Chapter 3 to approximate geographic access. The underlying assumption of those two approaches is if a fertility clinic exists inside the boundaries of the county or CBSA, then every person in the county or CBSA has geographic access to the clinic. Conversely, if a fertility clinic exists in a different county or CBSA than a person's residence, then that person is assumed to not have geographic access to the clinic. The CBSA geographic access assumption likely overestimates true geographic access in CBSAs where intra-CBSA travel times are prohibitively long, and the county geographic access assumption likely underestimates true geographic access in the case of highly connected counties near each other.

An alternative approach found in the literature that does not suffer from these limitations in approximating access is to use driving times between population centers. For example, Nangia et al. (2010b) mapped 60-minute driving times from ART clinics and urology specialists overlayed on male population at the county level. It found that the median state percentage of male population within a 60-minute drive of either a fertility center or urology clinic to be 73%, and the best regional access in the Northeast and Southern California, with the worst access in the Midwest to Northwest. A separate study by Nangia et al. (2010a) explored both female and male geographic access to ART clinics using driving times vs. geographic areas, finding the best geographic access in the Northeast and the worst in Alaska, Montana, Wyoming, and West Virginia. More recently, Gharani et al. (2015) modeled spatial accessibility to ART using a gravity model with travel times between census tracts along road networks. Though these studies were valuable contributions towards characterizing current geographic access to ART, only rough inferences could be made on where to site new clinics to improve geographic access because they did not consider the placement of new clinics to increase access.

The problem of placing new healthcare facilities to increase geographic access can be addressed using optimization methods. The methodology for optimally locating facilities along a road network using linear programming was first introduced by ReVelle & Swain (1970) with the formulation of the "Central Facilities Location." After decades of development in location analysis, Revelle & Eiselt (2005) summarized progress made in the field and summarized various location analysis models. These models included the *median* (ReVelle & Swain, 1970), *center* (Hakimi, 1964), and *location set covering problems* (Toregas et al., 1970). The *median* or *p*-median problem finds the *p* facilities that minimize the average travel distance or time per person, the *center* or *p*-center problem finds the *p* facility locations that minimize the largest demand node to facility distance, and the *location set covering problem* or the maximal covering location problem (MCLP) (Church & ReVelle, 1974) finds the *p* facilities that cover all demand points within a "prespecified distance standard" (Revelle & Eiselt, 2005).

An illustration of the MCLP is shown in Figure 4-1, with candidate facility sites selected such that the most demand can be covered within a reasonable service distance or travel time S.

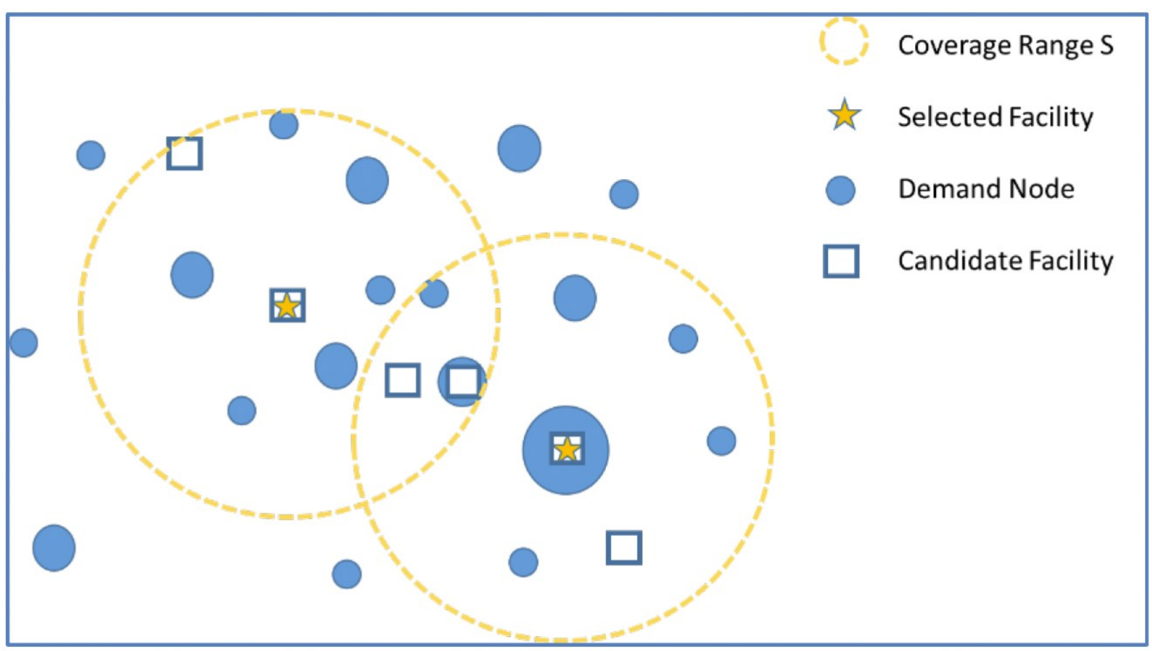


Figure 4-1. An illustration of locating facilities with the maximal covering location problem (MCLP), with demand nodes sized proportional to the amount of notional demand.

A few studies have characterized present geographic access to ART, but no studies have explored how adding new facilities would improve access. Though location analysis has been applied to optimally site facilities for an array of chronic diseases (Griffin et al., 2008), it has not been applied locate fertility clinics to treat the chronic disease of infertility. Additionally, optimal placement of healthcare facilities to improve racial equity in geographic access has been explored in the literature (Eaton et al., 1985), but it has never been explored for racial/ethnic geographic access to ART clinics. Because optimal facility location analysis has proven to yield valuable information concerning tradeoffs among objectives and possible Pareto improvements that enhance the system's performance in one or more objectives without substantially deteriorating performance in others, I believe that applying location analysis to optimize and expand fertility clinic access addresses a clear gap in the literature and is likely to provide valuable insights on where and why new facilities should be optimally sited.

4.1.3 Contribution

In this essay, I correct the above identified shortcoming of the literature by first optimizing coverage of new clinics with respect to population and demand via a multiobjective approach to examine tradeoffs between the two objectives of maximum population coverage and maximum demand coverage. I hypothesize (H10) new clinics will be placed when maximizing coverage of population or demand throughout the US with the exception of the Northeast, which was found in the literature to be overserved (Nangia et al., 2010a). Further, I hypothesize (H11) that as solutions along the Pareto frontier move from maximizing demand covered to maximizing population covered, solutions will move from IVF mandate states to non-IVF mandate states because state IVF mandates were found in Chapter 3 to be associated with higher ART demand. Next, to combat racial disparities in geographic access to ART, an analysis of two approaches to optimally siting new clinics is presented. Finally, I hypothesize (H12) that locations that optimize coverage of underserved racial and ethnic minorities will primarily be located to cover populations on Indian reservations because American Indian and Alaska Native alone was identified in Chapter 2 as the racial group that most lacks geographic access to ART.

4.1.4 Chapter scope

The scope of this essay (Chapter 4) is organized as follows. First, in the methods section (Section 4.2), I first present the optimization problem formulation for maximizing coverage of fertility clinics (Section 4.2.1). Then, I describe the search method for finding the multiobjective Pareto frontier in Section 4.2.2, and a population weighting method is outlined in Section 4.2.3 that accounts for racial disparities in geographic access to care. Findings are reported in the results section, including the locations of new clinics identified from the population and demand coverage objectives in Section 4.3.1 and 4.3.2, and the maximizing racial equity objectives in Section 4.3.3. The hypotheses posited above are tested in Section 4.3.4, and a case study is presented in the use of an ensemble of model recommendations to identify new clinic locations in Section 4.3.5. The discussion explores the significance of the essay's results with an examination of the novelty of the multiobjective approach taken in Section 4.4.1, and the agreement of different model recommendations found in the case study in Section 4.4.2. Finally, a summary is given focusing on the highlights of the essay.

4.2 Methods

This essay's methods are detailed in this section, starting with how the maximal covering location problem (MCLP) is applied to fertility clinic placement in Section 4.2.1. Next, the rationale for taking a multiobjective approach is presented along with the search method used to find the set of non-inferior solutions between the two objectives in Section 4.2.2. Finally, a potential approach to expand racial equity in geographic access to ART clinics is formulated in Section 4.2.3.

4.2.1 Maximal covering location problem (MCLP) applied to fertility clinic locations

The MCLP (Church & ReVelle, 1974) formulation was employed to derive solutions that provide optimal fertility clinic geographic access. The MCLP formulation is most appropriate for the problem of providing access to infertility treatments. A *center* formulation does not incorporate population, so it would simply site clinics in geographic centers between the most unserved counties to minimize the distance between clinic locations and counties. A *median* formulation would minimize the average distance/time traveled per person (ReVelle & Swain, 1970), so that would likely be more appropriate than the *center* formulation.

But to best expand ART treatment *access*, I believe it is more valuable to maximize the number of people within a reasonable travel time to a clinic versus minimizing the average travel time per person. Minimizing the average travel time per person would place clinics near population centers and be more *convenient* to more people, but fewer people would ultimately be served within a reasonable travel time. The results in Chapter 2 (Section 2.3.2.1) showed that 13 million women (20% of the reproductive-age population) do not live in an area with an ART clinic. With so many women without geographic access to care, I argue that the best way to expand geographic access is to prioritize *access* over *convenience*. Therefore, an adapted MCLP formulation is shown in Equation 4-1, with the parameter A_i representing either the county ART demand predicted in Chapter 3 or the county female reproductive-age population dependent on which objective (population covered vs. demand covered) is maximized.

Equation 4-1. The maximal covering location problem (MCLP) applied to fertility clinic siting.

$$Maximize \sum_{i=1}^{m} A_i Y_i$$

$$Subject\ to:$$

$$Y_i \le \sum_{j \epsilon N_i} X_j, \quad \forall\ counties\ i = 1, \dots, m$$

$$\sum_{j=1}^{n} X_j = p$$

$$X_j\ binary,\ \forall\ j = 1, \dots, n \text{ and } Y_i\ binary,\ \forall\ i = 1, \dots, m$$

$$A_i\ is\ the\ demand\ for\ ART\ or\ population\ in\ county\ i$$

$$N_i\ is\ the\ set\ of\ candidate\ sites\ j\ within\ driving\ time\ S\ of\ county\ i$$
$$= \{\, j \mid D_{ij} \in S \,\}, \text{ the covering set of } i$$

$$Y_i = 1\ if\ county\ i\ is\ covered\ by\ at\ least\ one\ facility,\ 0\ otherwise$$

$$X_j = 1\ if\ a\ facility\ is\ placed\ at\ candidate\ site\ j,\ 0\ otherwise$$

$$X_j = 1, \forall\ counties\ j\ with\ existing\ clincs$$

The best counties (as defined by the coverage objective) to place new ART clinics to expand geographic access were selected using the MCLP formulation. The MCLP is formulated to optimize coverage of populations at point locations, but counties are polygons not points, so point approximations of counties were utilized. The population weighted center of each county was the point location chosen to best represent the location of each county in the US. These points were calculated by the US Census Bureau using 2010 county boundaries and population estimates. The center of population is calculated as the point where a geographic area “would balance if weights of identical size were placed on it so that each weight represented the location of one person,” and where the geographic area was represented as a “weightless, rigid, and flat (no elevation effects) surface” (U.S. Census Bureau, 2010). The population weighted mean location is a good approximation for where most people in a county live because it minimizes the average error between where people live and the point approximation.

Travel times between county centers of population were calculated using the North American roads network included with ArcGIS's 2019 Business Analyst. Driving times between each county's population center and every other county's population center were calculated. In some cases, the population center of a county did not fall on the network of roads. For example, this occurred in oddly shaped county-equivalent areas on the Aleutian Islands in Alaska. In these rare cases, the closest point along the road network was searched for and used as the point representing the county's location for travel time calculations. Optimization models were formulated in R (v 4.1.2) using the R Optimization Infrastructure (v 0.8.0) through the OMPR package to model mixed integer linear programs (MILPs) (v 0.8.1), and the Gurobi solver (v 9.1.2) was used to solve the MILPs.

4.2.2 Balancing objectives: maximizing population covered vs. maximizing demand covered

If poor geographic access was the only barrier to care, the only approach needed would be to optimize coverage of female reproductive-age population. However, many barriers exist to accessing fertility treatments. In 2015, the ASRM's Access to Care Summit identified the following barriers: sociocultural, geographic, insurance, financial, economic, outreach to underserved populations and regions, and male infertility education (American Society for Reproductive Medicine, 2015b). Because environmental exposures have been shown to affect fecundity (Schettler et al., 2000), demand for ART, which can potentially overcome infertility due to environmental exposures, is likely also influenced by environmental factors. Additionally, the large time-cost of infertility treatments (Wu et al., 2013) is likely a barrier to care for infertile men and women who lack the paid time off necessary for treatment. The strongest rationale against using solely a maximization of population covered approach is the consideration of the long-term financial viability of clinics. If a clinic is placed to optimally cover an unserved population that does not have sufficient demand for ART to cover costs, then the clinic will ultimately close, and no one will be covered as a result of the large capital and human effort invested in opening a failed clinic. Therefore, demand for ART must also be considered when deciding where to locate new clinics to expand geographic access to care.

If the primary purpose of operating fertility clinics was to make the most profit, a good approach might be to optimize coverage of demand for ART. However, placing new clinics based solely on a maximization of demand covered requires a great deal of trust in the underlying model of demand. Location-specific demand for ART is not well understood primarily because infertility diagnoses are not tracked nationally. National surveys are conducted to gather rates of impaired fecundity among all women and infertility among married couples but only at the national level. The National Survey of Family Growth estimates these rates from a sample of roughly 10k women and men (National Center for Health Statistics, 2020). The national impaired fecundity rate could be applied coarsely to the reproductive-age population at the county level to estimate the population with the need for ART, but optimized clinic placement based solely on this metric this would yield identical recommendations for clinic placement as the population-only model.

Therefore, optimization of clinic locations should balance the two objectives of covering population and demand. Methods exist for eliciting weights from decision makers to allow for a single optimal solution to be found that optimizes the weighted sum of the multiple objectives. These methods require both elicitation from relevant people (managers, experts, or stakeholders) and that “weights to be proportional to the relative value of unit change” in their objective (Hobbs, 1980). The objective combining approach could be appropriate if there is a single organization who desires to select a single solution to target investments, but independent parties with different organizational objectives are likely to have different priorities among the objectives. Additionally, there could be non-quantifiable objectives known as "hidden objectives" (Schilling et al., 1982) that could play a role, reinforcing the rationale to defer the decision of a best solution to the various independent decision makers. Therefore, a multi-objective approach was taken that presents a range of solutions defining a tradeoff curve between both objectives, informing deliberations and negotiations among affected parties about the right balance between favoring population coverage and demand coverage.

This tradeoff curve between objectives is known as the "Pareto frontier" or the "non-inferior set" (Cohon, 2004), which contains multiple non-inferior/Pareto-optimal/non-dominated solutions, as opposed to the optimal solutions yielded by single-objective optimization. For non-inferior solutions along the Pareto frontier, an improvement in one objective requires one or more other objectives to be worsened. It is not, in general, possible to definitively find all solutions along the Pareto frontier, but the frontier can be estimated using different search methods. For this study, the constraint method was the search method used to estimate the set of non-inferior solutions of the fertility clinic placement problem. In general, the constraint method optimizes one objective while systematically using bounding constraints on the other objectives (Cohon, 2004).

The process of generating non-inferior solutions used here proceeds as follows. In the case of fertility clinic placement, first the non-inferior solutions with the best and worst demand covered are found. Next, the parameter Δ is calculated as:

$$\frac{(best\ demand - worst\ demand)}{M} = \Delta \text{ with } M = 100.$$

Then, $for\ i\ =\ 1\ to\ M$, the $population\ objective$ is maximized subject to the normal MCLP formulation with an additional constraint added in each iteration of:

$demand\ objective \geq\ worst\ demand + i * \Delta$.

Larger values of M allow the constraint method to produce a better estimate of the full non-inferior solution set because larger values of M make Δ smaller, reducing the chance that a non-inferior solution between added constraints is skipped over; however, larger values of M require more computation (Cohon, 2004).

4.2.3 Expanding racial equity in geographic access to ART clinics

Chapter 2 found that racial groups White alone and American Indian and Alaska Native alone have worse geographic access to ART compared with the geographic access for all US races, and that every other racial category has better access to ART compared with all US races. Placement of new clinics to reduce the inequities in geographic access to care should attempt to correct for this imbalance, so one approach to correcting this racial inequity in geographic access was to calculate a weighted population in each county. The weighted population in a county is calculated based on the disparity in geographic access as shown in Equation 4-2, and the resultant weights are detailed in Table 4-1.

Equation 4-2. Weighting scheme for accounting for disparity in geographic access to ART showing how the weighted population in a county is calculated.

$$weight_{race} = 1 - (\%\ with\ geographic\ access\ to\ ART_{race} - \%\ with\ geographic\ access\ to\ ART_{all\ races})$$

$$weighted\ population_{county} = \sum_{race} weight_{race} * population_{race,county}$$

Table 4-1. Weighting scheme accounting for racial disparities in equitable geographic access to ART

Access to ART	White alone	Black or African American alone	American Indian & Alaska Native alone	Asian alone	Native Hawaiian & Other Pacific Islander alone	Some other race alone	Two or more races
At least one main or satellite clinic	79%	88%	65%	94%	87%	89%	85%
Difference in geographic access (compared to all US races, 82%) to at least one main or satellite clinic	-3%	6%	-17%	13%	6%	7%	3%
Weight accounting for disparity in geographic access	103%	94%	117%	87%	94%	93%	97%

However, if the purpose of new clinic placement is to impact racial disparities in utilization of ART, an alternate approach is to prioritize only racial groups that also underutilize ART and choose weights proportional to the lack of geographic access among those groups. Because the only racial group that is geographically underserved, based on my CBSA-based analysis, that also underutilizes ART is American Indian and Alaska Native alone, it follows that the best strategy to reduce racial disparities in utilization of ART is to place clinics with the sole objective of maximizing coverage of American Indians and Alaska Natives.

4.3 Results

This section details the results found when applying optimization to the problem of placing new fertility clinics. The locations for five new clinics per objective are shown in Section 4.3.1 for the two objectives of maximizing female reproductive-age population covered and maximizing predicted ART demand covered. Next, the Pareto frontier between the two objectives is explored in Section 4.3.2 to understand how decision space is affected along the Pareto frontier. Then, results from applying optimization to improve racial equity in access to ART are detailed in Section 4.3.3. Finally, hypotheses H10, H11, and H12 proposed in the introduction of this essay are tested in section 4.3.4, and a case study is presented in which an ensemble of model recommendations spanning the dissertation are spatially combined in Section 4.3.5 to be hopefully less wrong and more useful.

4.3.1 Optimizing coverage of demand and population from new ART clinics

The MCLP optimization was solved to maximize coverage of the female reproductive-age population in every US county (3,142 in total), where "coverage" is defined as having a facility in a county where its center of population is within a 60-minute drive of the covered counties' centers of population). The counties with existing fertility clinics were mathematically required to have fertility clinics via added constraints, and an additional *N* counties were selected for optimal placement by the MCLP algorithm. A base case of *N* = 5 was considered to allow for later comparison (in Section 4.3.5) with the locations found in Chapter 2 (Section 2.3.3.1). The optimal solution when maximizing coverage of population from 5 new facilities found a total covered US population of 57.06M women with a corresponding demand covered of 950.8k ART cycles. The optimal solution when maximizing coverage of demand found a covered population of 57.01M women with a corresponding demand covered of 952.5k ART cycles. The map displaying these solutions is shown in Figure 4-2.

New clinics were placed in different states dependent of the objective maximized. New clinics were placed in Missouri, Georgia, Kentucky, North Carolina, and Michigan when maximizing coverage of population, but when maximizing coverage of demand, the optimal solution changed slightly. Under the alternate objective, the locations in Missouri and North Carolina are again selected, the clinic placed in Georgia shifts slightly south to an adjacent county. The clinics in Kentucky and Michigan were not selected under this objective and were placed instead in Illinois and New York.

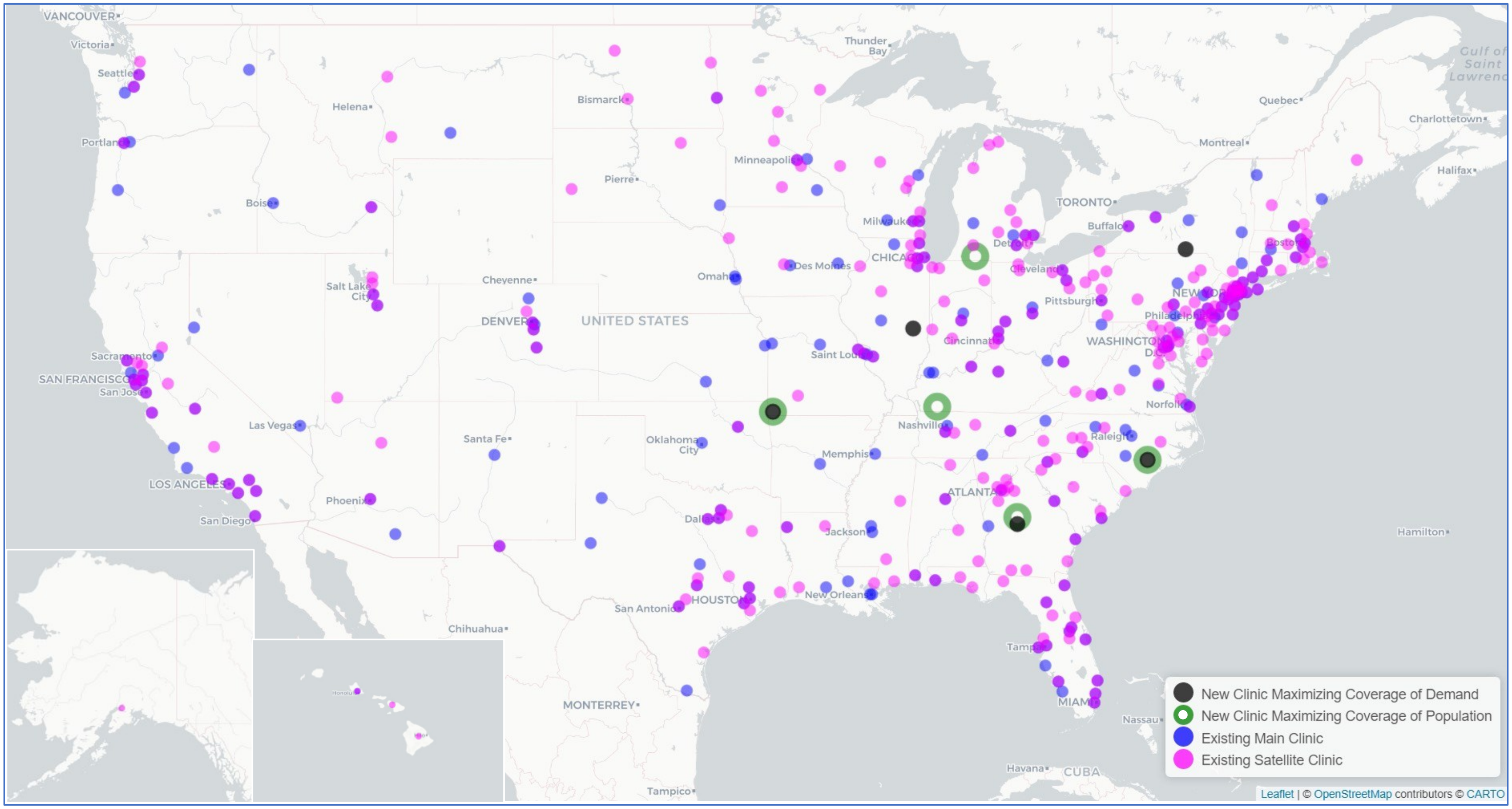


Figure 4-2. Counties with existing main and satellite clinics are shown as blue and pink dots respectively. Purple dots represent counties with both clinic types. Five new clinic locations maximizing coverage of population are shown as green circles, and five new clinic locations maximizing coverage of demand are shown as black dots

4.3.2 Multiobjective Pareto frontier between covering demand and population

The constraint method was then employed, finding 7 non-inferior solutions from 100 searches, each with 5 new additional facilities. The non-inferior solutions forming the Pareto frontier are shown in Figure 4-3. Moving between solutions along the Pareto frontier reduces the value of one objective in exchange for an increased value of the other objective.

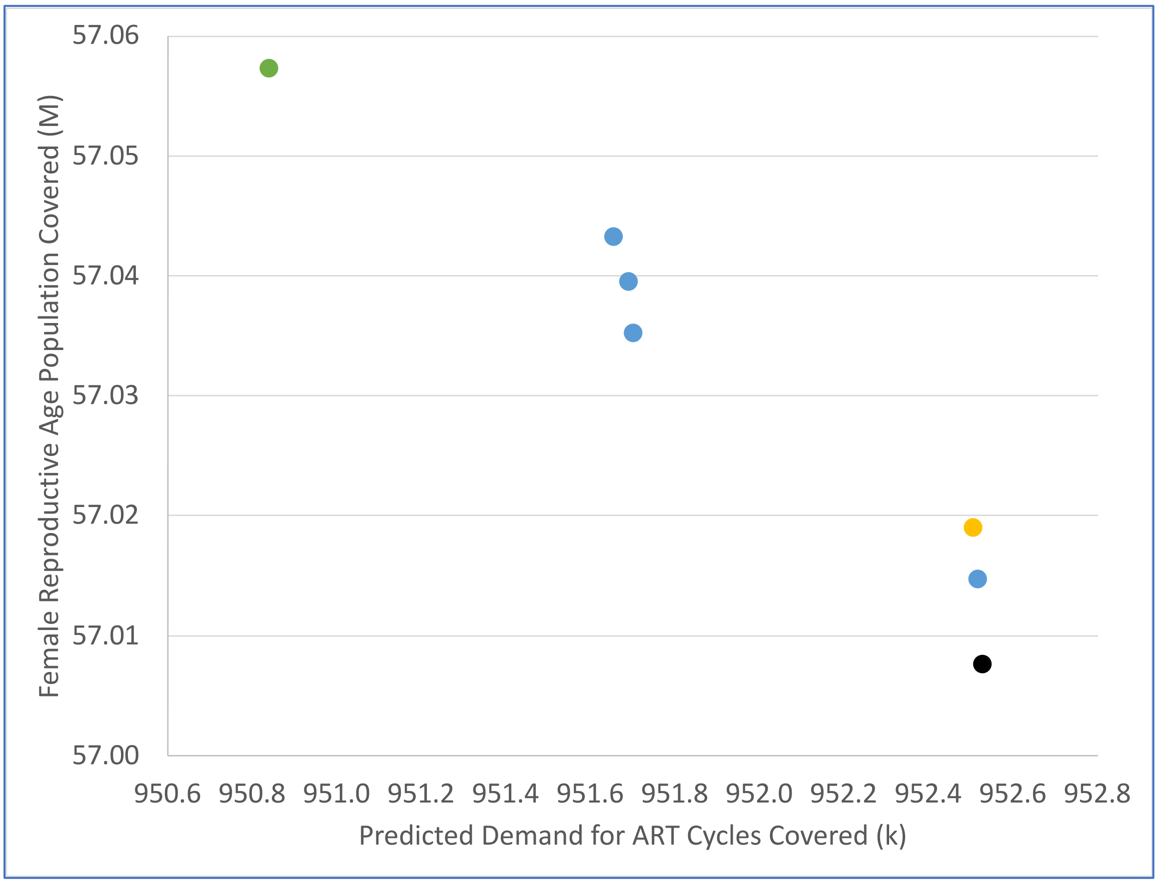


Figure 4-3. Seven solutions along the Pareto frontier between competing objectives of reproductive-age population covered and demand for ART cycles covered from placing 5 new ART clinics. The solution shown in green maximizes population covered, the solutions in blue are other non-inferior solutions along the Pareto frontier, the solution in orange is the visually selected balanced solution or "knee in the curve," and the black solution maximizes demand covered.

From the population covered vs. demand covered tradeoff curve, a balanced solution between both objectives was identified. This balanced solution is known as the “knee in the curve,” which is the cutoff point of diminishing returns, where the marginal benefit to one objective declines greatly and is unlikely to be worth the cost to the other objective. This “knee” point was selected visually with 57.02M population covered and 952.5k demand covered. Starting from the maximize coverage of demand only solution (black point in the figure), swapping clinics in Houston County, GA and Coles County, IL for clinics in Bibb County, GA and Champaign County, IL will yield the balanced solution (orange point in the figure), resulting in an increased coverage of 11,000 reproductive-age women, compared with the demand-only solution at a cost to demand coverage of only 21 ART cycles. Starting from the maximize coverage of population only solution (green point in the figure), 17,000 more ART cycles can be covered in exchange for reducing coverage of 38,000 women. This is achieved by swapping placement of clinics selected from Todd County, KY and St. Joseph County, MI for Tioga County, NY and Champaign County, IL, to yield the balanced solution (orange point in the figure).

The three solutions (maximize population covered, maximize demand covered, and the balanced solution, which are the green, black, and, orange points, respectively in the figure) were explored to understand the changes in decision space as the prioritization of objectives was changed. Two locations (McDonald County, MO and Duplin County, NC) were found to be stable solutions that were selected for every solution on the Pareto curve. One of two adjacent counties in GA (Bibb and Houston) were selected for every solution on the Pareto curve and when combined are considered a pseudo-stable solution. The balanced solution (orange) also selected Tioga County, NY and Champaign County, IL. Coles County, IL was selected as a location for a new clinic in the maximize coverage of demand-only solution (black). St. Joseph County, MI and Todd County, KY were selected as locations for new clinics in the maximize coverage of population-only solution (green). These locations mapped are shown in Figure 4-4, and the estimated coverage of population and demand for each new clinic location is detailed in Table 4-2.

Table 4-2 Identified new clinics' estimated coverage of female reproductive-age population and predicted demand for ART cycles

New Clinic County	Female Reproductive-Age Population in Cities Covered	Estimated Demand for ART Cycles Covered
Bibb, Georgia	113,487	1,121
Todd, Kentucky	110,914	887
St. Joseph, Michigan	104,461	888
McDonald, Missouri	164,064	2,109
Duplin, N. Carolina	134,025	1,061
Tioga, New York	90,405	1,703
Champaign, Illinois	86,686	1,740
Houston, Georgia	109,230	1,131
Coles, Illinois	79,544	1,751

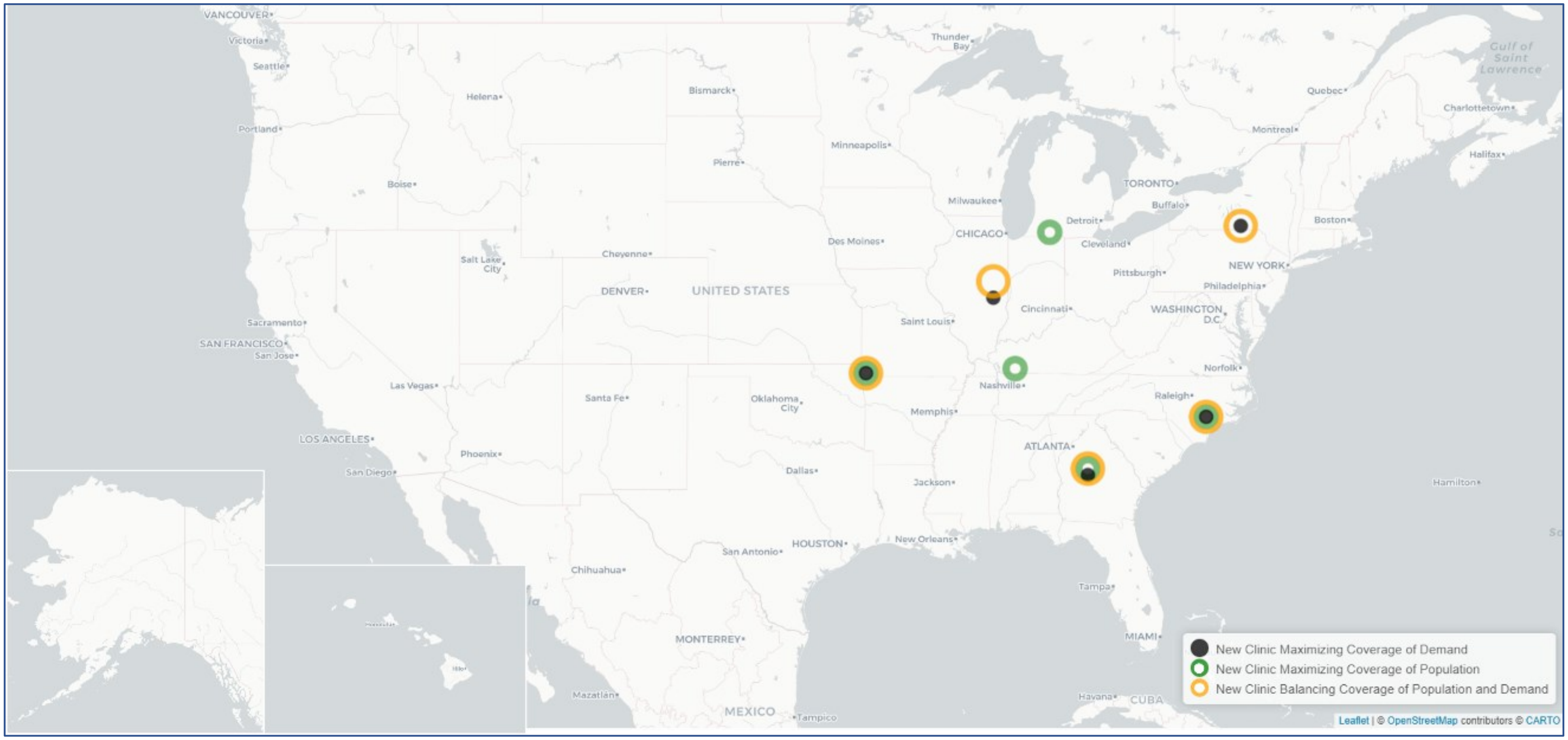


Figure 4-4. Optimized locations of five new fertility clinics dependent on the objective. Locations maximizing population covered are shown as green circles, locations maximizing demand covered are shown as black dots, and locations balancing coverage of population and demand are shown as orange circles.

The changes to decision space while traveling along the Pareto frontier is now analyzed in more detail. Starting with the solution maximizing coverage of population, the following counties were selected by the MCLP algorithm: Bibb County Georgia, Todd County Kentucky, St. Joseph County Michigan, McDonald County Missouri, and Duplin County North Carolina. As the amount of demand covered increased and population covered decreased, various non-inferior combinations of counties were selected by the MLCP algorithm, ending with the solution maximizing coverage of demand, selecting the following counties: McDonald County Missouri, Duplin County North Carolina, Tioga County New York, Houston County Georgia, and Coles County Illinois. This progression across the Pareto frontier in decision space is shown in Table 4-3.

Table 4-3. Decision space along the Pareto frontier trading off maximized coverage of demand for maximized coverage of population. Colored cells represent the county at the left selected by MCLP for an ART clinic for the demand and population coverage shown at the top. Green cells are counties selected to maximize coverage of population, blue cells are other non-inferior solutions along the Pareto frontier, orange cells are counties selected to balance coverage of population and demand, and black cells are counties selected to maximize coverage of demand.

Non-inferior solutions	Green point in Figure 4-3				Orange point in Figure 4-3		Black point in Figure 4-3
Demand	950,839	951,653	951,690	951,701	952,506	952,516	952,527
Population	57,057,295	57,043,239	57,039,520	57,035,263	57,019,011	57,014,754	57,007,612
Bibb, Georgia							
Todd, Kentucky							
St Joseph, Michigan							
McDonald, Missouri							
Duplin, N. Carolina							
Tioga, New York							
Champaign, Illinois							
Houston, Georgia							
Coles, Illinois							

4.3.3 Optimizing for improved racial equity in geographic access to ART

Results from two approaches addressing racial equity in geographic access are described here: a weighted population approach, and an approach that sites new clinics to emphasize service to groups that are historically underserved.

As described in Section 4.2.3 earlier, a weighting scheme was devised to account for racial disparities in geographic access to ART, and a maximal coverage optimization was performed to identify locations where the racially weighted population could be best covered. This analysis found the optimal solution accounting for racial equity in geographic coverage to be only a minor change from the solution maximizing coverage of female reproductive-age population of all US racial groups. Only the clinic located in North Carolina was changed from Duplin County, NC to Pender County, NC. This minor change resulted in an increase in the racially weighted equity objective by only 1,130 people at a cost of reducing coverage of reproductive-age women of all racial groups by 6,023. The comparison of the two solutions was mapped and is shown in Figure 4-5.

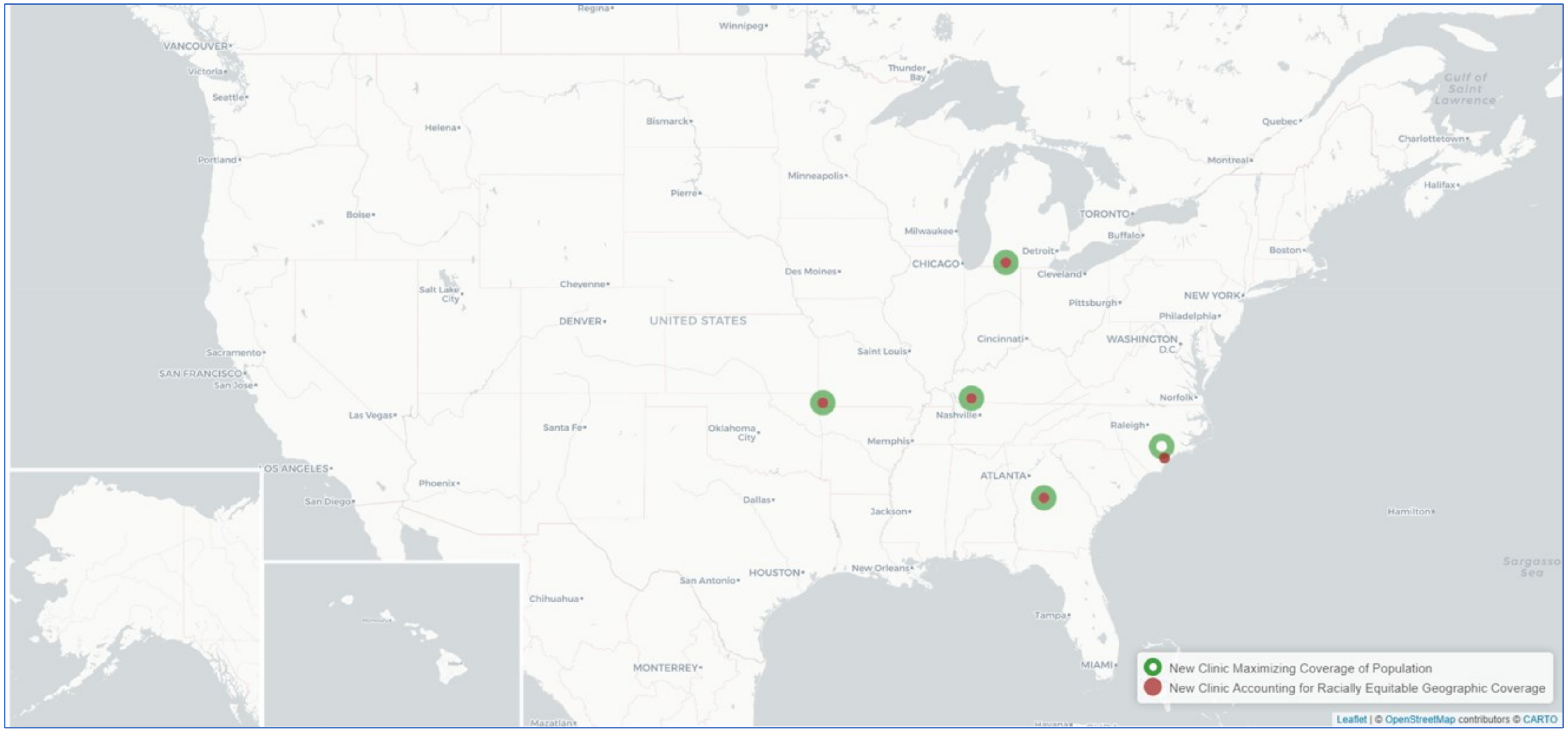


Figure 4-5. Comparison of five new clinics maximizing coverage of female reproductive-age population in green circles and weighted population in brown dots to account for racial inequities in geographic access to ART. Circles in green are five new clinics selected to maximize coverage of female reproductive-age population, and brown dots are five new clinics selected to maximally cover population weighted by race to account for existing racial disparities in geographic access.

It could be argued that geographic placement of new clinics should prioritize all ethnic and racial groups that underutilize ART services. The racial/ethnic groups with lower rates of ART utilization include Black, Hispanic, and American Indian/Alaska Native (Dieke et al., 2017). However, would adding a new ART clinic in an area with large numbers of people from racial/ethnic groups who have historically low ART utilization actually result in improved utilization? An analysis from Chapter 2 (Figure 2-8) found that African American and Hispanic populations already have better geographic access to ART than all US races and ethnicities (when using CBSAs to define access); therefore, geographic barriers to access ART are likely not driving the lack of utilization for those racial/ethnic groups, and other barriers such as economic, cultural, or educational are likely more impactful on utilization. Improving geographic access to care for African American or Hispanic populations who already have better geographic access compared with all US races is likely a less efficient way to improve utilization rates of those groups.

Because American Indians and Alaska Natives are the only racial/ethnic group that both underutilizes ART and has poor geographic access to ART, the second of the two approaches taken here to improve equitable utilization of ART via the placement of new ART clinics was to exclusively maximize coverage of American Indians and Alaska Natives. This maximization was performed using the MCLP model using county-level data, and the resultant five new clinic locations is shown in Figure 4-6 along with the best coverage of population solution and for comparison the previously shown solution using the weighted population accounting for inequities in racial geographic access to care.

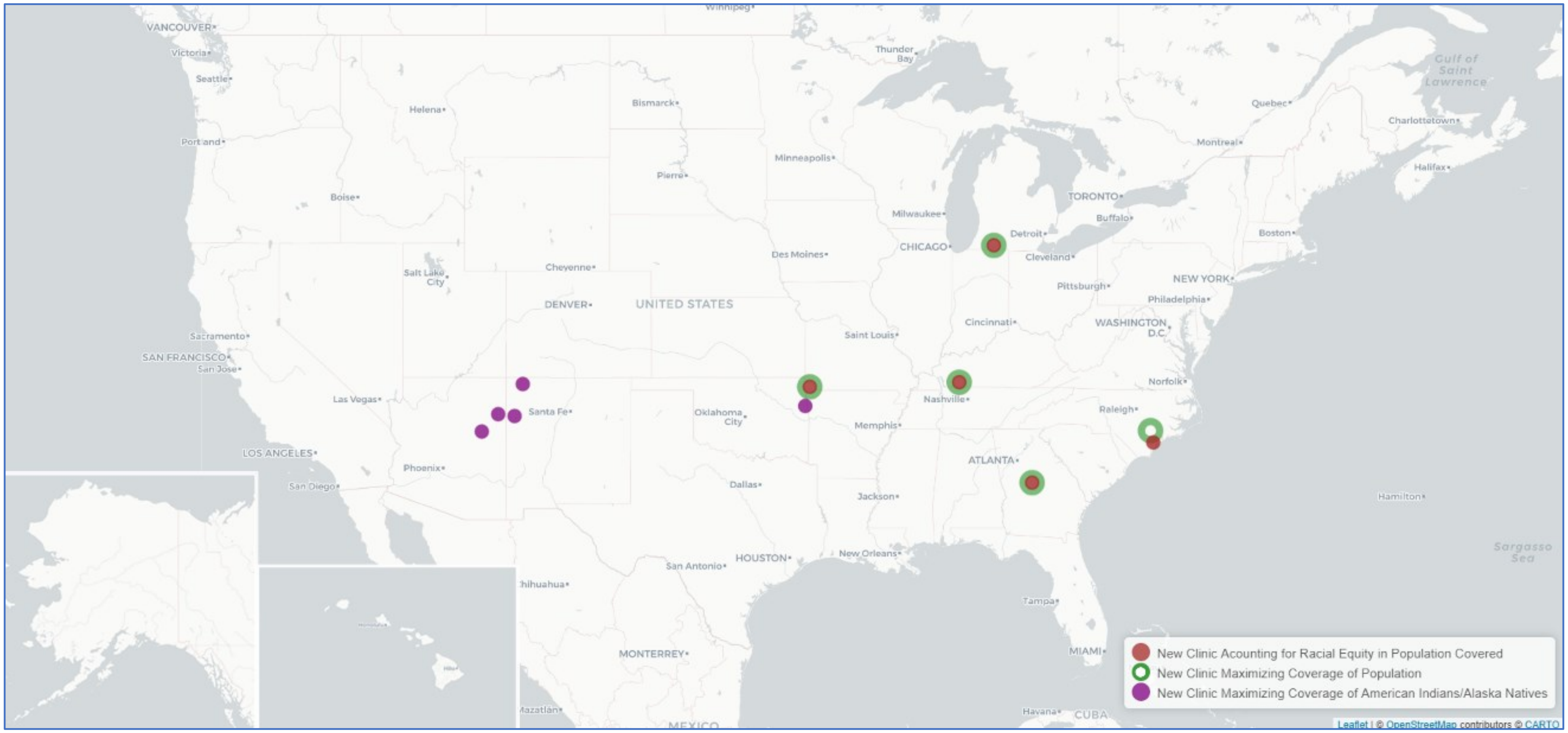


Figure 4-6. New clinic placement accounting for racial equity in geographic access to care and utilization of ART. Circles in green are five new clinics selected to maximize coverage of female reproductive-age population, brown dots are five new clinics selected to maximally cover population weighted by race to account for existing racial disparities in geographic access, and purple dots are five new clinics selected to maximize coverage of American Indians and Alaska Natives.

The locations found that maximize coverage of American Indian and Alaska Natives are: Apache County Arizona, Navajo County Arizona, McKinley County New Mexico, San Juan County New Mexico, and Adair County Oklahoma. This solution differs slightly from the solution found in Chapter 2 that used CBSAs to approximate geographic access vs. driving times with MCLP optimization in this chapter. The top-five CBSAs found with the most American Indians and Alaska Natives are: Gallup New Mexico, Lumberton North Carolina, Farmington New Mexico, Show Low Arizona, and Tahlequah Oklahoma. Since county names and CBSA names do not generally match, the identification of similar locations was achieved through mapping both the best counties and the best CBSAs. This map (Figure 4-7) shows that the MCLP solution and CBSA solution largely agree for how to best increase geographic access for American Indians and Alaska Natives.

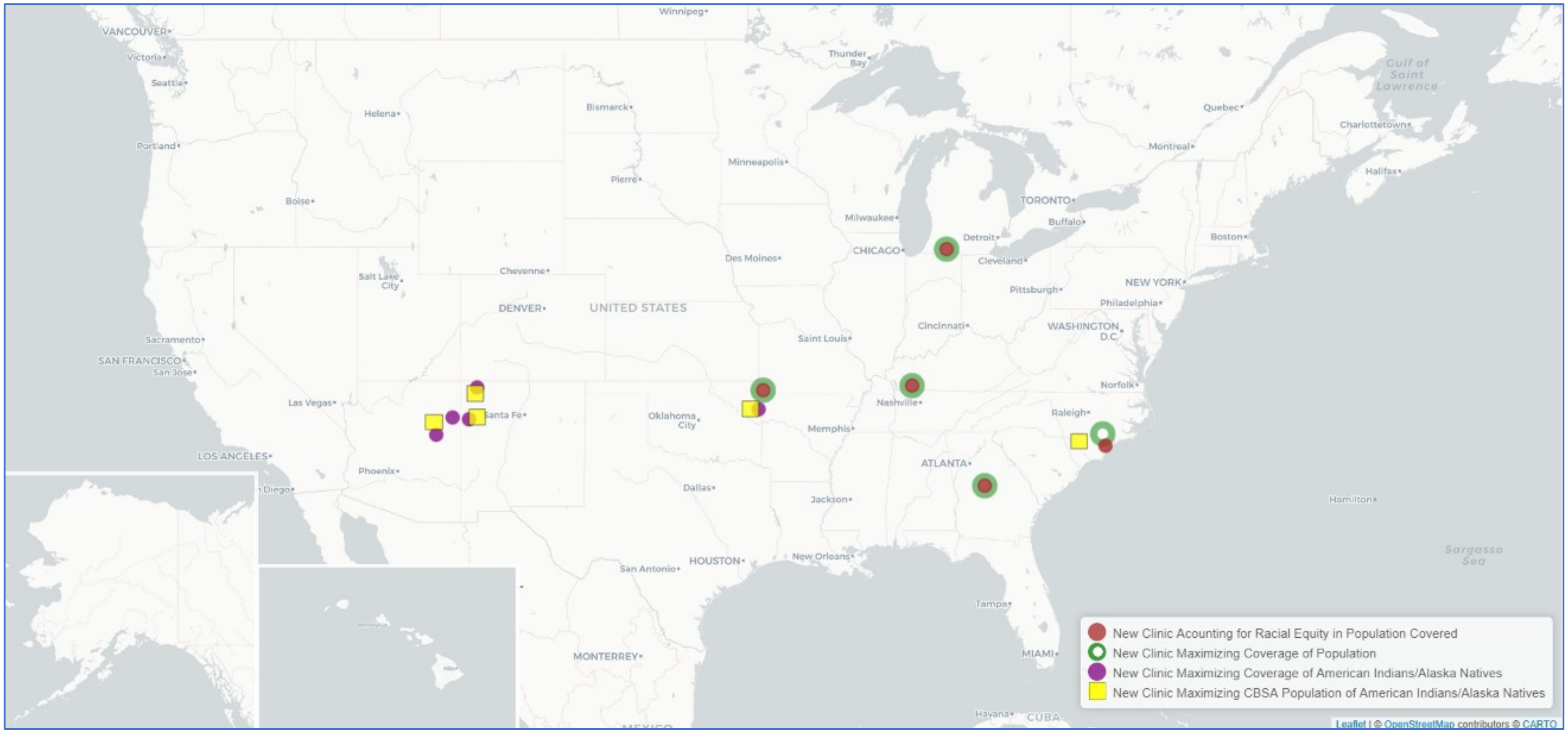


Figure 4-7. New clinic placement from different modeling approaches to improve racial equity in geographic access to care and utilization of ART. Circles in green are five new clinics selected to maximize coverage of female reproductive-age population, brown dots are five new clinics selected to maximally cover population weighted by race to account for existing racial disparities in geographic access (first method of this section). Purple circles are five new clinics selected to maximize coverage of American Indians and Alaska Natives via MCLP optimization (second method of this section), and yellow squares are five new clinics selected as the top five CBSAs based on American Indian/Alaska Native population (Section 2.3.3.2, in Chapter 2)

Another interesting finding is that the stable solutions identified in Section 4.3.2 that maximize coverage of both population and demand (Duplin, NC and McDonald, MO) are near two locations that maximize geographic access for American Indian/Alaska Native populations. There are relatively large population centers of underserved American Indians and Alaska Natives nearby in Lumberton, NC and Tahlequah, OK, with American Indian/Alaska Native populations of 53k and 16k people respectively. Not only are these stable solutions good locations for covering all races/ethnicities with large demand estimated, they also help reduce racial inequities in geographic access to ART.

4.3.4 Tests of hypotheses

In Section 4.1.3, above, three hypotheses were offered. The hypotheses and the results of the analysis are summarized as follows:

H10: New clinic locations chosen when maximizing coverage of population or demand will occur across the entire US: *Contradicted* (see Section 4.3.4.1, below)

H11: As solutions along the Pareto frontier move from maximizing demand covered to maximizing population covered, solutions will move from IVF mandate states to non-IVF mandate states: *Confirmed* (see Section 4.3.4.2, below)

H12: Locations that optimize coverage of underserved racial and ethnic minorities will primarily be located to cover populations on Indian reservations: *Confirmed* (see Section 4.3.4.3, below)

4.3.4.1 H10 - Spatial distribution of optimized new clinic locations

It was hypothesized that new clinics would be optimally placed throughout the US except for in the Northeast, but this hypothesis was contradicted with new clinics being placed exclusively in the Eastern half of the US and one new clinic placed in New York. The placement of clinics in the Eastern half of the US was surprising based on the finding from Nangia et al. (2010a) that the northeastern US was overserved, and underserved states included Alaska, Montana, Wyoming, and West Virginia. But the optimization model was not performed at the coarse state-level resolution; optimization was done at the county level, so it was able to identify highly populated counties that are near areas with poor geographic access that are just out of reach of the 60-minute drive coverage standard. For example, Macon, GA covered by clinics placed in Bibb or Houston Counties has 31,089 unserved reproductive-age women, but it is about a 90-minute drive from the closest fertility clinic in Fayette County, GA.

Additionally, the optimization was able to identify clever sites that are centrally located relative to two or more population or demand centers that would not seem from first glance to be good places, but the MCLP model was able to identify facility locations with the most coverage potential. For example, McDonald County Missouri is in a rural area and has a limited female reproductive-age population of only 4,098, but it is advantageously located near the cities of Joplin Missouri, Rogers Arkansas, Springdale Arkansas, and Fayetteville Arkansas, allowing it to reach 164k women and 2,109 estimated ART cycles in counties within a 60-minute drive.

4.3.4.2 H11 - Change in allocation of new clinics along the Pareto frontier between IVF and non-IVF mandate states

I hypothesized that as solutions along the Pareto frontier moved from maximizing demand covered to maximizing population covered, solutions would move from IVF mandate states to non-IVF mandate states because state IVF mandates were found in Chapter 3 to be associated with higher ART demand. This was found to be the case (detailed in Table 4-3) with two new facilities in the IVF mandate states of New York and Illinois being switched to non-IVF mandate states in Michigan and Kentucky when switching between the objectives of maximizing demand covered and maximizing population covered.

4.3.4.3 H12 - Locations to optimize coverage of underserved racial and ethnic minorities

Finally, I hypothesized that locations that optimize coverage of underserved racial and ethnic minorities would primarily be located to cover populations on Indian reservations because primarily American Indians and Alaska Natives were identified in Chapter 2 to lack geographic access to care. This hypothesis was strongly supported by the optimization results (shown in Figure 4-6) that sited the five new clinics to maximize coverage of American Indians and Alaska Natives exclusively in counties containing Indian reservations in Arizona, New Mexico, and Oklahoma.

4.3.5 Case study in using an ensemble of model recommendations to inform placement of new ART clinics

The use of optimization to identify locations of new fertility clinic locations has a fundamental modeling difference from the statistical regression models of clinic presence used in Chapter 2 and Chapter 3. That difference is the use of optimization to maximize coverage vs. statistically predicted likelihood of clinic presence regardless of population covered other than as a property of the CBSA or county. Both modeling methods have their own strengths and limitations. The statistical prediction of clinic presence approach could be described as near-sighted because it is exclusively driven by characteristics of each CBSA or county without information from nearby areas considered. Nearby counties/CBSAs could potentially provide geographic access to a county/CBSA without a clinic if transit times were short enough. Conversely, the optimization approach could be described as far-sighted because it focuses exclusively on maximizing coverage without factoring in suitability of a chosen county/CBSA. For example, adequate staffing, available facilities, internet connectivity, and power reliability are not considered in the optimal siting of a clinic according to the MCLP algorithm. These factors are not explicitly considered in the statistical models, but the predictors used in the statistical models could allow for these factors to be captured implicitly.

Incremental tweaks to both the statistical and optimization methods could be applied to make them less near-sighted or less far-sighted, but another option is to spatially combine recommendations of the models into an ensemble of recommendations. The aphorism that "all models are wrong, but some are useful" (Box, 1979) inspired the decision to incorporate recommendations of each model to hopefully produce a recommendation that is less wrong and more useful. If different modeling approaches recommend the same locations, those locations would likely be promising places for a new clinic. Therefore, each modeling approach from Chapters 2, 3, and 4 was used to produce a recommendation of five clinics each.

For the CBSA statistical modeling approach from Chapter 2, main clinic and satellite clinic presence were modeled separately, then the likelihood of either clinic type was calculated by applying probability theory described in Equation 4-3.

Equation 4-3. Probability theory calculation of the likelihood of either clinic type being present in a CBSA or county.

$$P(Either\ Clinic) = 1 - (1 - P(Main\ Clinic)) * (1 - P(Satellite\ Clinic))$$

This expression assumes that the probabilities are independent: the presence or absence of a main clinic does not affect the probability of presence of a satellite clinic, and vice versa.

The five CBSAs without a clinic but with the highest modeled likelihood of either clinic type were chosen as the CBSA statistical model recommendation. These locations are shown in Table 4-4.

Table 4-4. The recommended top five CBSAs without fertility clinics based on the modeled likelihood of either clinic type.

Geographic Area (CBSA) Name	Likelihood Main Clinic	Likelihood Satellite Clinic	Likelihood Either Clinic
Stockton, CA Metro Area	93%	64%	97%
Lakeland-Winter Haven, FL Metro Area	72%	42%	84%
Fayetteville-Springdale-Rogers, AR Metro Area	60%	40%	76%
Los Alamos, NM Micro Area	24%	53%	64%
Santa Cruz-Watsonville, CA Metro Area	31%	39%	58%

For the county-based statistical modeling approach from Chapter 3, main and satellite clinic presence were also modeled in separate statistical models, and again the likelihood of either clinic being located in a county was calculated using probability theory (shown in Equation 4-3). Then, the five counties without a clinic but with the highest modeled likelihood of either clinic type were chosen as the county statistical model recommendation. These locations are shown in Table 4-5.

Table 4-5. The recommended top five counties without fertility clinics based on the modeled likelihood of either clinic type.

County	State	Likelihood Main Clinic	Likelihood Satellite Clinic	Likelihood Either Clinic
Denton	Texas	78%	84%	96%
Ramsey	Minnesota	87%	68%	95%
Prince George's	Maryland	37%	90%	93%
Richland	South Carolina	80%	71%	91%
Lancaster	Nebraska	64%	68%	90%

Finally, from the optimization modeling approach from Chapter 4, three recommendations were considered from the MCLP models that: maximized population covered, maximized demand covered, and balanced the coverage of population and demand. The five counties selected by the MCLP model that optimally expand coverage of female reproductive-age population were chosen as the recommendation that maximizes population covered. These locations are shown in Table 4-6.

Table 4-6. The recommended five counties without fertility clinics that optimally expand coverage of female reproductive-age population.

County	State	Method
Duplin	North Carolina	Population Coverage Maximized
McDonald	Missouri	Population Coverage Maximized
Bibb	Georgia	Population Coverage Maximized
Todd	Kentucky	Population Coverage Maximized
St. Joseph	Michigan	Population Coverage Maximized

The five counties that optimally expanded coverage of demand were chosen as the recommendation that maximizes demand covered. The ART demand estimate was produced from the statistical models in Chapter 3, so these recommendations are essentially a hybrid between an optimization and statistical approach with the optimization model using the statistical predictions as inputs. These locations are shown in Table 4-7.

Table 4-7. The recommended five counties without fertility clinics that optimally expand coverage of demand for ART.

County	State	Method
Duplin	North Carolina	Demand Coverage Maximized
McDonald	Missouri	Demand Coverage Maximized
Houston	Georgia	Demand Coverage Maximized
Coles	Illinois	Demand Coverage Maximized
Tioga	New York	Demand Coverage Maximized

Lastly, the compromise “knee in the curve” solution selected from the multiobjective MCLP modeling, balancing population covered vs. demand covered was chosen as the recommendation balancing coverage of population and demand. These locations are shown in Table 4-8.

Table 4-8. The recommended five counties without fertility clinics that balance expanding coverage of population and demand for ART.

County	State	Method
Duplin	North Carolina	Balanced Coverage of Population and Demand
McDonald	Missouri	Balanced Coverage of Population and Demand
Bibb	Georgia	Balanced Coverage of Population and Demand
Champaign	Illinois	Balanced Coverage of Population and Demand
Tioga	New York	Balanced Coverage of Population and Demand

The synthesis of the recommendations from Chapter 2, Chapter 3, and this chapter (Chapter 4) was performed by mapping the recommended five geographic areas from each modeling approach in Table 4-4 through Table 4-8. This map with these recommendations compiled is shown in Figure 4-8. The combination of each modeling approach in a geospatial context yields many interesting insights. First, there are many potential locations available across the US to increase geographic access to fertility care. No single modeling approach was concentrated in any one metropolitan or geographic region. Second, two stable solutions were found by the optimization approach when maximizing coverage of demand or population at McDonald County Missouri and Duplin County North Carolina. Additionally, a pseudo-stable solution was found in central Georgia in the adjacent counties of Bibb and Houston. Finally, a consensus was found from the recommendations of the CBSA statistical modeling approach and all optimization objectives for one location in Northwest Arkansas/Southwest Missouri.

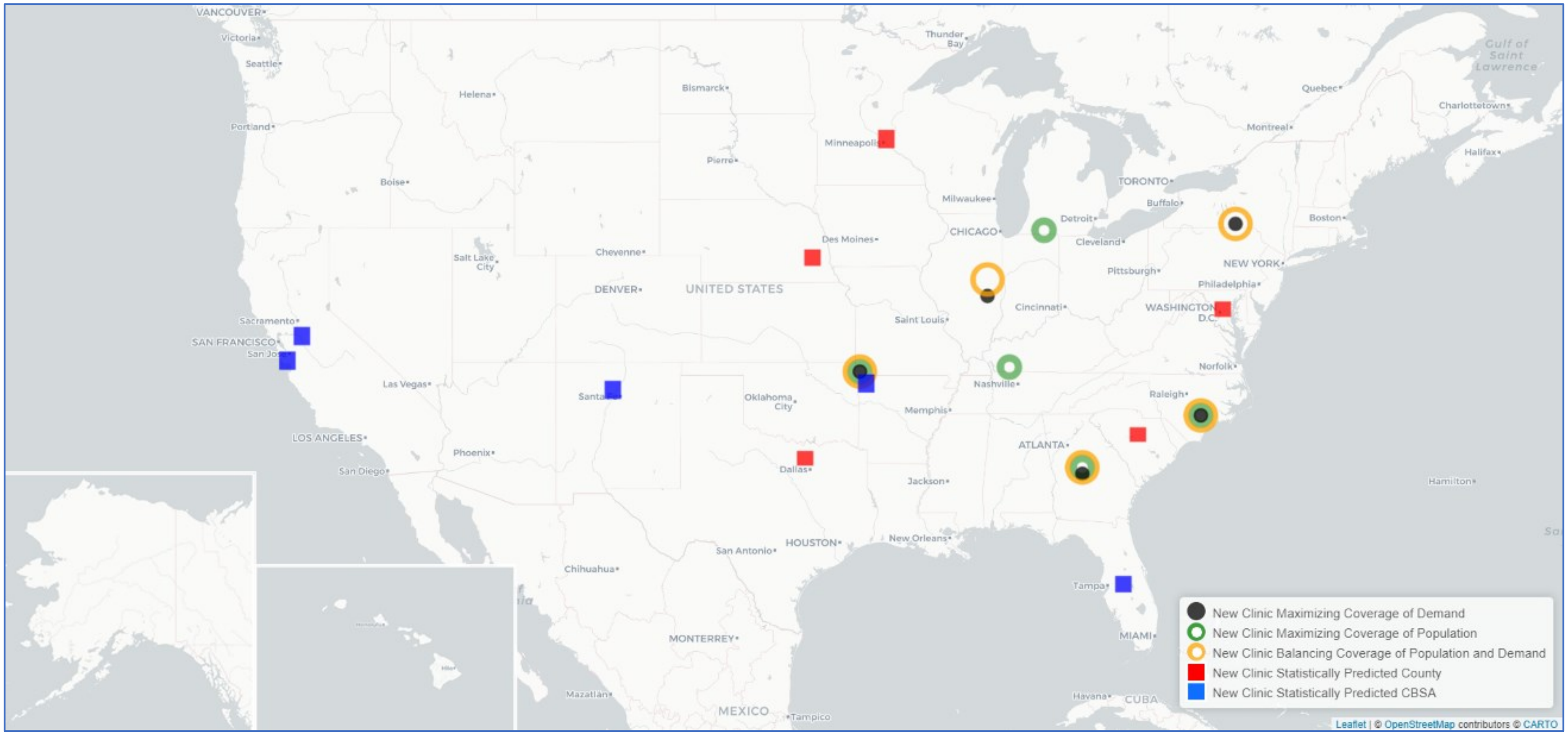


Figure 4-8. Combined recommendations from models in Chapters 2, 3, and 4. Each modeling method recommended the five best locations for a new clinic using different methodologies. Black circles maximize coverage of demand, green circles maximize coverage of population, and orange circles balance coverage of population and demand. Squares are the new clinics selected by statistical logistic regression modeling of county clinic presence in red and CBSA clinic presence in blue.

4.4 Discussion

In this section, several broad insights derived from this chapter's optimization results are discussed. First, the novel multiobjective approach presented in this essay is explored in Section 4.4.1 for its utility in improving the social benefit of general for-profit facilities. Then, this essay's case study is considered in 4.4.2 to examine potential reasons why the different model recommendations for new clinics from each chapter of the dissertation either agreed or disagreed. Finally, limitations of this essay's methods are covered with potential mitigations for future studies in Section 4.4.3.

4.4.1 For-profit facility locating with social benefit by exploring the Pareto frontier between competing objectives of maximizing demand covered vs. population covered

The approach presented in this essay to explore the Pareto frontier between competing objectives of maximizing demand covered and maximizing population covered is unique because it presents an analytical framework for finding socially better locations for profit maximizing businesses to locate. A review of multiobjective facility locating by Current et al., (1990) summarized that covering of demand in the literature is exclusively focused on public facilities, whereas profit maximization problems tend to focus on maximizing the rate of return. No studies found in the literature reviewed have focused on improving the social good of a for-profit facility by selecting a solution along the Pareto frontier other than the one expected to maximize profit.

Outside of other factors such as facility operation costs that vary spatially, ART practices seeking to maximize profit from their clinics choose locations that will have the highest demand for ART cycles. Based on the analysis in Chapter 2, this profit maximizing approach appears to have been taken historically, resulting in co-located ART clinics in areas with proven markets rather than clinic locations being selected to reach the most underserved potential patients. If ART clinics were placed like public facilities, it could be argued that new clinics should be placed to maximize population covered rather than demand, but this will yield suboptimal locations with respect to profit maximization if demand and population are not perfectly correlated. The approach taken in this essay was to explore the Pareto frontier between the two coverage objectives where it was learned when selecting new clinic locations that it was possible to concede an incremental amount of demand covered (21 cycles) in exchange for a substantial number (11,000) of reproductive-age women gaining geographic access to ART.

This novel approach has the potential to be applied outside the healthcare domain to locate for-profit facilities in ways that improve their societal benefit in exchange for reduced estimated profit. For example, food deserts are places where people do not live within a reasonable distance to grocery stores and those people lack the means to travel the distances required to these stores (Beaulac et al., 2009). Grocery stores are not public facilities, and just like fertility clinics, they would be drawn to locate in areas where their profits will be maximized, not where they can provide groceries to the most people. This tension between providing geographic access for the most people to a facility and the profit that facility will generate exists in all cases where demand and population are not equivalent. A model could be derived to address this issue, but it would require far more geographic disaggregation than the county-level optimization model used in this thesis.

Policies could be adopted to encourage for-profit businesses to locate in less profitable areas. For example, tax incentives have been used to attract businesses to invest in underserved areas, such as the opportunity zones created from the passage of the Tax Cuts and Jobs Act of

2017. This law provided targeted tax incentives designed to spur investment in geographic areas historically lacking investment (IRS, 2021; Tax Cuts and Jobs Act, 2017). Similar "ART Opportunity Zones" could be created with tax incentives for the opening of new ART clinics in underserved areas.

Another approach would be to tweak the tax code to allow for ART practices to deduct from their taxes an estimate of their annual profit given up by operating in locations that are less profitable but cover more people. This proposed taxation model is similar to the one used to incentivize conservation of natural resources through conservation easements in which landowners agree to not develop land they own in exchange for tax deductions (Rissman et al., 2007). Multiobjective optimization could serve as a decision aid in selecting these alternate more socially beneficial locations, and it could also be used to estimate the "social benefit" tax deduction.

4.4.2 Concurrence of model recommendations in case study

Concurrence was found between the optimization models and the CBSA statistical clinic presence model but not with the county statistical presence model. The county statistical modeling recommendations did not corroborate with the recommendations from the other modeling approaches likely because the CBSA statistical and optimization methods could be described as more far-sighted, and the county statistical could be described as more near-sighted. Dependent on a decision maker's view on how far couples are willing/able to travel to seek care, the county statistical recommendations could be discounted or valued. For example, Denton County, TX (home to University of North Texas) does not have a fertility clinic in the county, but there are fertility clinics in nearby counties such as Dallas County. A decision maker considering Denton County for a new clinic could conclude that its geographic access is sufficient based on the specific locations of ART clinics in nearby counties, but another decision maker could conclude that geographic access in Denton County is not sufficient and would be a good location to open a new ART clinic.

4.4.3 Limitations

The statistical prediction of ART demand from Chapter 3 was able to estimate ART demand using a range of predictors in addition to population. However, data on ART demand are only available from the counties that have ART clinics with embryology labs reporting to the CDC (4% [n=120] of all counties) to predict ART demand for the remaining 96% [n = 2,787] of counties without IVF cycles reported to the CDC. Therefore, confidence in the results of the model that optimized demand coverage in choosing new clinic locations should be tempered by considering the underlying demand model's limited sample size. The best way to overcome this limitation would be for the CDC to track infertility diagnoses nationally. The National Survey for Family Growth's (National Center for Health Statistics, 2020) annual survey is not extrapolated to produce local estimates, and the roughly 10,000 people it surveys are not sufficient to produce accurate estimates of ART demand vis-à-vis infertility or impaired fecundity.

The MCLP model operates using point representations of populated places. The Center of Population of each county (U.S. Census Bureau, 2010) was used as the best point representation of counties, which are polygons. This location in the county could be located further from a different county nearby than the driving time standard of 60-minutes and thus be uncoverable, but there could be parts of the county that are within the driving time standard. Choosing smaller geographic areas than counties would reduce the error associated with this modeling approach. For example, using zip codes, census tracts, or census blocks would provide higher resolutions than counties and reduce this limitation, but fewer environmental and socioeconomic datasets are available for these smaller geographic areas. Smaller geographic areas could be used if exclusively maximizing population covered, but demand estimates generated in Chapter 3 were at the county scale. It is possible to apply county estimates to smaller geographic areas, but a translation with various assumptions would have to be performed.

The primary reason for using counties over a smaller geographic area was the synergistic benefits between Chapter 3 and Chapter 4. Counties were selected in Chapter 3 because of the rich environmental and socioeconomic estimates available from the County Health Rankings dataset (County Health Rankings & Roadmaps, 2021b). To my knowledge such a rich source of environmental and socioeconomic estimates does not exist for smaller geographic areas. With the selection of counties in Chapter 3, the statistical estimates of demand for ART of each county produced in Chapter 3 were used directly by the demand coverage optimization in Chapter 4. A final benefit is the statistical predictions for both demand and clinic presence from Chapter 3 at the *county-level* could be considered alongside optimization results for *counties* selected as optimal locations for clinics in Chapter 4.

A final limitation of the MCLP model is the 60-minute driving time assumption. It assumes that coverage is binary until the driving time between counties is greater than 60-minutes. A sensitivity analysis could be performed to test the impact on coverage estimates using a range of travel time standards. However, the creation of the ensemble of recommendations helps reduce the negative impacts of this limitation because the combination of several modeling approaches, each with different assumptions, allows for a more holistic picture to be generated that does not hinge on a single assumption.

4.5 Summary

Optimization methods were applied to locate new fertility clinics with the aim of identifying areas to optimally expand geographic access to care. The set-covering location analysis method maximal covering location problem (MCLP) was applied to identify five US counties to locate new fertility clinics. Two coverage objectives were considered in this optimization: maximizing coverage of female reproductive-age population and maximizing coverage of demand for ART. A novel concept in equitable geographic access to facilities was presented that posits that in cases where population is not a perfect analog of demand (as is the case for ART), solutions along the Pareto frontier between maximizing coverage of demand vs. maximizing coverage of population allows for balanced solutions to be found that allow for-profit businesses to select alternate locations from maximizing coverage of demand that are expected to reduce profits incrementally in exchange for large increases in social utility.

County-level demand for IVF was estimated using predictions from statistical regression models in Chapter 3 that used socioeconomic and environmental factors as predictors of the number of ART cycles in a county. Multiobjective optimization was then employed to understand the tradeoff between covering population and covering demand. Seven solutions were found along the Pareto-optimal frontier using the constraint method, and a balanced solution between both objectives was chosen at the point along the curve when the increase in population covered was no longer worth the reduction in coverage of demand. The three solutions found through optimization (maximizing coverage of population, maximizing coverage of demand, and the solution balancing coverage of both population and demand) were mapped to understand the changes to decision space in moving along the Pareto frontier between the two objectives.

Equitable racial and ethnic geographic access to ART was explored via the application of the MCLP optimal coverage algorithm. A population weighting scheme accounting for current racial inequities in geographic access was applied. Five new clinics were located to optimally cover the racially weighted population, but the solution accounting for racial equity in geographic access differed only slightly from the solution maximizing coverage of female reproductive-age population of all races.

Therefore, I also considered a second approach to addressing equity. In this approach, when considering the utilization of ART, American Indians and Alaska Natives were prioritized for optimal placement of new clinics. The MCLP algorithm was applied with the objective of maximizing coverage of their population, and solutions were found in Arizona, New Mexico, and Oklahoma. Currently, only 63% (n = 1.73M) of the American Indian/Alaska Native population was found within a 60-minute drive of an ART clinic. Optimally adding just five new clinics identified by the MCLP algorithm extends access within a 60-minute drive to a quarter of a million American Indians and Alaska Natives, bringing the percentage of their population with geographic access up to 72% (n = 1.98M). Clearly, locations with high utility for increasing access to American Indians and Alaska Natives exist to substantially reduce the racial disparities in geographic access to ART.

Lastly, to consider recommendations from a range of modeling approaches, the logistic regression statistical models from Chapter 2 and Chapter 3 were applied to recommend five locations from each chapter to place new clinics, and the three optimization derived solutions from Chapter 4 were used to recommend five locations for new clinics from each method. Each of these recommendations were combined geospatially to identify geographic areas of concurrence between each modeling approach. Model concurrence was found in recommending a new clinic in Southwest Missouri/Northwest Arkansas where both the CBSA statistical method and optimization methods recommended locating a new clinic. However, the approaches did not agree on locations for other clinic locations. This synthesis of modeling approaches in a geospatial context allows for the array of recommendations derived from the various modeling approaches presented in this dissertation to be compared in the same domain to allow each approach to contribute to answering the question of how to best expand geographic access to fertility care.

Chapter 5 Conclusion

5.1 Summary of conclusions

This dissertation has contributed much needed knowledge to the study of geographic access to Assisted Reproductive Technology (ART) in the United States. First, I created a comprehensive, open dataset of all fertility clinics in the United States. With this dataset, geographic access to care was estimated based on access to main ART clinics and satellite ART clinics, and geographic access was characterized based on race and ethnicity. Next, statistical modeling was employed to better understand socioeconomic and environmental drivers of demand for ART and the industry's clinic placement patterns. Finally, optimization was utilized to select ART clinic locations to maximize coverage of population and demand, and multiobjective optimization was employed to understand the tradeoff curve between prioritizing coverage of population or coverage of demand.

The first contribution was in the fundamental understanding of where fertility clinics are located. A single list containing the locations of all fertility clinics in the US did not exist prior to this work. The CDC's list of clinics includes primarily main clinics that have embryology labs, and SART's list of clinics includes satellite clinics, but only 80% of clinics in the 2018 reporting to the CDC are SART members. There is a critical need for the community to have a common dataset on which to perform various analyses. The data underpinning the insights derived from any analysis must be both comprehensive and accurate. Additionally, by compiling a comprehensive, open dataset and then using it to perform subsequent analysis allowed its accuracy to be confirmed. For example, when the analyses in this dissertation indicated promising areas for new clinics, those areas were then manually confirmed with a web-search for clinics in the recommended location to confirm that a fertility clinic was not missed when the dataset was compiled.

A geospatial analysis was then conducted to produce an accounting of access to care to understand the patterns of main and satellite clinic placement and their impact on practices. An estimated 72% of reproductive-age women were found to live in an area with a main clinic that has an embryology lab. It was discovered that most satellite clinics are located near another practice's main clinic and do not extend geographic access to care. The 22% of satellite clinics that are located in an area without a main clinic, provide access to ART to approximately 8% (5.1M) of the female reproductive-age population. Therefore, it was concluded that the primary role of satellite ART clinics is to increase patient convenience and compete for patients over expanding geographic access to care. This finding is consistent with Hotelling's Law in economics that found competing sellers "cluster unduly" rather than locate in the socially optimal locations (Hotelling, 1929). However, satellite clinics were found to be positively associated with practices performing more IVF cycles, so they must be playing an advantageous service role in attracting patients and providing them enhanced convenience, reducing the high time-costs of ART and increasing the number of completed ART cycles.

The US reproductive-age population is underserved with respect to ART with only a quarter of estimated demand met for ART in the US. If the US's current insurance-based medical system was tweaked to mandate coverage of IVF nationwide and paid time off for treatment was universal, the problem of optimally locating ART facilities would likely be more easily solved using only maximizing reproductive-age population covered. Paid time-off to receive medical treatment is not required in the US. Some states and counties have mandates for paid sick leave, but it varies based on employer nationwide. The Family and Medical Leave Act of 1993 allows employees to take up to 12 weeks of job-protected unpaid leave for treatment of a serious illness of themselves or their family members, but it does not include infertility in its section on chronic conditions and requires the condition to cause “incapacity” (The Family and Medical Leave Act of 1993, 1993). This lack of inclusion of infertility as a chronic condition in FMLA may result in some employers determining infertility a serious illness and others not. There is a single court ruling on the matter that determined that an employee of a health insurance company who was terminated for taking time off for IVF treatments didn’t qualify for FMLA protection because her IVF treatments did not require more than three consecutive days of treatment (U.S. District Court, 2008). Because there are so many barriers to accessing ART, with lack of job protected leave being just one of them, deciding where to place a clinic based on female reproductive-age population alone would fail to capture demographic trends resulting in more demand for ART in certain areas over others.

Chapter 3 created models of demand and clinic presence to better understand the socioeconomic and environmental drivers of ART demand; however, the primary goal of Chapter 3’s statistical modeling was prediction of ART demand for use in Chapter 4’s optimization. Because existing clinics operate in a market-based, profit driven system, models capturing why existing clinics currently operate where they do give the best estimate of where to locate new clinics that will be economically viable. Opening a new clinic in an area that is not economically viable will ultimately result in the clinic closing and no improvement to geographic access to care.

The statistical model of ART demand found socioeconomic factors to be more important than environmental factors. Additionally, state IVF insurance mandates were highly impactful in predicting demand but not the presence of clinics in a county. Predictions of demand were generated for every county in the US, to produce the first ever estimate of ART demand not produced solely from reproductive-age population. Clinic locations were also statistically predicted nationwide, and counties were identified where no clinic operates, but the models forecasted the presence of either a main or satellite clinic. These locations represent the areas where statistical models capturing industry patterns in locating clinics indicate promising locations to operate new clinics. Counties indicated for new clinic locations included less affluent counties near large cities, but also in counties containing large state-run universities.

In Chapter 4, optimization methods from location analysis were applied to overcome the geographic boundary limitations of the models used in Chapter 2 and Chapter 3. The statistical models from Chapter 2 and Chapter 3 were highly focused on individual geographic areas and their characteristics without accounting for neighboring CBSAs or counties. Therefore, in Chapter 4, the ability to travel between counties was accounted for in the MCLP optimization models that maximized coverage within a 60-minute drive of county population centers of two objectives: reproductive-age populations and ART demand predicted from Chapter 3. This approach identified two stable solutions optimal for both objectives in McDonald County, MO and Duplin County, NC and one pseudo-stable solution in Houston/Bibb Counties in GA. Multiobjective optimization was then employed to understand the tradeoff curve between the two objectives, and a balanced solution was identified that allowed for 11k more women to be covered in exchange for a reduction in only 21 ART cycles of demand.

Recommendations for new clinics locations from models from Chapters 2, 3, and 4 were combined in the spirit of George Box’s adage that “all models are wrong, but some are useful” (Box, 1979). Each modeling approach had its own limitations. The CBSA-level models from Chapter 2 could be overly optimistic in their assumption that an ART clinic anywhere in a CBSA translates into geographic access for all residents of a CBSA. Conversely, the county-level models from Chapter 3 could be overly pessimistic with the assumption that people living nearby but ultimately outside a county with an ART clinic do not have geographic access. The optimization models from Chapter 4 used a reasonable travel time assumption of 60-minutes, but this could be too high or too low dependent on what amount of time a decision maker believes is reasonable for most people to travel for ART treatments. Therefore, recommendations from models from each chapter were combined geospatially to identify areas of agreement where hopefully the aggregate recommendation is less wrong and more useful.

This exercise ultimately identified one geographic area with consensus from both the statistical and optimization modeling in Southwest Missouri/Northwest Arkansas. This area is both optimal in its coverage of female reproductive-age population and estimated demand for ART, and it is also an area found from the CBSA statistical modeling. Moreover, Arkansas is an IVF insurance mandate state that was found in Chapter 3 to be a significant predictor of increased ART demand. This concurrence of modeling approaches provides strong support towards investing the millions of dollars needed to establish and operate a new fertility clinic in this area that will likely be economically viable.

The last analysis performed was to apply the MCLP model to optimize coverage of geographically underserved American Indians and Alaska Natives from the opening of five new clinics. Results from this optimization were overlayed geospatially with the findings in Chapter 2 that identified the five best CBSAs based on American Indian and Alaska Native population. Generally, the same geographic areas containing tribal lands in Arizona and New Mexico were identified from both the simple CBSA ranking approach and the MCLP optimization approach, but one location in Lumberton County, NC was identified in the CBSA ranking approach that was not selected in the MCLP optimization. These results provide a compelling rationale for the expansion of geographic access to ART via the Indian Health Service (IHS) facilities on tribal lands rather than simply expanding economic access by deeming ART treatments of infertility to be medically necessary. The long travel times would make completion of an IVF cycle difficult for residents of tribal lands who would still have to travel a few hours to reach the closest ART clinic for regular monitoring.

5.2 Future work

There are promising areas of future work in this domain from expanding the understanding of fertility clinic access to other statistical and optimization models that could be employed to gain additional insights. First, a limitation of the work from Chapter 2 was that all satellite ART clinics were treated equally despite the services provided at each satellite clinic in fact varying widely. Future work should try to better understand the services offered at satellite clinics. Formal surveys of fertility clinics and the location-specific services they provide would allow for a more detailed analysis of geographic access to care based on the various service levels available at satellite clinics.

Next, alternate statistical models could be created to explore the various socioeconomic and environmental health drivers of ART demand and clinic presence patterns. Focusing on the role of socioeconomic status could allow for a better understanding of the threshold requirements for opening and operating ART clinics in the US. Additionally, individual metrics contributing to overall county health rankings were used in the county level statistical modeling in Chapter 3, but the relationship between ART demand and clinic presence was not modeled based on the rolled-up county health ranking. It would be interesting to better understand the relationship between this single health ranking and demand for ART and the presence or lack of

ART clinics. Finally, future studies could model ART demand and clinic presence as a function of political voting behavior and religious affiliation. Though these factors do not have an impact on fecundity, they could be good predictors of ART demand and could be explored.

The location analysis methods employed to maximize coverage should be explored in greater depth. For example, a sensitivity analysis could be performed on the MCLP population and demand models to better understand how the coverage benefit diminishes as more clinics are added. Approaches applying alternate location analysis optimization methods could also be employed to gain additional insights. The MCLP method was chosen for the models in Chapter 4 because it was believed most appropriate for a healthcare application, but industry fertility clinic placement appears to follow Hotelling's Law whereby clinics cluster. Alternate location analysis methods like *p*-median may recommend alternate locations potentially closer to industry practices. A multiobjective approach could be taken to understand how clinics are placed along the Pareto frontier between the best MCLP objective and the best *p*-median objective. Additionally, the hierarchical service location problem (Moore & ReVelle, 1982) could also be utilized to optimize the placement of new main and satellite clinics with different reasonable service driving time for each clinic type.

Finally, additional analysis in this domain would be needed in the future if the federal government allocates funding to increase ART access for Native Americans or members of the military. If congressional appropriation allocates funds to open new ART clinics on tribal lands, the MCLP model could be leveraged to select specific IHS facility locations in which to open new ART clinics and embryology labs to best reach underserved American Indians and Alaska Natives at a higher spatial resolution than the counties used in Chapter 4. Members of the military/TRICARE insurance holders are able to use military-run ART clinics at a lower cost than clinics available to the general population ($4-7k at military clinics vs. $25-30k at private clinics) (Peck, 2019). Also, veterans and service members may qualify for free IVF if their infertility was caused by their military service, contingent on continued congressional funding. However, there are only six US military-run ART clinics worldwide. If funding is allocated to open additional military-run ART clinics, the best locations to place those new clinics will need to be determined.

Finally, a future longitudinal study could be performed to examine if new ART clinics are opened in the future at the recommended locations from this dissertation. Due to the two-year lag in reporting of IVF cycles to the CDC, there are new clinics established yearly that could be opened in areas identified by this work independently or as a result of the publication of this dissertation. Future studies should determine if these predicted locations were economically viable locations to operate and what factors led to the opening of new clinics in certain locations.

References

Advanced Fertility Center of Chicago. (2021). *Infertility Treatment & In Vitro Fertilization - IVF - Insurance Coverage Issues*. https://advancedfertility.com/fertility-treatment/affording-care/fertility-insurance/

American Society for Reproductive Medicine. (2012). *Gestational Carrier (Surrogate)*. https://www.reproductivefacts.org/news-and-publications/patient-fact-sheets-and-booklets/documents/fact-sheets-and-info-booklets/gestational-carrier-surrogate/

American Society for Reproductive Medicine. (2014a). *Can I Freeze My Eggs to Use Later If I'm Not Sick?* https://www.reproductivefacts.org/news-and-publications/patient-fact-sheets-and-booklets/documents/fact-sheets-and-info-booklets/can-i-freeze-my-eggs-to-use-later-if-im-not-sick/

American Society for Reproductive Medicine. (2014b). Fertility Rights and Responsibilities. In *Fact Sheet*. https://www.reproductivefacts.org/news-and-publications/patient-fact-sheets-and-booklets/documents/fact-sheets-and-info-booklets/fertility-rights-and-responsibilities/

American Society for Reproductive Medicine. (2015a). *Female Cancers, Cryopreservation, and Fertility*. https://www.reproductivefacts.org/news-and-publications/patient-fact-sheets-and-booklets/documents/fact-sheets-and-info-booklets/female-cancers-cryopreservation-and-fertility/

American Society for Reproductive Medicine. (2015b). *WHITE PAPER: Access to Care Summit*. https://www.asrm.org/globalassets/asrm/asrm-content/news-and-publications/news-and-research/press-releases-and-bulletins/pdf/atcwhitepaper.pdf

ASRM. (2020). *Utah Coverage Mandate*. ReproductiveFacts.Org. https://www.reproductivefacts.org/resources/state-infertility-insurance-laws/states/utah/

Augood, C., Duckitt, K., & Templeton, A. A. (1998). Smoking and female infertility: a systematic review and meta-analysis. *Human Reproduction (Oxford, England)*, *13*(6), 1532–1539.

Beaulac, J., Kristjansson, E., & Cummins, S. (2009). A systematic review of food deserts, 1966-2007. In *Preventing Chronic Disease* (Vol. 6, Issue 3).

Bennett, V. L., Eaton, D. J., & Church, R. L. (1982). Selecting sites for rural health workers. *Social Science & Medicine*, *16*(1), 63–72. https://doi.org/https://doi.org/10.1016/0277-9536(82)90424-5

Ben-Shlomo, Y. (2005). Real epidemiologists don't do ecological studies? In *International Journal of Epidemiology* (Vol. 34, Issue 6). https://doi.org/10.1093/ije/dyi242

Biggers, J. D. (2012). IVF and embryo transfer: Historical origin and development. In *Reproductive BioMedicine Online* (Vol. 25, Issue 2). https://doi.org/10.1016/j.rbmo.2012.04.011

Bock, D. E., Velleman, P. F., & de Veaux, R. D. (2007). *Stats: Modeling the World*. Pearson/Addison-Wesley. https://books.google.com/books?id=puxUAAAAYAAJ

Bolúmar, F., Olsen, J., Rebagliato, M., Sáez-Lloret, I., & Bisanti, L. (2000). Body mass index and delayed conception: A European multicenter study on infertility and subfecundity. *American Journal of Epidemiology*, *151*(11), 1072–1079. https://doi.org/10.1093/oxfordjournals.aje.a010150

Box, G. E. P. (1979). Robustness in the strategy of scientific model building. In *Robustness in statistics* (pp. 201–236). Elsevier.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C., & Olshen, R. (1984). *Classification and regression trees*. Chapman and Hall/CRC.

Brownlee, J. (2019). *A Tour of Machine Learning Algorithms*. https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/

Buck Louis, G. M., Dmochowski, J., Lynch, C., Kostyniak, P., McGuinness, B. M., & Vena, J. E. (2009). Polychlorinated biphenyl serum concentrations, lifestyle and time-to-pregnancy. *Human Reproduction*, *24*(2), 451–458. https://doi.org/10.1093/humrep/den373

Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data*. Cambridge University Press.

Cao, X.-N., Shen, L.-J., Wu, S., Yan, C., Zhou, Y., Xiong, G., Wang, Y.-C., Liu, Y., Liu, B., Tang, X.-L., Guo, M., Liu, D.-Y., Long, C.-L., Sun, M., He, D.-W., Lin, T., & Wei, G.-H. (2017). Urban fine particulate matter exposure causes male reproductive injury through destroying blood-testis barrier (BTB) integrity. *Toxicology Letters*, *266*, 1–12. https://doi.org/10.1016/j.toxlet.2016.12.004

Carlsen, E., Giwercman, A., Keiding, N., & Skakkebaek, N. E. (1992). Evidence for decreasing quality of semen during past 50 years. *British Medical Journal*, *305*(6854). https://doi.org/10.1136/bmj.305.6854.609

Carré, J., Gatimel, N., Moreau, J., Parinaud, J., & Léandri, R. (2017a). Does air pollution play a role in infertility?: A systematic review. *Environmental Health: A Global Access Science Source*, *16*(1), 1–16. https://doi.org/10.1186/s12940-017-0291-8

Carré, J., Gatimel, N., Moreau, J., Parinaud, J., & Léandri, R. (2017b). Influence of air quality on the results of in vitro fertilization attempts: A retrospective study. *European Journal of Obstetrics and Gynecology and Reproductive Biology*, *210*, 116–122. https://doi.org/10.1016/j.ejogrb.2016.12.012

Casey, J. A., Gemmill, A., Karasek, D., Ogburn, E. L., Goin, D. E., & Morello-Frosch, R. (2018). Increase in fertility following coal and oil power plant retirements in California. *Environmental Health: A Global Access Science Source*, *17*(1), 1–10. https://doi.org/10.1186/s12940-018-0388-8

CDC. (2020a). *2018 ART Success Rates*. https://www.cdc.gov/art/artdata/index.html

CDC. (2020b). *2018 Assisted Reproductive Technology Fertility Clinic Success Rates Report*. https://www.cdc.gov/art/reports/2018/fertility-clinic.html

CDC. (2021). *Infertility FAQs*. https://www.cdc.gov/reproductivehealth/infertility/index.htm

CDC. (2022). *STDs & Infertility*. https://www.cdc.gov/std/infertility/default.htm

Cha, A. E. (2018). How religion is coming to terms with modern fertility methods. *The Washington Post*. https://www.washingtonpost.com/graphics/2018/national/how-religion-is-coming-to-terms-with-modern-fertility-methods/

Chambers, G. M., Sullivan, E. A., Ishihara, O., Chapman, M. G., & Adamson, G. D. (2009). The economic impact of assisted reproductive technology: a review of selected developed countries. *Fertility and Sterility*, *91*(6), 2281–2294. https://doi.org/10.1016/j.fertnstert.2009.04.029

Chavarro, J. E., Rich-Edwards, J. W., Rosner, B. A., & Willett, W. C. (2007). Diet and lifestyle in the prevention of ovulatory disorder infertility. *Obstetrics and Gynecology*, *110*(5), 1050–1058. https://doi.org/10.1097/01.AOG.0000287293.25465.e1

Chirgwin, K. D., Feldman, J., Muneyyirci-Delale, O., Landesman, S., & Minkoff, H. (1996). Menstrual Function in Human Immunodeficiency Virus-Infected Women Without Acquired Immunodeficiency Syndrome. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *12*(5). https://journals.lww.com/jaids/Fulltext/1996/08150/Menstrual_Function_in_Human_Immunodeficiency.8.aspx

Choe, S. A., Jun, Y. B., Lee, W. S., Yoon, T. K., & Kim, S. Y. (2018). Association between ambient air pollution and pregnancy rate in women who underwent IVF. *Human Reproduction*, *33*(6), 1071–1078. https://doi.org/10.1093/humrep/dey076

Church, R., & ReVelle, C. (1974). The maximal covering location problem. *Papers of the Regional Science Association*, *32*(1), 101–118. https://doi.org/10.1007/BF01942293

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1). https://doi.org/10.1177/001316446002000104

Cohon, J. L. (2004). *Multiobjective programming and planning* (Vol. 140). Courier Corporation.

Cole, D. C., Wainman, B., Sanin, L. H., Weber, J. P., Muggah, H., & Ibrahim, S. (2006). Environmental contaminant levels and fecundability among non-smoking couples. *Reproductive Toxicology*, *22*(1), 13–19. https://doi.org/10.1016/j.reprotox.2005.12.001

Collins, J. (2002). An international survey of the health economics of IVF and ICSI. *Human Reproduction Update*, *8*(3), 265–277. https://doi.org/10.1093/humupd/8.3.265

Collins, J., Evers, H., Golombok, S., Hannaford, P., Jacobs, H. S., la Vecchia, C., Macunovich, D. J., Crosignani, P. G., Devroey, P., Diedrich, K., Farley, T., Gianaroli, L., Liebaers, I., Persson, J., Quartarolo, J. P., Ragni, G., Tarlatzis, B., van Steirteghem, A., Diczfalusy, E., & Rubin, B. L. (2001). Social determinants of human reproduction. *Human Reproduction*, *16*(7), 1518–1526. https://doi.org/10.1093/humrep/16.7.1518

Comptroller General of the United States. (1976). *Investigation of Allegations Concerning Indian Health Service*. https://www.gao.gov/assets/hrd-77-3.pdf

County Health Rankings & Roadmaps. (2021a). *2021 County Health Rankings National Data*. https://www.countyhealthrankings.org/sites/default/files/media/document/2021 County Health Rankings Data - v1.xlsx

County Health Rankings & Roadmaps. (2021b). *2021 Measures*. https://www.countyhealthrankings.org/explore-health-rankings/measures-data-sources/2021-measures

Cousineau, T. M., & Domar, A. D. (2007). Psychological impact of infertility. *Best Practice and Research: Clinical Obstetrics and Gynaecology*, *21*(2), 293–308. https://doi.org/10.1016/j.bpobgyn.2006.12.003

Crawford, S., Boulet, S. L., Jamieson, D. J., Stone, C., Mullen, J., & Kissin, D. M. (2016). Assisted reproductive technology use, embryo transfer practices, and birth outcomes after infertility insurance mandates: New Jersey and Connecticut. *Fertility and Sterility*, *105*(2), 347–355. https://doi.org/10.1016/j.fertnstert.2015.10.009

Current, J., Min, H., & Schilling, D. (1990). Multiobjective analysis of facility location decisions. *European Journal of Operational Research*, *49*(3). https://doi.org/10.1016/0377-2217(90)90401-V

Dieke, A. C., Zhang, Y., Kissin, D. M., Barfield, W. D., & Boulet, S. L. (2017). Disparities in assisted reproductive technology utilization by race and ethnicity, United States, 2014: a commentary. *Journal of Women's Health*, *26*(6), 605–608. https://doi.org/10.1089/jwh.2017.6467

Diprose, W. K., Buist, N., Hua, N., Thurier, Q., Shand, G., & Robinson, R. (2020). Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator. *Journal of the American Medical Informatics Association*, *27*(4). https://doi.org/10.1093/jamia/ocz229

Donabedian, A. (1973). *Aspects of medical care administration: specifying requirements for health care*. Harvard University Press.

Dunson, D. B., Colombo, B., & Baird, D. D. (2002). Changes with age in the level and duration of fertility in the menstrual cycle. *Human Reproduction*, *17*(5), 1399–1403. https://doi.org/10.1093/humrep/17.5.1399

Dutney, A. (2007). Religion, infertility and assisted reproductive technology. In *Best Practice and Research: Clinical Obstetrics and Gynaecology* (Vol. 21, Issue 1). https://doi.org/10.1016/j.bpobgyn.2006.09.007

Eaton, D. J., Daskin, M. S., Simmons, D., Bulloch, B., & Jansma, G. (1985). Determining Emergency Medical Service Vehicle Deployment in Austin, Texas. *Interfaces*, *15*(1), 96–108. https://doi.org/10.1287/inte.15.1.96

Edwards, R. G., Steptoe, P. C., & Purdy, J. M. (1980). Establishing full-term human pregnancies using cleaving embryos grown in vitro. *BJOG: An International Journal of Obstetrics & Gynaecology*, *87*(9), 737–756. https://doi.org/10.1111/j.1471-0528.1980.tb04610.x

Ehrlich, S., Williams, P. L., Missmer, S. A., Flaws, J. A., Ye, X., Calafat, A. M., Petrozza, J. C., Wright, D., & Hauser, R. (2012). Urinary bisphenol A concentrations and early reproductive health outcomes among women undergoing IVF. *Human Reproduction*, *27*(12). https://doi.org/10.1093/humrep/des328

Embryo Options. (2022). *Disposition Options*. https://embryooptions.com/disposition-options.html

Eskenazi, B., Warner, M., Marks, A. R., Samuels, S., Needham, L., Brambilla, P., & Mocarelli, P. (2010). Serum dioxin concentrations and time to pregnancy. *Epidemiology*, *21*(2), 224–231. https://doi.org/10.1097/EDE.0b013e3181cb8b95

Ethics Committee of the American Society for Reproductive Medicine. (2015). Human immunodeficiency virus (HIV) and infertility treatment: a committee opinion. *Fertility and Sterility*, *104*(1), e1–e8. https://doi.org/10.1016/j.fertnstert.2015.04.004

Fix, E., & Hodges, J. L. (1989). Discriminatory analysis. Nonparametric discrimination: Consistency properties. *International Statistical Review/Revue Internationale de Statistique*, *57*(3), 238–247. https://doi.org/1403797

Frank, L. D., Andresen, M. A., & Schmid, T. L. (2004). Obesity relationships with community design, physical activity, and time spent in cars. *American Journal of Preventive Medicine*, *27*(2), 87–96. https://doi.org/10.1016/j.amepre.2004.04.011

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189–1232. https://www.jstor.org/stable/2699986

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, *38*(4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2

Fronstin, P. (2012). Self-insured health plans: state variation and recent trends by firm size. In *EBRI Notes* (Vol. 33, Issue 11). https://www.shrm.org/ResourcesAndTools/hr-topics/benefits/Documents/EBRI_Notes_11_Nov-12.Slf-Insrd1.pdf

Gambone, J. C., Segars, J. H., Cedars, M., & Schlaff, W. D. (2015). Fellowship training and board certification in reproductive endocrinology and infertility. *Fertility and Sterility*, *104*(1), 3–7. https://doi.org/https://doi.org/10.1016/j.fertnstert.2015.04.039

Gaskins, A. J., Hart, J. E., Mínguez-Alarcón, L., Chavarro, J. E., Laden, F., Coull, B. A., Ford, J. B., Souter, I., & Hauser, R. (2018). Residential proximity to major roadways and traffic in relation to outcomes of in vitro fertilization. *Environment International*, *115*(February), 239–246. https://doi.org/10.1016/j.envint.2018.03.029

Gesink Law, D. C., Klebanoff, M. A., Brock, J. W., Dunson, D. B., & Longnecker, M. P. (2005). Maternal serum levels of polychlorinated biphenyls and 1,1-dichloro-2,2- bis(p-chlorophenyl)ethylene (DDE) and time to pregnancy. *American Journal of Epidemiology*, *162*(6), 523–532. https://doi.org/10.1093/aje/kwi240

Gharani, P., Stewart, K., & Ryan, G. L. (2015). An enhanced approach for modeling spatial accessibility for in vitro fertilization services in the rural Midwestern United States. *Applied Geography*, *64*(October), 12–23. https://doi.org/10.1016/j.apgeog.2015.08.005

Government Studies & Systems, inc. (1980). *Methods to Determine Geographic/population Boundaries for Specific Health Services: An Annotated Bibliography*. Department of Health and Human Services, Public Health Service, Health Resources Administration, Bureau of Health Planning. https://books.google.com/books?id=yGTNMiUzwdwC

Grady, R., Alavi, N., Vale, R., Khandwala, M., & McDonald, S. D. (2012). Elective single embryo transfer and perinatal outcomes: a systematic review and meta-analysis. *Fertility and Sterility*, *97*(2), 324–331. https://doi.org/10.1016/j.fertnstert.2011.11.033

Green, L. C. (1975). North America's Indians and the Trusteeship Concept. *Anglo-American Law Review*, *4*(2), 137–162. https://doi.org/10.1177/147377957500400202

Griffin, P. M., Scherrer, C. R., & Swann, J. L. (2008). Optimization of community health center locations and service offerings with statistical need estimation. *IIE Transactions*, *40*(9), 880–892. https://doi.org/10.1080/07408170802165864

Hakimi, S. L. (1964). Optimum Locations of Switching Centers and the Absolute Centers and Medians of a Graph. *Operations Research*, *12*(3), 450–459. https://www.jstor.org/stable/168125

Hamilton, B. H., & McManus, B. (2012). The effects of insurance mandates on choices and outcomes in infertility treatment markets. *Health Economics*, *21*(8), 994–1016.

Hammoud, A. O., Gibson, M., Stanford, J., White, G., Carrell, D. T., & Peterson, M. (2009). In vitro fertilization availability and utilization in the United States: a study of demographic, social, and economic factors. *Fertility and Sterility*, *91*(5), 1630–1635. https://doi.org/https://doi.org/10.1016/j.fertnstert.2007.10.038

Hanke, S. (2019). Elizabeth Warren's Plan to Keep American Indians Wards of the State. *Forbes*. https://www.forbes.com/sites/stevehanke/2019/09/15/elizabeth-warrens-plan-to-keep-american-indians-wards-of-the-state/?sh=736b324964e6

Harley, K. G., Marks, A. R., Chevrier, J., Bradman, A., Sjödin, A., & Eskenazi, B. (2010). PBDE concentrations in women's serum and fecundability. *Environmental Health Perspectives*, *118*(5), 699–704. https://doi.org/10.1289/ehp.0901450

Harris, J. A., Menke, M. N., Haefner, J. K., Moniz, M. H., & Perumalswami, C. R. (2017). Geographic access to assisted reproductive technology health care in the United States: a population-based cross-sectional study. *Fertility and Sterility*, *107*(4), 1023–1027. https://doi.org/10.1016/j.fertnstert.2017.02.101

Hauser, R. (2006). The environment and male fertility: Recent research on emerging chemicals and semen quality. *Seminars in Reproductive Medicine*, *24*(3), 156–167. https://doi.org/10.1055/s-2006-944422

Hauser, R., & Sokol, R. (2008). Science linking environmental contaminant exposures with fertility and reproductive health impacts in the adult male. *Fertility and Sterility*, *89*(2 SUPPL.), 59–65. https://doi.org/10.1016/j.fertnstert.2007.12.033

Hayford, S. R., & Morgan, S. P. (2008). Religiosity and fertility in the United States: The role of fertility intentions. *Social Forces*, *86*(3). https://doi.org/10.1353/sof.0.0000

Henne, M. B., & Bundorf, M. K. (2008). Insurance mandates and trends in infertility treatments. *Fertility and Sterility*, *89*(1), 66–73. https://doi.org/https://doi.org/10.1016/j.fertnstert.2007.01.167

Hobbs, B. F. (1980). A comparison of weighting methods in power plant siting. *Decision Sciences*, *11*(4), 725–737.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, *39*(153), 41–57. https://doi.org/10.2307/2224214

Idrovo, A. J. (2011). Three criteria for ecological fallacy. In *Environmental Health Perspectives* (Vol. 119, Issue 8). https://doi.org/10.1289/ehp.1103768

Indian Health Service. (2018). *Can PRC pay for your referral medical care? Find out in 3 stages.* https://www.ihs.gov/sites/prc/themes/responsive2017/display_objects/documents/PRC-ProcessHandout.pdf

Indian Health Service. (2021a). *Agency Overview*. https://www.ihs.gov/aboutihs/overview/

Indian Health Service. (2021b). *Infertility*. https://www.ihs.gov/womenshealth/reproductivehealth/infertility/

Institute of Medicine. (1993). *Access to health care in America*. National Academies Press. https://doi.org/10.17226/2009

Inter-American Court of Human Rights. (2012). *Inter-American Court of Human Rights Case of Artavia Murillo et al. ("In Vitro Fertilization") v. Costa Rica Judgement of November 28, 2012*. http://hrlibrary.umn.edu/iachr/C/257-ing.html

IRS. (2021). *Opportunity Zones*. https://www.irs.gov/credits-deductions/businesses/opportunity-zones

Jain, T. (2006). Socioeconomic and racial disparities among infertility patients seeking care. *Fertility and Sterility*, *85*(4), 876–881. https://doi.org/10.1016/j.fertnstert.2005.07.1338

Jain, T., & Hornstein, M. D. (2005). Disparities in access to infertility services in a state with mandated insurance coverage. *Fertility and Sterility*, *84*(1), 221–223. https://doi.org/10.1016/j.fertnstert.2005.01.118

Janitz, A. E., Peck, J. D., & Craig, L. B. (2019). Racial/ethnic differences in the utilization of infertility services: a focus on American Indian/Alaska natives. *Maternal and Child Health Journal*, *23*(1), 10–18. https://doi.org/10.1007/s10995-018-2586-y

Jarvis, E. (1852). On the supposed increase of insanity. *American Journal of Psychiatry*, *8*(4), 333–364. https://doi.org/10.1176/ajp.8.4.333

Jasanoff, S., & Metzler, I. (2018). Borderlands of Life: IVF Embryos and the Law in the United States, United Kingdom, and Germany. *Science, Technology, & Human Values*, *45*(6), 1001–1037. https://doi.org/10.1177/0162243917753990

Katerine Keisler-Starkey, & Lisa Bunch. (2021). *Health Insurance Coverage in the United States: 2020*. https://www.census.gov/library/publications/2021/demo/p60-274.html

Katz, P., Showstack, J., Smith, J. F., Nachtigall, R. D., Millstein, S. G., Wing, H., Eisenberg, M. L., Pasch, L. A., Croughan, M. S., & Adler, N. (2011). Costs of infertility treatment: results from an 18-month prospective cohort study. *Fertility and Sterility*, *95*(3), 915–921. https://doi.org/10.1016/j.fertnstert.2010.11.026

Kearney, M. S., Levine, P. B., & Pardue, L. W. (2021). *The Puzzle of Falling US Birth Rates Since the Great Recession*. National Bureau of Economic Research. https://doi.org/10.3386/w29286

Kindig, D. A., & Movassaghi, H. (1989). The adequacy of physician supply in small rural counties. *Health Affairs*, *8*(2). https://doi.org/10.1377/hlthaff.8.2.63

Klein, J., & Sauer, M. v. (2001). Assessing fertility in women of advanced reproductive age. *American Journal of Obstetrics and Gynecology*, *185*(3). https://doi.org/10.1067/mob.2001.114689

Klemetti, R., Raitanen, J., Sihvo, S., Saarni, S., & Koponen, P. (2010). Infertility, mental disorders and well-being–a nationwide survey. *Acta Obstetricia et Gynecologica Scandinavica*, *89*(5), 677–682. https://doi.org/10.3109/00016341003623746

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). Springer. https://link.springer.com/book/10.1007/978-1-4614-6849-3

Kuhn, M., Weston, S., Keefer, C., & Coulter, N. (2012). Cubist models for regression. *R Package Vignette R Package Version 0.0*, *18*. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.398.3360&rep=rep1&type=pdf

Kushnir, V. A., Barad, D. H., Albertini, D. F., Darmon, S. K., & Gleicher, N. (2017). Systematic review of worldwide trends in assisted reproductive technology 2004-2013. *Reproductive Biology and Endocrinology : RB&E*, *15*(1), 6. https://doi.org/10.1186/s12958-016-0225-2

Lawrence, J. (2000). The Indian Health Service and the Sterilization of Native American Women. *American Indian Quarterly*, *24*(3), 400–419. https://doi.org/10.1353/aiq.2000.0008

Levin, K. A. (2006). Study Design VI - Ecological Studies. *Evidence-Based Dentistry*, *7*(4), 108. https://doi.org/10.1038/sj.ebd.6400454

Lintsen, A. M. E., Pasker-de Jong, P. C. M., de Boer, E. J., Burger, C. W., Jansen, C. A. M., Braat, D. D. M., & van Leeuwen, F. E. (2005). Effects of subfertility cause, smoking and body weight on the success rate of IVF. *Human Reproduction*, *20*(7), 1867–1875. https://doi.org/10.1093/humrep/deh898

Louis, G. M. B., & Platt, R. W. (2011). *Reproductive and perinatal epidemiology*. Oxford University Press.

Lunenfeld, B., van Steirteghem, A., & participants, on behalf of all. (2004). Infertility in the third millennium: implications for the individual, family and society: Condensed Meeting Report from the Bertarelli Foundation's Second Global Conference*. *Human Reproduction Update*, *10*(4), 317–326. https://doi.org/10.1093/humupd/dmh028

Maheshwari, A., Griffiths, S., & Bhattacharya, S. (2011). Global variations in the uptake of single embryo transfer. *Human Reproduction Update*, *17*(1), 107–120.

Mallampati, D., Simon, M. A., & Janiak, E. (2017). Evolving state-based contraceptive and abortion policies. In *JAMA - Journal of the American Medical Association* (Vol. 317, Issue 24). https://doi.org/10.1001/jama.2017.6103

Martin, J. A., Hamilton, B. E., Osterman, M. J. K., Driscoll, A. K., & Mathews, T. J. (2017). *Births: final data for 2015*.

McGarity, M. Z. (2021). *Fertility Clinic Finder*. ArcGIS Online. https://gisanddata.maps.arcgis.com/apps/webappviewer/index.html?id=3e073183ee824e2dae5900da1774ba4b

McGarity, M. Z., Herndon, C. N., Harris, J. A., & Hobbs, B. F. (2021). *Data and code associated with publication: Impact of satellite clinics on geographic access to assisted reproductive technology services in the United States* (p. V1). https://doi.org/10.7281/T1/QXKY4B

McGarity, M. Z., & Hobbs, B. F. (2021). *Fertility clinic locations and details in the United States* (p. V1). Johns Hopkins University Data Archive. https://doi.org/10.7281/T1/XZ0CGJ

Meeker, J. D., Ehrlich, S., Toth, T. L., Wright, D. L., Calafat, A. M., Trisini, A. T., Ye, X., & Hauser, R. (2010). Semen quality and sperm DNA damage in relation to urinary bisphenol A among men from an infertility clinic. *Reproductive Toxicology*, *30*(4). https://doi.org/10.1016/j.reprotox.2010.07.005

Mendola, P., Sundaram, R., Louis, G. M. B., Sun, L., Wallace, M. E., Smarr, M. M., Sherman, S., Zhu, Y., Ying, Q., & Liu, D. (2017). Proximity to major roadways and prospectively-measured time-to-pregnancy and infertility. *Science of the Total Environment*, *576*, 172–177. https://doi.org/10.1016/j.scitotenv.2016.10.038

Menken, J., Trussell, J., & Larsen, U. (1986). Age and infertility. *Science*, *233*(4771), 1389–1394. https://doi.org/10.1126/science.3755843

Mitchell, C. G. (2022). *Re: Final Determination Letter FOIA Case 22-027*. Indian Health Service.

Moore, G. C., & ReVelle, C. (1982). The hierarchical service location problem. *Management Science*, *28*(7), 775–780.

Mumford, S. L., Flannagan, K. S., Radoc, J. G., Sjaarda, L. A., Zolton, J. R., Metz, T. D., Plowden, T. C., Perkins, N. J., DeVilbiss, E. A., Andriessen, V. C., A C, P. S., Kim, K., Yisahak, S. F., Freeman, J. R., Alkhalaf, Z., Silver, R. M., & Schisterman, E. F. (2021). Cannabis use while trying to conceive: a prospective cohort study evaluating associations with fecundability, live birth and pregnancy loss. *Human Reproduction (Oxford, England)*, *36*(5). https://doi.org/10.1093/humrep/deaa355

Nachtigall, R. D. (2006). International disparities in access to infertility services. *Fertility and Sterility*, *85*(4), 871–875. https://doi.org/10.1016/j.fertnstert.2005.08.066

Nangia, A. K., Likosky, D. S., & Wang, D. (2010a). Access to assisted reproductive technology centers in the United States. *Fertility and Sterility*, *93*(3), 745–761. https://doi.org/https://doi.org/10.1016/j.fertnstert.2008.10.037

Nangia, A. K., Likosky, D. S., & Wang, D. (2010b). Distribution of male infertility specialists in relation to the male population and assisted reproductive technology centers in the United States. *Fertility and Sterility*, *94*(2), 599–609. https://doi.org/10.1016/j.fertnstert.2009.02.012

National Center for Health Statistics. (2020). *About the National Survey of Family Growth*. https://www.cdc.gov/nchs/nsfg/about_nsfg.htm

National Survey of Family Growth. (2021). *2015-2019 NSFG Key Statistics*. https://www.cdc.gov/nchs/nsfg/keystatistics.htm

Newhouse, J. P. (1990). Geographic access to physician services. *Annual Review of Public Health*, *11*(1), 207–230.

NHSC. (2021). *How to Meet NHSC Site Eligibility Requirements*. https://nhsc.hrsa.gov/sites/eligibility-requirements

Noble, W. S. (2006). What is a support vector machine? In *Nature Biotechnology* (Vol. 24, Issue 12). https://doi.org/10.1038/nbt1206-1565

NYU Langone Health. (2021). *Egg Freezing & Embryo Banking*. https://nyulangone.org/locations/fertility-center/in-vitro-fertilization-egg-freezing-embryo-banking/egg-freezing-embryo-banking

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *New England Journal of Medicine*, *375*(13). https://doi.org/10.1056/nejmp1606181

Odisho, A. Y., Nangia, A. K., Katz, P. P., & Smith, J. F. (2014). Temporal and geospatial trends in male factor infertility with assisted reproductive technology in the United States from 1999-2010. *Fertility and Sterility*, *102*(2), 469–475. https://doi.org/10.1016/j.fertnstert.2014.05.006

Office of Disease Prevention and Health Promotion. (2022). *Access to Health Services*. https://www.healthypeople.gov/2020/leading-health-indicators/2020-lhi-topics/Access-to-Health-Services

Ogliari, K. S., de Faria Coimbra Lichtenfels, A. J., de Marchi, M. R. R., Ferreira, A. T., Dolhnikoff, M., & Saldiva, P. H. N. (2013). Intrauterine exposure to diesel exhaust diminishes adult ovarian reserve. *Fertility and Sterility*, *99*(6), 1681-1688.e2. https://doi.org/10.1016/j.fertnstert.2013.01.103

Patrizio, P., Albertini, D. F., Gleicher, N., & Caplan, A. (2022). The changing world of IVF: the pros and cons of new business models offering assisted reproductive technologies. *Journal of Assisted Reproduction and Genetics*, 1–9.

Peck, A. (2019, August 22). *IVF helps military couples grow families*. Military Families Magazine. https://militaryfamilies.com/military-health/ivf-helps-military-couples-grow-families/

Penchansky, R., & Thomas, J. W. (1981). The concept of access: Definition and relationship to consumer satisfaction. *Medical Care*, *19*(2). https://doi.org/10.1097/00005650-198102000-00001

Penn State Health. (2022). *Reproductive Endocrinology and Infertility*. https://www.pennstatehealth.org/services-treatments/ob-gyn-womens-health/reproductive-endocrinology-infertility

Peredo, A. M., & McLean, M. (2006). Social entrepreneurship: A critical review of the concept. *Journal of World Business*, *41*(1). https://doi.org/10.1016/j.jwb.2005.10.007

Peretz, J., Vrooman, L., Ricke, W. A., Hunt, P. A., Ehrlich, S., Hauser, R., Padmanabhan, V., Taylor, H. S., Swan, S. H., Vandevoort, C. A., & Flaws, J. A. (2014). Bisphenol A and reproductive health: Update of experimental and human evidence, 2007-2013. In *Environmental Health Perspectives* (Vol. 122, Issue 8). https://doi.org/10.1289/ehp.1307728

Pew Research Center. (2013). *Abortion Viewed in Moral Terms: Fewer See Stem Cell Research and IVF as Moral Issues*. https://www.pewforum.org/wp-content/uploads/sites/7/2013/08/Morality-of-abortion-8-15-for-pdf.pdf

Pham, H., Odisho, A., Nangia, A. K., Sandlow, J., Herndon, C., & Smith, J. F. (2014). Male infertility and access to assisted reproductive technology (ART) in the USA: a mapping approach. *Fertility and Sterility*, *102*(3, Supplement), e345–e346. https://doi.org/http://dx.doi.org/10.1016/j.fertnstert.2014.07.1167

Pickett, K. E., & Wilkinson, R. G. (2007). Child wellbeing and income inequality in rich societies: Ecological cross sectional study. *British Medical Journal*, *335*(7629). https://doi.org/10.1136/bmj.39377.580162.55

Practice Committee of the American Society for Reproductive, & Technology, P. C. of the S. for A. R. M. (2013). Criteria for number of embryos to transfer: a committee opinion. *Fertility and Sterility*, *99*(1), 44–46. https://doi.org/10.1016/j.fertnstert.2012.09.038

Quinlan, J. R. (1992). Learning with continuous classes. *5th Australian Joint Conference on Artificial Intelligence*, *92*, 343–348. https://doi.org/10.1142/9789814536271

Quinlan, J. R. (1993). *C4.5: programs for machine learning*. Morgan Kaufmann Publishers. http://www.rulequest.com/see5-unix.html

Ramlau-Hansen, C. H., Thulstrup, A. M., Nohr, E. A., Bonde, J. P., Sørensen, T. I. A., & Olsen, J. (2007). Subfecundity in overweight and obese couples. *Human Reproduction*, *22*(6), 1634–1637. https://doi.org/10.1093/humrep/dem035

Raphael, S., & Stoll, M. A. (2001). Can Boosting Minority Car-Ownership Rates Narrow Inter-Racial Employment Gaps? *Brookings-Wharton Papers on Urban Affairs*, *2001*(1). https://doi.org/10.1353/urb.2001.0013

RDocumentation. (2022a). *findCorrelation: Determine highly correlated variables*. Caret Package. https://www.rdocumentation.org/packages/caret/versions/6.0-90/topics/findCorrelation

RDocumentation. (2022b). *preProcess: Pre-Processing of Predictors*. Caret Package. https://www.rdocumentation.org/packages/caret/versions/6.0-90/topics/preProcess

Redman, E. (1973). *The Dance of Legislation: An Insider's Account of the Workings of the United States Senate*. University of Washington Press. https://books.google.com/books?id=bbCcyVFuO8QC

RESOLVE. (2018). *INFERTILITY COVERAGE BY STATE (2018)*. https://web.archive.org/web/20181231035253/https://resolve.org/what-are-my-options/insurance-coverage/infertility-coverage-state/

RESOLVE. (2021a). *EMPLOYEE SEEKING FERTILITY INSURANCE COVERAGE*. https://resolve.org/what-are-my-options/insurance-coverage/coverage-at-work/employee-seeking-fertility-insurance-coverage/

RESOLVE. (2021b). *HEALTH INSURANCE 101*. https://resolve.org/what-are-my-options/insurance-coverage/health-insurance-101/

RESOLVE. (2021c). *INFERTILITY COVERAGE BY STATE (2021)*. https://web.archive.org/web/20211120032317/https://resolve.org/what-are-my-options/insurance-coverage/infertility-coverage-state/

RESOLVE. (2022). *Military Personnel and Infertility*. https://resolve.org/what-are-my-options/military-personnel-infertility/

ReVelle, C., Bigman, D., Schilling, D., Cohon, J., & Church, R. (1977). Facility location: a review of context-free and EMS models. *Health Services Research*, *12*(2), 129–146. https://pubmed.ncbi.nlm.nih.gov/885725

Revelle, C., & Eiselt, H. (2005). Location analysis: A synthesis and survey. *European Journal of Operations Research*, *165*, 1–19. https://doi.org/10.1016/j.ejor.2003.11.032

Revelle, C., Marks, D., & Liebman, J. C. (1970). An Analysis of Private and Public Sector Location Models. *Management Science*, *16*(11), 692–707. http://www.jstor.org/stable/2628903

ReVelle, C., & McGarity, A. E. (1997). *Design and operation of civil and environmental engineering systems*. John Wiley & Sons.

ReVelle, C., & Swain, R. (1970). Central Facilities Location. *Geographical Analysis*, *2*(1), 30–42. https://doi.org/10.1111/j.1538-4632.1970.tb00142.x

Rissman, A. R., Lozier, L., Comendant, T., Kareiva, P., Kiesecker, J. M., Shaw, M. R., & Merenlender, A. M. (2007). Conservation easements: Biodiversity protection and private use. *Conservation Biology*, *21*(3). https://doi.org/10.1111/j.1523-1739.2007.00660.x

Rosenthal, M. B., Zaslavsky, A., & Newhouse, J. P. (2005). The geographic distribution of physicians revisited. *Health Services Research*, *40*(6p1), 1931–1952. https://doi.org/10.1111/j.1475-6773.2005.00440.x

Rosenthal, M. M., & Frederick, D. (1984). Physician maldistribution in cross-cultural perspective: United States, United Kingdom, and Sweden. *Inquiry*, *21*(1). https://www.jstor.org/stable/29771614

Schettler, T., Solomon, G., Valenti, M., & Huddle, A. (2000). *Generations at Risk*. The MIT Press.

Schilling, D. A. (1976). Multiobjective and Temporal Considerations in Public Facility Location. *ProQuest Dissertations and Theses*, 127. https://search.proquest.com/docview/302814576?accountid=188395

Schilling, D. A., McGarity, A., & ReVelle, C. (1982). Hidden attributes and the display of information in multiobjective analysis. *Management Science*, *28*(3), 236–242.

Schnabel, L. (2021). Secularism and Fertility Worldwide. *Socius*, *7*. https://doi.org/10.1177/23780231211031320

Schwartz, D., Mayaux, M. J., Spira, A., Moscato, M. L., Jouannet, P., Czyglik, F., & David, G. (1983). Semen characteristics as a function of age in 833 fertile men. *Fertility and Sterility*, *39*(4), 530–535. https://doi.org/10.1016/S0015-0282(16)46946-3

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human Computer Studies*, *146*. https://doi.org/10.1016/j.ijhcs.2020.102551

SHRM. (2015). *2015 Employee Benefits A Research Report by the Society for Human Resource Management (SHRM)*. https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/Documents/2015-Employee-Benefits.pdf

Smith, G. A. (2021). *About Three-in-Ten U.S. Adults Are Now Religiously Unaffiliated*. https://www.pewforum.org/2021/12/14/about-three-in-ten-u-s-adults-are-now-religiously-unaffiliated/

Society for Assisted Reproductive Technology. (2021). *Find an IVF Clinic*. https://www.sartcorsonline.com/members/Search

Stentz, N. C., Griffith, K. A., Perkins, E., Jones, R. D., & Jagsi, R. (2016). Fertility and Childbearing Among American Female Physicians. *Journal of Women's Health (2002)*, *25*(10), 1059–1065. https://doi.org/10.1089/jwh.2015.5638

Stephen, E. H., & Chandra, A. (2006). Declining estimates of infertility in the United States: 1982-2002. *Fertility and Sterility*, *86*(3). https://doi.org/10.1016/j.fertnstert.2006.02.129

Swan, S. H., Brazil, C., Drobnis, E. Z., Liu, F., Kruse, R. L., Hatch, M., Redmond, J. B., Wang, C., Overstreet, J. W., Carter, B. S., Kelly, D. J., Stewart, S. L., Simmons, T. M., Treece, C., Swerdloff, R. S., Lumbreras, L., Villanueva, S., Diaz-Romero, M., Victoroff, A., … DeJonge, C.

(2003). Geographic differences in semen quality of fertile U.S. males. *Environmental Health Perspectives*, *111*(4), 414–420. https://doi.org/10.1289/ehp.5927

Therneau, T. M., & Atkinson, E. J. (1997). *An introduction to recursive partitioning using the RPART routines*. Technical report Mayo Foundation. https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf

Thornton, R. (1984). Cherokee population losses during the Trail of Tears: A new perspective and a new estimate. *Ethnohistory*, 289–300. https://doi.org/482714

Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1970). The Location of Emergency Service Facilities. *Operations Research*, *19*(6), 1363–1373. https://doi.org/10.1287/opre.19.6.1363

Tu, J. v., & Ko, D. T. (2008). Ecological studies and cardiovascular outcomes research. In *Circulation* (Vol. 118, Issue 24). https://doi.org/10.1161/CIRCULATIONAHA.107.729806

United Nations. (1948). *Universal Declaration of Human Rights*. https://www.un.org/en/about-us/universal-declaration-of-human-rights

U.S. Census Bureau. (2010). *Centers of Population*. https://www.census.gov/geographies/reference-files/time-series/geo/centers-population.2010.html

U.S. Census Bureau. (2019a). *Age and Sex, 2015-2019 American Community Survey (ACS) - All Metropolitan Statistical Areas/Micropolitan Statistical Areas within United States*. https://data.census.gov/cedsci/table?q=United States&g=0100000US.310000&tid=ACSST5Y2019.S0101&hidePreview=true

U.S. Census Bureau. (2019b). *Age and Sex in CBSAs*. https://data.census.gov/cedsci/table?q=age%20and%20sex&g=0100000US%243100000&tid=ACSST1Y2019.S0101

U.S. Census Bureau. (2019c). *Hispanic or Latino Origin in CBSAs* (p. ACS 5-Year Estimates). https://data.census.gov/cedsci/table?q=hispanic&g=0100000US%243100000&tid=ACSDT1Y2019.B03003

U.S. Census Bureau. (2019d). *Hispanic or Latino Origin in Urban Areas* (p. ACS 5-Year Estimates). https://data.census.gov/cedsci/table?q=hispanic%20&g=0100000US%244000000&tid=ACSDT5Y2019.B03003

U.S. Census Bureau. (2019e). *Historical Income Tables: Households*. https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html

U.S. Census Bureau. (2019f). *Race in CBSAs* (p. ACS 5-Year Estimates). https://data.census.gov/cedsci/table?q=United%20States&g=0100000US%243100000&tid=ACSDT5Y2019.B02001

U.S. Census Bureau. (2019g). *Race in Urban Areas*. ACS-5-Year Estimates. https://data.census.gov/cedsci/table?q=United%20States&g=0100000US%244000000&tid=ACSDT5Y2019.B02001

U.S. Census Bureau. (2021). *Core-Based Statistical Areas*. https://www.census.gov/topics/housing/housing-patterns/about/core-based-statistical-areas.html

U.S. Centers for Medicare & Medicaid Services. (2016). *10 Important Facts about Indian Health Service and Health Insurance*. https://www.cms.gov/Outreach-and-Education/American-Indian-Alaska-Native/AIAN/Downloads/10-Important-Facts-About-IHS-and-Health-Care-.pdf

U.S. Conference of Catholic Bishops. (2018). *Ethical and Religious Directives for Catholic Health Care Services, Sixth Edition*. https://www.usccb.org/resources/ethical-religious-directives-catholic-health-service-sixth-edition-2016-06_0.pdf

Indian Removal Act, Session 1 Chapter 148 (1830). https://calhum.org/files/uploads/program_related/TD-Indian-Removal-Act-1830.pdf

Employee Retirement Income Security Act of 1974, Pub. L. No. Public Law 93–406 (1974). https://www.congress.gov/93/statute/STATUTE-88/STATUTE-88-Pg829.pdf

Indian Health Care Improvement Act, Pub. L. No. Public law 94-437 (1976). https://www.congress.gov/bill/94th-congress/senate-bill/522

Public Law 94-521, (1976). https://www.govinfo.gov/link/statute/90/2460

Fertility Clinic Success Rate and Certification Act of 1992, Pub. L. No. Public Law 102-493, 106 (1992). https://www.congress.gov/102/statute/STATUTE-106/STATUTE-106-Pg3146.pdf

The Family and Medical Leave Act of 1993, (1993). https://www.ecfr.gov/current/title-29/subtitle-B/chapter-V/subchapter-C/part-825

Tax Cuts and Jobs Act, Pub. L. No. Public Law No: 115-97 (2017). https://www.congress.gov/bill/115th-congress/house-bill/1/text

U.S. Dept. of Veteran Affairs. (2019). *In Vitro Fertilization Treatment*. Community Care. https://www.va.gov/COMMUNITYCARE/programs/veterans/ivf.asp

U.S. District Court. (2008). *Culpepper v. Bluecross Blueshield of Tennessee, Inc.* https://casetext.com/case/culpepper-v-bluecross-blueshield-of-tennessee

U.S. Office of Personnel Management. (2021). *Healthcare COMPARE 2021 PLANS*. https://www.opm.gov/healthcare-insurance/healthcare/plan-information/compare-plans/fehb

Valerio, C., Vargas, K., & Raventós, H. (2017). IVF in Costa Rica. *JBRA Assisted Reproduction*, *21*(4), 366–369. https://doi.org/10.5935/1518-0557.20170060

van Empel, I. W. H., Dancet, E. A. F., Koolman, X. H. E., Nelen, W. L. D. M., Stolk, E. A., Sermeus, W., D'Hooghe, T. M., & Kremer, J. A. M. (2011). Physicians underestimate the importance of patient-centredness to patients: a discrete choice experiment in fertility care. *Human Reproduction*, *26*(3), 584–593. https://doi.org/10.1093/humrep/deq389

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer New York. https://books.google.com/books?id=CzwmBQAAQBAJ

Verter, V., & Lapierre, S. D. (2002). Location of preventive health care facilities. *Annals of Operations Research*, *110*(1–4). https://doi.org/10.1023/A:1020767501233

Waller, L., & Gotway, C. (2004). *Applied Spatial Statistics for Public Health Data*. Wiley-Interscience. https://doi.org/10.1198/jasa.2005.s15

Weber, A. (1962). *Theory of the Location of Industries.* University of Chicago Press.

Wilcox, A. J. (2010). *Fertility and pregnancy: an epidemiologic perspective*. Oxford University Press.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical Machine Learning Tools and Techniques. In *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier Science. https://doi.org/10.1016/c2009-0-19715-5

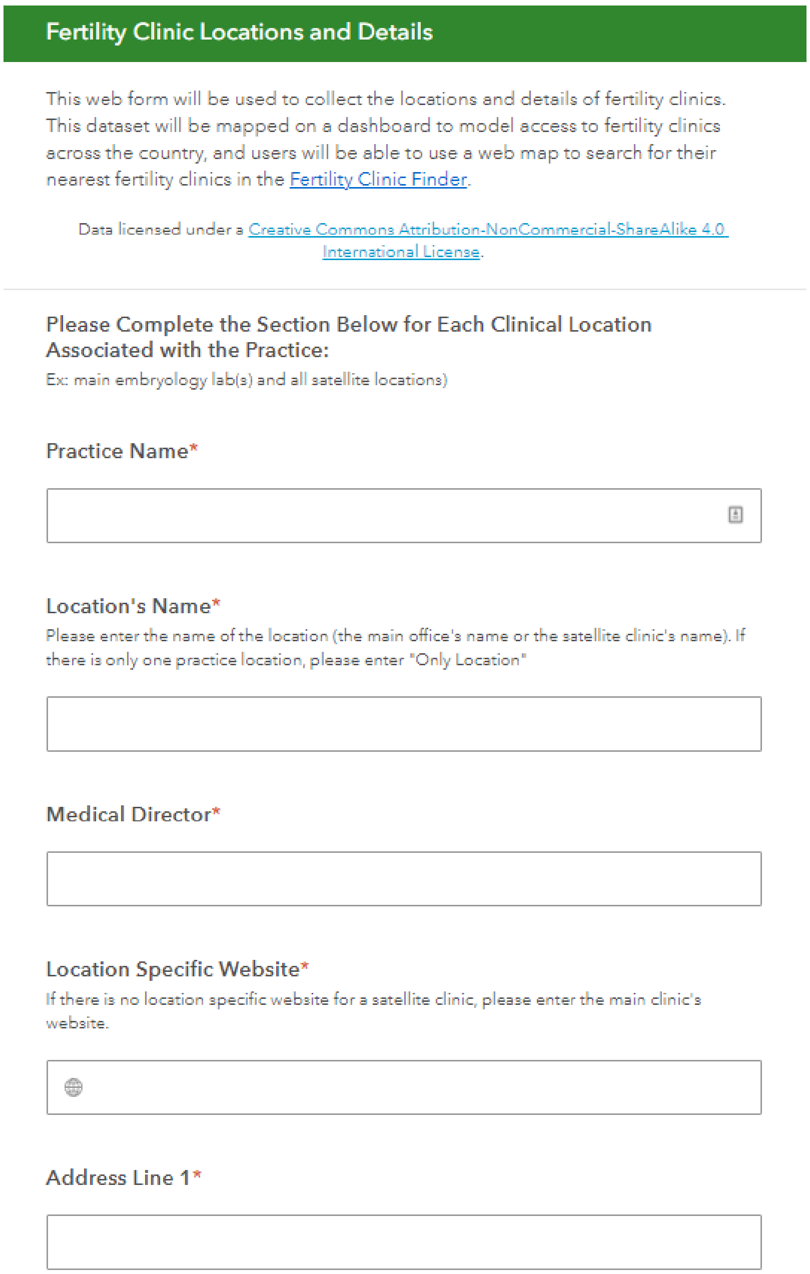
Wu, A. K., Elliott, P., Katz, P. P., & Smith, J. F. (2013). Time costs of fertility care: The hidden hardship of building a family. *Fertility and Sterility*, *99*(7), 2025–2030. https://doi.org/10.1016/j.fertnstert.2013.01.145

Yin, P., & Mu, L. (2012). Modular capacitated maximal covering location problem for the optimal siting of emergency vehicles. *Applied Geography*, *34*. https://doi.org/10.1016/j.apgeog.2011.11.013

Zegers-Hochschild, F., Adamson, G. D., de Mouzon, J., Ishihara, O., Mansour, R., Nygren, K., Sullivan, E., & van der Poel, S. (2009). The international committee for monitoring assisted reproductive technology (ICMART) and the world health organization (WHO) revised glossary on ART terminology, 2009. *Human Reproduction*, *24*(11), 2683–2687.

Appendix A Webpage for collecting fertility clinic locations and details

Screenshot of ArcGIS Survey 123 page for collecting location and details of fertility clinics using the website https://arcg.is/j8Hrf.



Fertility Clinic Locations and Details

This web form will be used to collect the locations and details of fertility clinics. This dataset will be mapped on a dashboard to model access to fertility clinics across the country, and users will be able to use a web map to search for their nearest fertility clinics in the Fertility Clinic Finder.

Data licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Please Complete the Section Below for Each Clinical Location Associated with the Practice:

Ex: main embryology lab(s) and all satellite locations)

Practice Name*

Location's Name*

Please enter the name of the location (the main office's name or the satellite clinic's name). If there is only one practice location, please enter "Only Location"

Medical Director*

Location Specific Website*

If there is no location specific website for a satellite clinic, please enter the main clinic's website.

Address Line 1*

Address Line 2

City*

State / District*

Zip Code*

Country*

USA

Phone Number

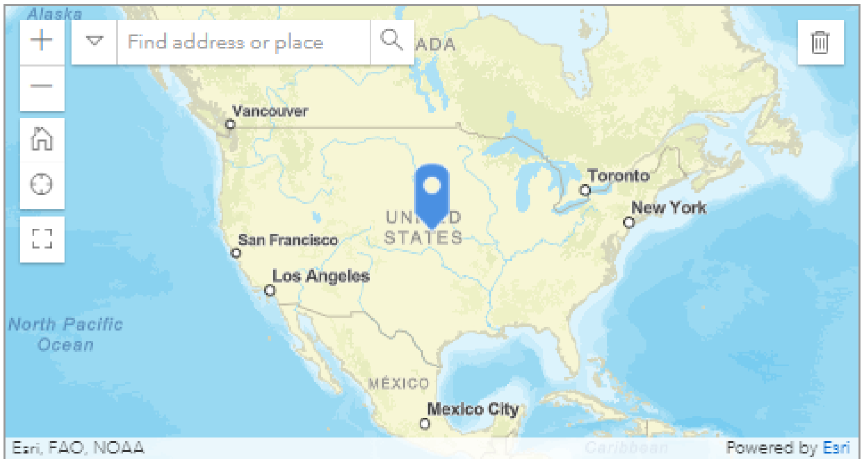
Fax Number

Days of the Week this Location is Open*

- [ ] Monday
- [ ] Tuesday
- [ ] Wednesday
- [ ] Thursday
- [ ] Friday
- [ ] Saturday
- [ ] Sunday
- [ ] Open By Appointment

Geographic Location*

Please use the search bar below to enter the location's address so that the location is captured accurately on the map.



Lat: 39.68385 Lon: -98.26173

What Type of Clinic is at this Location?*

○ Main Clinic
○ Satellite Clinic

Is there an Embryology Lab at this Location?*

○ Yes
○ No

What Services are Provided at this Location?*

Only check boxes if the procedure below is performed physically at this location.
For example, if egg retrievals and embryo transfers are not performed at this location, do not check the box for IVF.

- [ ] Consultations
- [ ] Blood Work
- [ ] Ultrasound Cycle Monitoring
- [ ] Intrauterine Insemination (IUI)
- [ ] In Vitro Fertilization (IVF)
- [ ] Semen Analysis
- [ ] Saline Infusion Sonohysterogram (SHG or SIS)
- [ ] Hysterosalpingogram (HSG)

Is Use of this Clinic Restricted? (ex: Only members of Armed Forces and their families can use this clinic)*

Yes

No

Other Details

1000

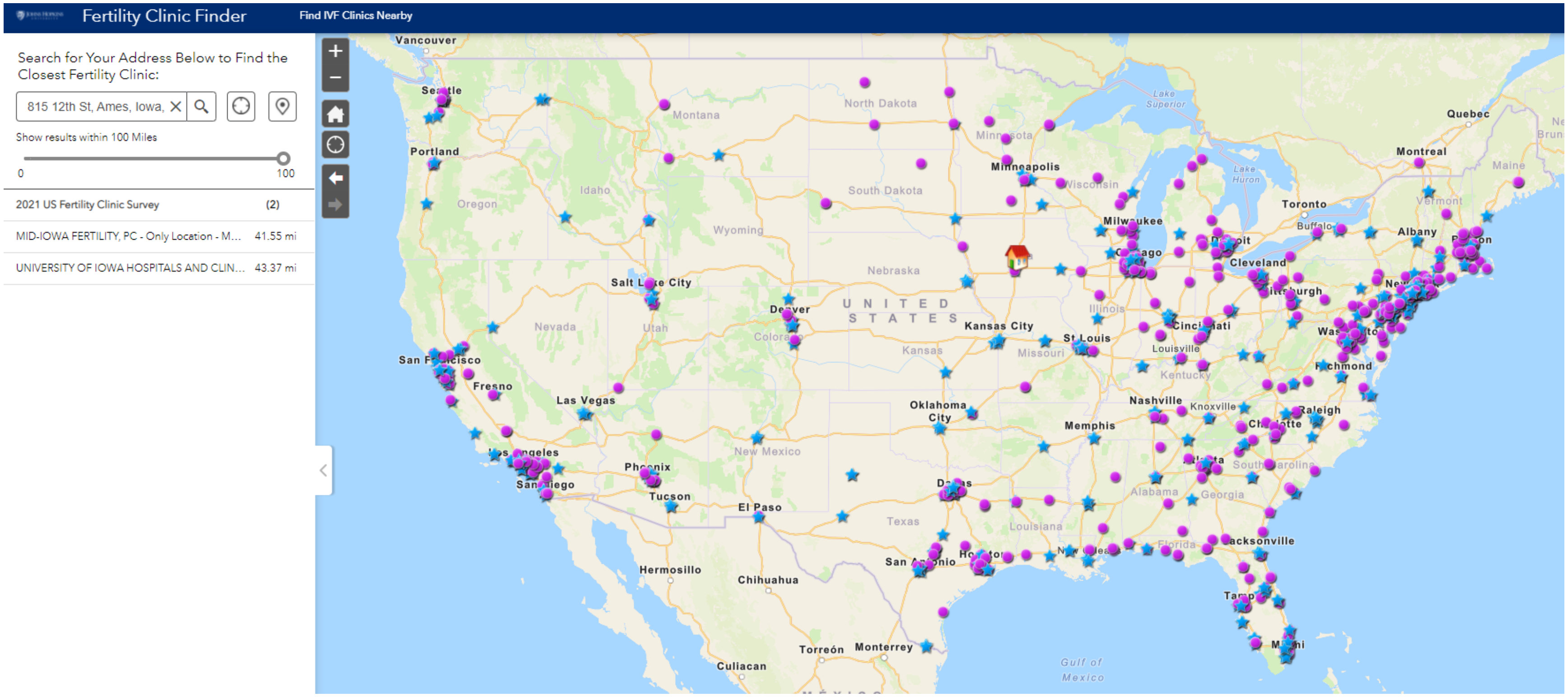
Submit

Appendix B Screenshot of the Fertility Clinic Finder web application

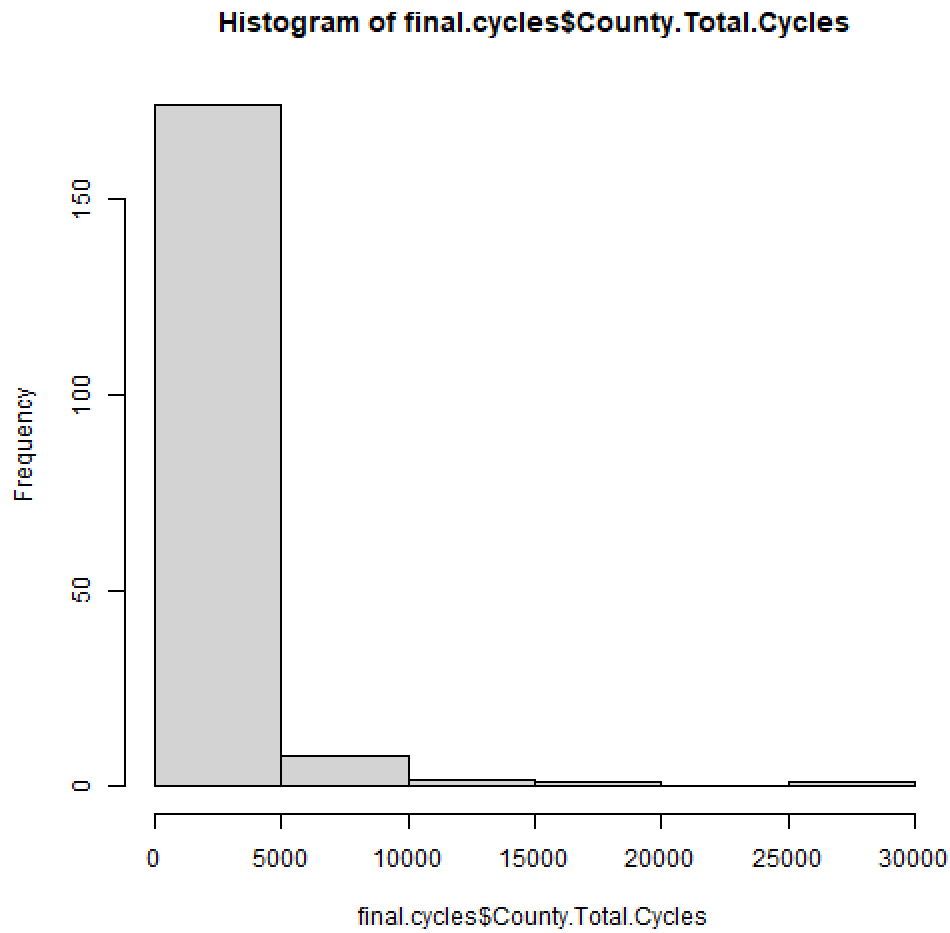
Available at the following links:

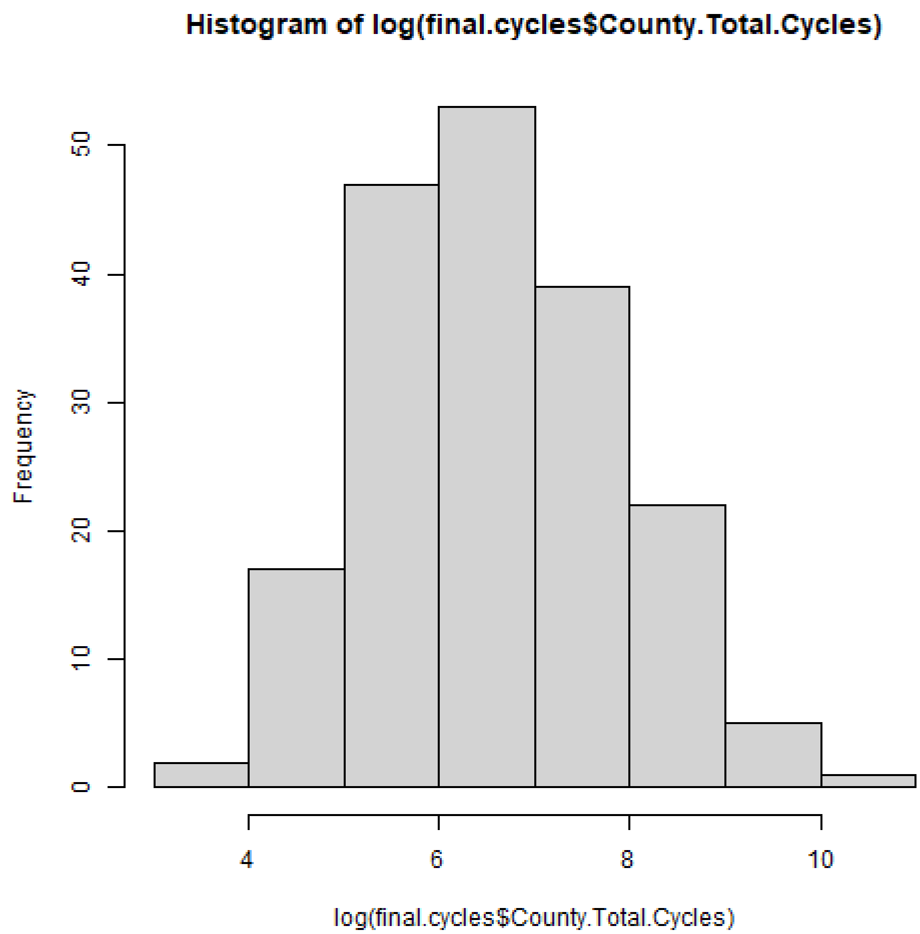
https://gisanddata.maps.arcgis.com/apps/webappviewer/index.html?id=3e073183ee824e2dae5900da1774ba4b

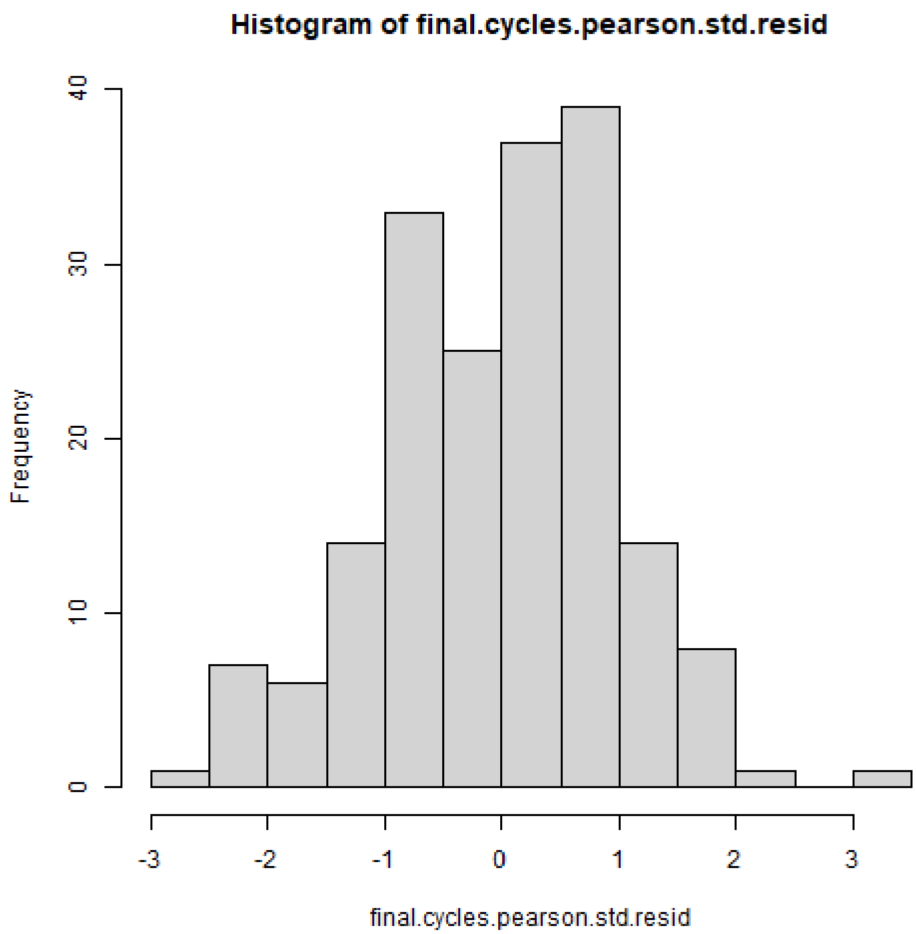
https://tinyurl.com/yxpdsdmu or using QR code to the right:



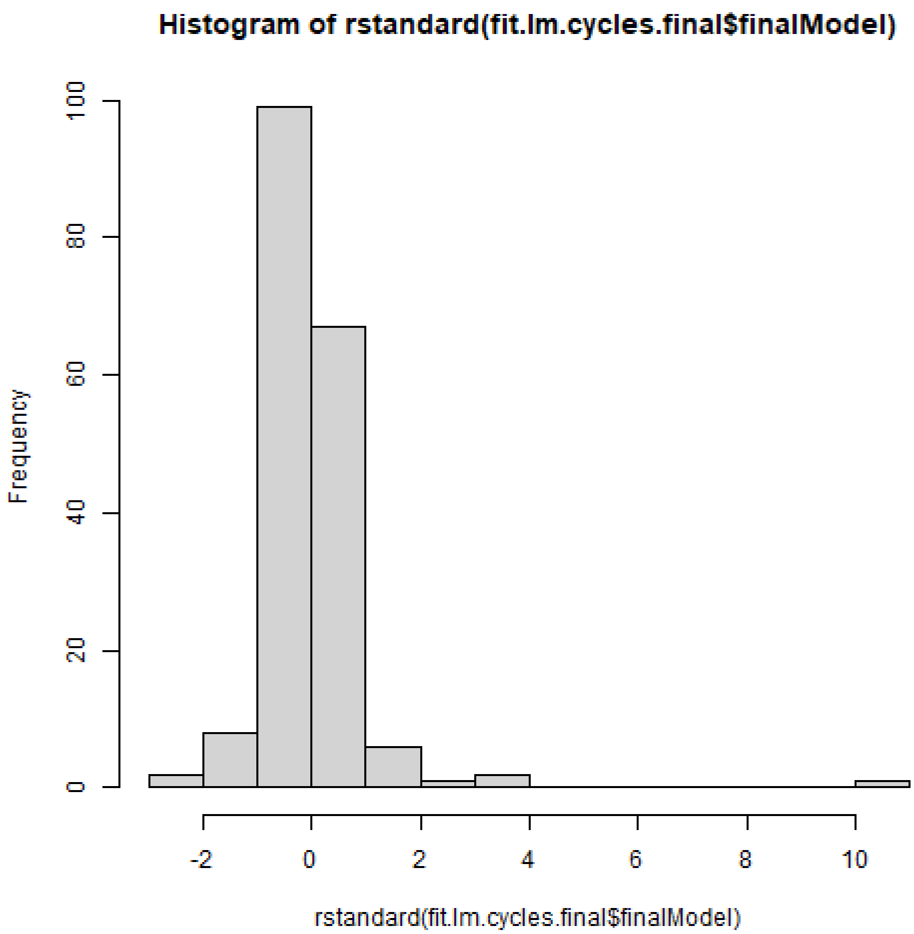
Appendix C Exploratory histograms of demand and clinic presence modeling

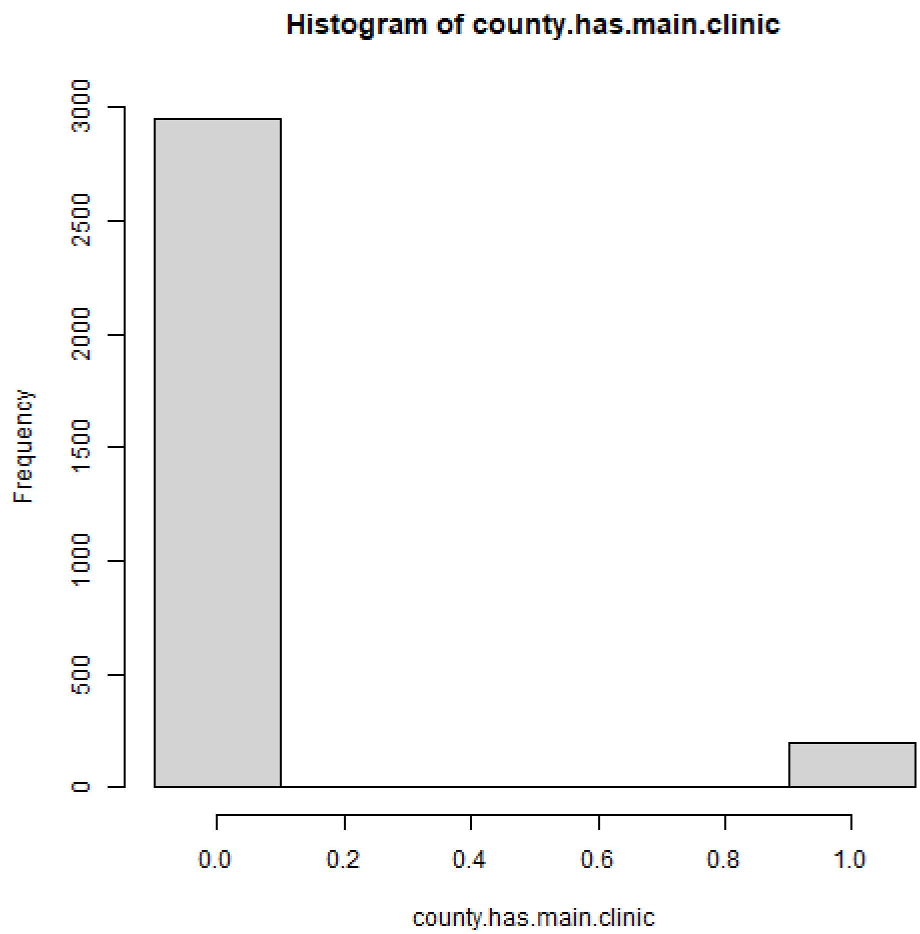


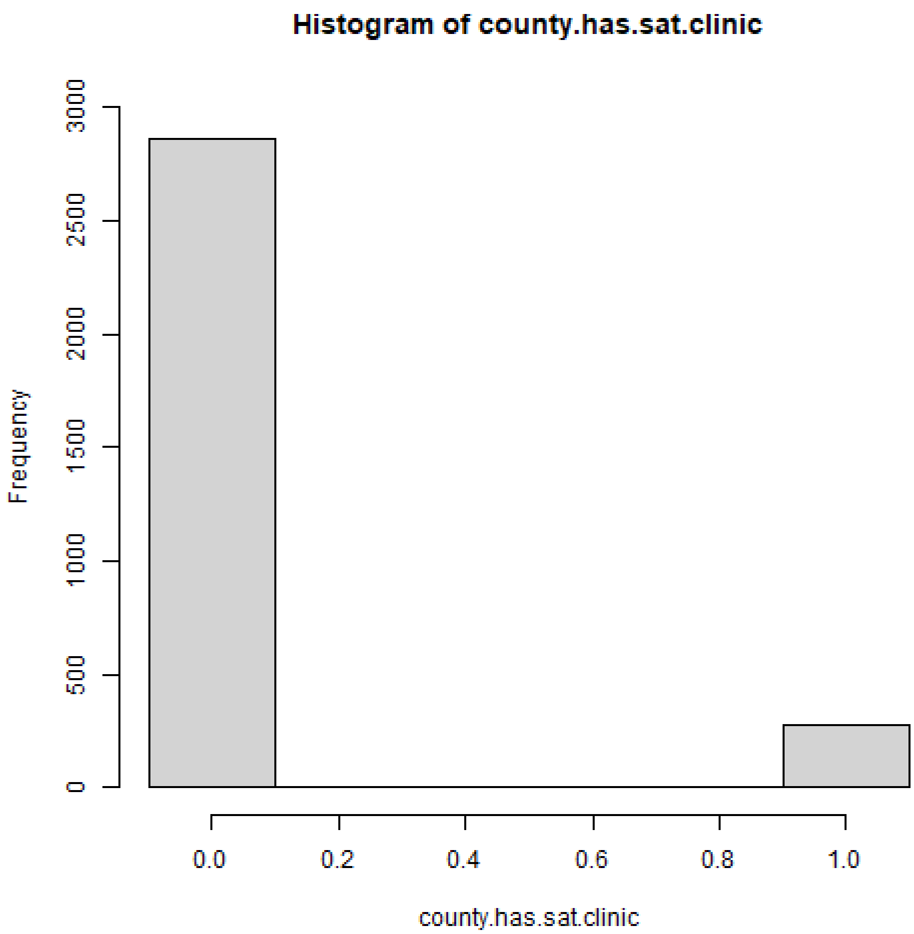

Histogram of log(final.cycles$County.Total.Cycles)
Frequency
0
10
20
30
40
50
4
6
8
10
log(final.cycles$County.Total.Cycles)



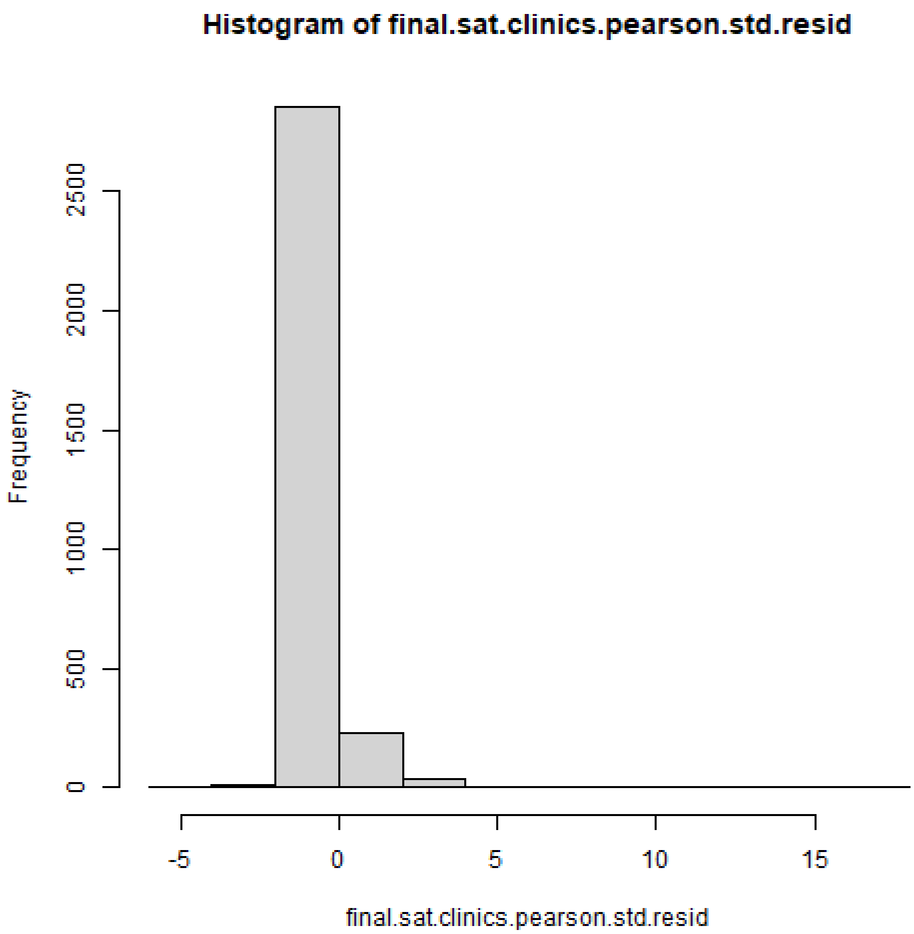
Histogram of final.cycles.pearson.std.resid
Frequency
0
10
20
30
40
-3
-2
-1
0
1
2
3
final.cycles.pearson.std.resid

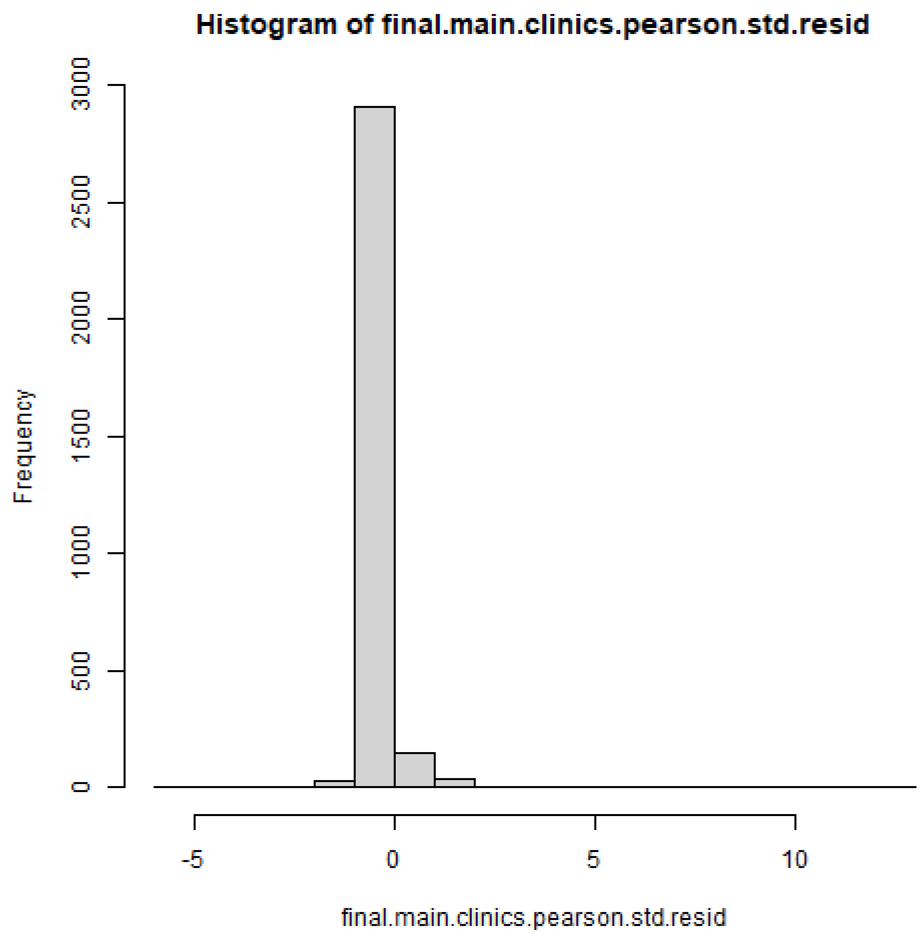

Histogram of rstandard(fit.lm.cycles.final$finalModel)
Frequency
0
20
40
60
80
100
-2
0
2
4
6
8
10
rstandard(fit.lm.cycles.final$finalModel)

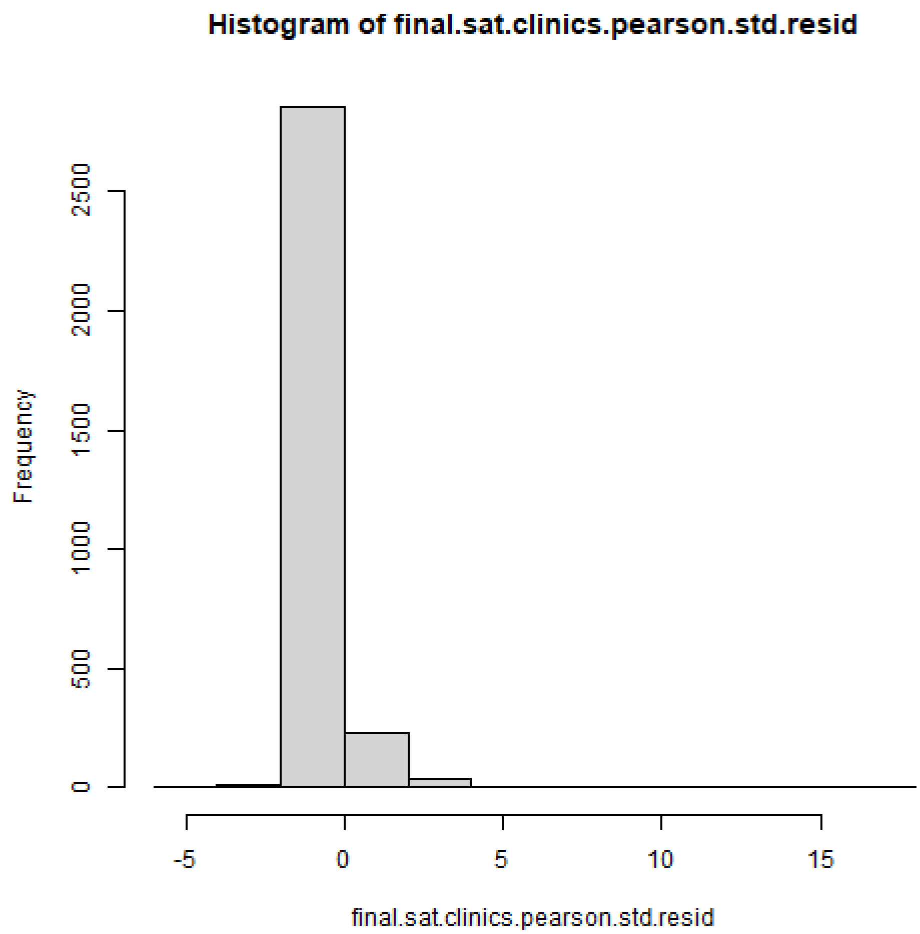

Histogram of county.has.main.clinic
Frequency
0
500
1000
1500
2000
2500
3000
0.0
0.2
0.4
0.6
0.8
1.0
county.has.main.clinic



Histogram of county.has.sat.clinic
Frequency
0
500
1000
1500
2000
2500
3000
0.0
0.2
0.4
0.6
0.8
1.0
county.has.sat.clinic


Histogram of final.sat.clinics.pearson.std.resid
Frequency
0
500
1000
1500
2000
2500
-5
0
5
10
15
final.sat.clinics.pearson.std.resid


Histogram of final.main.clinics.pearson.std.resid
Frequency
0
500
1000
1500
2000
2500
3000
-5
0
5
10
final.main.clinics.pearson.std.resid


Histogram of final.sat.clinics.pearson.std.resid
Frequency
0
500
1000
1500
2000
2500
-5
0
5
10
15
final.sat.clinics.pearson.std.resid