Empirical Studies of Discrete Choice Models in Health, Fertility, and Voting

Ali Hashemi

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics

Djavad Salehi-Isfahani, Chair
Wen You, Co-chair
Suqin Ge
Kevin J. Boyle

March 30, 2011
Blacksburg, Virginia

Keywords: Discrete Choice Modeling, Health, Fertility, Voting
Copyright 2011, Ali Hashemi

Empirical Studies of Discrete Choice Models in Health, Fertility, and Voting

Ali Hashemi

(ABSTRACT)

Almost everything that we do involves a choice. In recent years there has been a growing interest in the development and application of quantitative statistical methods to study choices made by individuals with the purpose of gaining a better understanding of how choices are made and also to predict future choice responses. In many fields, the choices made by individuals will determine the effectiveness of policy. Understanding what drives peoples choices and how these choices may change is critical for developing successful policy. Discrete choice modeling provides an analytical framework with which to analyze and predict how peoples choices are influenced by their personal characteristics and by the different attributes of the alternatives available to them.

In an ideal situation we would build discrete choice models using information from choices that people are observed to make, i.e., revealed preference (RP) information. From these data we can quantify the influence of particular variables in the real choice context; for example, how important is price in the decision to travel by train? There are, however, potential problems with these data. There might not be enough variation of the explanatory attributes; for example little price variation across alternatives. Furthermore, several attributes might be highly correlated e.g. price and quality. But the most important of all is the fact that it is not possible to observe choices for alternatives that do not yet exist; for example new programs and technologies. In cases where the data limits the information provided by real choices it may be appropriate to collect stated preference (SP) data, which is information on preferences provided from hypothetical choice situations.

This dissertation provides several applications of discrete choice modeling using both raveled preferences and stated preference. Unlike the last two chapters which deal with the revealed preference, the first Chapter, uses stated preference data. This Chapter evaluates the impact of several attributes of monetary incentives on the decision of patients to participate in a new weight loss program. Since this program does not exist yet, revealed preference data were not available and stated preference data were collected. The attributes of interest in this study include magnitude, timing and form of payment. The goal is to see what level and what combination of these attributes provides greater impact on the reach of the program. We also account for preference heterogeneity by using a random parameter framework.

Chapter 2 discusses another application of discrete choice models in event history models (also called survival analysis). In these type of models, the main goal is to use the history of happening an event to learn more about the effect of different factors on the probability of occurrence. The event of interest in our case is the birth. We use the birth history of rural women and try to model their decision to give birth over time. The ultimate goal is to evaluate the effect of health clinics and family planning program on this decision.

The final Chapter considers the application of discrete choice modeling in an electoral framework. The 2005 presidential election in Iran is used to model the decisions of Iranian voters. Using this revealed preference data we try to learn more about the main factors evolved in both participation and in the candidate selection.

Acknowledgments

This document, beyond the accumulation of years of study, reflects my relationship with many generous and inspiring people I have met since beginning my graduate work. I would like to thank their contribution in this journey:

To my gracious mentor and advisor, Djavad Salehi-Isfahani, who demonstrates that social change is central to intellectual work and, as such, scholars have a responsibility to use the privileges of academia to imagine and create a better world. I have been inspired by his passion, hard work and depth of understanding of economic and social issues.

To Wen You, my co-chair, who kindly advised me through a new field of research. I have been amazed by her discipline, hard work, and strength in managing both academic life and personal life as a young mother.

To my committee members, Suqin Ge and Kevin Boyle, for their encouraging words, thoughtful criticism, and time and attention during busy semesters. To Christophere Parmeter, who although was not a member of my committee, but I always benefited from his valuable comments and advises on technical issues.

To the Department of Economics staff for assisting me with the administrative tasks necessary for completing my doctoral program: Amy Stanford, and Sherry Williams

To my dearest friend, Saeideh, who was always there cheering me up and stood by me through the good times and bad.

And finally, to my family: my father, Mojtaba, for showing me the joy of intellectual pursuit ever since I was a child; my mother, Tahereh, for raising me with her caring and gently love; and my supportive and caring siblings, Rezvan, Samineh, and Javad.

I would like to thank everybody who was important to the successful realization of thesis, as well as expressing my apology that I could not mention personally one by one.

Contents

List of Figures

List of Tables

Chapter 1

What, When, and How Much: The search for financial incentive designs that can achieve wider reach for incentivized weight loss program

[1]

[1] This is a joint work with Wen You, Kevin J. Boyle, Paul Estabrooks, Christopher F. Parmeter and Barbara Kanninen.

1.1 Introduction

The obesity epidemic contributes to substantial economic, social, and personal losses (Mokdad et al. 1999, Wellman and Friedberg 2002). Its rapid and consistent growth is accompanied by a host of chronic diseases (e.g., asthma, atherosclerosis, depression, hyperinsulinaemia, hypertension, sleep apnea, and Type 2 diabetes) (Must et al. 1999). The resulting direct medical expenditures are 9% of U.S. annual medical expenditures (Finkelstein et al. 2003) and obesity causes 27% of the rise in per capita health care expenditures (inflation adjusted) (Thorpe et al. 2004). Furthermore, obese adults on average spend 48% more inpatient days per year than normal-weight adults (Thompson, et al., 2001), experience more absenteeism (Aldana and Pronk 2001, Tsai et al. 2005), are more vulnerable to short-term disability (Arena et al. 2006), and exhibit more decreased productivity (presenteeism) in the workplace (Ricci and Chee 2005, Burton et al. 2005). In aggregation, the annual costs attributable to obesity alone are estimated to be $73.1 billion (Finkelstein et al. 2010). Therefore, the federal government and employers are more and more impelled to invest in means to encourage cost effective weight loss programs and the most common ways to increase program reach and adherence are through financial incentives (Zamosky 2010).

Financial incentives have been utilized in numerous preventive care studies (see reviews such as (Kane et al. 2004b, Kane et al. 2004a, Wall et al. 2006, Paul-Ebhohimhen and Avenell 2008). Different aspects of incentives are documented: incentive forms such as insurance co-pay, coupon, gift or cash (Cherkin et al. 1990, Lillard et al. 1986, Kenkel 1994); the monetary value (Doody et al. 2003, Helpern et al. 2002); the certainty of the incentive; the contingency conditions (i.e., whether the payment is tied to participation or to certain outcomes) (Higgins et al. 2004, Curry et al. 1991, Donatelle et al. 2000); and the timing of payment. The reviews by (Kane et al. 2004b, Kane et al. 2004a) found financial incentives to be effective in achieving desirable outcomes in 73% of 111 randomized control preventive care trials they reviewed. However when it comes to more complex preventive cares such as weight loss programs or smoking cessations which require greater investments of participants' time and effort, no clear conclusion can be made about the effectiveness of financial incentives. This conclusion is confirmed by a systematic review of the effect of financial incentives in the behavioral treatment of obesity (Paul-Ebhohimhen and Avenell 2008). Their meta-analysis found that incentives did not have a statistically significant effect on short-term weight loss or on long-term weight maintenance.

Although these reviews do find that financial incentives are consistently positively related to initial program participation rates, there is a dearth of information on the perceptions of potential program participants related to the type (What), timing (When), or magnitude (How much) of incentives for weight loss. For example, no studies employing financial incentives provided a rationale to justify the incentive designs and incentive levels employed. The equivocal findings on the potential of financial incentives to stimulate and maintain weight loss may also be a consequence of the incentives employed being too small to offset money and time costs of participants (Kane et al. 2004). Overall, a practice uniformly absent in previous incentive-based intervention work is a specific definition of the financial incentive, the intended purpose, and how it was hypothesized to help the participants (Kane et al. 2004a, Paul-Ebhohimhen and Avenell 2008). Without systematic quantitative research, incentive choices will continue to be arbitrary and difficult to describe from the perspective of reach, effectiveness, cost, and maintenance-all important aspects for future community and clinic-level adoption (Glasgow et al. 2002, Glasgow et al. 2004, Glasgow and Emmons 2007, Finkelstein et al. 2007, Volpp et al. 2008).

This study takes the first step to fill the literature paucity through formally evaluating, within an economic framework, the potential effectiveness of the magnitude, type, and timing of financial incentives in stimulating weight loss program participation in overweight and obese adults. To avoid potentially masking the effectiveness of financial incentives, our evaluation is based on the platform of a minimal weight loss program structure similar to those that have demonstrated small weight loss in other populations (i.e., $\prec$.4 lbs/week) (Black and Cameron 1997, Heshka et al. 2003, Volpp et al. 2008).

1.2 Related Literature

Individuals do not always make rational decisions specially when it comes to their health habits. Smoking, misuse of drugs and alcohol and obesity are just a few examples of irrational decision making. To influence the the health habits, economists traditionally have sought to correct the externalities. Taxation, setting regulations and providing information are all part of different strategies recommended by economists. Although effective but these strategies do not seem to be enough. Drug and alcohol misuse are still prevalent issues and obesity is spreading. New methods are needed to help individuals when they are willing to change their irrational health habits but lack of willpower is a barrier. Slinder (2008) reviews the recent development in economics literature when the neoclassical model breaks down and suggests that economists should borrow the idea of incentive-based approaches from the psychology literature as a base for developing population-based policies. Positive incentives have been used in several studies of smoking cessation (Shoptaw et al. 2002), drug and alcohol reduction (Epstein et al. 2003, Kosten et al. 2003, Petry et al. 2004) and weight loss (see Section 1.1) and have been found to be surprisingly powerful motivator of change (Lussier et al. 2006). The reason for effectiveness of the conditional payments, which usually are small in magnitude when compared to the potential gains from correcting a health habit in long term, could be in factors that are not considered in the traditional neoclassical model: immediacy and certainty of payments conditioning on behavior change. Individual might be present biased and do not discount as economists assume (Sindelar 2008).

Economists have recently found the incentive based programs of interest (Charness and Gneezy 2009, Sindelar 2008). The conditional cash transfer programs like the PROGRESSA in Mexico, or Mayor Bloomberg's poverty alleviation program in New York are recent examples of incentive based programs developed to influence individuals behaviors.

In this study, we use a specific approach, stated preference method, to elicit individuals preference for different attributes of monetary incentive in an incentive-based weight loss program. Stated preference techniques have been widely used in marketing, transportation and environmental economics (e.g., (Louviere et al. 2000, Ryan and Hughes 1997)). In recent years it has also been increasingly popular in the field of health economics. Studies have used stated preference to model choice of health plans (Cunningham et al. 1999, Booske et al. 1995, Harris et al. 2002) and choice of medication (Bingham et al. 2001), to estimate patient preferences for different aspects of health care services (Moayyedi et al. 2002, Chakraborty et al. 1993, Chakraborty et al. 1994, Ryan and Hughes 1997, Wordsworth and Scott 2002), to assess preferences in doctor-patient relationship (Vick and Scott 1998) and establish general practitioner preference for different practice jobs (Wordsworth et al. 2004).

1.3 Conceptual Framework and Model Specification

The analysis of perceived program participation is based on responses to the weight loss program choice questions. The random utility model provides the conceptual framework where an individual's utility associated with each weight loss program package is assumed to be comprised of systematic (i.e., observed) and random (i.e., unobserved) elements and depends on the attributes of the program's incentive component. Specifically, individual i's utility associated with program package j in choice scenario s is:

$$U_{ijs} = \beta' x_{ij} + \varepsilon_{ijs}$$
$$i = 1, ..., I, j = 1, 2, 3s = 1, 2, 3, 4$$

where x_{ij} is a vector of the relevant incentive attributes for program package option j (magnitude, type and timing), β is a vector of associated preference parameters and ε_{ijs} is the random error term. If a program choice j is chosen, it is assumed that program package j yields the maximum utility among available options ($j \in J$). Assuming an extreme value distribution of errors, the probability of choosing program option j is the standard conditional logit model (Greene, 2000). This model was estimated for our full sample of respondents as well as a variety of demographic splits (such as BMI status and gender) to provide an initial look at preference heterogeneity.

However, conditional logit model contains some restrictive assumptions that are questionable: a. it assumes that choices made by an individual are not correlated over the S repeated choice scenarios; b. it assumes fixed parameters which implies that all individuals have the identical preference. To relax these assumptions, we also estimated mixed logit models (or random-parameters logit models) which allows some utility parameters (β) to vary randomly over individuals (Revelt and Train 1998, Train 1998, Train 2003). The utility parameters are drawn from a density $f(\beta|\theta)$ with parameters θ (typically a mean and a variance).
Conditional on the unknown β_i, the probability of individual i making a given sequence of choices over S choice scenarios is the product of conventional logit formulas:

$$L_{is}(\beta_i) = \prod_{s=1}^{S} \frac{e^{\beta_i' x_{js}}}{\sum_{j=1}^{J} e^{\beta_i' x_{js}}} \tag{1.1}$$

In our case, S equals 4 and J equals 3. The unconditional probability for the random parameter logit models is the integral of L_{is} over all values of β_i according to $f(\beta|\theta)$:

$$P_{is}(\theta) = \int L_{is}(\beta_i) f(\beta_i|\theta) d\beta_i. \tag{1.2}$$

By specifying a distribution for β, θ can be estimated through a simulated maximum likelihood procedure (Train 2003). Here we report results where the random parameters are assumed to be distributed normally.

Similar to (Paterson et al. 2008) we specify utility for individual i associated with program j in choice scenario s in additive form as:

$$U_{ijs} = \beta_{iA}A_i + \beta M_i + \beta_{iP}F_j + \beta_{iT}T_j + \varepsilon_{ijs}, \tag{1.3}$$

where A_i is a alternative-specific constant (ASC) indicating a program alternative was chosen (i.e., the individual did not opt out) and equals to 1 when either program was chosen and 0 when the individual chose neither, M is the financial incentive magnitude of the program, F is a dummy variable capturing the form of the incentive payment and T is a dummy variable capturing the timing of the payment. Our omitted categories from F and T are the form of insurance co-pay waiver and pay at 12 months, respectively. Detail descriptions of the incentives and the associated levels are in the next section.

1.4 Study Design

We administered a mail survey to 1,500 adults (300 normal weight and 1,200 overweight or obese) to elicit responses to attribute-based choice (ABC) questions that investigate subjects' willingness to participate in a minimal weight loss program when coupled with a range of potential incentive components[2]. ABC question is a derivate of economic stated-preference methods which are tools used to systematically investigate individuals' preferences for programs (e.g., weight loss interventions) and program attributes (e.g., financial incentive designs). The merit of stated-preference methods is their ability to elicit preference information through carefully constructed scenarios.

ABC questions have been commonly used to elicit individuals' preferences for different treatment and prevention programs by asking respondents to choose their preferred program from a set of programs defined by differences in their attributes (Hall et al. 2002, Sassi et al. 2005, Harris et al. 2002, Pol and Cairns 2001). The results of the ABC analysis will enable researchers to determine the relative importance of the different attributes in influencing perceived program participation decisions. Therefore, this study can provide the much needed rationale for constructing specific incentive framework and/or using different incentives on different population groups in the context of the minimal weight loss program. The results of this study will further inform full-scale randomized control trials through calibrating the incentive components and it is the crucial step recommended by other studies (Stevens et al. 2007).

1.4.1 Program Description and Incentive Attributes

Fundamental to ABC method is the use of surveys in which alternative (hypothetical) program/policy scenarios are described by varying levels of attributes and/or costs. Respondents are asked to choose

[2]The normal weight survey participants were told to complete the survey as if they actually need a weight loss program.

their preferred option across the range of alternatives. In this study, we asked the survey participants to indicate their willingness to participate in a 3-month minimal weight loss program. The minimal weight loss program was described to them as including monthly program weigh-in visits, two booster weight-in visits (6 and 12 month), an initial face-to-face consultation with a dietitian, personalized eating and exercising plans and tracking tools, and weekly telephone support calls (See appendix for details). We chose the minimal weight loss program as the platform for evaluation instead of the intensive ones in order to avoid potentially masking the effectiveness of financial incentives. Survey participants were informed that losing 5% of body weight for overweight and obese individuals will greatly reduce the risk of chronic disease. This condition was used as the basis for a single incentive reward contingency: if program participants achieve or maintain at least 5% weight loss when compared to their initial weight they will receive the incentive reward.

The program scenarios are defined in terms of different combinations of financial incentive attributes. These attributes include: a. Monetary value of the incentive earned per weigh-in (M) (e.g., $ amount), b. Form of incentive (F) (e.g., cash or in-kind such as prepaid grocery gift cards, pre-paid gym passes and health care cost waivers), c. Timing of incentive payment (T) (e.g., immediate or delayed). Each scenario represents two incentive package options and participants were asked to indicate which program package they would choose or to opt out. Before the finalization of the survey materials, two rounds of focus group meetings were conducted with participants from the targeted population to identify the appropriate range for, and intervals between, levels of each attribute. The list of attributes and levels used in the final survey materials is shown in Table 1.1 and ABC question example is shown in the Appendix.

1.4.2 Experimental Design, Survey Development and Implementation

ABC method is basically the utility-theoretic approach to model binary responses which is based on utility differences (McFadden 1974, Train 2003). Therefore, it is important to make sure that all possible binary pairings of each categorical attribute are included in the study design. Furthermore, if considering possible order effects, all comparisons must be included in all possible orders. For our study, we have a total of three attributes and 38 possible pairings (20 pairings for magnitude, 12 pairings for type, and 6 pairings for timing). Trying to include all 38 possible pairings would lead to a very large design ($20 \times 12 \times 6$). We opted to take a randomized approach to include various magnitudes in the design treating the magnitude attribute as a continuous variable. This approach results in a slight amount of correlation among attribute differences, but it keeps the design a manageable size (12×6).

Once all the pairings for each attribute are defined, the next step is to combine the attribute pairings into full choice sets so that each of the pairings are included in a balanced way (meaning they are included the same number of times across the choice questions) and in an uncorrelated way (meaning they are combined with the other attributes in different, nonsystematic ways, i.e. orthogonally). We did this by finding a 12×6 fractional factorial matrix. The 72-line matrix we used comes from Warren Kuhfeld's on-line design catalog (SAS, 2009). We then assigned magnitudes randomly to each of the alternatives in each choice set. The final design of ABC questions is comprised of 72 choice sets (i.e., scenarios) which equals to 144 program packages. We divided these 72 scenarios into 18 survey versions with 4 scenarios each. These 18 survey versions were then randomly assigned to participants.

Besides four ABC questions, each survey includes questions to collect information about the individual's socio-economic characteristics, physical activity level, healthy eating behavior, intrinsic and extrinsic motivations, weight and height, and general health status.

All survey participants were recruited throughout Virginia Commonwealth via random digital dialing. The phone recruitment and survey mailing were conducted through Virginia Tech Center for Survey Research. The individuals were screened over the phone to ensure phone recruitment eligibility (i.e., should be of at least 18 yrs old). Then individuals provided self-reported weight and height over the phone and research staffs used this information to calculate Body Mass Index (BMI) and screened them further: the recruitment goal is to have 300 normal weight adults and 1,200 overweight or obese adults agreed to participate in the survey over the phone. A total of 1,500 mail surveys were mailed out in three rounds over the months of October and November 2009 and February 2010 to those eligible participants who consented over the phone. No follow-up calls and post cards were used. We received 863 complete and usable surveys. The completion rate is 57.5%. The surveys were then coded by trained research staff and data were entered and double-checked for human errors.

1.5 Results

In the data analysis that follows we first treat all participants as equivalent before stratifying them across various demographic variables. For example, we may expect that obese and overweight individuals may have different preferences for exercise than those normal weight individuals and this could be reflected in our estimated parameters inside the utility function and similar expectation can be tested across gender as well.

1.5.1 Data Description

Table 1.2 reports summary statistics of survey respondents' characteristics. The sample was dominated by female respondents (67%). Approximately 84% of respondents are Caucasian and 11% are African American. Both male and female subpopulations shared a similar race pattern. With the average age of 55 years in total sample, most of the respondents aged more than 40 years old. On average the subsample of males tend to be older than their female counterparts by about 2 years. A majority of the respondents have earned at least some college education. Approximately half (48%) of female respondents were not working at the time of survey either because they were retired, unemployed, students or unpaid home workers. This ratio is lowered to 40% in the male subsample. As expected, on average, male respondents earn more than female respondents. Nearly half of the male respondents (49%) were overweight and 35% were obese. In contrast, in the female subsample we notice a higher ratio of obese (41%) respondents and a lower ratio of overweight

(35%). In sum, male respondents tend to be older, earn higher level of income, have higher education and have higher overweight prevalence rate.

1.5.2 Fixed Parameter Conditional Logit Estimates

We begin our analysis by estimating standard conditional logit models. These results are presented in Table 1.3 (Model 1 and Model 2). Model 1 is for the full sample while Model 2 examines preference differences across gender and BMI status. Several results are immediate regardless of specification (i.e., similar for both Model 1 and 2). First, both models show negative coefficients associated with ASC indicating that individuals require certain compensation to select a program relative to the status quo, i.e., exercising/eating healthfully is costly to potential program participants. Second, both model estimates the same positive influence of financial incentive magnitude on the program participation probability (0.014) which indicates that the probability of an individual select a program, regardless of the options, increases as the overall magnitude of the incentive increases.

Focusing on our attribute specific results we see that individuals generally prefer to be paid either in cash or in gift cards (which essentially acts as cash) relative to either a prepaid gym membership or a health insurance copay waiver. Furthermore, while the difference between receiving a gift card or cash is statistically insignificant, these effects are roughly 5 times larger than that of receiving a gym pass, which is statistically indistinguishable from the copay. Thus, individuals respond to financial incentives well, provided the interface (form) is appropriate. Lastly, the pay at each weigh-in option is the preferred one in the timing category while pay upfront in the 3-month weigh-in is the least favorable one. The upfront payment only reward the one-time weigh-in (3-month weigh-in) and no incentives are associated with the other two booster weigh-ins while the other two timing options attach incentives to all three weigh-ins. Therefore, results indicate that individuals would prefer to be paid more often in the program which may be related to motivation and desire to be recognized.

Model 2 examines the preference heterogeneity across gender and BMI status through adding gender and BMI two-way and three-way interaction terms with ASC. All interaction terms are statistically significant which shows considerable gender and BMI status heterogeneity in preferences with regard to perceived participation decision. The inclusion of two-way and three-way interactions enables the comparison between gender and weight status groups (a total of 6 groups). Table 1.4 presents the magnitudes and p-values for the differences among these groups. It shows clearly that the perceived participation decision preference is statistically significantly different when comparing normal weight male with all the other five groups. Normal weight male is statistically less likely to participate in the program than other groups. Furthermore, obese male is more likely to participate compared to overweight male. No further significant preference heterogeneity is found through interaction models after controlling for the incentive attributes.

To further investigate the relationship among all these attributes, we calculated how much is needed to offset the "cost" of participating in the program under different incentive framework. Holding other attributes constant (i.e., type and timing), magnitude of the incentive per weigh-in that gives the same utility level as the one without program participation will answer this question. Model 1 fixed parameter model provide average information across the sample. For the baseline program that pays the individual at the end of the 12-month booster weigh-in in the form of the co-pay waiver, the individual will receive higher utility (i.e., he/she is more likely to choose to participate in the program) if the magnitude of the incentive per weigh-in is $\succ$\$80.43. This dollar amount can drop to \$18.64 if the program incentive framework is to pay participants cash instead of co-pay waiver at the end of the 12-month weigh-in. This shows the how different the effective incentive amount can be under different forms in terms of program reach. Model 2 shows different compensation needs by gender and weight status and provides further insights to program customization. For example, the break-even compensation magnitude for the baseline program drops from 80.43 to 75.50 for an obese female. This \$75.50 will drop further to \$14.07 if the program is to pay cash instead of co-pay waivers.

1.5.3 Random Parameters Logit Estimates

The conditional logit model with interaction terms allowed some degree of heterogeneity (across gender and weight status). In order to further examine the heterogeneity within gender and weight status, we estimated random parameter logit models with and without interactions as well (i.e., Model 3 and 4 in Table 1.3 respectively). For the model without interactions, there is vast improvement in model fit moving from the fixed parameter conditional logit model to random parameter logit model: the pseudo-likelihood ratio increases from 0.09 to 0.3. Smaller improvement in fit is found in the model with interactions: the ratio increases from 0.1 to 0.2. The standard deviations for all of the parameters' distributions in Model 3 and Model 4 are statistically significant which confirmed the importance of the preference heterogeneity that goes beyond gender and weight status.

First, Model 3 shows the estimated mean of the distribution of the random coefficients for program participation (i.e., ASC) is -0.052 and not statistically significant. However, the standard deviation of this parameter distribution is significant and of the magnitude of 4.804. Coupled with the normal distribution assumption we imposed, the mean and standard deviation for ASC suggests that half of the target population has a positive perceived participation preference. This means that more than half of the target population would choose to participate regardless of any variation of the incentive attributes.

Second, both Model 3 and Model 4 confirm that individuals are more likely to participate if they receive a higher value of the incentive (i.e., significant 0.022 and 0.026 respectively). Third, cash and gift cards are also found to be preferred to gym passes and insurance co-pay waivers, as well as being statistically indistinguishable from one another at the mean. However, random parameter logit model enables us to examine the coefficient distribution which shows the extent to which individuals differ in their preferences for financial incentive attributes. Given the normal distribution assumption imposed, the estimated mean (1.188) and standard deviation (0.696) associated with the cash coefficient in Model 3 show that about 4.4% of the target population would prefer co-pay waivers to cash (i.e., the coefficient of cash will be negative). For the gift card coefficient distribution, it can be shown that about 9.5% of the target population would prefer co-pay waivers to gift cards (i.e., the coefficient of gift card will be negative). These percentages increases to 31% and 34% when gender and weight status break down is added (Model 4).

Model 3 confirms the average preference towards payment at each weigh-in and dislike of pay at 3-month option. However, standard deviations associated with timing attributes are not significant which shows no significant timing preference heterogeneity across gender and weight status. One obvious difference comparing Model 4 and Model 2 is that the coefficient associated with ASC is much larger in magnitude in Model 4 (-3.339 vs. -1.840). This large discrepancy shows that allowing perceived participation preference to differ across and within gender and weight status results in larger compensation need in the normal weight male group. Model 4 results confirm the findings on the incentive attribute influence on participation decision in Model 2 as well. Cash and gift card are preferred to either pre-paid gym pass or co-pay waiver of equal value and pay at each weigh-in is the most preferred timing option. Allowing within gender and weight status heterogeneity shows larger variation in attribute preference: the much larger standard deviations in Model 4 compared to Model 3. Furthermore, timing preference heterogeneity is found within gender and weight classification group (the standard deviations are significant in Model 4).

Model 4 also enables calculations of the compensation magnitude needed to offset the "cost" of participation for different gender and weight classification groups. Considering the preference heterogeneity within groups, the break-even compensation magnitude needed for an obese female to say yes to the baseline program increases from $75.50 (Model 2) to $133.50 (Model 4). This increase reflects the impact of the wide dispersion found in the participation decision preference in these random parameter models. This $133.50 drops to $85.58 if the payment form changes from co-pay waiver to cash. As shown in Table 1.3, random parameter logit models show the similar findings as fixed parameter logit models: Normal weight male is statistically less likely to participate in the program compared to other groups and obese male is more likely to say yes to the program compared to overweight male.

Both conditional logit and random parameter logit show differences in coefficient estimates across gender. However, this coefficient differences may be just due to scale differences in the random part of the utility (i.e., variance differences). Therefore, we adapted Swait and Louviere's test procedure (Swait and Louviere 1993) to test whether the differences in coefficient estimates are due to gender heterogeneity or due to scale differences. We tested whether the parameter estimates are equal while allowing the scale factor to differ across gender groups. The test statistic is a likelihood ratio test. For conditional logit model, we got test statistic of 43.18 with a p-value of 0.000. For random parameter logit model, we got test statistic of 127.03 with a p-value of 0.000. Therefore, both tests reject the null hypothesis that the coefficient estimates across gender groups are equal after accounting for scale differences. Therefore, these tests further confirm that the preference heterogeneity does exist beyond gender differences and this minimal weight-loss program will be more effective in reach if the program can customize the incentive component by gender.

1.5.4 The Incentivized Minimal Weight-Loss Program Reach Prediction

Table 1.5 reports the average predicted probability that individuals in the given gender and weight status group will say yes to the incentivized minimal weight-loss program we presented. The results are calculated based on the random parameter logit model with weight status and gender interactions (Model 4) since model results and tests we performed all confirm that this full heterogeneity model is more appropriate. We are interested in those incentive attribute combinations that have $\succ$ 50% chance to reach the individual (i.e., the group average probability that the individual will say yes to this program package is $\succ$ 50%). Examining each group, it is clear that the winner package is the combination of incentive payment at each weigh-in with the magnitude of $98 and in the form of either cash or gift card. Individuals' preference toward immediate reward is confirmed again. Furthermore, the probability of program participation increases as the incentive magnitude increases within each group.

Let us fix the timing of payment at each weigh-in (the last four columns of Table 1.5) and examine them closely. It shows that $98 per weigh-in achieves the highest probability with cash or gift card for obese male (94%). If payment is in the form of cash or gift card, the probability will not exceed 50% for normal weight male if the $ amount is at $55 or less and will not exceed 50% for overweight male if the $ amount is at $5 or less. It only takes $5 to achieve $\succ$ 50% probability of program participation for females regardless of weight status. Examining each row in Table 5 (i.e., the monetary value is fixed for each row) shows that cash and gift card stimulate the highest participation probability regardless of payment timing.

Comparing the upper part in Table 1.5 with the lower part, we can see the preference heterogeneity across gender. Normal weight and overweight females are in general more likely to participate in the program (i.e., regardless of payment timing and form, table 1.5 shows more average probabilities that are $\succ$ 50% and the minimum amount of incentives that achieve the $\succ$ 50% probability are smaller). This trend is reversed when comparing obese male and obese female. The group average participation probabilities for obese males are generally higher than obese females and more incentive packages can promote higher than 50% participation probabilities among obese males.

Examining preference heterogeneity within gender but across weight status (i.e., comparing rows within either upper or lower part of Table 1.5, it shows that males generally shows more interests in the program as the weight status increases. However, this trend is not clear when we look at females. Furthermore, normal weight females even exhibits higher preference to the program compared to overweight and obese females. Normal weight or overweight females in general can accept any payment timing of any form at lower magnitudes compared to normal weight or overweight males. This trend is reversed when the comparison is made between obese females and obese males.

1.6 Discussion

Obesity is costly both to individuals and to the society. Participation rate (i.e., reach) of weight-loss programs has been relatively low which limits the potential public health impacts. Financial incentives are common tools employed in weight loss programs to increase program reach and adherence. Studies that have examined the effectiveness of financial incentives suggest that financial incentives significantly increase participation, but not weight loss (Kane et al. 2004b, Kane et al. 2004a, Wall et al. 2006, Paul-Ebhohimhen and Avenell 2008). However, these findings may be driven by the lack of a systematic selection of the magnitude, form, and timing of the financial incentives employed (Campbell et al. 2000). This study fills the gap of the literature by systematically evaluating the effectiveness of financial incentives in optimizing the reach of a minimal weight-loss program. Three attributes of the financial incentives were studied (i.e., magnitude, timing and form of the payments) while the reward contingency condition was fixed (i.e., the individual will be eligible for financial incentive reward if he/she achieves and maintains 5% weight loss compared to the baseline weigh-in). The weight-loss program was kept at minimal (i.e., no intensive counseling, group interaction etc.) to avoid confounding the financial incentive effectiveness.

Economic ABC methods were employed to elicit individual preference towards different incentive attributes. Fixed parameter conditional logit and random parameter logit models were estimated. Results show that individual preference towards financial incentive attributes differs not only across gender and weight status but also beyond them. Financial incentive magnitude clearly has positive impact on the perceived program participation. For a given magnitude and timing, incentive payments in the form of cash or gift card are favored over pre-paid gym pass and insurance co-pay waiver. This shows that the flexibility of incentive payment plays an important role in the program reach. Individuals' preference toward immediate reward and recognition shows clearly through the popularity of payment at each weigh-in. Examining these three incentive attributes systematically shows that a well-calibrated incentive component will go a long way in achieving high program reach while keep costs manageable. Results show that similar program reach can be achieved with lower amount of payment per weigh-in if the payment is in the form of cash or gift card and the timing is at each weigh-in. Furthermore, males and females exhibit different preference. Normal weight and overweight females are more likely to participate in the program compared to the normal weight and overweight males. However, obese males are more likely to say yes to the program compared to obese females. Customizing the financial incentive component by gender and weight status accordingly will help achieving the optimal program reach.

Due to the hypothetical nature of the survey data we collected, our study is limited to understand the influence of incentive attributes on potential program reach and cannot answer the impacts on program effectiveness on actual weight-loss outcomes and maintenance. Meanwhile, the reach optimal attributes combination identified in this study may not be the packages that optimize the program effectiveness. Despite this limitations, this study is the first one to systematically investigate the interaction among different incentive attributes and the program participation. Results provide the much needed data for calibrating financial incentive components in weight-loss interventions.

Appendix A

Section B. In this section we describe a weight loss program and we would like to learn your thoughts on the program.

We want you to consider a weight-loss program that has been shown to be effective and does not require a lot of time or meetings. This is a 3-month program to get people started on weight loss. The program, when followed closely, helps people lose 1 to 2 lbs per week, and successful participants can lose between 12 and 24 lbs over the 3-month period. The program components are:

One-on-one meeting with a dietitian: Participants have an initial 1-hour meeting with a registered dietitian to develop a plan for eating and exercising to help them lose weight and keep the weight off.

Eating Plan: Participants will receive a workbook that includes the detailed eating plan discussed with the dietitian. The eating plan will be tailored specifically to the participants' needs and will include recipes to help participants to achieve their goals. The program will include information on appropriate calorie intake and focus on balancing protein, fat, and carbohydrates.

Physical Activity Plan: Participants will receive a workbook that includes the detailed physical activity plan discussed with the dietitian. The physical activity plan will be tailored specifically to the participants' ability and needs. The program will include regular physical activity that could range from walking 5 or more days per week and simple strength exercises to more moderate and vigorous activity and weight training.

Tools to Keep track of eating: The program will include tracking tools (available online and in papers) for foods and calories eaten. Participants will be able to enter in the food they eat in a simple way and create meals. These tools will allow participants to review progress and plan for the future.

Tools to Keep track of physical activities: The program will include tracking tools for monitoring physical activity. Participants will be able to enter the amount and type of activities they do every day and get a record of the number of calories expended. These tools will allow participants to review progress and plan for the future and are available online and in papers.

Coaching Calls: During the 3-month program, participants will receive weekly coaching calls to help them keep up with their diet and exercise plans.

Program Weigh-ins: Participants will be asked to weigh-in at the program location (for example a local clinic) once a month during the 3-month program.

Booster Weigh-ins: After the program ends, participants will be asked to return for two follow-up weigh-ins (6-month and 12-month) with the goal of tracking the weight maintenance progress.

Figure 1: Program description in the questionnaire

Section C. In this section we describe incentives for participating in the weight-loss program and would like to know your opinions on the incentives.

To encourage participation and motivate weight loss sometimes incentive packages are added to this type of program.

Losing 5% of body weight for overweight and obese individuals will greatly reduce the risk of chronic disease. For all of the examples we provide below, participants will receive incentives if they achieve or maintain **a 5% weight loss** (about 10 pounds for a 200 pound person) when compared to their **initial** weight. For example, a participant joins the program and weights 200 lbs in the initial weigh-in; this participant weighs 190 lbs in the 3-month weigh-in which means he/she achieves the 5% weight loss goal. If he/she weighs no more than 190 lbs in the two booster weigh-ins, he/she achieves the 5% weight loss maintenance goal.

In our examples that follow, we will ask you to consider several different incentives are being considered:

A. <u>Monetary Value</u> of Incentives: $5, $24, $55, $72, or $98.

B. <u>Form</u> of the Incentives: cash, pre-paid grocery gift cards, gym passes, or waivers of copays for doctor visits.

C. <u>Timing</u> of the Incentive Payments:

- received only at the end of the program weigh-in (3-month) and no incentives for the two booster weigh-ins (6-month and 12-month);
- earned and received at each of the 3-month, 6-month and 12-month weigh-ins; or
- earned at each of the 3-month, 6-month and 12-month weigh-ins but only received at the 12-month weigh-in.

Figure 2: Description of incentives in the questionnaire

9. Please consider the following two weight loss programs.

	Program A	**Program B**
Monetary value of incentive received per weigh-in	$24	$55
Form of incentive	Gym passes	Waivers of copays for doctor visits
Timing of incentive payment	Pay at the last booster weigh-in (12-month)	Pay at the end of the program weigh-in (3-month) and no rewards for the other two booster weigh-ins
Total Maximum Amount Earned	$24 + $24 + $24 = $72 worth of gym passes	$55 + $0 + $0 = $55 worth of copay waivers

Which weight-loss program would you choose to participate in?
(PLEASE CHECK ONE BOX)

☐1 I would choose Program A.

☐2 I would choose Program B.

☐3 I would not choose either program.

Figure 3: Sample of a choice situation in the questionnaire

Table 1.1: Financial Incentive Attribute Levels

Attribute	Level	Wording
Monetary value	1	$5
	2	$24
	3	$55
	4	$72
	5	$98
Form of incentive	1	Cash
	2	Grocery gift cards
	3	Gym passes
	4	Waivers of copays for doctor visits
Timing of payment	1	Pay at the last booster weigh-in (12-month)
	2	Pay at the end of the program weigh-in (3-month)
	3	Pay at each weigh-in

Table 1.2: Summary Statistics

	Percent of male respondents (n=273)	Percent of female respondents (n=581)	Percent of total respondents (n=863)
Race			
white	87	83	84
black	7	13	11
other	6	4	5
Age (in years)			
18-29	4	5	5
30-39	4	11	9
40-49	19	20	20
50-59	31	25	27
≥ 60	42	38	39
Education			
high school and less	27	28	28
some college	24	32	29
college graduate	25	22	23
postgrad degree	24	18	20
Employment status			
unemployed	40	48	46
part-time employed	5	13	10
full-time employed	55	39	44
Income			
30k or less	15	22	20
30k-75k	30	42	38
75k-120k	32	22	25
120k or more	22	14	17
BMI			
Normal	16	25	22
Overweight	49	35	39
Obese	35	41	39

Table 1.3: Estimation Results

	Conditional Logit		Mixed Logit			
	Model (1)	Model (2)	Model (3)		Model (4)	
			mean	sd	mean	sd
asc	-1.126**	-1.840**	-0.052	4.804**	-3.339**	
	(0.109)	(0.181)	(0.272)	(0.307)	(9.52)	
monetary value	0.014**	0.014**	0.022**		0.026**	
	(0.001)	(0.001)	(0.001)		(15.25)	
Pay form (Base: Copay waiver)						
cash	0.865**	0.860**	1.188**	0.696**	1.246**	2.586**
	(0.080)	(0.075)	(0.105)	(0.233)	(8.28)	(10.62)
gift card	0.900**	0.882**	1.244**	0.952**	1.242**	2.974**
	(0.085)	(0.075)	(0.109)	(0.215)	(7.40)	(11.42)
gym pass	0.150	0.137	-0.125	1.926**	-0.224	2.876**
	(0.086)	(0.078)	(0.134)	(0.219)	(1.27)	(10.36)
Timing of payment (Base: payment at 12 month)						
3 month	-0.279**	-0.281**	-0.430**	0.450	-0.789**	2.005**
	(0.064)	(0.063)	(0.090)	(0.254)	(5.35)	(9.39)
each weigh-in	0.391**	0.397**	0.636**	0.000	0.491**	1.721**
	(0.061)	(0.060)	(0.090)	(0.430)	(3.96)	(10.33)
Gender Interaction						
asc*female		0.745**			1.697**	
		(0.185)			(4.76)	
BMI interaction (Base: Normal individuals)						
asc*overweight		0.607**			1.336**	
		(0.186)			(3.73)	
asc*obese		0.917**			1.892**	
		(0.199)			(5.09)	
asc*overweight*female		-0.561*			-1.529**	
		(0.225)			(3.63)	
asc*obese*female		-0.879**			-2.024**	
		(0.234)			(4.71)	
Log likelihood	-3,326	-3,231	-2,644		-2,984	
Pseudo-Rsquared	0.09	0.1	0.3		0.2	
Observations	10,029	9,726	10,029		9,726	

Robust standard errors in parentheses

* significant at 5%; ** significant at 1%

Table 1.4: Participation Preference Heterogeneity across Gender and Weight Status

Conditional Logit Model							
		Normal Weight		Overweight		Obese	
		Male	Female	Male	Female	Male	Female
Normal Weight	Male	-	0.745	0.607	0.791	0.917	0.783
			(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
	Female		-	-0.138	0.046	0.172	0.038
				(0.313)	(0.718)	(0.267)	(0.755)
Overweight	Male			-	0.184	0.31	0.176
					(0.149)	(0.046)	(0.153)
	Female				-	0.126	-0.008
						(0.072)	(0.946)
Obese	Male					-	-0.134
							(0.351)
	Female						-

Random Parameter Logit Model							
		Normal Weight		Overweight		Obese	
		Male	Female	Male	Female	Male	Female
Normal Weight	Male	-	1.697	1.336	1.503	1.892	1.565
			(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
	Female		-	-0.361	-0.193	0.196	-0.131
				(0.130)	(0.367)	(0.443)	(0.534)
Overweight	Male			-	0.168	0.557	0.23
					(0.450)	(0.034)	(0.294)
	Female				-	0.389	0.062
						(0.106)	(0.746)
Obese	Male					-	-0.327
							(0.169)
	Female						-

Table 1.5: Incentivized Weight-Loss Program Reach Percentage Prediction

		Pay at 12 month				Pay at 3 month				Pay at each weigh-in			
	Monetary incentive	Cash	Gift card	Gym pass	Copay waiver	Cash	Gift card	Gym pass	Copay waiver	Cash	Gift card	Gym pass	Copay waiver
Normal male	5	12%	12%	3%	4%	6%	6%	0%	2%	19%	19%	5%	0%
	24	19%	19%	5%	6%	9%	9%	2%	3%	27%	27%	8%	10%
	55	34%	34%	11%	13%	19%	19%	5%	6%	46%	45%	16%	19%
	72	44%	44%	15%	19%	27%	26%	8%	9%	57%	56%	23%	27%
	98	61%	61%	26%	31%	41%	41%	14%	17%	72%	72%	37%	42%
Overweight male	5	35%	35%	11%	13%	20%	19%	0%	7%	47%	46%	17%	0%
	24	47%	47%	17%	20%	28%	28%	8%	10%	59%	59%	25%	29%
	55	66%	66%	31%	36%	47%	47%	17%	20%	76%	76%	42%	48%
	72	75%	75%	41%	47%	58%	58%	24%	28%	83%	83%	53%	59%
	98	86%	86%	58%	63%	73%	73%	38%	44%	91%	91%	69%	74%
Obese male	5	48%	48%	18%	21%	30%	30%	0%	11%	60%	60%	26%	0%
	24	60%	60%	26%	30%	41%	41%	14%	17%	71%	71%	36%	42%
	55	77%	77%	44%	49%	61%	61%	26%	31%	85%	85%	56%	62%
	72	84%	84%	55%	60%	71%	71%	36%	41%	90%	90%	66%	71%
	98	91%	91%	70%	75%	82%	82%	52%	58%	94%	94%	80%	83%
Normal female	5	43%	43%	15%	18%	26%	26%	0%	9%	56%	55%	22%	0%
	24	56%	56%	22%	26%	36%	36%	12%	14%	67%	67%	32%	37%
	55	74%	74%	39%	45%	56%	56%	23%	27%	82%	82%	51%	57%
	72	81%	81%	50%	56%	66%	66%	31%	36%	88%	88%	62%	67%
	98	89%	89%	66%	71%	79%	79%	47%	53%	93%	93%	76%	80%
Overweight female	5	39%	39%	13%	15%	22%	22%	0%	8%	51%	51%	19%	0%
	24	51%	51%	19%	23%	32%	32%	10%	12%	63%	63%	28%	33%
	55	70%	70%	35%	40%	51%	51%	19%	23%	79%	79%	46%	52%
	72	78%	78%	45%	51%	62%	62%	27%	32%	85%	85%	57%	63%
	98	88%	87%	62%	67%	76%	76%	42%	48%	92%	92%	73%	77%
Obese female	5	40%	40%	13%	16%	23%	23%	0%	8%	52%	52%	20%	0%
	24	52%	52%	20%	24%	33%	33%	10%	13%	64%	64%	29%	34%
	55	71%	71%	36%	41%	53%	53%	20%	24%	80%	80%	48%	54%
	72	79%	79%	47%	52%	63%	63%	28%	33%	86%	86%	59%	64%
	98	88%	88%	63%	68%	77%	77%	44%	49%	92%	92%	74%	78%

Bibliography

Aldana, S. G. and N. P. Pronk (2001). Health Promotion Programs , Modifiable Health Risks , and Employee Absenteeism. *Journal of Occupational and Environmental Medicine 43*, 36–46.

Arena, V. C., K. R. Padiyar, W. N. Burton, and J. J. Schwerha (2006). The impact of Body Mass Index on short-term disability in the workplace. *J Occup Environ Med 48*, 1118–1124.

Bingham, M. F., F. R. Johnson, and D. Miller (2001). Modeling Choice Behavior for New Pharmaceutical Products. *Response 4*(1), 32–44.

Black, D. R. and R. Cameron (1997). Self-administered interventions : a health education strategy for improving population health. *Health Education Research 12*(4), 531–545.

Booske, B. C., F. Sainfort, and A. S. Hundt (1995). Eliciting Consumer Preferences for Health Plans. *Health Services Research 34*(4), 839–854.

Burton, W. N., C. Y. Chen, D. J. Conti, A. B. Schultz, G. Pransky, and D. W. Edington (2005). The association of health risks with on-the-job productivity. *J Occup Environ Med 47*, 769–777.

Campbell, M., R. Fitzpatrick, A. Haines, A. L. Kinmonth, P. Sandercock, D. Spiegelhalter, and P. Tyrer (2000). Framework for design and evaluation of complex interventions to improve health. *BMJ 321*, 694–696.

Chakraborty, G., R. Ettenson, and G. Gaeth (1994). How consumers choose health insurance. *Journal of Health Care Marketing 14*(1), 21.

Chakraborty, G., J. G. Gary, and M. Cunningham (1993). Understanding consumers ' preferences for dental service. *Journal of Health Care Marketing 13*(3), 48.

Charness, G. and U. Gneezy (2009). Incentive to Exercize. *Econometrica 77*(3), 909–931.

Cherkin, D., L. Grothaus, and E. Wagner (1990). The effect of office visit copayments on preventive care services in an HMO. *Inquiry 27*(1), 24–38.

Cunningham, M., G. Gaeth, J. C, and C. G (1999). Using choice-based conjoint to determine the relative importance of dental benefit plan attributes. *J Dent Educ 63*(5), 391–399.

Curry, S. J., E. H. Wagner, and L. C. Grothaus (1991). Evaluation of Intrinsic and Extrinsic Motivation Interventions With a Self-Help Smoking Cessation Program. *Journal of Consulting and Clinical Psychology 59*(2), 318–324.

Donatelle, R. J., S. L. Prows, D. Champeau, R. J. Donatelle, S. L. Prows, D. Champeau, and D. Hudson (2000). Randomised controlled trial using social support and financial incentives for high risk pregnant smokers : Significant Other Supporter (SOS) program Randomised controlled trial using social support and financial incentives for high risk pregnant smokers. *Tobacco Control 9*(suppl 3), III67–III69.

Doody, M. M., A. S. Sigurdson, D. Kampa, K. Chimes, B. H. Alexander, E. Ron, R. E. Tarone, and M. S. Linet (2003). Randomized Trial of Financial Incentives and Delivery Methods for Improving Response to a Mailed Questionnaire. *American Journal of Epidemiology 157*(7), 643–651.

Epstein, D. H., W. E. Hawkins, L. Covi, A. Umbricht, and K. L. Preston (2003). CognitiveBehavioral Therapy Plus Contingency Management for Cocaine Use: Findings During Treatment and Across 12-Month Follow-Up. *Psychol Addict Behav 17*(1), 73–82.

Finkelstein, E., I. Fiebelkorn, and G. Wang (2003). National medical spending attributable to overweight and obesity: how much, and who's paying? *HEALTH AFFAIRS-MILLWOOD VA THEN BETHESDA MA- 22*(3; SUPP), 3219.

Finkelstein, E. A., M. Dibonaventura, S. M. Burgess, and B. C. Hale (2010). The Costs of Obesity in the Workplace. *Journal of Occupational & Environmental Medicine 52*(10), 971–976.

Finkelstein, E. A., L. A. Linnan, D. F. Tate, and B. E. Birken (2007). A Pilot Study Testing the Effect of Different Levels of Financial Incentives on Weight Loss Among Overweight Employees. *Occupational and Environmental Medicine 49*(9), 981–989.

Glasgow, R. E., S. S. Bull, C. Gillette, L. M. Klesges, and D. A. Dzewaltowski (2002). A Review of Recent Reports with Emphasis on External Validity. *American Journal of Preventive Medicine 23*(1), 62–69.

Glasgow, R. E. and K. M. Emmons (2007). How Can We Increase Translation of Research into Practice ? Types of Evidence Needed. *Integration The Vlsi Journal 28*, 413–33.

Glasgow, R. E., L. M. Klesges, D. A. Dzewaltowski, S. S. Bull, and P. Estabrooks (2004). The Future of Health Behavior Change Research : What Is Needed to Improve Translation of Research Into Health Promotion Practice ? *Annals of Behavioral Medicine 27*, 3–12.

Hall, J., P. Kenny, M. King, J. Louviere, R. Viney, and A. Yeoh (2002). Using stated preference discrete choice modelling to evaluate the introduction of varicella vaccination. *Health Economics 465*(April), 457–465.

Harris, K., J. Schultz, and R. Feldman (2002). Measuring consumer perceptions of quality differences among competing health benefit plans. *Journal of Health Economics 21*, 1–17.

Helpern, S. D., P. A. Ubel, J. A. Berlin, and D. A. Asch (2002). Randomized Trial of $5 Versus $10 Monetary Incentives, Envelope Size , and Candy to Increase Physician Response Rates to Mailed Questionnaires. *Medical Care 40*(9), 834–839.

Heshka, S., J. W. Anderson, R. L. Atkinson, F. L. Greenway, J. O. Hill, S. D. Phinney, and K. Miller-kovach (2003). Weight Loss With Self-help Compared With a Structured Commercial Program: A Randomized Trial. *JAMA 289*(14), 1792–1798.

Higgins, S. T., S. H. Heil, L. J. Solomon, I. M. Bernstein, J. P. Lussier, R. L. Abel, M. E. Lynch, and G. J (2004). A pilot study on voucher-based incentives to promote abstinence from cigarette smoking during pregnancy and postpartum. *Nicotine & Tobacco Research 6*(6), 1015–1020.

Kane, R., P. Johnson, R. Town, and M. Butler (2004a). A structured review of the effect of economic incentives on consumers' preventive behavior. *American journal of preventive 27*, 327–352.

Kane, R., P. Johnson, R. Town, and M. Butler (2004b). Economic incentives for preventive care. *Evid Rep Technol Assess (Summ) 101*, 17.

Kane, R. L., P. E. Johnson, and R. J. Town (2004). Economic Incentives for Preventive Care, AHRQ Publication No. 04-E024-2.

Kenkel, D. S. (1994). The demand for preventive medical care. *Applied Economics 26*(4), 313–325.

Kosten, T., A. Oliveto, A. Feingold, J. Poling, K. Sevarino, E. Mccance-katz, S. Stine, G. Gonzalez, and K. Gonsai (2003). Desipramine and contingency management for cocaine and opiate dependence in buprenorphine maintained patients. *Drug and Alcohol Dependence 70.*

Lillard, L., W. Manning, C. Peterson, N. Lurie, G. Goldberg, and C. Phelps (1986). *Preventive medical care: standards, usage, and efficiency.* Santa Monica, CA: Rand Corporation.

Louviere, J., D. Hensher, and J. Swait (2000). *Stated choice methods: analysis and applications.* Cambridge Univ Pr.

Lussier, J. P., S. H. Heil, J. A. Mongeon, G. J. Badger, and S. T. Higgins (2006). A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction 101*, 192–203.

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behaviour*, pp. 105–142. New York: Academic Press.

Moayyedi, P., M. Wardman, J. Toner, M. Ryan, and S. Duffett (2002). Establishing patient preferences for gastroenterology clinic reorganization using conjoint analysis. *European Journal of Gastroenterology & Hepatology 14*, 429–433.

Mokdad, A. H., M. K. Serdula, W. H. Dietz, B. A. Bowman, J. S. Marks, J. P. Koplan, and P. Page (1999). The Spread of the Obesity Epidemicin the United States , 1991-1998. *October 282*(16), 1991–1994.

Must, A., J. Spadano, E. Coakley, A. Field, G. Colditz, and W. Dietz (1999). The disease burden associated with overweight and obesity. *Jama 282*(16), 1523.

Paterson, R. W., K. J. Boyle, C. F. Parmeter, J. E. Neumann, and P. D. E. Civita (2008). HETEROGENEITY IN PREFERENCES FOR. *Health Economics 1377*(January), 1363–1377.

Paul-Ebhohimhen, V. and A. Avenell (2008). Systematic review of the use of financial incentives in treatments for obesity and overweight. *Obes Rev 9*, 355–367.

Petry, N. M., J. Tedford, M. Austin, C. Nich, K. M. Carroll, B. J. Rounsaville, and N. M. Petry (2004). Prize reinforcement contingency management for treating cocaine users : how low can we go , and with whom ? *Addiction 99*, 349–360.

Pol, M. V. D. and J. Cairns (2001). Estimating time preferences for health using discrete choice experiments. *Social Science & Medicine 52*, 1459–1470.

Revelt, D. and K. Train (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. *The Review of Economics and Statistics 80*, 647–657.

Ricci, J. A. and E. Chee (2005). Lost productive time associated with excess weight in the US workforce. *J Occup Environ Med 47*, 1227–1234.

Ryan, M. and J. Hughes (1997). Using conjoint analysis to assess women's preferences for miscarriage management. *Health Economics 6*(3), 261273.

Sassi, F., D. McDaid, and W. Ricciardi (2005). Conjoint analysis of preferences for cardiac risk assessment in primary care. *International journal of technology assessment in health care 21*(02), 211218.

Shoptaw, S., E. Rotheram-fuller, X. Yang, D. Frosch, D. Nahom, M. E. Jarvik, R. A. Rawson, and W. Ling (2002). Smoking cessation in methadone maintenance. *Addiction 97*, 1317–1328.

Sindelar, J. L. (2008). Paying For Performance : The Power Of Incentives Over Habits. *Health Economics 451*, 449–451.

Stevens, J., D. Taber, D. Murray, and D. Ward (2007). Advances and Controversies in the Design of Obesity Prevention Trials&ast. *Obesity 15*, 2163–2168.

Swait, J. and J. Louviere (1993). The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing 30*(3), 305–314.

Thorpe, K., C. Florence, D. Howard, and P. Joski (2004). The impact of obesity on rising medical spending. *HEALTH AFFAIRS-MILLWOOD VA THEN BETHESDA MA- 23*, 283283.

Train, K. (1998). Recreation Demand Models with Taste Differences over People. *Land Economics 74*(2), 230–239.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge Univ Pr.

Tsai, S. P., J. K. Wendt, F. S. Ahmed, R. P. Donnelly, and T. R. Strawmyer (2005). Illness Absence Patterns Among Employees in a Petrochemical Facility: Impact of Selected Health Risk Factors. *Journal of Occupational and Environmental Medicine 47*, 838–846.

Vick, S. and A. Scott (1998). Agency in health care . Examining patients preferences for attributes of the doctor patient relationship. *Journal of Health Economics*, 587–605.

Volpp, K. G., L. K. John, A. B. Troxel, L. Norton, J. Fassbender, and G. Loewenstein (2008). Financial Incentive Based Approaches for Weight Loss Trial, A Randomized Trial. *JAMA 300*(22), 2631–2637.

Wall, J., C. Mhurchu, T. Blakely, A. Rodgers, and J. Wilton (2006). Effectiveness of monetary incentives in modifying dietary behavior: a review of randomized, controlled trials. *Nutrition reviews 64*(12), 518531.

Wellman, N. S. and B. Friedberg (2002). Causes and consequences of adult obesity : health , social and economic impacts in the United States. *Asia Pacific J Clin Nutr 11*, S705S709.

Wordsworth, S. and A. Scott (2002). Ultrasound scanning by general practitioners : is it worthwhile ? *Journal of Public Health Medicine 24*(2), 88–94.

Wordsworth, S., D. Skå tun, A. Scott, and F. French (2004). Preferences for general practice jobs : a survey of principals and sessional GPs. *Br J Gen Pract 54*(May), 740–746.

Zamosky, L. (2010). Employers ready to raise the stakes for health incentives.

Chapter 2

From health service delivery to family planning: the changing impact of health clinics on fertility in rural Iran

[1]

[1] This is a joint work with Djavad Salehi-Isfahani and is under review in the Economic Development and Cultural Change.

2.1 Introduction

Iran's fertility decline in the 1990s is something of a record in the annals of demographic transition. In its recent survey of world fertility decline, the *Economist* noted that Iran's fertility decline is "about as fast as it can happen." Although urban fertility is now below replacement, the more impressive aspect of Iran's transition is fertility decline in rural areas. Whereas attempts to lower rural fertility before the revolution had failed (Aghajanian 1995), after the revolution the drop in rural fertility broke the world record: in the span of 15 years the fertility fell from about 8 to 2 births.

This impressive decline in fertility in Iran is usually identified with the family planning program launched in 1989 (Abbasi-Shavazi et al. 2009). Mehryar et al. 2001 attribute the decline, which they call a 'miracle', to this program. Before then, having discontinued the Shah's program,[2] the government had shown no interest in family planning, and adopted pro-natal policies that promoted larger families. Following the surprising results of the 1986 census that revealed a record rate of population growth of 3.9 percent since 1976 (3.6 percent not including Afghan and Iraqi refugees) the government reversed its pro-natal policies and adopted family planning. The baby boom occurring around the time of the revolution in 1979 had caused unusually large cohorts of children entering primary school in the mid 1980s, forcing schools into two and three shifts. The timing of the family planning program and the turnabout from pro- to anti-natal policies in 1989 was also influenced by the fact that the war had ended the year before, shifting the government's attention to millions of youth in need of schooling and in search of jobs. The family planning program was thus an integral part of the reconstruction effort pursued by the Rafsanjani administration that took power in 1989. Rural areas figured prominently in the reconstruction effort, because of their support for the war effort and because their disadvantaged position symbolized the Shah's policies (Salehi-Isfahani 2009a, Salehi-Isfahani 2009b). In 1985, as part of its rural development strategy, the government had begun the rural Health Network System (HNS), which expanded the construction of rural health houses rapidly around the country, and was later used as the institutional and physical base for the family planning program. In this paper we evaluate the impact of exposure to the health clinic services and particularly this program on rural fertility.

Understanding the role of health clinics in the spectacular decline in rural fertility in Iran is important as this program gains respect as a model for other developing countries (Boonstra 2001). The evidence of the causal impact of Iran's program on rural fertility is still being accumulated. Salehi-Isfahani et al. 2010 employ a difference-in-differences method to measure the impact of the program and find that 4 to 20% of the decline in rural fertility can be attributed to the presence of rural health clinics. While their methodology is strong in terms of identification, their measure of fertility, the village-level child-woman ratio, is an imprecise measure of fertility because it is susceptible to changes in child mortality and migration.

In this paper we use micro data from the 2000 Iran Demographic and Health Survey (IDHS) to examine the response of women of given characteristics – e.g., age and education – to exposure to the program. In particular, we are interested in how access to health houses (availability of birth control), schools, and other village infrastructure has affected the timing of births of different

[2] For a description of this program, see (Moore 2007); for its effect on fertility, see (Raftery et al. 1995).

parities and women of different ages and education. Our work builds on the village level study of Salehi-Isfahani et al. 2010 by identifying the impact of the family planning program at the individual level and by parity, thus giving us a more precise notion of program impact. Our findings support their general results, that the expansion of family planning services contributed to lower fertility in rural Iran. However, measuring the impact of exposure to clinics for different births in different time periods shows even a weaker effect than what they estimated. We could not see any significant impact for the first two births. The main effect of the program is limited to the third birth. It is worth mentioning that we also tried higher order births (up to sixth birth). Although we were able to capture a sharper decline in fertility of these parities in the last two periods after the introduction of family planning program, we could not observe a significant impact of exposure to health clinics on these births. This could be related to the higher level of censoring for higher order births. Thus we focus on the first third births.

To identify program effect, we take advantage of the geographic spread of health houses across rural areas. Concern for endogenous placement arises from the tendency of administrators to give priority to areas in greater need of program services. Although Iran's program did not have this particular bias –areas with lower fertility received their clinic earlier– it was not totally random either. As Salehi-Isfahani et al. 2010 show, Iran's family program proceeded from areas with better local infrastructure, such as roads, electricity and water, to less developed areas. To control for placement endogeneity, we use fixed effects at the district level, which removes the effect of all time-invariant factors including those unobservables which affected both placement and fertility.

We use micro data derived from Iran's 2000 Demographic and Health Survey (IDHS) to estimate a hazard model of birth by parity. IDHS includes data on about 114,000 women of childbearing age, of whom 54,000 are rural. We are able to link a particular birth for each woman to the level of program exposure at her district of residence at the time of birth. We measure exposure to the program by the number of health clinics per 1000 women of childbearing age (each health house is designed to serve approximately 1500 women).

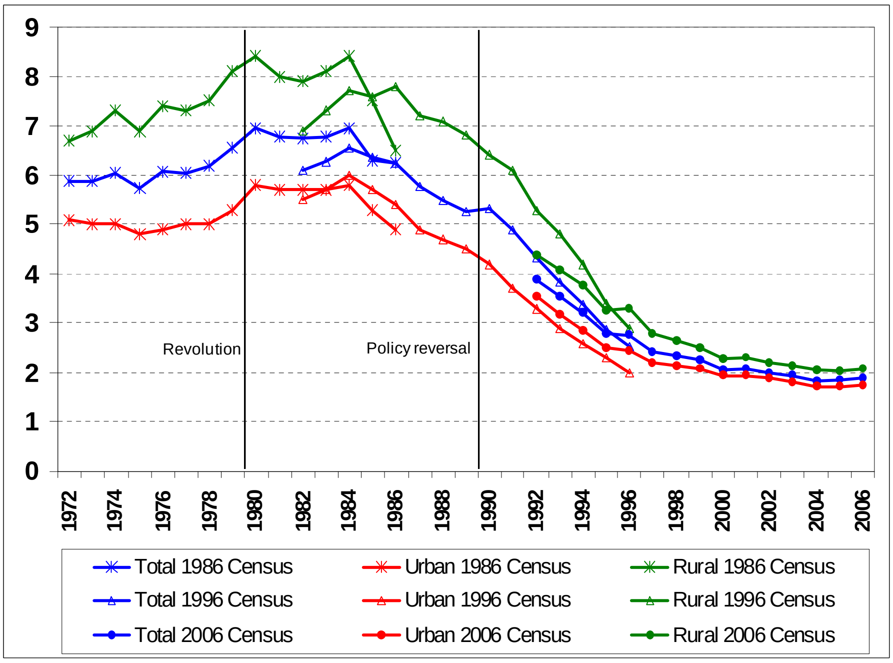
In the rest of this paper we first review the evidence for fertility transition, showing the extent of rural fertility decline and its timing. We then describe Iran's family planning program in more details. In section 2.4 we present the methodology used in this study, especially how we deal with the issue of identification of program effect. In section 2.5 we introduce the data we use in this study. Section 2.6 presents the results of discrete-time hazard regressions of births by parity and section 2.7 is the conclusion.

2.2 Fertility transition

A well-established literature, reviewed in Abbasi-Shavazi et al. 2009, shows that fertility transition in Iran began sometime in the mid 1980s and was completed less than two decades later. Estimates of the fertility transition in rural Iran have been derived from census data by Abbasi-Shavazi and McDonald 2006, which is reproduced in Figure 2.1. Two interesting observations can be made with the help of this graph. First, the decline in rural fertility has been faster than urban fertility, so that the gap in fertility between the two areas was all but closed by 2006. In a period of about 15 years the average number of birth per rural woman, the total fertility rate

(TFR), declined from more than 7 to about 2, compared to a decline of about 4 births for urban women. The fact that such an impressive decline has occurred in Iran's rural areas, where families are more conservative and less educated, is probably the reason why most observers have assumed a causal link between the rural program and fertility decline. There are certain aspects of the family planning program that were national in impact, such as the removal of the implicit child subsidies, and there are aspects that are specific to rural areas, most importantly, active delivery. The distinction between active and passive service delivery is based on whether the clinics offer their services to those who seek the service or whether health workers seek out individual women. In rural Iran, health workers based in the clinics visited each woman of child bearing age at least once a year to take note of her health and reproductive history (Abbasi-Shavazi et al. 2009). The narrowing of the fertility gap may be attributed to these aspects of the program but not to the shared aspects.

Figure 2.1: Total Fertility Rates, 1972-2006



Source: Abbasi-Shavazi et al. 2009

Second, as Figure 2.1 shows, fertility decline in both rural and urban areas began before the introduction of family planning. This suggests that other factors besides the family planning program were acting on fertility at the time. In 1989 there were several changes taking place simultaneously that could have affected fertility. The rationing of basic goods introduced during the war which had benefited larger families were on their way out. There was talk of ending of other subsidies, mainly in schooling and health, for higher parity children. By 1993, a law limited the benefits from various government programs to three children (Roudi-Fahimi 2002). Educational opportunities in rural areas were expanding, raising the expectation of higher returns to schooling for the rural population, which could limit the desire for larger families (Becker 1992).

Our focus in the empirical section of this paper is on married women. In principle, family planning programs influence both the prevalence of marriage as well as marital fertility. However, in Iran, there are strong social and legal taboos against relations between men and women outside marriage. So, we expect that a large part of the program effect is to be searched among married women. Abbasi-Shavazi et al. 2009 show that of the decline of 3.71 births during 1986-96, 3.11 births, or 84 percent, was due to decrease in marital fertility. The implication of this for our results is that we do not expect family planning to have had a strong effect on the timing of the first birth (see section 2.6).

Iran's program appears to have affected higher order births only (Abbasi-Shavazi et al. 2009). In Iran, especially in rural areas, marriage is a rite of passage to adulthood, which is completed with the birth of the first. In Figure 2.2 we present the average number of years a couple wait between marriage and first birth, first and second, and second and third birth. The waiting time to first birth does not show a particular trend; in particular there is no increase after family planning went into effect. The time to second and third birth show a distinct upward trend, consistent with the decline in fertility over time. In our empirical results we allow for differences in program effect by parity.

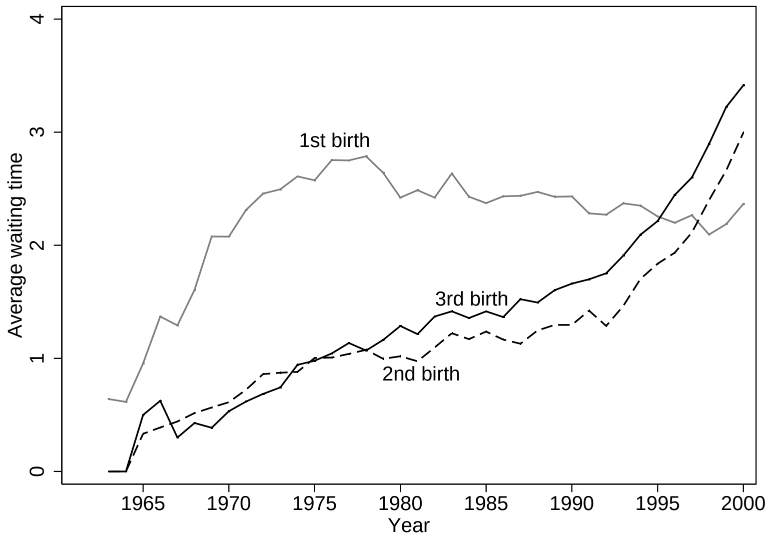


Figure 2.2: Average waiting time by parity

2.3 Family Planning program

Health care and family planning are delivered to Iran's rural areas through the Health Network System (HNS), at the core of which are rural health clinics.[3] Health house construction in rural

[3]We use the terms health clinics and health houses interchangeably. For descriptions of the HNS see (Roudi-Fahimi 2002, Abbasi-Shavazi et al. 2009, Salehi-Isfahani et al. 2010).

areas started in early 1970s with few health clinics built before the revolution. Figure 2.3 shows the pace of the construction of rural health houses in Iran. The network expanded gradually before the revolution but the major jump in health house construction happened in 1985 after that ministry of health was mandated by legislature to improve the rural health infrastructure, and reached its pick in 1989, when the family planning program became law. By 2005, about 90% of the country's rural population (20.4 million individuals living in 4.2 million households) had been covered by the program [4]. The construction program implemented in one district in each of the provinces and then expanded to other districts within each province. The selection was not random and according to officials in charge of the program, placement was influenced by the local infrastructure and the availability of educated potential health workers (Salehi-Isfahani et al. 2010). These local characteristics have also a direct impact on the fertility decision of women residing in that area. We control for these district level characteristic and also imply a fixed effect estimator to deal with the problem of non-random placement. See Section 2.4 for more details.

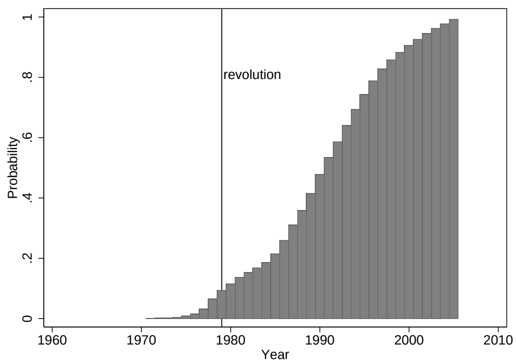


Figure 2.3: Timing of health house establishment

An important feature of Iran's rural health program is that health houses existed for a number of years before family planning was introduced, so we are able to estimate the effect of exposure to health houses in two distinct periods, with and without family planning services. We believe that this unique feature of Iran's program adds to the strength of identification. We thus distinguish between the periods before and after 1989, when health houses started provide family planning services. To further understand how the effect of health houses has changed over time, we split each period (before and after 1989) into two. The base period (1966-78), which serves as the reference period in the estimation, is before the revolution, when the family planning under the Shah was still in effect (Moore 2007). The circumstances surrounding rural family planning differs after the revolution from the earlier period in two ways. First, during this period the government suspended the existing family planning and adopted pro-natal policies with economic incentives through rationing of basic necessities that treated children equal to adults, and state encouragement of larger families. Second, during this period the revolutionary government began a rural reconstruction effort which included massive investments in rural infrastructure, especially in health house con-

[4]From total of 15,000 health clinics, 8% were built before revolution, 30% since revolution until policy reversal at 1989, 33% since the start of family planning program lunch until 1995, and 17% from 1995 to 2000

struction. About 45 percent of all the health houses in existence by the year 2000 were constructed during this period. The important difference between this period and the two following periods is that before 1989 the mandate of the rural health delivery system did not include family planning.

We divide the post 1989 period when family planning went into effect into two five year periods, 1989-1994 and 1995-2000. Our interest in doing so is to capture the effect of the increased national information campaign for smaller families, which may have had its own effect on incentives for smaller families separate from the services provided by the local clinics. Not only the state propaganda against large families intensified as the years went by, the government made its intentions to penalize large families official by enacting a law in 1993 that prevented the government from providing subsidies to the fourth and higher parity children.

2.4 Empirical model

We use a duration model to estimate the effect of access to health houses on fertility. Our outcome of interest is the probability of a birth in each year since marriage or a previous birth. In order to estimate this probability we construct birth histories for individual women using retrospective information in IDHS. Our unit of observation is thus woman-year, for which we know whether or not a birth occurs. We do this separately for each parity (or birth order). For the first birth, we follow each woman from the time of marriage until she gives birth to her first child, and for the second birth from the time of the first birth, and so on. For some women and some parities we do not observe a birth occurring before the survey year but one could have occurred after that, so the interval is open. The hazard regression method we employ is designed to deal with these censored observations. We choose to define the unit of time as a year rather than a month because we do not have confidence on accurate reporting of month of birth for older children. This allows us to use logistic regression, also known as *proportional odds model*, which is the common method for estimating discrete-time hazard models.[5]

The binary birth outcome each year is modeled as a function of own age and education and husband's education, as well as community characteristics, such as access to electricity, piped water, schools, etc. We estimate the following logistics equation:

$$ln\left[\frac{P(B_{ijt}=1|X_{ijt},COV_{jt},Z_{jt},\mu_j)}{P(B_{ijt}=0|X_{ijt},COV_{jt},Z_{jt},\mu_j)}\right] =$$
$$\sum_{k=1}^{3}\tau_k T_k + \delta COV_{jt} + \beta[COV_{jt} * \sum_{k=1}^{3} T_k] + X_{ijt}\alpha + Z_{jt}\gamma + \mu_j + \sum_{p=1}^{5}\psi_p S_p + \epsilon_{ijt},$$

where the dependent variable, B_{ijt}, takes the value of 1 if the woman i in district j gives birth at year t, and 0 otherwise. T_k's are three dummy variables indicating the period during which the birth occurs. The reference period is 1966-78 (the pre-revolution period), the dummy variables represent periods 1979-88, 1989-1994, and 1995-2000. The latter two periods are when family planning was in effect. COV_{jt} is the the measure of health house coverage at time t in district j and reports the number of health houses per 1000 rural women in each district at each year.

[5] cloglog hazard model yields similar estimates.

We expect the effect of health houses to vary depending on the period. Program impact will be reflected in the difference between the coefficients of interaction terms before and after the start of family planning. X_{ijt} are personal characteristics, such as age and education, which directly impact the fertility decision. Z_{jt} are time varying local characteristics, such as access to schools, electricity, and piped water. Time-invariant local characteristics such as religious and cultural norms are captured by μ_j. We also include spell dummies (S_p's) to account for temporal dependence of hazard.

Estimation of the program effect is complicated by the fact that distribution of health houses across the country was not completely random, even though placement did not favor high fertility areas, as programs are likely to do (Salehi-Isfahani et al. 2010). Local infrastructure and availability of qualified health workers played an important role in the placement decision, thus introducing factors in the decision to build a health house that are also correlated with fertility. For instance, paved roads, electricity, and the presence of educated women who would staff the health clinic, were obviously important considerations in building a health house in a particular location. The likely effect of these observable infrastructure factors (Z_{jt}) on placement is not detrimental because we can account for them by conditioning the probability of birth on the characteristics of the district of residence. There are also unobservable factors (μ_j), such as social norms that influence individual choice in adoption of contraceptive use, that can affect the administration's decision to implement the program in an area. The influence of unobservable factors on placement can be viewed as an omitted variable problem. Omitting the characteristics which affect placement from the fertility equation, leads to estimation bias (Strauss and Thomas 1995). To the extent that placement of health houses is not random, μ_j will be correlated with the random disturbance (ϵ_{ijt}), which causes estimation bias. The standard practice to deal with time-invariant unobserved factors is to use fixed-effects estimator to difference them out from the regression (Rosenzweig and Wolpin 1986; Pitt et al. 1993; Gertler and Molyneaux 1994; Angeles et al. 2005). The main advantage of this approach is that it takes out all unobservable time-invariant characteristics. However, a drawback of the fixed-effects procedure is that it also eliminates time-invariant observable factors, so it could be inefficient.

2.5 Data

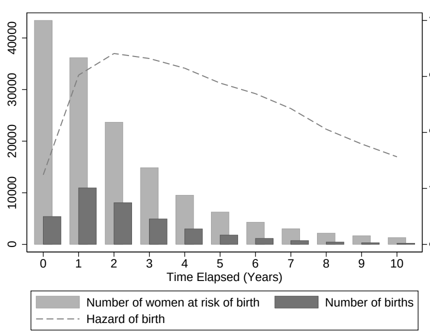
The individual level data for this study are from Iran's Demographic and Health Survey of 2000 (IDHS). IDHS is a nationally representative survey of 113,913 households in 28 provinces (plus city of Tehran). In each province, 4000 households (2000 in rural and 2000 in urban areas) selected and 90,201 ever-married women 10-49 were interviewed. In this study, the sample is restricted to the rural areas (43,279 ever-married women age 15-49)[6]. IDHS contains detailed information on fertility, family planning practices, socioeconomic characteristics of women 10-49, and also the district of residence.[7] Our working sample consists of 43,279 ever-married women aged 15-49 in

[6]From the total number of 43,813 rural women aged 10-49, we dropped 173 observations whose age were less than 15. We also dropped 365 observations whose age at marriage was missing or incorrect (age at marriage less than 10 years old).

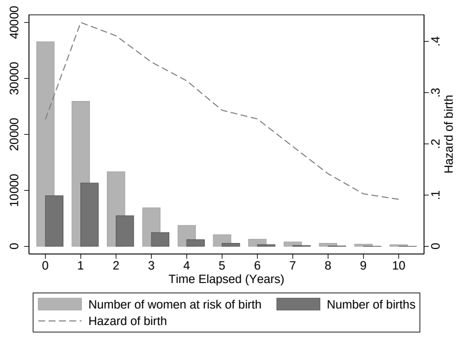
[7]For further details on the survey see (Aghajanian and Mehryar 1999).

2000 living in 193 districts.

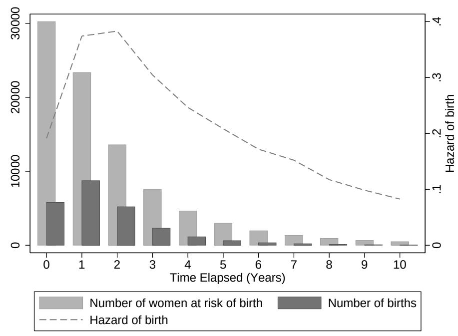
The dependent variable is an indicator variable showing whether a woman has given birth in a particular year. There are 676,994 woman-years in the retrospective data set, with a total of 161,903 births. Figure 2.5 shows the number of women at risk of birth and number of birth occurred at each spell year for each parity. The second y-axis in these graphs reports the hazard, which is simply the ratio of these two numbers at each spell year.



(a) 1st birth



(b) 2nd birth



(c) 3rd birth

Figure 2.4: Hazard rates for different parities

Table 2.1 presents descriptive statistics for the variables included in the empirical model. The left section of the table reports woman's characteristics at the survey year, 2000. The right section of this table shows the summary statistics for the restructured data set in which the unit of observation is woman-year. The restructured sample includes information on 676,994 woman-year observations of which 22% is about women who are waiting for the first birth, 14% is about those who already have a child and are waiting for the second birth, and finally 13% is about those who are waiting for the third birth.

The individual level variables include education, age, husband's education at year 2000, household's wealth, information about womans's migration in the last five years, the mortality of the last birth, and whether the woman has at least one son. The last two variables are included to take into account situations where replacement birth or sex preference are additional motives in the fertility decision.

Education is included in the model using a set of dummy variables for three education categories; illiterate, primary, and secondary levels. The survey does not report educational attainment of women at the time of each birth. To construct a time varying variable for years of schooling, we follow Angeles et al. 2005 by assuming that each woman entered school at age 6 and gained one year of schooling each year until she graduated. For instance, a woman who completed primary education (5 years of schooling) is recorded as having 3 years of education by age 9, 4 years by age 10, and 5 years from age 11 onwards. All but 1% of the married women in our sample have completed their education[8]. From the total of 43,279 ever married women of in our sample, nearly half (45%) are illiterate, 37% have completed their primary education and one in five of them (about 20%) have secondary and higher education. Husbands are generally more educated than their wives.

Our sample is limited to the women in child bearing ages (15 to 49). The average age of women in the sample is about 32 years. The average age at marriage which is the age at which women are introduced to the sample is about 17 years old.

We use an index based on ownership of various assets as a proxy for household wealth. Using the method of principal component analysis (Filmer and Pritchett 1998, Rutstein and Johnson 2004) we combine 15 asset indicators for private cars, TV, radio, refrigerator, and the like, into a single measure of household wealth.[9] The resulting index varies between -5 to +5 with the average of about -1.

Migration presents a problem for our estimation. Women who move between rural districts or between rural and urban areas may have had access to different facilities, including family planning services, than their current place of residence indicates. IDHS contains only limited information on immigration. Only for women who have moved in the last five years, the survey reports the number of years of residence in the same village. For these women, we also know where they came from (whether they have moved from rural/urban areas of the same province or another province). Fortunately, only about 11.6 percent of women in the sample had moved in the last five years, and of these only about half had moved from an urban area. Since access to education, health and family planning services differ more between rural and urban areas than between rural areas, it is the latter group that really concerns us. Very likely these women had better access to family planning services where they came from, and perhaps continued to have the same level of access after they moved to the rural district. We control for this characteristic in our estimation and find that the hazard of birth for those who moved in the last five years are significantly lower than those who did not move. This is true for the first and second parities. For the third birth, however, there is no significant difference between those who moved and those who did not.

Information on program placement includes data on year of health house establishment for 15,070 villages in which at least one health clinic were present by 2005. To obtain community characteristics, such as the presence of schools, mosque, electricity, and piped water, we use data on village facilities over time.[10] Since district is the smallest geographical division that can be identified from

[8]Using the reported highest attainment from the survey does not change the results significantly.

[9]The complete list of assets is: number of rooms per household member, car, motorcycle, TV, radio, refrigerator, telephone, bicycle, electricity, piped water, natural gas, central heating, shower, latrine, and hygienic toilet.

[10]These data are available only for the following years: 1950, 1966, 1973, 1976, 1981, and 1986. To

the IDHS, we aggregated the village characteristics to district level by defining the total number of villages in a district which uses that service (electricity, piped water) or facility (middle school, mosque) per 1000 rural residents of district. The district rural population comes from 1986 census. Health house coverage is also calculated in district level and shows the number of health houses per 1000 rural women aged 15-49. The average of this measure for our woman-year sample is about 2 health clinics per 1000 women. We also calculated a variable indicating the proportion of total population in each district who live in the mountainous areas. We used this variable as an indicator for ease of access to other areas nearby the village of residence. Since it did not show any effect on the hazard of birth in none of parities we dropped it from the model.

As noted earlier, spell dummies were included to take into account the temporal dependence of hazard function. We set as reference category for spell dummies the year of marriage or of a birth. We clock the hazard for ten periods after the event.

2.6 Estimation Results

Tables 2.2 report the estimation results using logit. Columns 1-3 present the simple logit results, in which we control for both individual and community characteristics. These results are likely to be inconsistent since they are affected by endogeniety of placement. Columns 4-6 of the same table report the results of the fixed-effects model, using 193 district-level dummy variables which remove the effect of unobservable community characteristics which affected both placement and fertility decisions.

Because the same woman is observed multiple times, we report Huber-White standard errors which correct for the correlation between observations for the same woman. The number of observations we report in these tables refer to the number of woman-years for each parity. We report our results by parity order, for birth 1 through 3. For higher birth orders we have fewer observations and face greater censoring, so we limit our analysis to these three births.

We discuss our findings with reference to the fixed effects estimates, which are our consistent estimates and are therefore preferred. The coefficient of the period dummies capture the change in fertility over time. The hazard of first birth after the revolution has increased relative to before, but there is no change after that. In other words, the post-revolution periods all show a decline in the waiting time between marriage and the first birth, which may reflect increased emphasis on motherhood and family formation after the Islamic Revolution. Whatever impact the change in the social atmosphere of the country brought for fertility does not seem to go beyond the first birth. The hazard of the second and third birth do not change in pro-natal period relative to the base period. This is consistent with the hypothesis put forward by Salehi-Isfahani and Murphy 2004 who argue that the baby boom of the early years of the revolution was the result of a change in tempo rather than completed fertility. The hazard of the second and third births decline in periods 2 and 3, reflecting the sharp drop in fertility after the introduction of family planning program.

construct community characteristics for the entire period of interest, from 1966 to 2000, we set the value of these variables for later years to the latest value available.

The magnitude of this drop raises with the birth order and continues for higher order births beyond the third birth. These results are consistent with earlier studies (Abbasi-Shavazi et al. 2009) which claim that the fertility decline was higher in higher order births. However, to see the impact of exposure to the program on fertility we need to look at the coefficients of health clinic coverage and its interaction with period dummies.

The effect of health clinic coverage for each period is the sum of the coefficient of coverage in the base period plus the coefficient of the relevant interaction term. In the pre-revolution period, the exposure to clinics shows a positive effect on the hazard of birth in all parities. This maybe related to several factors. Very limited number of health clinics existed before the revolution and there could be huge selection on the placement of these clinics. Also the family planning program before the revolution focused mainly on the urban middle class families rather than rural families (Abbasi-Shavazi et al. 2009). After the revolution and during the pro-natal period, the magnitude of the impact of exposure to health clinics on fertility drops but still is positive and significant for the first and second births. As mentioned earlier, during this period clinics services were limited to mother and child health care. It seems that the impact of clinics health services on improving the general level of health and reducing mortality can be the main reason behind their positive effect on fertility during the pro-natal period. In the next two periods in which family planning was in place, the impact of coverage on the hazard of the first two births becomes insignificant. This suggests that exposure to health clinic services (which included family planning services in these periods) did not affect the hazard of the first and second births. For the third birth, the initial positive effect of health clinics, disappears in the pro-natal period and then becomes negative and significant in the family planning period. We associate this change in the impact of exposure to clinic services between pro-natal and family planning periods to the introduction of family planning program. However, the observed negative effect washes out moving from the first five-year period of family planning to the second one. This implies that the variation in coverage can only partially explains the decline in the fertility. It also suggests that the magnitude of the impact of health clinic coverage on the fertility is small comparing to the decline taking place nationally.

The estimates of the coefficients of individual and community characteristics are consistent with previous studies in the literature. The effect of age on fertility is non-linear, but increasing for the entire child bearing ages. This finding is most likely due to the fact that women who marry young wait longer to have their first birth compared to those who marry later. Similarly, in higher order births, younger women can wait longer since they have longer fertile periods ahead of them. The effect of education on the hazard of birth are as expected and as predicted by the literature on economics of fertility (Willis 1973, Martin and Juarez 1995, Schultz 1994, Schultz 2001). Our results show that increase in the woman's own education decreases the hazard of birth. Having some secondary education has a rather large impact on the hazard of birth compared to primary education for the second and third births. However, less than 10 percent of women-years are with some secondary level of education. Interacting woman's education with the exposure variable shows that woman's education has strengthened the negative impact of health clinics on fertility. In contrast to the woman's education which is an indicator of opportunity cost of having children, husband's education can be interpreted as an indicator of household wealth. This difference in the role of woman and husband education is clearly reflected in our results. Husband's education generally shows a positive effect on the hazard of birth for all parities. The only exception is the negative impact of secondary education on the hazard of the third birth.

We also control for household assets (excluding homes, but including the number of rooms, car, and key appliances), which, like husband education, has a positive effect on the hazard of first birth. However, the effect is not significant for the second birth and turns negative for the third. This suggests that, in rural Iran, the first child is a normal good, so that it is more likely to occur in more wealthy couples. Women who moved in the last five years show lower hazard of first and second birth and similar hazard for the third birth compared to those who did not move. As expected, mortality of last child shows a positive and significant effect for all parities which shows the desire for replacement births when a family loses a child. Another interesting result is the existence a significant son preference in rural households. Our results show that women with no son have higher hazard of birth in all parities.

The effects of district-level characteristics on the hazard of birth do not show a clear pattern of influence on hazard of birth in different parities. However, in general, access to middle school and electricity can be associated with lower hazard of birth. Since better access to school lowers the cost of education and access to electricity generally associates with higher level of economic development of the local area, both these variables have their predicted effect. Interestingly, the effect of piped water is positive, indicating perhaps that, *ceteris paribus*, better health conditions increase fertility. The presence of a mosque shows a positive effect for the first and the second births but shows no effect for the third birth. We put mosque in the regression expecting it to indicate religiosity, and be positively correlated with fertility. To the extent that access to mosque is a good index of religiosity, this result shows the higher hazard of birth in more religious areas.

The last set of estimates are the coefficients of temporal dummy variables. As noted earlier, these coefficients constitute the baseline hazard function (see Figure 2.5). The hazard of each birth are highest in the second year and decline later.

2.7 Conclusion

The dramatic decline in fertility in rural Iran, and its coincidence with the implementation of a nationwide family planning, has led to speculations about cause and effect. There is considerable interest among policy makers in developing countries to learn how the shift in rural Iran from one of the highest fertility levels to replacement level in less than two decades took place. Among the innovative features of Iran's program is the construction of rural health clinics that provided active delivery of family planning services to all women in a geographic area. In this paper we take advantage of the variation in exposure to this program due to the timing of the construction of these clinics across Iran to measure the impact of the program. We matched data on the birth histories of about 50,000 rural women obtained from Iran's 2000 DHS with information about the coverage of rural clinics in their district of residence.

Since the placement of the program was not entirely exogenous and variables such as the level of education and availability of basic services (roads, electricity, and piped water) in a rural area influenced when a clinic was established there, we have to account for these factors. We control for the observable variables that influenced both fertility and placement. We also use fixed effects at the district level that help eliminate all time-invariant unobserved variables that affect program placement.

We use a discrete hazard model to estimate the effect of exposure to family planning on the probability of births of different parities. We are able to observe a small effect only on the third birth. We distinguish between the impact of the clinics during their different phases of operation. We show that in the 1980s, when the government did not offer family planning and focused on mother and child health care, exposure to the clinics was associated with a general increase in the hazard of birth in particular in the first and second births. However, during the period in which the active family planning services was offered, we see a significant negative effect on the hazard of the third birth.

In line with the conclusions of similar studies of the impact of family planning in other countries, we find a limited impact of health clinics in the case of Iran. However, it would be wrong to conclude from our findings that government policy, in reversing itself in 1989 to promote family planning, was inconsequential. As with other family planning programs, there are two government policies that occur at once: an initiative to subsidize –in Iran's case provide for free–birth control technology, as well as a signal to change policies that affect the costs and benefits of high fertility. While this study found only weak evidence of the effect of technology provision, we should not conclude that the substantial change in the calculus of the benefits of large families that the reversal of policy in 1989 entailed was not a big factor in bringing down rural fertility. For one thing, the revolutionary government was at the time, as it is still to a lesser extent, a big player in the economy. It provided child subsidies through the war time rationing system as well as in provision of health and education services. By indicating that it was no longer willing to support large families, it could have influenced family decisions significantly. For another, it was, again to a much greater extent in the 1980s as it is today, a government credible to the lower social strata, so its new theme of *"smaller families, better lives,"* broadcast ubiquitously through state media and billboards, was particularly effective. Finally, its pro-poor stance may have raised, especially for the lower social strata, the expected returns to child education, not an unreasonable expectation in view of the fact that in 2005 the county would elect its first rural born president, Ahmadinejad. Coupled with the anticipated rise in child costs the quantity-quality substitution could have played a large role.

Table 2.1: Descriptive Statistics

Women characteristics at survey year (2000)			Woman-year sample (1966-2000)		
Variable	Mean	Std. Dev.	Variable	Mean	Std. Dev.
Education			**Education**		
Illiterate	0.45	0.50	Illiterate	0.61	0.49
Primary	0.37	0.48	Primary	0.31	0.46
Secondary	0.18	0.38	Secondary	0.09	0.28
Husband's education			**Husband's Education**		
Illiterate	0.30	0.46	Illiterate	0.42	0.49
Primary	0.41	0.49	Primary	0.40	0.49
Secondary	0.29	0.45	Secondary	0.18	0.38
No. of marriages			Coverage	1.83	1.68
1	0.96	0.19			
+2	0.04	0.19	**Community characteristics**		
			Middle school	0.19	0.33
Marriage status			Electricity	0.82	1.89
Married	0.96	0.19	Pipedwater	0.86	3.02
Divorced/Widowed	0.04	0.19	Mosque	1.20	2.87
Age	32.02	9.08	Wealth index	-0.98	1.65
Age at marriage	17.44	3.72	Moved in the last 5 years	0.07	0.26
Currently student	0.01	0.10	Previous child born dead	0.06	0.25
			Has no son	0.35	0.48
Immigration status (last 5 years)					
Moved from urban	0.06	0.24	**Number of woman-year spells**		
Moved from rural	0.06	0.23	1st birth	153,397	
Did not move	0.88	0.32	2nd birth	93,207	
			3rd birth	89,212	
Number of women	43,279				
Number of districts	193		Total	676,994	

Table 2.2. Estimates from standard logit and fixed-effects regressions

	Simple logit			Fixed-effects logit		
	1st birth (1)	2nd birth (2)	3rd birth (3)	1st birth (4)	2nd birth (5)	3rd birth (6)
Time periods (1966-1978 excluded)						
Pro-natal period (1979-1988)	0.242**	0.043	-0.014	0.247**	0.035	-0.013
	(0.022)	(0.030)	(0.032)	(0.023)	(0.033)	(0.035)
FP period I (1989-1994)	0.227**	-0.328**	-0.393**	0.266**	-0.308**	-0.325**
	(0.033)	(0.047)	(0.046)	(0.037)	(0.065)	(0.052)
FP period II (1995-2000)	0.162**	-0.996**	-1.291**	0.185**	-0.952**	-1.230**
	(0.043)	(0.051)	(0.060)	(0.047)	(0.061)	(0.069)
Coverage	0.157**	0.139*	0.172**	0.177**	0.127*	0.139*
	(0.043)	(0.058)	(0.059)	(0.045)	(0.062)	(0.063)
(Pro-natal period)*coverage	-0.111*	-0.066	-0.126*	-0.135**	-0.083	-0.149*
	(0.044)	(0.060)	(0.060)	(0.045)	(0.062)	(0.062)
(FP period I)*coverage	-0.113*	-0.111	-0.146*	-0.150**	-0.128*	-0.175**
	(0.045)	(0.061)	(0.060)	(0.046)	(0.064)	(0.064)
(FP period II)*coverage	-0.151**	-0.132*	-0.128*	-0.180**	-0.158*	-0.154*
	(0.045)	(0.060)	(0.062)	(0.047)	(0.063)	(0.065)
Age						
age	0.534**	0.163**	0.106**	0.530**	0.162**	0.110**
	(0.013)	(0.015)	(0.017)	(0.013)	(0.015)	(0.017)
age sqrd	-0.011**	-0.004**	-0.003**	-0.011**	-0.004**	-0.003**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education (illiterate excluded)						
primary	-0.101**	-0.063*	-0.136**	-0.114**	-0.033	-0.077*
	(0.022)	(0.030)	(0.031)	(0.023)	(0.034)	(0.032)
secondary	-0.035	-0.545**	-0.738**	-0.069	-0.459**	-0.623**
	(0.038)	(0.047)	(0.056)	(0.040)	(0.049)	(0.060)
primary*coverage	-0.028*	-0.067**	-0.076**	-0.010	-0.070**	-0.094**
	(0.011)	(0.014)	(0.013)	(0.011)	(0.016)	(0.014)
secondary*coverage	-0.075**	-0.029	0.012	-0.058**	-0.054**	-0.023
	(0.016)	(0.019)	(0.023)	(0.017)	(0.021)	(0.025)
Husband's education (illiterate excluded)						
primary	0.044**	0.074**	0.035	0.043*	0.081**	0.048*
	(0.017)	(0.021)	(0.021)	(0.017)	(0.021)	(0.021)
secondary	-0.017	0.039	-0.122**	-0.035	0.031	-0.120**
	(0.021)	(0.025)	(0.027)	(0.022)	(0.026)	(0.028)
Wealth index	0.028**	-0.027**	-0.054**	0.040**	-0.011	-0.030**
	(0.004)	(0.006)	(0.006)	(0.005)	(0.006)	(0.007)
Moved in the last 5 years	-0.123**	-0.151**	-0.094**	-0.123**	-0.134**	-0.041
	(0.021)	(0.027)	(0.036)	(0.022)	(0.028)	(0.037)
Mortality of previous child		0.647**	0.719**		0.651**	0.724**
		(0.037)	(0.039)		(0.038)	(0.040)
Has no son		0.036*	0.152**		0.035*	0.150**
		(0.016)	(0.020)		(0.016)	(0.020)
District level characteristics						
middle school	-0.123	-0.425**	-0.397**	-0.018	-0.588**	0.007
	(0.068)	(0.076)	(0.079)	(0.115)	(0.153)	(0.165)
electricity	-0.016	-0.022*	-0.007	-0.023	-0.111**	-0.017
	(0.009)	(0.010)	(0.009)	(0.022)	(0.027)	(0.030)
piped water	0.031**	0.044**	0.035**	-0.011	0.191**	0.015
	(0.008)	(0.009)	(0.010)	(0.025)	(0.032)	(0.037)
mosque	-0.010	0.014	0.009	0.068**	0.108**	0.034
	(0.007)	(0.008)	(0.009)	(0.016)	(0.021)	(0.022)

Table 2.2. (continued)

	Simple logit			Fixed-effects logit		
	1st birth	2nd birth	3rd birth	1st birth	2nd birth	3rd birth
	(1)	(2)	(3)	(1)	(2)	(3)
Spell years (0 is excluded)						
1 year	1.048**	1.040**	1.126**	1.060**	1.052**	1.141**
	(0.019)	(0.019)	(0.022)	(0.019)	(0.019)	(0.022)
2 years	1.163**	1.117**	1.379**	1.192**	1.144**	1.414**
	(0.021)	(0.024)	(0.026)	(0.021)	(0.024)	(0.026)
3 years	1.068**	1.053**	1.233**	1.109**	1.096**	1.285**
	(0.023)	(0.033)	(0.033)	(0.024)	(0.033)	(0.034)
4 years	0.929**	0.977**	1.113**	0.978**	1.033**	1.180**
	(0.027)	(0.044)	(0.043)	(0.028)	(0.045)	(0.044)
5 years	0.745**	0.720**	1.021**	0.800**	0.780**	1.094**
	(0.033)	(0.060)	(0.056)	(0.033)	(0.061)	(0.056)
6 years	0.607**	0.632**	0.929**	0.664**	0.697**	1.012**
	(0.039)	(0.080)	(0.072)	(0.039)	(0.081)	(0.072)
7 years	0.437**	0.307**	0.863**	0.495**	0.366**	0.951**
	(0.046)	(0.106)	(0.092)	(0.047)	(0.107)	(0.093)
8 years	0.204**	-0.091	0.580**	0.268**	-0.038	0.665**
	(0.057)	(0.141)	(0.119)	(0.058)	(0.142)	(0.120)
9 years	0.025	-0.497**	0.444**	0.086	-0.450*	0.531**
	(0.068)	(0.184)	(0.148)	(0.069)	(0.185)	(0.150)
10+ years	-0.370**	-1.148**	-0.286*	-0.290**	-1.096**	-0.215
	(0.055)	(0.151)	(0.143)	(0.055)	(0.150)	(0.144)
constant	-7.759**	-2.673**	-2.199**	-7.842**	-2.828**	-2.461**
	(0.136)	(0.173)	(0.215)	(0.149)	(0.189)	(0.229)
Log likelihood	-77,322	-51,785	-45,263	-76,956	-51,517	-44,949
Pseudo R^2	0.07	0.10	0.12	0.07	0.11	0.13
Observations	148,886	90,904	87,491	148,886	90,904	87,491

Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Bibliography

Abbasi-Shavazi, M. J. and P. McDonald (2006). Fertility decline in the islamic republic of iran, 1972-2000. *Asian Population Studies 2*(3), 217–237.

Abbasi-Shavazi, M. J., P. McDonald, and M. Hosseini-Chavoshi (2009). *The Fertility Transition in Iran: Revolution and Reproduction*. New York: Springer.

Aghajanian, A. (1995). A new direction in population policy and family planning in the Islamic Republic of Iran. *Asia-Pacific Population Journal 10*(1), 320.

Aghajanian, A. and A. H. Mehryar (1999). Fertility, contraceptive use and family planning program activity in the islamic republic of iran. *International Family Planning Perspectives 25*(2).

Angeles, G., D. K. Guilkey, and T. A. Mroz (2005). The determinants of fertility in rural Peru: Program effects in the early years of the national family planning program. *Journal of Population Economics 18*, 367–389.

Becker, G. S. (1992). Fertility and economy. *Journal of Population Economics 5*(3), 185–201.

Boonstra, H. (2001). Islam, women and family planning: A primer. *The Guttmacher Report on Public Policy 4*(6).

Filmer, D. and L. Pritchett (1998, October). Estimating wealth effects without expenditure data - or tears : with an application to educational enrollments in states of india. Policy Research Working Paper Series 1994, The World Bank.

Gertler, P. J. and J. W. Molyneaux (1994). How economic development and family planning programs combined to reduce indonesian fertility. *Demography 31*(1), 33–63.

Martin, T. C. and F. Juarez (1995). The impact of women's education on fertility in latin america: Searching for explanations. *International Family Planning Perspectives 21*(2), 52–80.

Mehryar, A. H., B. Delavar, G. A. Farjadi, M. Hosseini-Chavoshi, M. Naghavi, and M. Tabibian (2001). Iranian miracle: How to raise contraceptive prevalence rate to above 70 percent and cut TFR by two-thirds in less than a decade? Paper presented at the IUSSP General Population Conference, Salvador, Brazil.

Moore, R. (2007). Family planning in iran, 196079. In W. C. Robinson and J. A. Ross (Eds.), *The Global Family Planning Revolution: Three Decades of Population Policies and Programs*, pp. 33–58. Washington, DC: The World Bank.

Pitt, M. M., M. R. Rosenzweig, and D. Gibbons (1993). The determinants and consequences of the placement of government programs in indonesia. *World Bank Economic Review 7*, 31948.

Raftery, A. E., S. M. Lewis, and A. Aghajanian (1995). Demand or ideation? evidence from the Iranian marital fertility decline. *Demography 32*(2), 159–182.

Rosenzweig, M. R. and K. Wolpin (1986). Evaluating the effects of optimally distributed public programs: child health and family planning interventions. *American Economic Review 76*(3), 470–482.

Roudi-Fahimi, F. (2002). Iran's family planning program: Responding to a nation's needs. MENA policy brief, Population Reference Bureau, Washington, DC.

Rutstein, S. O. and K. Johnson (2004). The dhs wealth index. dhs comparative reports no. 6. Technical report, Calverton, Maryland: ORC Macro.

Salehi-Isfahani, D. (2009a). Poverty, inequality, and populist politics in Iran. *Journal of Economic Inequality 15*(1), 5–24.

Salehi-Isfahani, D. (2009b). The Revolution and the rural poor. *Radical History Review 2009*(105), 139–144. Special issue: The Iranian Revolution Turns Thirty.

Salehi-Isfahani, D., M. J. Abbasi, and M. Hosseini-Chavoshi (2010). Family planning and fertility decline in rural Iran: The impact of rural health clinics. *Health Economics 19*(S1), 159–180.

Salehi-Isfahani, D. and R. D. Murphy (2004). Labor market flexibility and investment in human capital. Working paper, Department of Economics, Virginia Tech.

Schultz, T. P. (1994). Human capital, family planning, and their effects on population growth. *The American Economic Review Papers and Proceedings 84*(2), 255–260.

Schultz, T. P. (2001). The fertility transition: Economic explanations. *SSRN eLibrary*.

Strauss, J. and D. Thomas (1995, March). Human resources: Empirical modeling of household and family decisions. In H. Chenery and T. Srinivasan (Eds.), *Handbook of Development Economics*, Volume 3 of *Handbook of Development Economics*, Chapter 34, pp. 1883–2023. Elsevier.

Willis, R. J. (1973). A new approach to the economic theory of fertility behavior. *The Journal of Political Economy 81*(2), S14–S64.

Chapter 3

Voter Choice in 2005 Iranian Presidential Election

3.1 Introduction

Elections play an important role in the political system of the Islamic Republic of Iran. However, until early 90's the relatively small number of qualified candidates made the outcomes of these elections rather predictable. The role of electoral competition became increasingly visible in 1990s when the turmoil of the early years of the Revolution and the eight-year war with Iraq had ended.

Interest in understanding voter behavior emerged initially after the surprise landslide victory of the reformist candidate, Mohammad Khatami, in the 1997 presidential election. Subsequent elections produced strong support for the reformists (about 70-80% with turnout rates of over 60%) for some years, suggesting a comfortable lead for reformaists against conservatives. In this sense, the unexpected landslide victory of Mahmoud Ahmadinejad in the final round of Iran's two-round runoff presidential election, after finishing second in initial round (see Table 3.2), was another turnabout in Iran's politics that highlighted the need to know more about politics in Iran. Figure 3.1 compares district vote shares for the conservative and reformist candidates in the aformentioned presidential elections. It clearly shows how voters moved from reformists towards conservatives in the two consequative elections of 1997 and 2005.

This study estimates a model of voter choice for the 2005 Iranian presidential election. Its objective is to characterize the relative importance of traditional bases of Iranian voter choice. I employ the district-level results of this election to shed light on the socio-economic factors that contributed to Ahmadinejad's popularity among Iranian voters. The main econometric issue is that the individual voting decisions are unobservable. Only aggregate vote shares of districts are available. To resolve this problem, a demand estimation method which has been used in the Industrial Organization literature is employed. This method enables me to use aggregate data to infer discrete individual-level relationships of interest when individual level data are not available. A multinomial logit model is introduced to estimate individual probability of voting for each candidate. Then a search routine is employed that finds the value of parameters which make these aggregate predictions as close as possible to the observed vote shares. The empirical analysis uses district level data from the 2005 Iranian presidential election.

The results demonstrate the importance of both voter and location of vote characteristics and considerable differences in the bases of voter choice. District level demographic have strong explanatory power. Female ratio, urbanization and minority ratio seems to play a negative role in participation in the election. The positive effect of income and urbanization on Ahmadinejad's votes contradicts with common thought that the poor and rural voters supported him. He also did not find much support from women and older voters. The effect of local economic measures (e.g. unemployment rate and gini coefficient) is about the same for both candidates.

The paper is structured as follows. Section 3.2 provides a brief overview of Iranian electoral political system. In Section 3.3, Iran's political structure after the Islamic revolution is briefly reviewed. Model specifications are given in Section 3.4. Section 3.5 contains a description of the data as well as the estimation results. Conclusion is provided in Section 3.6.

3.2 Review of Literature

Several models of presidential voting have been proposed in the literature over the years (Niemi 2001) including Colmbia sciological model, social-psychological attitudinal models (Michigan model), economic forecasting models, political psychology models, group-based sociological models, and rational choice models. Columbia, Michigan and the rational choice model are the main schools of thought in studying voting behaviors and will be discussed in more details in the following section.

The Columbia researchers employed a consumer preference model and explained the 1940 US presidential election with a sociological model (Lazarsfeld 1968). They found that a person's socio-economic status (education, income and class), religion, and place of residence (rural, urban) were strongly related to the person's vote. Social group factors accounted for most of the differences in voting. Yet this model does not try to explain *why* people voted like they did. It does not consider the political aspects of an election. And it's not useful in explaining change across elections, since different parties are elected in different years even if the social characteristics don't change much (Niemi and Weisberg 1976).

The social-psychological model (also known as the Michigan Model, see Beck 1986, Campbell et al. 1960) emphasizes the distinction between long-term factors, such as social demographics, partisanship, and ideology, that are determined before the election year and short-term factors that are more election specific, such as issues, candidates, and the campaign. This approach has become the main model for empirical studies of voting. A person's identification with a party became the core of this model (Campbell et al. 1960). Sociological background characteristics (ethnicity, race, region, religion, and the like), social status characteristics (education, occupation, class), and parental characteristics (class, partisanship) all affect the person's choice of party identification. Party identification influence the person's evaluation of the candidates and the issues. These evaluations along with campaign incidents, conversations with family and friends affect the vote.

The other model of voting which has become popular is a rational voter model. This approach emphasizes the strategic calculations candidates and voters make as they face elections. The basic (Downs 1957) model, for example, views politics as a battle along a single liberal/conservative dimension, with citizens voting for the party that is closest to their own position on that dimension. Thus the rational-choice model focuses our attention on ideology as a factor in elections, both in terms of how the candidates position themselves ideologically and the extent to which voters view the election in liberal/conservative terms.

Table 3.1 summarizes the analysis methods, data, and dependent variables of these models employed in analyzing the 2000 U.S presidential election. One common analysis procedure is to use logit (or probit) analysis to explain voting behavior. However, two other techniques are also illustrated: a discrete-choice model that also includes voting for minor-party candidates and a bivariate probit model to analyze the turnout and vote decisions simultaneously (Box-Sterffensmeier et al. 2004). Weisberg and Hill 2004 employ what might be termed a multi-stage logit analysis, in which they analyze one stage of predictors at a time and then put together all the significant predictors in a single equation at the end. Kanthak and Norrander 2004 show how interaction terms can be used to test whether men and women give different weight to issues. Jacoby's analysis tests whether voters with different levels of political sophistication rely on different factors in making the their vote choice. Norpoth 2004 oves to time-series analysis in some of his exploration of the role of economic performance in voting. Table 3.1 also shows the types of variable used in the literature. Partisanship, not surprisingly, is most of studies, and ideology is used in nearly all. Most use issue variables, though the exact set differs from one analysis to another. Candidate personality variables are included as predictors in some studies, though others avoid those measures. In the next section some common debates about the preceding models are discussed.

3.2.1 Common issues in the voting studies

Modern voting studies rely heavily on survey research (Niemi 2001). There are certainly other ways studying elections, but the most direct and often the most valid way of understanding why people vote as they do is to ask them. An important recent development in voting studies has been the increased use of experiments (Lodge et al. 1989, McKelvey and Ordeshook 1985). This increased use of experiments is expected to continue, even while political surveys remain the dominant mode of studying elections.

All voting analysis approaches are subject to selection problem. This is because of the fact that researchers just observe the voting decision of individuals who decided to vote in first place. Empirical voting research is primarily concerned with the relation between various political, demographic, and psychological characteristics and political preference. Nonvoters, however, also have preferences, but they do not vote. If preference is measured by vote, then data on preference are missing for nonvoters. If turnout and preference are unrelated, there should be no bias in estimating a model of preference based on the subsample of voters whose preference is observed. To the degree that there are common factors determining both turnout and preference, turnout is a source of selection problem (Dubin and Rivers 1989). The customary practice in voting studies has been to analyze turnout and vote choice separately (Campbell et al. 1960, Downs 1957, Wolfinger and Rosenstone 1980, Teixeira 1992). The voting electorate, however is not a random subsample of the voting age population. Voters are known to be older, more educated, and likely to be married than nonvoters (Wolfinger and Rosenstone 1980). The selection problem even is more complicated in US elections because of its two-stage nature of participation which reflects the procedure of first registering and then voting among the subsample of the electorate that is registered (Timpone 1998). The two-step procedure is adapted in several studies of individual-level behavior (Erikson 1981, Jackson 1996).

The last common issue in this literature is the choice between logit or probit model in explaining voting behavior. The most commonly used estimation technique for these models has been Multinomial Logit (MNL). However there are substantial drawbacks associated with the use of MNL because it assumes that the disturbances are independent across alternatives (Independence of Irrelevant Alternatives or IIA for short). IIA holds when the ratio of the probability of choosing choice j to the probability of choosing choice k is independent of the set of alternatives available (Train 2003). If individuals view two candidates as having similar attributes in some unmeasured way then the disturbances will be correlated. Under this scenario MNL is an inappropriate estimator (Dows and Endersby 2004).

Since the MNL specification is tractable and simple to estimate, it is included in many commercial software packages and is used regularly in electoral research. Multinomial Probit (MNP) does not impose the independence assumption and advances in computer technology make its estimation increasingly practical. Thus, one might reasonably argue that MNP should be the benchmark methodology in the study of voter choice in electoral studies (Alvarez and Nagler 1994, Alvarez and Nagler 1995, Alvarez and Nagler 1998, Schofield et al. 1998, Lacy and Burden 1999, Quinn et al. 1999) However, Dows and Endersby 2004 compare two models and argue that the simpler logit is often preferable to the more complex probit for the study of voter choice. Also they claim that imposing the IIA property for which the logit model is been criticized, is neither relevant nor particularly restrictive for most applications. Dow and Endersby conclude that in the study of voter choice in multiparty /candidate elections these concerns are exaggerated and, under most circumstances, logit estimation performs as well or better than MNP.

In this paper we employ a Columbia type model which does not rely on the survey data. This model balances structural foundations with inherently aggregate election results and uses district level data from the presidential election. Same as Columbia approach (Lazarsfeld 1968) a consumer preference model is employed. Each party presented a product to the public and that people considered these competing products throughout the advertising campaign, carefully considering alternative, and then stepped into the booth on election day to record their final choices.

Following Dows and Endersby 2004, the individual voters' utility maximization problem is specified using a multinomial logit model that allows for heterogeneity in preferences across voters and districts. The impact of demographic characteristics of voters and districts on the individual's utility level is analyzed. To deal with the selection problem, the model developed in this paper allows for the inclusion of abstention into the voters' choice set. The main idea is that "*If you choose not to decide, you still have made a choice*"(Lee et al. 1980). In our model, abstention is assumed as the reference alternative which its utility is normalized to zero and all other utilities are expressed based on the reference group's utility.

3.3 Background

The political structure under the Islamic Republic is unique in a sense that it is a combination of both theocratic and democratic institutions. Presidents and members of parliament are selected by the public through elections but there are other institutions like the guardian council–which is in charge of vetting the candidates–that are not directly elected by the public.

There are no influential parties currently active in the political sphere of Iran. However, based on their positions on different issues such as extent of power they accept for the Leader, foreign policy, social issues, etc, political figures can be divided into two main groups, conservatives vs. reformists. Groups that have taken conservative positions favor strength for the Leader, while others identified as reformists have typically prefer greater role for elected offices. In the social sphere, conservatism has meant greater emphasis on extensive enforcement of Shia religious rules and norms, segregation and maintenance of relative status of genders, and opposition to Western culture, whereas reformist have advocated individual rights, tolerance of diversity, and improvements in women's relative status. In the economic arena, the issues have been reliance on markets vs. the state and the extent of emphasis on redistribution vs. growth.

The ninth presidential election was held on June 17 and June 24, 2005 for which 46,786,418 electors were eligible to vote. A total of 29,400,857 and 27,959,253 valid ballots were cast in round 1 and round 2 respectively. From seven candidates in round one, three were supported by the conservatives and the other four by the reformists (see Table 3.2). Two winners of the first round, Rafsanjani and Ahmadinejad, are introduced in more detail in the following section.

Akbar Rafsanjani was viewed as a leading figure in the Revolution of 1979 and one of the clerics closest to the founder of the Islamic Republic, Ayatollah Khomeini. He had served as the Speaker of the Parliament during 1980-1989. He was seen as the most important political power broker in the country with elitist tendencies and socially relatively liberal views. When he took office as President in 1989, Iran's economy had endured a significant decline during the Revolution and Iran-Iraq War. Markets were under extensive government controls and trade had declined to historically low levels. He announced a liberalization policy, which initially produced a rapid recovery. His reforms coincided with an increase in oil revenues that helped imports expand quickly and supported investment, infrastructure development, and recovery. However, the liberalization was quickly extended to risky areas, such as the opening of capital account with inadequate supervision, soon leading the economy into a major macroeconomic crisis in 1993-1994. The result was high inflation and a sharp slowdown during Rafsanjani's second term, 1993-1997. Meanwhile, the country's high economic and political risks limited private sector activity and investments were largely confined to projects controlled by individuals and organizations close to the centers of power. Naturally, Rafsanjani was associated with unequalizing market liberalization and with limited concern for the lower income groups.

The second candidate in the first round, Mahmoud Ahmadinejad, had started as a student activist before the Revolution and had accumulated revolutionary credentials via his activities among the Basij Militia during Iran-Iraq war. He then rose through administrative ranks by serving as district governor in West Azerbaijan the late 1980s and, after a short period of time at the Ministry of Higher Education, as Governor of Ardabil province in the 1990s. He also served as faculty member at the Science and Technology University in Tehran. In 2004, He was selected as the Mayor of Tehran by the city council which was dominated by the his conservative supporters.

As predicted, Rafsanjani won the first round but with a much smaller share than he had expected, 22.0 percent of the votes followed by Ahmadinejad (20.3 percent). In the runoff Ahmadinejad were able to beat Rafsanjani (see Table 3.2).

Table 3.3 reports the vote shares in all provinces. Province of origin seemed particularly important in round 1 of the election. Rafsanjani came from Kerman in the south and dominated the race there. This was clearly the case for Karrubi in Lorestan in southwest, Qalibaf in Khorasan in the northeast, and Larijani in Mazandarn in the north. (See Table 3.3.) Mehralizadeh is an ethnic Azeri and was ranked first in all three Azeri-speaking northwestern provinces. Ahmadijenad grew up in Tehran and captured the lead in all central provinces (Tehran, Markazi, Qom, Isfahan, Yazd, and Semnan).

In round 2 Rafsanjani was able to win only the poorest province, Sistan and Baluchestan, with 55 percent of the vote, losing even his native Kerman to Ahmadinejad (see Table 3.3). He received high shares in Kermanshah (49.18%) and Lorestan (49.01%) provinces which, like Sistan, have large ethnic and religious minority populations. Sistan's outcome in the election shows the complicated interplay of regional politics and economic issues and the need for multivariate analysis of the results. As the poorest province, some would expect it to have gone to Ahmadinejad, but as a province with a large Baluchi population who are Sunni, it seems to have placed more weight on Rafsanjani's relatively more liberal approach to social issues compared to Ahmadinejad's promise of redistribution.

Table 3.3 also reports the abstention rates which varies significantly across provinces. Turnouts in elections have always been an issue in the Islamic Republic as indicators of support for the system, in part because candidates for election are vetted by the Guardian Council. In 2005 presidential election, there were calls for a boycott from prominent individuals in Iran, including the Nobel Laureate Shirin Ebadi. Looking at the result of both rounds (Table 3.3), it seems that the boycott movement was more appealing to the ethno-linguistic minorities such as Arabs, Baluchs, Kurds, and Turkmans. Kurdistan, West-Azerbayjan and East-Azerbayjan which are top three provinces with abstention rate of more than 50% in both rounds, are all ethinic minorities. Although the participation rate was lower than most previous presidential elections in the country, it was quite reasonable by international standards. Figure 3.2 illustrates the district level abstention rates and vote shares of both candidate in round 1 and 2.

3.4 Model Specification

3.4.1 Voter's Utility Function

The individual voter's problem is specified using a discrete choice model of consumer behavior in the context of a differentiated products market. A market is defined as an electoral district and a product is defined as candidate. The size of the market is specified as the number of eligible voters in a given district. A candidate will be defined by a dummy variable (i.e. the voter's utility is defined on the candidate per se rather than on the characteristics of the candidate). Each voter will be assumed to have characteristics (observable and unobservable) that influences the voting decision. The indirect utility of voter i for candidate j in district can be modeled as

$$U_{ijr} = \beta_{ij} + Z_r\alpha_j + \varepsilon_{ijr}$$
$$i = 1, ..., I, j = 1, 2, r = 1, ..., 30$$

where β_{ij} is a vector of individual-specific taste coefficients. Z_r is a k-dimensional vector of characteristics of district r and ε_{ijr} is the stochastic error term distributed *iid* Type I extreme value across candidates and voters. β_{ij} can be decomposed as

$$\beta_{ij} = \beta_j + D_i\pi_j + v_i\gamma_j$$

- β_j represent a candidate specific intercept, representing the general appeal (or political equity) of candidate j across all voters.
- D_i is a p-dimensional vector of known voter's demographic characteristics.
- π_j is a k-dimensional vector of demographic coefficients representing candidate j's appeal to a particular demographic constituency, relative to other candidates.
- v_i denotes a one dimensional vector of latent scores capturing unobserved deviation in voter preferences relative to the population average, assumed to be independent identically distributed standardized normal.
- γ_j represents the effect of the heterogeneity in voters' preference for each candidate.

Combining these two equations we have,

$$U_{ijr} = \beta_j + D_i\pi_j + v_i\gamma_j + Z_r\alpha_j + \varepsilon_{ijr} \quad (3.1)$$

The specification is completed with the introduction of an *outside option*: the voters may decide not to vote for any of the candidates and abstain in the election[1]. The standard practice is to set $U_{i0r} = 0$ (that is we subtract U_{i0r} from each choice). Though this is notationally convenient we should keep in mind that the utilities from the various choices are now actually the difference in utility between the choice of the particular candidate and the not-voting option.

A voter may condition her own preference to the preferences revealed by those spatially nearby. To capture the effect of location of vote on the voter's preference, the Z_r term which include information about the district such as population, ratio of religious and ethnic minorities, attachment to political parties and district-specific measures of the economy, specifically unemployment rate, gini coefficient added to the model. To absorb a province's predisposition towards a specific political party, the first round vote share of the conservative party is further included.

The demographic variables include voter gender (female coded 1), age (voting age), education (recorded as years of schooling), urban or rural (urban coded 1). I use the log of real per capita expenditure as a proxy for income.

The probability that choice j is made, is

$$Prob(U_{ijr} > U_{ilr}), \forall l \neq j. \quad (3.2)$$

This probability is given by the well-known Multinomial Logit model, widely used in modeling voter choice behavior (Dows and Endersby 2004, Paap et al. 2005) Given individual choices are not observed, aggregate candidate-district level vote shares are derived. The vote share for candidate j in district r is

$$s_{jr}(Z_r) = \int_{A_{jr}} \frac{exp(\beta_{ij} + Z_r\alpha_j)}{1 + \sum_{l=1}^{2} exp(\beta_{il} + Z_r\alpha_l)} \phi(v_i)dv_i, \quad (3.3)$$

[1] I also group invalid votes as abstain. As just %1 of total eligible voters cast invalid votes, making them as a spearate alternative does not make any difference in the results.

A_{jr} represents the group of individuals in district r who voted for candidate j. v_i is distributed as standard normal. The econometric procedure given in Berry (1994), Berry et al. (1995) and Nevo (2000) is used to estimate the parameters of the model.

3.4.2 Estimation

The dataset contains information at the district level, individual decisions are not observed. Because of lack of variation in candidate characteristics we are not able to identify the voters preference for those characteristics. Thus we are looking at the candidate per se. The number of votes each candidate received at each of 292 districts are observed. Abstention incorporated to the model so vote shares are calculated as the percentage of vote received in total eligible voters (not the number of vote cast) in each district. The parameters of the model are estimated using the two-step procedure (Berry et al. 1995, Berry 1994, Nevo 2000, Ackerberg et al. 2007) which has been used for estimating discrete-choice models of demand by market-level price and quantity data. The first step of this procedure is designed to obtain the estimated aggregate vote shares of each candidate in each district. The probability of voting for each candidate is estimated using (3.3). Typically this integral is intractable. Consequently we use simulation to obtain an approximation of it. I.e. we take N_r pseudo random draws from $\phi(v_i)$ and compute

$$\hat{s}_{jr}(Z_r) = \frac{1}{N_r}\sum_{i=1}^{N_r} \frac{exp(\beta_{ij} + Z_r\alpha_j)}{1 + \sum_{l=1}^{2} exp(\beta_{il} + Z_r\alpha_l)} .I\left\{v_i \in A_{jr}\right\},$$

where N_r is the number of observations in each district and I is the usual indicator function. Parameters can be estimated by minimizing the difference between these estimated aggregate vote shares and the observed voter shares.

The second step is a standard method of moments step that finds the value of β that makes the distribution of the estimated vote shares as close as possible to that of actual vote shares. Let the vector of observed shares be s^o_{jr}, where $j = 1, 2$ and r is the district index. Denote ξ_{jr} as

$$\xi_{jr} = s^o_{jr} - \hat{s}_{jr}$$

I employ a search routine[2]that finds the value of β which makes these aggregate predictions as close as possible to the observed vote shares (i.e. making ξ_{jr} as close as possible to zero). Results from this model are provided in the empirical section.

[2]Average of demographic variables in each district, and the interaction between average income and average age with candidate dummies have been used as instruments. Standard errors have been calculated using a bootstrap with 100 iterations.

3.5 Data and Estimation Results

3.5.1 Data

The election results are published by the Ministry of Interior (MOI). For two runoff candidates, information on both rounds is available at the district[3] level. For each district, the total number of votes of Ahmadinejad and Rafsanjani, total number of votes cast, the number of eligible voters, and the number of invalid ballots are reported. For the rest of candidates in round 1, the same information is available at the province level. I use the first round information to calculate the partisanship index defined as conservative vote share of each province in the first round of election.

There are a number of districts with participation rates well above 100 percent[4]. The reason for this outcome is that voters are not restricted to vote in their residency and also elections take place on Fridays (the weekend) when some people leave town. The possibility of moving voters has led to higher than 100% pariticipation rates in some distritcs. This is not a notable factor in most districts. To avoid problems that this pattern of voting causes for our estimation process, we exclude those five districts with participation rates of more than 100 percent.

District-level characteristics such as population, female ratio, urbanization data are obtained from census data of 2006, made available by the Statistical Center of Iran (SCI). Household Expenditure and Income Survey (HEIS) of 2004 is used to obtain demographics and individual characteristics such as age, gender (a value of 0 represents a male voter and 1 represents a female voter), education, urban which records an urban or rural voter (taking a value of 1 when the voter is urban and 0 otherwise), income and some district level measures of the local economy such as unemployment rate and gini coefficient. Voting-age population[5] is used for calculating all the population ratios. Individuals under age fifteen, are omitted from the dataset. Mean years of schooling is computed based on the education levels and varies as 1 to 5 years for primary school, 6 to 8 years for middle school, 9 to 12 years for high school and 15 years for university.

Unfortunately there is no measure of religious and ethnic minority[6] in the HEIS. The only available source for estimating these two important variables in district level is a village level survey which just cover the rural areas of Iran. There are questions about the religion and language in this survey. The average ratio of non-shia rural individuals in the each district is used as a proxy for religious minority. Similarly, the ratio of non-persian speakers is defined as a measure of ethnic minority. Complete data is available for 292 districts. The descriptive statistics for our sample of 71,728 individuals (292 districts, 30 provinces) are given in Table 3.4.

---

[3]In Iran's administrative divisions, districts are subdivisions of provinces. At the time of 2005 election, Iran was subdivided into 30 provinces and 325 districts

[4]The participation rates of these five districts (round1, round2) is as following:
Shemiranat (797%,840%), Rey (215%,216%), Robatkarim (130%,131%), Damavand (92%,100%) and Manoojan (112%,108%)

[5]Voting-age at the time of election was 15 but in January 2007 the parliament passed a bill to raise the country's voting age from 15 to 18

[6]Religious minorities include sunni muslim, Baha"i' Faith, Christianity, Judaism and Zoroastrianism. Ethnic minorities include Kurds, Balochis, Turkmens, Lors and Arabs

Table 3.5 shows the average and standard deviation of share of each alternative for different quartiles of demographic groups. As the individual votes are not observable the reported numbers in this table are averages of district shares. The average abstention rate of districts in the first quartile of education is more than 10% higher than the districts with more educated eligible voters and this average . This shows a positivie relation between education and participatinf in the election. Average vote share of both candidate are increasing with the level of district average edcuation. Urbanization does not show a clear effect. Districts with more urban population show higher abstension rate. Income has the same effect on participation as education. The participation rate is higher in wealthier districts. Interestingly, Ahmadinejad also benefited from districts with higher level of average income. Unemployment show a positive effect on all the alternative shares. People in district with higher average rate of unemployment participated more and supported both candidates. As expected minority rate plays a negative role in participation in the election. Districts with higher minority rate show less tendency to participate and also less support for Ahmadinejad and higher support for Rafsanjani.

3.5.2 Results

Estimation was done in Matlab and Stata. Table 3.6 reports the estimation results for three models. Model A is an ordinary least square (OLS) regression which can be described as,

$$log(\frac{Candidate\ vote\ share}{Abstain\ share}) = Z_r\alpha_j + \epsilon_{rj}$$

Where Z is a k-dimensional vector of known demographic profile of district r. This model works with aggregate (district) level data. Robust standard errors of model A is reported. Model B is the multinomial logit structure which employs both the individual level and district level information. This model takes into account the effect of voting location on the voter's preference (district level heterogeneity). Method of moment estimates and bootstrap standard errors are shown. Marginal effects are also reported in the last column. All the coefficients are reported with respect to the excluded option (Abstention).

The negative sign of coefficients in the Table 3.6 can be interpreted as the negative effect of corresponding variable on participation. Looking at the OLS results (Model A) reveals that districts with older voter population, higher ratio of female voters, higher share of urban voters experienced a relatively lower participation rates. On the contrary, increase in the overall educational attainment of the voters in a district rises participation rates. Worse income distribution is also associated with the higher participation rates. Unemployment and population size do not seem to be significant. District's average level of income and the relative size of minority population show different effect for two candidates. These variables were associated with a negative and significant effect on the vote share of the conservative candidate but not significant effect on the share of reformist candidate. The share of eligible voters in wealthier districts and districts with bigger minority population that chose to abstain rather than vote for Ahmadinejad was significantly higher than the other districts.

Comparing the magnitude of coefficients for both candidate indicates that the positive effect of districts' average education on the vote share of Rafsanjani is twice as big as its effect on the

Ahmadinejad's vote share. The ratio of female voters had also a greater negative effect on the vote share of Ahmadinejad comparing the Rafsanjani. The effect of other factors are about the same for both candidate.

The rest of Table 3.6 reports the full MM model results that uses both individual level and district level data and allows for the effect of location of vote on voting decision. The interpretation of marginal effect in nonlinear models is easier than the interpretation of estimates. I also report them in the last two columns. By few exceptions, the estimated coefficients of this model is usually consistent with the OLS results. Below, the results of this model are explained in more detail.
As discussed in the OLS results, age has negative effect in participation. Younger eligible voter were more motivated to participate. One reason for this could be the higher costs of going to polling stations for elderly people (e.g., for health reasons). Comparing preference over candidates, young voter seems to have leaned towards Ahmadinejad. Interestingly, Rafsanjani made an effort to appeal to the youth. However being from an older generation seems to have muted his message.

The educational attainment raises participation rates. Education also seems to have very strong positive effect on the Rafsanjani's vote share which is interesting because as is mentioned before, Ahmadinejad had served as a faculty member in one the famous universities of the capital. Given this, he should have been able to communicate and connect to the students, faculties, and generally more educated voters better than the other candidate which was from an older generation. But this did not happen and more educated voters leaned towards Rafsanjani. This shows that Rafsanjani could collect votes of other reformist candidate, Moin, who had served as the minister of science during the Khatami's presidency and had a good reputation among the higher educated voters.

Female voters preferred not to vote rather than to vote for Ahmadinejad. They did not support Rafsanjani either. One reason for lower participation rate among women could be their response to the boycott call by by Shirin Ebadi, Iranian famous lawyer and peace nobel prize winner of 2005. She is known for her long time efforts in defending the women rights in Islamic republic. She is now leading a feminist campaign against the gender discrimination in the Islamic judiciary system of the country. It seems that female voters could not find any hope for change in the either of the candidates's agenda. This negative effect was more severe for Ahmadinejad. This result shows a strong opposition for Ahmadinejad from the female voters.

Same as OLS results, urban areas and more populous districts show negative effect on participation. The negative effect of urban is greater for Rafsanjani showing that he was more appealed to rural voters. This is consistent with the positive effect of voter's income on the probability of voting for Ahmadinejad. The income effect for Rafsanjani is not significant.

One of the important consideration regarding regional impact on voting is the effect of minority groups. Ethnolinguistic diversity has long been a major issue in Iranian politics. It seems that the boycott call appealed to these groups, in particular religious minorities. They were also less likely to vote for Ahmadinejad and leaned towards Rafsanjani.

Neither of both candidates have significantly benefited from the district level measures of the economy. Unemployment and inequality played nearly the same positive role for Ahmadinejad and Rafsanjani. Districts with more unemployed population and worse distribution of income have higher participation rate probably hoping for change in the current bitter situation. But they did not clearly preferred any of the candidates.

3.6 Concluding Remarks

Understanding citizens' electoral behavior represents a fundamental step in the analysis of democratic institutions. The contribution of this paper is to employ a demand estimation framework for analyzing the eligible voter's participation and voting decision in elections at the same time. This framework has been applied to study empirically the issue of turnout and voting in the 2005 presidential election of Iran, using individual-level demographic data and district- level election data. It has been shown that even without observing the individual voting decisions, the model is able to utilize individual characteristics and connect them to the aggregate level election outcomes. Moreover, the estimated model has been used to quantify the relationship between a variety of citizen's characteristics and the election outcome. Based on the model's findings women, minorities, and citizens of major and populous districts were main groups with low turnouts. On the other hand, more educated individuals and districts with higher unemployment and higher inequality participated more in the election. Ahmadinejad did not receive much support from female voters. He also did not appeal among minorities and the poor. He benefited from the support of younger voters and urban areas with higher levels of income. Elite and more educated individuals voted in favor of Rafsanjani. He also was preferred to Ahmadinejad among minorities, women and rural voters. One of the interesting findings of this study is the equal role of local economy on the voting for each candidate.

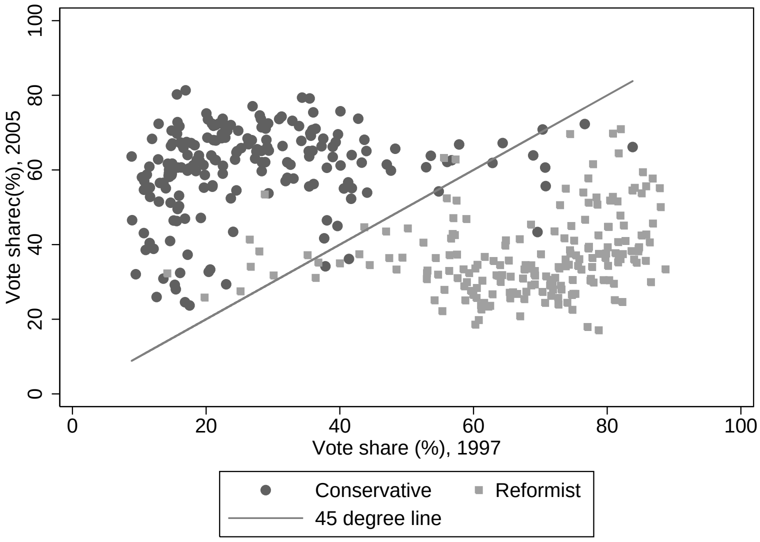


Figure 3.1: Vote shares in 1997 and 2005 presidential elections

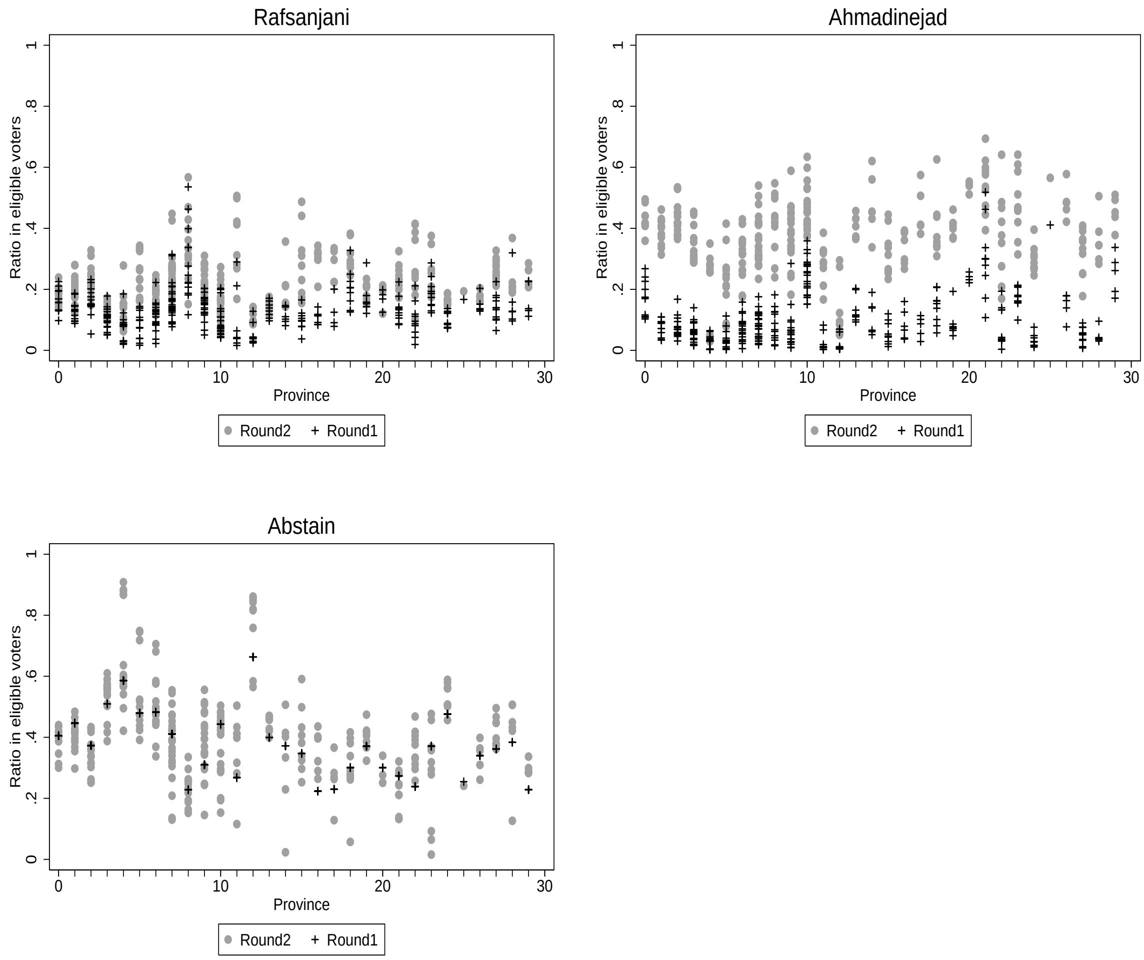


Figure 3.2: Ratio of three alternatives in total number of eligible voters

Table 3.1: Analysis, methods and variables employed in analyzing the 2000 U.S presidential election

Author(s)	Weisberg Hill	Norpoth	Kessel	Box-steffensmier et al.	Jacoby	Stanley & Niemi	Kanthak & Norrander	Kimball	Finkel & Freedman	Burden
Data	1952-2000 NES	1992-2000 VNS exit polls historical statistics	1952-2000 NES	2000 American politics study	2000 NES	1952-2000 NES	1952-2000 NES	1980-2000 NES & aggregate statistics	1992-2000 NES	VNS exit polls & aggregate statistics
Analusis Method	Multi-stage logit	Logit & time series	Probit	Bivariate probit	Logit & interactions	Logit	Regression & interactions	Logit	Logit	Discrete choice
	Major party vote	Major-party vote	Major-party vote	Voting turnout and major-party vote	Major-party vote	Party identification	Candidate theremometer difference	Split-ticket voting	Voting turnout	Vote, including minor parties
Partisanship	yes	yes		yes	yes		yes	yes		yes
Economy	yes	yes		yes	yes		yes			yes
Campaign finance				yes						
Ideology	yes				yes		yes	yes		yes
Issues	yes				yes		yes			
Social issues	yes						yes			yes
Candidate personality	yes			yes	yes			yes		
Demographics	yes			yes		yes	yes		yes	yes
Gender	yes			yes		yes	yes		yes	yes
Race	yes					yes			yes	yes
Region	yes					yes		yes		
Religiosity	yes					yes			yes	
Age	yes			yes		yes			yes	yes
Marriage	yes			yes					yes	yes
Education	yes			yes					yes	yes
Income									yes	
Union						yes			yes	
Length of						yes			yes	yes
Residence									yes	
Home owner									yes	
Sexual Oreintation										yes

Table 3.2: Share of candidates in rounds one and two (percent)

	Round 1	**Round 2**
Rafsanjani (R)	22.00	36.75
Ahmadinejad (C)	20.33	63.25
Karoubi (R)	18.00	-
Moin (R)	14.49	-
Qalibaf (C)	14.50	-
Larijani (C)	6.11	-
Mehralizadeh (R)	4.58	-
Participation rate	70.81	68.90

R: Reformist, C: Conservative

Source: Ministry of Interior

Table 3.3: 2005 Presidential election results, Round 1 & Round 2

	Round 1								Round 2		
Province	Rafsanjani	Ahmadi-nejad	Karoubi	Moin	Qalibaf	Larijani	Mehr-alizadeh	Abstain	Rafsanjani	Ahmadi-nejad	Abstain
Markazi	24.76	27.97	18.08	11.35	12.43	2.99	2.43	40.52	28.89	71.11	38.61
Gilan	21.41	14.81	20.26	18.12	17.05	4.98	3.38	44.66	33.89	66.11	40.96
Mazandaran	23.58	12.04	7.80	11.22	8.83	35.14	1.40	37.29	34.16	65.84	34.16
E.Azarbaijan	20.56	15.16	9.32	14.54	9.34	2.15	28.94	50.97	30.30	69.70	53.67
W.Azarbaijan	19.10	9.49	12.58	18.52	17.81	1.95	20.56	58.61	36.86	63.14	67.24
Kermanshah	19.06	9.76	35.45	14.86	16.06	3.07	1.74	47.98	49.18	50.82	52.96
Khuzestan	21.81	15.87	36.72	10.12	10.10	3.99	1.37	48.25	37.01	62.99	50.31
Fars	22.81	13.73	30.94	12.29	15.48	3.47	1.27	41.09	41.45	58.55	37.59
Kerman	41.47	11.16	13.19	4.57	9.68	19.10	0.84	22.81	48.61	51.39	23.54
R.Khorasan	20.95	15.00	11.83	12.91	34.84	3.14	1.33	31.00	35.06	64.94	38.26
Isfahan	14.84	45.61	11.18	11.17	11.29	4.18	1.73	44.27	26.41	73.59	37.80
Sistan	18.04	5.55	8.96	55.72	7.98	2.90	0.85	26.80	55.05	44.95	35.92
Kurdestan	15.54	6.43	32.02	26.73	14.08	2.24	2.95	66.34	46.07	53.93	77.28
Hamadan	22.25	24.66	27.56	10.67	9.23	3.03	2.59	39.96	28.18	71.82	44.68
Bakhtiari	16.26	24.84	20.50	13.21	17.50	6.32	1.38	37.21	26.95	73.05	31.83
Lorestan	15.27	8.79	55.51	6.78	8.85	3.93	0.87	34.73	49.01	50.99	39.18
Ilam	14.04	11.20	37.59	19.56	14.21	2.35	1.05	22.37	47.22	52.78	35.82
Kohkiloyeh	17.99	11.02	30.91	16.33	16.74	6.51	0.50	22.99	37.82	62.18	26.23
Bushehr	23.96	20.26	24.14	16.86	11.55	2.02	1.22	30.07	42.83	57.17	29.93
Zanjan	24.69	20.81	14.02	15.31	15.92	5.10	4.14	37.11	31.07	68.93	39.90
Semnan	24.80	34.84	9.20	9.44	13.17	7.18	1.38	30.05	26.75	73.25	27.96
Yazd	17.20	38.66	12.83	13.35	14.76	2.06	1.14	27.35	31.53	68.47	22.10
Hormozgan	12.60	13.36	29.57	25.61	4.22	13.03	1.61	23.86	38.72	61.28	35.89
Tehran	25.58	30.13	8.34	13.02	12.33	4.94	5.66	37.11	37.47	62.53	27.23
Ardebil	20.06	7.16	11.32	14.10	22.32	1.63	23.41	47.61	35.55	64.45	53.61
Qom	22.40	55.15	5.44	5.99	5.55	2.35	3.11	25.46	25.51	74.49	24.11
Qazvin	21.90	23.81	16.40	13.74	15.56	3.63	4.96	34.01	25.90	74.10	33.78
Golestan	22.19	8.10	27.62	22.38	12.49	6.04	1.18	36.20	45.24	54.76	40.80
N.Khorasan	20.38	6.64	25.92	10.81	28.97	4.89	2.38	38.38	39.25	60.75	40.72
S.Khorasan	20.04	35.59	9.70	13.75	17.17	2.00	1.74	22.84	32.77	67.23	29.84
Total	22.00	20.33	18.00	14.49	14.50	6.11	4.58	39.25	36.75	63.25	40.92

Table 3.4: Data Summary

	Share of votes		2nd round support	
	Average (%)	Std. dev. (%)	Ahmadinejad	Rafsanjani
First round				
Abstention	39.51	0.09		
Ahmadinejad	12.29	0.08	x	
Rafsanjani	13.30	0.04		x
Karoubi	10.89	0.07		x
Moin	8.76	0.06		x
Mehralizade	2.77	0.04		x
Larijani	3.69	0.05	x	
Qalibaf	8.77	0.05	x	
Second round				
Abstention	41.28	0.15		
Ahmadinejad	36.88	0.12		
Rafsanjani	21.84	0.08		
Demographic	Average	Std. dev.	# of observations	
Age	34.19	15.59	71,728	
% female	0.51	0.50	71,728	
% urban	0.46	0.50	71,728	
Education	6.23	4.87	71,728	
Log income	13.02	0.94	71,728	
gini	0.37	0.06	292	
% unemployment	0.13	0.08	292	
% religious minority	0.13	0.29	292	
% ethnic minority	0.30	0.37	292	
% conservative support	0.38	0.12	292	
Population	576,213	1,317,956	292	

Table 3.5: Average shares (%) for different groups of eligible voters

quartile	Abstention (%)	Ahmadinejad (%)	Rafsanjani (%)
Education quartile:			
1	49.87 (18.44)	30.48 (12.82)	19.65 (9.75)
2	40.87 (13.02)	37.74 (11.23)	21.39 (8.32)
3	37.23 (12.73)	39.68 (11.21)	23.09 (8)
4	37.14 (9.66)	39.64 (7.73)	23.22 (7.12)
Urban ratio quartile:			
1	39.3 (13.26)	37.35 (10.94)	23.35 (9.06)
2	43.6 (15.95)	34.77 (11.41)	21.63 (9.07)
3	41.19 (14.9)	36.43 (11.28)	22.38 (8.98)
4	41.01 (14.57)	38.99 (12.14)	20 (6.03)
Income quartile:			
1	43.39 (15.66)	33.85 (12.2)	22.76 (8.55)
2	42.17 (16.3)	37.6 (12.85)	20.23 (8.67)
3	39.76 (13.28)	38.55 (10.29)	21.69 (8.74)
4	39.79 (13.3)	37.54 (10.06)	22.67 (7.66)
Unemployment quartile:			
1	45.24 (18.03)	35.88 (13.19)	18.88 (7.26)
2	41.08 (13.04)	36.86 (9.68)	22.06 (10.11)
3	40.52 (13.68)	37.05 (11.86)	22.42 (8.39)
4	38.26 (12.89)	37.74 (11.13)	23.99 (6.94)
Ethnic minority quartile:			
1	41.14 (13.12)	38.74 (9.94)	20.12 (7.21)
2	34.39 (10.03)	43.57 (7.93)	22.04 (7.7)
3	40.91 (10.11)	36.95 (8.66)	22.14 (7.79)
4	48.78 (19.97)	28.15 (13.18)	23.07 (10.56)

Table 3.6: Model estimates

	Model A		Model B			
	OLS		Mult. Logit			
	District level		Individual & district level			
			Estimate		Marginal effect	
	Ahmadinejad	Rafsanjani	Ahmadinejad	Rafsanjani	Ahmadinejad	Rafsanjani
Age	-0.030	-0.052	-4.01	-1.82	-0.483	0.170
	(0.013)	(0.015)	(0.79)	(0.49)	(0.248)	(0.227)
Education	0.139	0.219	5.60	7.49	0.158	0.443
	(0.040)	(0.040)	(1.15)	(1.3)	(0.438)	(0.272)
Female	-6.540	-4.488	-2.43	-1.85	-0.214	0.000
	(1.622)	(1.545)	(0.64)	(0.71)	(0.135)	(0.093)
Urban	-1.008	-1.160	-2.89	-3.45	-0.124	-0.172
	(0.271)	(0.282)	(0.65)	(0.69)	(0.201)	(0.117)
Unobserved	-	-	-2.43	-2.77	-0.118	-0.126
			(1.43)	(0.91)	(0.162)	(0.092)
Income	-0.415	-0.293	1.32	-0.08	0.230	-0.149
	(0.172)	(0.173)	(0.35)	(0.24)	(0.123)	(0.129)
Religious minority	-1.234	-0.364	-1.30	-0.55	-0.161	0.061
	(0.196)	(0.189)	(0.65)	(0.24)	(0.082)	(0.077)
Ethnic minority	-0.346	-0.085	-0.52	-0.31	-0.055	0.012
	(0.148)	(0.160)	(0.19)	(0.13)	(0.03)	(0.025)
Population	-0.005	-0.007	-1.79	-1.19	-0.177	0.025
	(0.004)	(0.005)	(0.57)	(0.43)	(0.101)	(0.078)
Unemployment	0.651	0.823	1.42	1.59	0.071	0.071
	(0.481)	(0.485)	(0.53)	(0.63)	(0.093)	(0.053)
Gini	1.461	1.819	2.37	2.53	0.133	0.100
	(0.455)	(0.474)	(1.11)	(1.18)	(0.15)	(0.085)
Partisanship	1.905	1.252	2.20	1.44	0.219	-0.032
	(0.374)	(0.414)	(0.43)	(0.32)	(0.124)	(0.096)
Constant	6.897	4.328	-0.86	-1.78		
	(2.006)	(2.003)	(0.53)	(0.49)		
Observations	292	292				
R-sq	0.51	0.26				

Bibliography

Ackerberg, D., L. Benkard, S. Berry, and A. Pakes (2007). Econometric tools for analyzing market outcomes. In J. J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6A, Chapter 63. North-Holland Press.

Alvarez, M. R. and J. Nagler (1994, December). Correlated disturbances in discrete choice models: A comparison of multinomial probit models and logit models. Working Papers 914, California Institute of Technology, Division of the Humanities and Social Sciences.

Alvarez, R. M. and J. Nagler (1995). Economics, issues and the perot candidacy: Voter choice in the 1992 presidential election. *American Journal of Political Science 39*(3), 714–744.

Alvarez, R. M. and J. Nagler (1998). When politics and models collide: Estimating models of multiparty elections. *American Journal of Political Science 42*(1), 55–96.

Beck, P. A. (1986). *Political science: The science of politics*, Chapter Model choice in political science: The case of voting behavior research. New York: Agathon.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics 25*(2), 242–262.

Campbell, A., P. E. Converse, W. E. Miller, and D. E. Stokes (1960). *The American Voter*. University Of Chicago Press.

Downs, A. (1957). *An Economic Theory of Democracy*. New Yrok: Harper and Row.

Dows, J. and J. Endersby (2004). Multinomial probit and multinomial logit: a comparison of choice models for voting research. *Electoral Studies 23*, 107–122.

Dubin, J. A. and D. Rivers (1989). Selection bias in linear regression, logit and probit models. *Sociological Methods & Research 18*, 360–390.

Erikson, R. S. (1981). Why do people vote? because they are registered. *American Politics Quarterly 9*, 259–76.

Jackson, R. A. (1996). A reassessment of voter mobilization. *Political Research Quarterly 49*(2), 331–349.

Kanthak, K. and B. Norrander (2004). *Models of Voting in Presidential Elections: The 2000 U.S. election*, Chapter The Enduring Gender Gap. Stanford, CA: Stanford University Press.

Lacy, D. and B. C. Burden (1999). The vote-stealing and turnout effects of ross perot in the 1992 u.s. presidential election. *American Journal of Political Science 43*(1), 233–255.

Lazarsfeld, P. F. (1968). *People's Choice*. Columbia University Press.

Lee, G., A. Lifeson, and P. Neil (1980). *Freewill. Permanent Waves*. Mercury Records.

Lodge, M., K. M. McGraw, and P. Stroh (1989). An impression-driven model of candidate evaluation. *The American Political Science Review 83*(2), 399–419.

McKelvey, R. D. and P. C. Ordeshook (1985). Sequential elections with limited information. *American Journal of Political Science 29*(3), 480–512.

Nevo, A. (2000.). A practitioners guide to estimation of random-coefficients logit models of demand. *Journal of Economics and Management Strategy 9*, 513548.

Niemi, R. G. (2001). *Controversies in voting behavior*. Washington D.C.: CQ press.

Niemi, R. G. and H. F. Weisberg (Eds.) (1976). *Controversies in American Voting Behaviour*. W.H.Freeman & Co Ltd.

Norpoth, H. (2004). From primary to general election: A forecast of the presidential vote. *Political Science and Politics 37*, 737–740.

Paap, R., E. van Nierop, H. J. van Heerde, M. Wedel, P. H. Franses, and K. J. Alsem (2005). Consideration sets, intentions and the inclusion of "don't know" in a two-stage model for voter choice. *International Journal of Forecasting 21*(1), 53–71.

Quinn, K. M., A. D. Martin, and A. B. Whitford (1999). Voter choice in multi-party democracies: A test of competing theories and models. *American Journal of Political Science 43*(4), 1231–1247.

Schofield, N., A. D. Martin, K. M. Quinn, and A. B. Whitford (1998). Multiparty electoral competition in the netherlands and germany: A model based on multinomial probit. *Public Choice 97*, 257–293.

Teixeira, R. A. (1992). *The Disappearing American Voter*. Washington D.C.: Brookings Institution.

Timpone, R. J. (1998). Structure, behavior, and voter turnout in the united states. *The American Political Science Review 92*(1), 145–158.

Train, K. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Weisberg, h. F. and T. G. Hill (2004). The succession presidential election of 2000. In H. F. Weisberg and C. Wilcox (Eds.), *Models of voting in presidential elections: the 2000 U.S. election*, Chapter 2, pp. 27–49. Stanford: Stanford University Press.

Wolfinger, R. E. and S. J. Rosenstone (1980). *Who votes?* New Haven, CT: Yale University Press.