ABSTRACT

Title of dissertation: ESSAYS ON THE ECONOMICS OF WOMEN, WORK AND FAMILY

Lisa Jeanne Dettling
Doctor of Philosophy, 2013

Dissertation directed by: Professor Melissa S. Kearney
Department of Economics

This dissertation explores recent trends in female labor supply and fertility, focusing on the impact of two modern phenomena that affected women's work and family formation decisions: the introduction of residential high speed Internet and the recent housing boom and bust cycle. The first chapter provides an introduction and discussion. The second chapter investigates the impact of home Internet usage on women's labor supply. From shopping to telecommuting, home high speed Internet has affected where, when and how individuals conduct numerous activities, and many of these changes could plausibly alter labor supply decisions. Utilizing exogenous variation in Internet usage induced by supply-side constraints to residential broadband Internet access, I find that married women who use the Internet at home are more likely to participate in the labor force. In the third chapter, I explore the potential mechanisms explaining this increase in labor supply by examining data on Internet usage, telework and time use. I find that telework, job search and time saved in home production are all plausible mechanisms through which Internet usage affects female labor supply, and the ability to engage in part-time telework through an employer is the greatest contributor to the estimated effects.

In the fourth chapter, coauthored with Melissa Kearney, we examine the effect of the recent housing boom and bust cycle on fertility decisions. Recognizing that housing is a

major cost associated with child rearing, and assuming that children are normal goods, we hypothesize that an increase in house prices will have a negative price effect on current period fertility. This applies to both potential first-time homeowners and current homeowners who might upgrade to a bigger house with the addition of a child. On the other hand, for current homeowners, an increase in house prices will increase home equity, leading to a positive effect on birth rates. Employing data on Metropolitan Statistical Area (MSA)-level house prices, MSA-group level fertility rates, and MSA-group level home ownership rates, we find that indeed, short-term increases in house prices lead to a decline in births among non-owners and a net increase among owners.

ESSAYS ON THE ECONOMICS OF WOMEN, WORK AND FAMILY

by

Lisa Jeanne Dettling

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Professor Melissa S. Kearney, Chair/Advisor
Professor Judith K. Hellerstien
Professor Sergio Urzua
Professor Jessica Goldberg
Professor Joan Kahn

Dedication

To Jake Grover, whose support and encouragement made this possible.

Acknowledgments

I would especially like to thank my mentor, coauthor and friend Melissa Kearney, who provided countless hours of advice and support throughout this process and provided me with excellent opportunities to learn. I am also grateful for the excellent comments and suggestions I received from Judy Hellerstein and Sergio Urzua throughout my Ph.D, and especially while working on my job market paper. There are many others who provided advice and support throughout this process, including Jessica Goldberg, Lesley Turner, Belen Sbrancia, Soohyung Lee, Tim Moore, John Ham, Ethan Kaplan, and John Shea. I'd also like to thank participants from various seminars, particularly at the University of Maryland, for helpful comments and suggestions. Thank you also to Jessica Goldberg and Joan Kahn for serving on my dissertation committee. I thank the Economic Club of Washington and the training program at the Maryland Population Research Center for financial support. Finally, I would like to thank my friends and family for their help and encouragement.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The past century witnessed dramatic changes in the role of women in society, as women increasingly entered the labor market and birth rates plummeted. In 1910, the average 40 year old woman had 4 children and a 21 percent chance of working. In 1990, the average woman had 2 children and a 78 percent chance of working (Gibson, 2013). In the 1990s, however, a divergence in the long run trend began to emerge for certain groups of women: college educated women began having more children and their labor force participation rates began to decline (Shang and Weinberg, 2009; Macunovich, 2010). This change popularized the term "opting out," which referred to the hypothesis that highly skilled women were increasingly opting out of the labor force to be stay-at-home mothers.[1] While anecdotal evidence abounded, more rigorous empirical investigations into the veracity of this hypothesis found mixed results (Boushey, 2005; Herr and Wolfram, 2009; Macunovich, 2010). And in the early 2000s, participation rates for highly educated women with children began to rise again.[2]

This dissertation will explore two possible explanations for recent trends in labor supply and family formation: the diffusion of home Internet technology and the recent housing boom and bust cycle. Between 2000 and 2009, home high speed Internet usage increased from 5 percent to 74 percent. Moreover, home Internet is increasingly used on a regular basis for more numerous aspects of daily life, from working from home to saving time in home production tasks like shopping and paying bills. In chapters 2 and 3, I explore the impact of this technology on women's labor supply. I find that high speed Internet has led to an increase in labor supply among highly educated women with children, which can primarily be explained by use of home Internet to telework. The final chapter of this dissertation, co-authored with Melissa Kearney, investigates the impact of this recent housing cycle on

[1] See, for example Belkin (2003); Wallis (2004); Story (2005).

[2] Figure 2.4 displays trends in participation among college educated women with children since 2000.

fertility rates. Between 1997 and 2006, home prices rose 70 percent, and between 2006 and 2010, they fell 40 percent.[3] We show that the recent housing boom led to a sizable increase in aggregate birth rates –driven by home owners who presumably felt "richer"– while the subsequent decline led to a net decrease.

The second chapter of this dissertation investigates the impact of home Internet usage on female labor supply. From telework to socializing, home Internet technology has changed how, when and where individuals conduct a diverse range of activities, each of which can affect labor supply decisions in different ways. On the one hand, home Internet can reduce the costs of working by allowing individuals to work from home and engage in flexible scheduling. It can also reduce search frictions if the Internet is used to search for jobs, and save time in tasks like shopping and paying bills. These uses of Internet technology could lead to increases in labor supply. On the other hand, home Internet provides numerous new leisure and entertainment options, which may lead to reductions in labor supply. The net effect of Internet usage therefore depends on the extent to which individuals use Internet for each of those activities and the responsiveness of female labor supply along each margin.

The goal of the second chapter is to estimate the net effect of Internet usage on women's labor supply. To do so, I employ individual level data on home high speed Internet usage and labor market outcomes, which can be found in the Current Population Survey (CPS) between the years 2000 and 2009. The obvious obstacle in conducting such an exercise is that home high speed Internet usage is likely to be endogenous to work patterns, and the direction of the bias is not *ex ante* clear. On the one hand, monthly subscription fees may be cost prohibitive for individuals who do not work, leading a naive estimate to overstate the true impact of Internet use on labor supply. On the other hand, individuals who work may spend less time in the home and therefore place less value on a technology that is only used within the home. In this case, a naive estimate will understate the true impact of home

[3] Author's calculation from the Federal Housing Finance Agency House Price Index, based on a weighted average over the 154 metropolitan areas that comprise the sample used in the fourth chapter of this dissertation.

Internet use on labor supply.

To estimate the causal effect of Internet usage on labor supply I employ an instrumental variables (IV) strategy that isolates exogenous variation in high speed Internet usage induced by cross-sectional variation in supply side constraints to high speed Internet access. Namely, I exploit the fact that aspects of the underlying housing infrastructure –the fraction of the population residing in multiple family dwellings– affected the costs of provision for Internet service providers. This difference in costs meant that, all else equal, states with a greater concentration of the population residing in these types of dwellings received Internet access earlier. My identification strategy exploits these differences in trends, and the identifying assumption is that the fraction of the state population residing in multiple family dwelling would not have been correlated with trends in labor supply in the absence of broadband Internet diffusion.

The results of this analysis indicate that home high speed Internet usage has led to large and statistically significant changes in the labor supply of married women, both on the extensive and intensive margin. The results indicate the home high speed Internet usage leads to a 4.1 percentage point increase in participation. The aggregate effects, however, mask significant heterogeneity across groups and the effects are only found for highly educated women with children. Effects among childless women, less educated women, and men are found to be small and insignificant, and a falsification test indicates that there is no reduced form impact of the instrument on labor supply prior to the diffusion of high-speed Internet.

In the third chapter, I explore the mechanisms underlying the increase in labor supply induced by home use of high speed Internet. I begin the analysis by investigating summary data on the ways in which individuals use the Internet, both across groups and over time. I find that Internet users use the Internet for a wide range of tasks, including work-related activities, leisure activities and home production tasks. Then, I propose a theoretical model which incorporates all of the different ways individuals use the Internet into a labor supply decision-making framework. The model considers four broad categories of Internet usage:

telework, home production, job search and leisure, and creates predictions for the effects of Internet usage for each activity on labor supply decisions.

I empirically examine the role of Internet usage for telework, home production, job search and leisure in explaining the increases in participation and hours found in the second chapter by employing data on Internet usage, telework, employment histories and time use. Using those data, I conduct various additional analyses, including examining the relationship between telework patterns and Internet usage, studying the effect of Internet job search on durations of transitions from non-participation to participation, and comparing the amount of time Internet users and non-users spend in home production tasks. I conclude that telework, job search and time saved in home production are all plausible mechanisms explaining the results, and the ability to engage in part-time telework through an employer appears to be the greatest contributor to the estimated effects.

The fourth chapter of the dissertation, coauthored with Melissa Kearney, investigates how changes in Metropolitan Statistical Area (MSA)-level house prices impact household fertility decisions. The conceptual approach is based on an economic model of fertility that recognizes that changes in house prices potentially have offsetting effects on fertility. Recognizing that housing is a major cost associated with child rearing, and assuming that children are normal goods, we hypothesize that an increase in real estate prices will have a negative price effect on current period fertility. This applies to both potential first-time homeowners and current homeowners who might upgrade to a bigger house with the addition of a child. On the other hand, for current homeowners, an increase in MSA-level house prices will increase home equity, leading to a positive effect on birth rates. The net effect of house prices on aggregate birth rates will depend on individual's responsiveness along these margins and rates of home ownership.

We estimate the effect of movements in aggregate house prices on aggregate birth rates by employing data on MSA-level median home prices, MSA-group level fertility rates and MSA-group level home ownership rates. Our main analyses consists of a set of ordinary least squares (OLS) regressions of MSA-group fertility rates on MSA-level house prices. Importantly, we interact our measure of house prices with the measure of MSA-group level home ownership rates so that we are able to separately identify the hypothesized price and wealth effects which are expected to effect owners and renters differently. Controlling for MSA fixed effects, trends, and time-varying conditions, our analysis finds that short-term increases in house prices lead to a decline in births among groups with low levels of ownership and an increase among groups with high level of ownership.

The key to our empirical approach is to separately identify effects for owners and non-owners. Therefore, any alternative explanation (other than a causal effect of house prices) for these observed relationships must differentially impact owners and non-owners. To address the possibility that other local factors that drive both local area house prices and affect the fertility behavior of owners and renters differently are biasing our OLS estimates, we additionally implement an instrumental variables (IV) strategy that exploits exogenous variation in house price movements induced by variation across MSAs in their housing supply elasticity. Results using the instrumental variables strategy are qualitatively similar to the OLS results.

The main results of our analysis are statistically significant and economically meaningful. Employing our regression estimates in a straightforward simulation exercise, we find that a \$10,000 increase in home prices is associated with a 5 percent increase in fertility rates in MSA cells with 100 percent ownership rates. For MSA cells with zero percent home ownership rates, we estimate a corresponding decrease in fertility rates of 2.4 percent. Under the assumption of linear effects, these estimates suggest that all else held constant, the roughly \$108,000 average increase in house prices during the housing boom of 1997 to 2006 would have led to a 9 percent increase in births over that time.

The results of this dissertation have important demographic and policy implications. In the second chapter, I find that Internet technology is important for the labor supply decisions of women with families. This suggests that by freeing up time spent in home production and permitting flexible scheduling, Internet usage facilitates work-family balance. It also contributes to a literature on the impact of technological progress in the home on labor market outcomes. In the third chapter, I find that Internet technology has allowed women to increase their labor supply primarily via work from home opportunities. This speaks to the potential role of telework and flexible scheduling as policy levers to encourage participation. To date, there is limited evidence on the impact of telework and flexible scheduling policies in the workplace, and take-up is still low. This is despite the potential for these policies to benefit both employers and employees, and even increase aggregate productivity if individuals are more productive at home or these policies encourage the participation of productive workers.[4] The second and third chapters shed new light on the possible benefits of those policies by showing that the diffusion of high speed Internet (via enabling telework participation) has increased the participation of a group of highly skilled individuals.

In the fourth chapter, I find that the recent housing boom and bust cycle had a sizable impact on aggregate fertility rates. As an issue of economic demography, understanding the impact real estate cycles have on birth rates is interesting and informative. Moreover, our finding that house prices exert a larger influence on fertility rates than unemployment rates can inform future work on the demographic impact of expansions and contractions of the economy, and suggests that economic models of fertility should consider movements in the housing market separately from the labor market. Our work is also related to a literature on the nature of the demand for children, and provides that literature with a useful test of wealth effects on fertility decisions. Finally, our work speaks to a body of work on the propensity to consume out of housing wealth, and provides insight into another way individuals may "consume" out their home equity.

This dissertation has provided new insight into recent trends in fertility and labor supply for American women. In the second chapter, I find that the diffusion of home high speed

[4]For a discussion of current research on telework and flexible scheduling policies, see the report on Work Life Balance and Flexibility, published by the Council of Economic Advisers (2010)

Internet has increased labor supply for highly educated women with children. This finding is of particular interest given that much of the attention focused on this group in the early 2000s concerned the "opt out revolution" (Belkin, 2003; Wallis, 2004; Story, 2005). Yet, after a decline in participation in the 1990s, participation among highly educated women with children began to rise in the 2000s. I find that home Internet leads to significant increases in labor force participation for this group of women, and a back of the envelope calculation suggests that high speed Internet can explain 90 percent of the increase in female labor force participation between 2000 and 2010. The third chapter finds that Internet usage for telework and time saved in home production can explain the increase in participation that occurred due to the diffusion of home high speed Internet. This suggests that by allowing for workplace flexibility and lowering the costs of combining work and family, home Internet permitted mothers to "opt back in" to the labor force.

At the same time this change in labor market participation occurred, fertility rates among highly educated women climbed (Shang and Weinberg, 2009). This trend is potentially consistent with the "opt out revolution" hypothesis, or a reduction in the costs of combining work and family, but I leave it to future research to investigate those mechanisms. The late 1990s and early 2000s, however, was period in which there was a major change in economic conditions which could have differentially affected highly educated women: the dramatic rise in home prices. In the last chapter of this dissertation, I find that the recent housing boom led home owners to significantly increase their fertility. Since college educated women have relatively high rates of home ownership, this finding is a plausible explanation for the differential rise in fertility among this group.[5]

This dissertation has found that two major changes which occurred in recent decades –the introduction of high speed Internet and the housing boom and bust– have affected women's work and family formation decisions. These findings have important implications

[5] In the sample used in the fourth chapter of this dissertation, 53 percent of college graduates are home owners, as compared to 29 percent of high school dropouts. At these rates of ownership, our simulation results (presented in figure 4.3) imply that the net impact of a rise in house prices would be an *increase* in fertility rates for college graduates and a *decrease* in fertility rates for high school dropouts.

for understanding recent demographic trends and the possible impacts of policies that reduce the costs of combining work and family. They also open up several important avenues for future research. In the second and third chapters, I provide evidence that the introduction of high speed Internet increased the labor force participation of married women by enabling telework through an employer. Future research could uncover whether or not the effects of Internet usage are concentrated on this particular group because telework opportunities were already most readily available for that group, and whether extending telework opportunities to less educated women could have a similar effect on those groups of women in the future. More generally, research into the potential costs and benefits of telework –both for employees, employers, and society as a whole– will make a valuable contribution to policy discussion in those areas. In the fourth chapter, we find that the recent housing boom and bust had a sizable impact on fertility decisions. Future work could uncover whether fluctuations in real estate markets affect other aspects of family formation, including household formation, marriage, and divorce. These issues will be of particular interest in the future as the implications of the recent housing cycle are fully realized.

Chapter 2

Opting Back In: Home Internet Use and Female Labor Supply

2.1 Introduction

From shopping to telecommuting, the Internet has changed how, when, and where individuals do myriad activities. From 2000 to 2009, home high-speed Internet usage rose from 5 percent to 74 percent, a change that is remarkable for both its speed and depth of penetration.[6] Yet, scant evidence has been brought to bear upon the question of whether or not this technological revolution has had a labor market impact.[7] In this chapter, I investigate whether home Internet use has altered married women's labor supply, a group for whom technological progress in the home has been shown to affect labor supply in the past.[8] The conceptual approach is based on an economic model of time allocation which incorporates the diverse range of tasks for which the Internet is an input and indicates that Internet usage can have offsetting effects on labor supply. On the one hand, home Internet can reduce the time and monetary costs of working by allowing individuals to work from home. It can also reduce search frictions in the labor market and save time in home production tasks like shopping and paying bills, freeing up time to engage in market work. On the other hand, the Internet offers a wide range of entertainment options, which may lead to reductions in labor supply. The net effect of home Internet use on labor supply will depend on the extent to which individuals use Internet for each of these tasks and the responsiveness of individual labor supply along each margin.

[6] Author's calculation from the CPS. High-speed Internet diffused more quickly than most other technologies. See for example Faulhaber (2002, figure 10-1) comparing broadband diffusion to VCR and wireless phone diffusion, or Greenwood et al. (2005, figure 1) for appliance diffusion. Of these technologies, only microwave ovens diffused at close to a similar pace (3 percent to 60 percent from 1975 to 1986)

[7] The exception is a more focused literature on Internet job search (e.g., Kuhn and Skuterad, 2004; Stevenson, 2009; Kroft and Pope, 2010; Kuhn and Mansour, 2011; Brencic, 2012)

[8] Greenwood et al. (2005) show that the diffusion of home appliances such as washing machines and microwaves can account for a large portion of the dramatic increase in female labor supply in the twentieth Century. Follow up work by Cavalcanti and Tavares (2008); Coen-Pirani et al. (2008) and Cardia (2010) provides empirical evidence.

The goal of this chapter is to identify the causal impact of home Internet use on labor market outcomes. The main empirical strategy is to employ micro-level current population survey (CPS) data on self-reported home Internet usage and labor market outcomes. A central concern when attempting to identify the causal impact of home Internet use on labor supply is the endogeneity of Internet take up to work patterns, as reverse causality could lead ordinary least squares (OLS) estimates to either overstate or understate the causal relationship. If monthly subscription fees are cost prohibitive for individuals who do not work, OLS estimates will overstate the true impact of Internet use on labor supply. On the other hand, if individuals who work spend less time in the home and place less value on a technology that is only used within the home, OLS estimates will understate the true impact of home Internet use on labor supply.

To overcome the endogeneity of home Internet use to work patterns, I employ an instrumental variables (IV) strategy that exploits cross-state variation in supply-side constraints to high-speed Internet access. Unlike dial-up Internet, high-speed Internet installation required substantial investments by Internet service providers and access was neither immediate nor uniformly distributed across locations. From an Internet service provider's perspective, multiple family dwellings were easier and more profitable for installation. Motivated by this differential investment incentive, I show that geographic variation in the housing infrastructure –namely, the percent of the state population residing in a multiple family dwelling– can predict trends in Internet access and usage. Conditional on state and year fixed effects, a host of time-variant state-level labor and housing market indicators, and various individual-level demographic controls, the identification assumption is that the fraction of a state residing in multiple family dwellings would not have been correlated with subsequent trends in labor supply in the absence of broadband Internet diffusion.

Ordinary least squares estimates indicate a positive correlation between home Internet use and married women's labor supply and the instrumental variables estimates confirm this relationship is causal. The instrumental variables results indicate that a married women who uses the Internet at home is 18.6 percentage points more likely to be in the labor force, however, this estimate is best interpreted as a local average treatment effect (LATE). The magnitude of the IV estimate is indicative that the OLS estimate is biased downwards, which is consistent with non-working women being more likely to adopt the Internet. There is also substantial heterogeneity across sub-groups of women, and increases in participation are only found for more educated women and women with school-aged children. The impact on less educated women, childless women, and men is small and statistically insignificant. Moreover, a falsification test indicates that there is no reduced form impact of the instrument on labor supply prior to the diffusion of high-speed Internet. Finally, increases in hours and the propensity to work full time suggest home Internet increases labor supply on the intensive margin as well.

Aggregate trends in labor supply indicate participation rose among married women over the period studied, but the magnitude of those aggregate increases are much smaller than the effects that I estimate using linear IV. I reconcile this apparent disparity by conducting several exercises intended to gauge the external validity of the linear IV estimate. It is well known that linear IV can only identify a local average treatment effect (LATE) which applies to the population whose treatment status is affected by the instrument (Imbens and Angrist, 1994). In practice, membership in this subgroup (the "compliers") is unobservable and it is not possible extrapolate the estimated effects outside this unknown population without assuming homogeneous treatment effects. Therefore, I begin by estimating the size and characteristics of the compliant sub-population. I find that this group is both small and observably different from the overall population. Then, I decompose the LATE using a non-parametric local IV estimator to estimate marginal treatment effects (MTEs) across the population and calculate the weights that the linear IV estimand places on those MTEs (Heckman and Vytlacil, 2001, 2005; Heckman, Urzua, and Vytlacil, 2006). This exercise indicates there is significant heterogeneity in the treatment effect of Internet use and that individuals sort into Internet usage based on their expected gains from treatment. I find that the linear IV estimate applies to a minority of individuals with large expected and realized treatment effects of home Internet usage, and while the LATE estimate is 18.6 percentage points, the overall unweighted average effect is 4.1 percentage points. An analysis of the observable characteristics of individuals with the largest treatment effects indicates those individuals are primarily college educated women with children.

The main contribution of this chapter is to provide an empirical examination of how home Internet technology has affected labor market outcomes. The diverse range of tasks for which the Internet is an input suggests a plausible labor market impact, and to the best of my knowledge, this is the first work to consider the net effect of individual home Internet use on labor supply. This adds to a more focused literature that has considered the use of Internet technology as a tool for job search (Kuhn and Skuterad, 2004; Stevenson, 2009; Kuhn and Mansour, 2011; Kroft and Pope, 2010; Brencic, 2012). Further, this work contributes to the development of a broad understanding of how technological progress in the home affects economic outcomes. Similar to work by Greenwood et al. (2005), who study the diffusion of washing machines, microwaves, and other home technologies, I find a substantial impact of a home-based technology on female labor force participation.

The findings of this chapter also have important demographic and policy implications. Recent work has proposed that highly skilled women may be increasingly "opting out" of the labor force to be stay-at-home mothers (Boushey, 2005; Herr and Wolfram, 2009). Yet, between 2000 and 2010 labor force participation rose among married women (particularly among highly educated women with children) and remained constant for other groups.[9] The demonstrated heterogeneity in the effects of Internet usage reconcile well with this divergence in labor supply trends, and a back of the envelope calculation indicates that the diffusion of high speed Internet can explain 90 percent of the observed increase in participation for married women during this time period. The result that Internet technology is important for the labor supply decisions of women with families suggests Internet usage facilitates

[9]Note that labor force participation rose during this period, however, employment did not. Employment fell for married women (as it did among other groups) during the 2007-2009 recession.

work-family balance.

2.2 Related Literature

This work is related to three separate literatures in economics. First, this chapter is related to the literature examining the diffusion of household technologies in the twentieth century and their impact on female supply. The literature is motivated by the hypothesis that the introduction of modern appliances freed up women's time to engage in market work and can therefore explain the dramatic rise in female labor force participation in the twentieth century. Greenwood et al. (2005) investigate the theoretical relationship between appliance diffusion and female labor supply using a Becker (1965) time allocation model within a dynamic general equilibrium framework. In their simulation exercises, they find that the diffusion of time-saving appliances can explain 28 percentage points of the 51 percentage point increase in female labor force participation generated in the model. Coen-Pirani, Leon, and Lugauer (2008) and Cavalcanti and Tavares (2008) provide empirical evidence in favor of this mechanism at the individual and cross-country level, respectively. Cardia (2010), on the other hand, finds only modest evidence in favor of the theory using county-level data.[10] This chapter contributes to that literature because home Internet is a modern, time-saving household technology. Home Internet contrasts with the previously technologies studied because it is not only an input in home production tasks such as shopping and paying bills, but also leisure and market work.

This work is also closely related to a literature that examines various determinants of the joint labor supply and family formation decisions of women. Balancing the competing demands of work and family is thought to be partially responsible for the gender wage gap (e.g., Black et al., 2008; Bertrand et al., 2009; Sasser, 2005) and the relative lack of women in leadership positions (e.g., Bertrand and Hallock, 2001). Research suggests that family

[10]Cardia (2010) finds evidence that the diffusion of plumbing increased participation in some occupations. He does not find an impact of refrigerator diffusion.

friendly work policies and subsidized child care can improve the labor market outcomes of mothers (e.g., Ruhm, 2004; Baker and Milligan, 2008; Herr and Wolfram, 2009). Similarly, there is evidence that maternal labor supply is sensitive to the price of child care (e.g., Anderson and Levine, 2000; Gelbach, 2002; Baker et al., 2008; Cortes and Tessada, 2008). I find evidence that home Internet use, which can potentially permit more family-friendly flexible work schedules and decrease necessary child-care expenditures leads to increased female labor force participation. The finding that working from home could be important for women is also related to recent work by Black, Kolesnikova, and Taylor (2008), who find that long commute times discourage labor force participation by married women.

Finally, this chapter is related to work that analyzes the impact of technological change in the workplace. Conceptually and empirically, I provide evidence that part of the net effect of home Internet use on labor supply can be attributed to an increased ability to telework. By changing where and when work is conducted, the introduction of home Internet is therefore related to the introduction of other technologies which changed the nature of market work, most notably the literature on workplace computerization.[11] Weinberg (2000) provides evidence that by decreasing the physicality of market work, computers in the workplace increased the relative demand for female workers. Black and Spitz-Oener (2010) find that much of the narrowing of the gender wage gap can be attributed to computerization and technological change in the workplace. This chapter adds to that literature because it also finds that technology can be particularly beneficial for women.

2.3 Conceptual Framework

I analyze the theoretical impact of home Internet use on labor supply in a Becker (1965) model of time allocation, which will permit an examination of the effect of Internet use as

[11] To be clear, I do not presume to speak directly as to how Internet use *in the workplace* may have altered productivity (although it certainly may have). Therefore, this work is not directly related to the literature on skill biased technical change, which is reviewed by Violante (2008).

an input in leisure, home production, telework, and job search. In what follows, I briefly summarize the theoretical predictions of that model, which is described in detail in the third chapter of this dissertation. In the standard formulation of Becker's model, an agent derives utility from a set of commodities that are produced in home using a combination of inputs purchased in the market and the agent's own time. These commodities include home production goods like meals or a clean home, as well as leisure goods like watching television. As an example, purchased inputs may include raw ingredients and cooking utensils, which produce a meal in combination with time spent cooking and eating. The agent's time can be allocated between home and market work. Time in the market earns a wage, and this income can be used to purchase inputs for production. The agent's labor market participation decision reduces to a comparison of the value of his/her time spent working in the market (e.g., the market wage (w)), to the value of his/her time spent at home (e.g., the reservation wage ($w*$)).

The introduction of home Internet technology can change the calculus of labor supply decisions in two fundamental ways: it can alter the market wage (w) and/or it can alter the reservation wage (w*). Changes in the reservation wage occur because home Internet represents a new, more cost effective input in the production of home goods. The effect this will have on the reservation wage depends on whether the good in question is better classified as a home production good or a leisure good. For a home production good, a technological improvement in inputs will tend to reduce the time agents need to spend in those tasks, reducing the reservation wage. This is because the elasticity of substitution between time and inputs in production tends to be relatively high for home production goods (Aguiar and Hurst, 2007). Thus, a technological improvement in inputs used in the production of home production goods like shopping or paying bills will decrease time spent in those tasks, reducing the reservation wage, and encouraging market entry for those on the margin. The opposite will be true for leisure goods, where the elasticity of substitution between time and inputs tends to be smaller. In this case, Internet technology will tend to increase the time individuals spend in leisure, increasing the reservation wage, and reducing labor market participation.[12]

Home Internet can also affect the market wage (w). Suppose that in addition to the standard assumptions of the model there is a time cost to working, so that the net wage (w) is some fraction $(1-\rho)$ of the market wage $\overline{w}$, where ρ represents the fraction of work hours the agent must spend commuting to work. This can represent time spent in daily commute or in transit when an emergency or family obligation requires the agent to leave work (say, to meet a repairman). If home Internet technology permits individuals to work at home some or all hours, this will be captured by a reduction in ρ , and hence, an increase in w.[13] Suppose also that the market wage is imperfectly observed by job searchers due to search frictions in the labor market. If the Internet improves search technology, it may increase the propensity the agent observes a high market wage. In each case, an increase in the observed market wage will tend to increase participation for those on the margin.[14]

In sum, this framework provides an ambiguous prediction for the net effect of home Internet on labor supply which depends upon the extent to which individuals use Internet for work, job search, home production and leisure and the responsiveness of individual labor supply along each margin. Ultimately, the size and magnitude of the net effect of Internet usage on labor supply will be an empirical question. However, in conjunction with several stylized facts it is possible to speculate about the *relative* impact of home Internet use across different groups. It is a well established fact that, within married couples, the majority of home production tasks and child rearing is done by women, even conditional on both partner's employment (e.g., Bianchi et al., 2000). This suggests that the ability to save time in home production tasks should have a relatively larger impact on married women than men. Research has also shown that commute time negatively impacts the participation decisions

[12]The model has similar predictions for the effect of a change in the reservation wage on hours worked.

[13]Home Internet could also affect w by increasing productivity if telework reduces absenteeism and/or workplace distractions. Bloom et al. (2012) provide empirical evidence that telework can improve productivity.

[14]The model has more ambiguous predictions for the effect of a change in the net wage on hours worked.

of married women, which suggests a relatively larger impact of telework opportunities for this group (Black et al., 2008). If work from home reduces the costs of paid child care, then women with children, whose labor supply is known to be sensitive to the price of child care, will be particularly affected (Anderson and Levine, 2000; Gelbach, 2002; Cortes and Tessada, 2008). Telework is much more common among more educated individuals, thus, enhanced telework opportunities should have a relatively larger impact on more educated men and women.[15] Finally, female labor supply has historically been more elastic than male labor supply (especially on the extensive margin), suggesting women should exhibit a greater response to new technology than men (e.g., Heckman, 1993). Combining these facts suggests that the Internet should have the largest effect on married women, with the impact potentially differing by education and the presence of children.

2.4 Data and Empirical Strategy

The main empirical strategy used in this chapter is to relate individual home Internet usage to individual labor force participation, controlling for relevant demographic factors, state and year fixed effects, and time-varying state-level controls. To overcome the endogeneity of individual home Internet use to labor supply, I propose an instrumental variables strategy that exploits cross-state variation in supply-side constraints to high-speed Internet access which was induced by differences in state's housing infrastructure. In this section, I describe the main empirical approach and data sources.

2.4.1 Data and Baseline Specification

The primary data source for this chapter is the August 2000 and September 2001 Current Population Survey Computer and Internet Use Supplement, and the October 2003, 2007,

[15] In 2009, 65 percent of teleworkers held a college or post-graduate degree (WorldatWork, 2009).

and 2009 Current Population Survey School Enrollment supplements.[16] The main regressor of interest is home high-speed Internet use, which is a combination of an individual-level indicator for whether or not the individual is reported to use the Internet at home and a household level indicator for whether or not the household has a high-speed broadband Internet subscription. I focus on high-speed Internet, versus dial-up, for both conceptual and empirical reasons. First, high-speed Internet is expected to be a more effective replacement for earlier technologies in the production of many of the activities which are expected to affect labor supply decisions. For example, it has been argued that both telework and shopping online were simply not feasible using slower dial-up connections (Hausman et al., 2001; Bittlingmayer and Hazlett, 2002). Second, unlike dial-up Internet, the diffusion of broadband Internet was hampered by supply-side constraints, which is essential for the identification strategy.

The main empirical approach used in this study is to relate home Internet use to labor force participation for married women using a linear probability model (LPM) of the following form:

$$y_{ist} = HSI_{ist}\beta_1 + X_i\beta_2 + S_{st}\beta_3 + \theta_t + \eta_s + \epsilon_{ist} \tag{2.1}$$

Where y_{ist} is the labor force participation of individual i in state s in year t. HSI_{ist} is an indicator for whether or not the individual uses high-speed Internet in the home. The main outcome of interest is labor force participation, although I provide alternate specifications using employment status, family income and intensive measures of labor supply (full time status and usual hours worked per week) in section 2.5.4. All labor supply outcome variables refer to current labor supply and are asked the month of the CPS supplement. The sample is initially limited to married women aged 18-59 whose husband is present. The specification includes a vector of individual controls X_i, which absorb demographic differences in rates of

[16] I do not use the 1997 and 1998 Internet Supplements because it is not possible to separately identify high-speed from dial-up internet users in those survey years.

home Internet use and labor force participation. The demographic controls used are dummy variables for race, age category, education, spouse's education, categories for number and ages of children under 18, metropolitan area status, and central-city status. Education is determined by highest degree of schooling obtained and split into four categories: high school dropouts; high school graduate or GEDs; college dropouts and associates degrees; and Bachelors, masters, professional and Ph.D's. Age and number of children are in categories based on both the number of children (zero, one, two, or three or more) and given the number of children, whether children are under or over age 6 (or both). This flexibly controls for the number of children which need to be looked after, as well as whether any or all children are expected to be in school. Metropolitan area and central-city status are included to control for rural/urban differences in high-speed Internet access and labor supply. Central-city status is included separately from metropolitan area status because suburban and urban residents may have differential rates of participation and access. Spouse's education is included to proxy for non-labor income, as well as spouse's Internet take-up rates. Table 2.1 summarizes the CPS data.

The analysis also includes a full set of year and state fixed effects, as well a set of time-variant state-level controls. State fixed effect η_s and year fixed effects θ_t are included to ensure the estimated coefficient on HSI_{ist} is net of any time-invariant differences across states and national trends in Internet access and participation. Various state-level control variables were also selected to mitigate concerns that various aspects of the labor and housing market may be correlated with home Internet use and labor supply. Table 2.2 describes these variables and their sources. They include income per capita, housing prices, population density, unemployment rates, and average wages.[17] Unemployment, income per capita, and average wages control for differences across state and time in the labor market, while housing prices and population density control for differences across state and time in the housing and real estate market that may be correlated with work patterns and home Internet use.

[17] Wages are constructed for all employed workers. I alternatively include female-only and male-only wages as controls in table 2.8.

Although businesses adopted the Internet much earlier than households, one could still be concerned that Internet use at work might be correlated with local Internet access and work patterns. For example, job creation in "high tech" industries might increase the participation of individuals with Internet skills, who are also likely to have a high demand for home high-speed Internet access. Thus, I include two measures designed to control for local rates of Internet usage at work: "adoption" and "enhancement", which measure the share of the population in each year employed in industries which use the Internet for each of these purposes. These are constructed from the industry specific measures estimated by Forman, Goldfarb, and Greenstein (2003). Adoption refers to the percent of firms in an industry that use the Internet for any purpose, while enhancement refers to using Internet to enhance business, such as through commercial sales online. These measures are interacted with state-year-industry level employment rates to create a measure of state-year Internet adoption and enhancement rates at work. Finally, standard errors are adjusted for clustering at the state-year level.

2.4.2 Preliminary Linear Probability Model Estimates

Table 2.3 displays the results of estimating equation 2.1 under a variety of specifications. Column (1) presents the parsimonious specification with no demographic or state level control variables, and column (2) adds the state and year fixed effects. Columns (3)-(5) add the individual-level control variables. Column (3) adds the controls for age category, race/ethnicity, education and MSA/Central City status. Column (4) adds the controls for spouse's education, which is intended to proxy for non-labor income. Column (5) adds the controls for the presence and age of children, which were initially omitted since these may be endogenous to the labor supply decision. Finally, column (6) adds the state level controls. This is the preferred and most conservative specification. The results indicate that being a home Internet user is associated with an approximately five percentage point increase in the probability of participating in the labor force. After column (3), the addition of spouse's education, presence/age of children, and the state level controls do not fundamentally alter the point estimate. The sign of the control variables are generally as expected. Relative to a college educated individual, less educated individuals are less likely to participate in the labor force, and relative to having a college educated husband, a woman whose husband has less education is more likely to participate. This is consistent with spousal education proxying for non-labor income.

The LPM estimate suggests a positive correlation between home Internet use and labor supply. This estimate, however, should not be interpreted as causal evidence of the impact of home Internet use on participation because Internet users are not randomly selected and estimates are likely to be biased by reverse causality. The direction of the bias in the LPM estimate, however, is unclear. Individuals who work may be able to afford the monthly subscription costs of Internet access, leading LPM estimates to overstate the true causal impact of home Internet use. On the other hand, individuals who do not work and spend more time in the home may place a higher value on home Internet access and the stream of services it provides, leading OLS estimates to understate the true causal impact of home Internet use. I propose an instrumental variable strategy to overcome these possible sources of bias.

2.4.3 Proposed Instrument

The proposed instrument for this study exploits supply-side constraints to the availability of residential high-speed broadband Internet induced by differences across states in the housing infrastructure. The alternative to high-speed broadband Internet is dial-up Internet service, which was first offered to residential customers in 1995 after the privatization of the Internet. Dial-up services use existing phone lines and do not require any installation from the Internet service provider (ISP), so access is essentially universal and adoption is driven purely by consumer demand. Although most businesses invested in high-speed broadband

Internet in the early 1990s, the service was not offered to residential customers until the end of the decade. The term "broadband" refers to "advanced communications systems capable of providing high-speed transmission of services such as data, voice, and video over the Internet and other networks" (F.C.C., 2010). Transmission can be provided by a wide range of technologies, including digital subscriber lines (DSL), fiber optic cable, coaxial cable, wireless technology, and satellite (F.C.C., 2010). For the majority of the population, the ISP was an existing cable company transmitting through a cable modem or an existing phone company transmitting through DSL.

Broadband access was (and still is) not universal. Cable and phone companies had to retrofit existing lines to enable high-speed two-way traffic. For cable companies, this included laying new fiber-optic lines and installing expensive operating switches and servers (Faulhaber, 2002; Greenstein and Prince, 2007). DSL from phone companies used existing phone lines, but in some areas existing wiring was not of sufficient quality and needed to be upgraded and the phone companies were not generally aware of which areas would need upgraded until they arrived for installation (Grubesic and Murray, 2002; Faulhaber, 2002). Moreover, initial cable and DSL installations required a visit from a service representative (Faulhaber, 2002). There is a general consensus that these costs slowed availability and access did not keep up with consumer demand (Greenstein and Prince, 2007; Faulhaber, 2002). In fact, there is some evidence that availability was not driven by consumer demand for home high-speed Internet connections at all. Faulhaber (2002) argues that cable companies upgraded lines under pressure to keep up with satellite TV competition and the phone companies simply followed suit. News media reports suggest that potential subscribers often faced long wait times for installation.[18]

The fact that residential broadband deployment lagged consumer demand provides a potential source of exogenous variation in home Internet adoption. Information on access, however, would be problematic as an instrumental variable because ISPs may have partially

[18] See, for example, "Broadband: What Happened?" Businessweek 6/11/01 (Rosenbush et al., 2001)

responded to consumer demand for home Internet services in determining where to provide access. If trends in latent consumer demand for Internet services are systematically correlated with trends in labor supply, access itself would not satisfy the exclusion restriction. Instead, I take advantage of aspects of the existing local infrastructure which made it easier and more profitable to install high-speed Internet in some areas first, so that even given similar levels of consumer demand for Internet services, those areas received access earlier than others.

When installing residential high-speed Internet, existing wiring within a home or building does not generally need to be upgraded for either cable-based or DSL Internet. What needs to be upgraded is the wiring that connects the home or building's existing indoor lines to the ISP.[19] From the ISP's perspective, this made certain types of residences easier and more profitable for installation than others. In particular, apartment buildings and other multiple family properties, collectively referred to here as "multiple dwelling units" (MDUs), were preferable to single family homes. Figure 2.1 illustrates the difference between the two types of connections. For an MDU, each length of upgraded wiring that is installed will service multiple customers, allowing for economies of scale and making it easier and more cost effective to provide each potential customer with access. Moreover, since the ISP or MDU owner usually held the property rights to the "home run" wiring within the building, the ISP usually obtained (de facto) monopoly rights to service all families after installation, a further boost to the relative profitability of connecting MDUs.[20] With these differences

[19]Both cable-based and DSL broadband Internet service requires the installation of fiber-optic wiring, which provides high-speed Internet service up to a certain point, from which the signal travels over traditional coaxial cable or copper telephone wiring the rest of the way. These fiber-optic lines may reach the ISPs central office, some remote terminal in the neighborhood, the "curb", or the "demarcation point" (see figure 2.1). The main issue that prevented timely roll-out for the cable companies was capacity. Cable companies had installed some fiber lines in the 1980s to provide better cable service, but each additional customer on a single fiber line reduces the "downstream" capacity, meaning that multiple simultaneous users reduces speeds and could exhaust the system. Thus, in order to provide reliable, high-speed Internet service cable companies needed to add more fiber lines which came closer to residential consumer's homes. For DSL Internet from the phone companies, roll-out was prevented by the need to upgrade the existing telephone wiring, much of which was old and had been split too often to be capable of carrying high-speed two-way traffic. The key insight is that in either case, existing wiring within the home was of sufficient quality to provide individuals with access, while much of the wiring outside the home was not. More technical details on the history of the provision of residential high-speed Internet are available in the appendix to this chapter (section 2.9.2).

[20]When the MDU owner held the property rights on wiring, the ISP would often offer him a discounted personal connection or incentives to recruit tenants. Until 2007, if the ISP held the rights to the wiring

in mind, I propose that areas with more MDUs should have received Internet access earlier than areas with less MDUs.[21]

Data on MDU rates was constructed using data from the 2000 Decennial Census, which records population totals in different types of housing units based on the number of units in the structure. A recent Federal Communications Commission (FCC) ruling defines an MDU as "... a multiple dwelling unit building (such as an apartment building, condominium building or cooperative) and any other centrally managed residential real estate development (such as a gated community, mobile home park, or garden apartment); provided however, that MDU shall not include time share units, academic campuses and dormitories, military bases, hotels, rooming houses, prisons, jails, halfway houses, hospitals, nursing homes or other assisted living facilities." (47 C.F.R. § 76.2000, 2008). Unfortunately, it is not possible to perfectly map the Census data to this definition, since its not clear if a structure like a townhouse or duplex is part of a centrally managed development. Therefore, I looked at several reasonable definitions and chose the one with the most predictive power in the first stage, which was to define an MDU as any unit in a structure with 3 or more units and mobile homes.[22] Results using various alternative definitions are similar and available in appendix table 2.8. In practice, defining an MDU in this way means that MDUs constitute about 25 percent of privately occupied residences.

There are conceptual reasons to believe that places with a higher concentration of MDUs should have received Internet access earlier than other areas and this geographical variation in the diffusion pattern of Internet access motivates the instrumental variable strategy. A

he had a monopoly. In 2007, the FCC issued a ruling aimed at encouraging competition within MDUs by forcing ISPs who owned wiring to share with competing firms. In practice, there is some skepticism about whether this policy has been effective and there is anecdotal evidence that ISPs simply shifted property rights to the MDU owners (F.C.C., 2007).

[21] I am not the first to suggest this type of adoption pattern may exist. It has been noted that one reason for South Korea and Hong Kong's relatively high adoption rates is the propensity for the population to live in apartments (Hausman (2002), Stross (2011)). To the best of my knowledge, however, I am the first to formally propose use of this variation as an empirical strategy for isolating a causal effect of home Internet usage.

[22] Miscellaneous housing units (recreational vehicles, vans, boats, etc) as well as dormitories, jails, prisons, and hotels were excluded in the construction of this measure.

strong first stage relationship will confirm that MDU rates are indeed predictive of individual usage, but it is useful to first consider the relationship between MDU rates and Internet *access*, since that is the assumption underlying the proposed link between MDU rates and Internet *usage*. Figure 2.2 examines this link empirically by comparing state MDU rates to Internet access rates. Data on local high-speed Internet access rates can be obtained from the FCC form 477 filing data, which has information reported by Internet services providers on zip codes in which they have at least one high-speed Internet customer. This information was aggregated to the state-level and used to estimate the unconditional correlation between the share of a state living in an MDU in 2000 and the share of households in a state with high-speed Internet access, separately by year.[23] The data confirms the intuition behind the instrumental variables strategy: a 10 percentage point increase in the share of the state residing in an MDU is correlated with an increase in the state's high-speed Internet access rate of 3.2 percentage points in 2000. The magnitude of the correlation declines each year there after to a precisely estimated zero by 2006. This provides support for the hypothesis that state MDU rates predict the early arrival of Internet access and in the next section, I will show that the first stage diagnostics are indicative of a strong relationship between state MDU rates and trends in individual usage of the Internet over time.

The key identifying assumption for this instrumental variable strategy is that baseline state MDU rates would not have been systematically correlated with subsequent trends in labor supply in the absence of residential broadband Internet diffusion. Certainly, women with preferences for being stay-at-home mothers might choose to live in states with more single family homes. It is therefore imperative that the empirical specification include state fixed effects, so that the estimated relationship between MDU rates, home Internet use, and labor supply will be net of time-invariant differences across states in preferences for housing and work. Importantly, I also condition on individual-level metropolitan area status and central city status, so that the estimated effect will be net of differences in preferences

[23]State-level aggregation was used because finer levels of geographic detail is not available for all individuals in the CPS.

between urban, suburban and rural residents.

Any threat to the identification strategy must come from unobserved variation over time in sorting patterns across states which are related to MDU rates.[24] To capture possible sources of variation over time in the housing and labor market which might be correlated with MDU rates and labor supply, I include a host of time-variant state-level controls. These include population density, housing prices, unemployment rates, per capita personal income, average wages, and employment in Internet intensive industries. Net of population density, prices, and local labor market conditions, left-over variation in housing stocks across states can be expected to be a function of factors such as historical zoning ordinances, weather, and elevation, and the intuition behind the identification strategy is rooted in how these factors led to differential rates of Internet access *over time*.[25]

Other empirical work on the impact of Internet usage has also had to address the endogeneity of Internet take up to labor market outcomes. Kuhn and Skuterad (2004) and Kuhn and Mansour (2011) study the impact of Internet on unemployment durations. They control for Internet use for all activities to control for self-selection into Internet job search. Since I am interested in estimating the net impact of Internet use, this strategy is not feasible. Stevenson (2009) also studies the impact of Internet on job search. The author uses state take-up rates of telephones and automatic washing machines in 1960 to instrument for current state Internet usage patterns. The fundamental motivation for this strategy is correlation over time in state-level latent demand for technology (Skinner and Staiger, 2007). Unfortunately, if there is a contemporaneous technology diffusing alongside the Internet or

[24]For the instrument to valid, I must also assume that residential Internet diffusion does not have spillover effects on other aspects of the local market. For example, if the Internet led to large increases in telework and reductions in traffic, non-users labor supply might be altered by the decrease in traffic. This seems unlikely given the fact that full-time telework take-up is still quite low (2.3 percent) and has not increases substantially over the period in question. Similarly, telework take up has not been shown to be correlated with local commute times (Lister and Harnish, 2011).

[25]For example, consider two observably similar cities (similar population, income, average age, and cost of living): Washington, DC and Boston, MA. In Washington, the 1910 building height act does not permit residential buildings above 90 feet (roughly a height so that the capital building can be viewed throughout the city). Because of this law, there are fewer tall buildings in Washington DC. All else equal, this would be predicted to have a relative negative impact on subsequent trends in high-speed Internet diffusion.

if this latent demand is systematically correlated with trends in labor supply, the instrument will violate the exclusion restriction.[26]

2.4.4 Instrumental Variables Specification

The impact of home high-speed Internet use on labor supply is estimated using two stage least squares. The first stage is a linear probability model of the impact of the percent of the state residing in an MDU on HSI use:

$$HSI_{ist} = MDU_s\theta_t\gamma_1 + X_i\gamma_2 + S_{st}\gamma_3 + \theta_t + \eta_s + \nu_{ist} \tag{2.2}$$

HSI_{ist} is a dummy variable for whether or not individual i reports using the Internet in a household with broadband in state s and year t. MDU_s is the percent of the state's 2000 population which resides in a housing unit that is classified as an MDU. MDU_s is expected to affect home Internet use via the timing of Internet access, since places with a higher MDU_s are expected to receive Internet access earlier. Therefore, it is interacted with the vector of year fixed effects θ_t to allow the impact to vary separately by year. The instrument is the vector $MDU_s\theta_t$.[27] The main effect of MDU_s only varies at the state level and is perfectly correlated with the state fixed effects, thus, it cannot be included in the model. The second

[26]Several recent working papers using non-U.S. data have also proposed identification strategies that exploit supply-side constraints to access, although their strategies would not be applicable in the U.S.. Bhuller et al. (2011) study the impact of the Internet on sex crimes. They use cross-sectional variation in a publicly funded broadband roll-out program in Norway. Falck, Gold, and Heblich (2012) study the impact of the Internet in German elections. They exploit a technological limitation of DSL provision which creates a kink in accessibility at a precise distance from the central office of the telephone company. While this technological feature of DSL provision is also apparent in the U.S., DSL has a considerably lower market share in the U.S. (around 30 percent, as opposed to over 95 percent in Germany). Since this kink is not present in cable-based broadband technology, it would be expected to have little predictive power for overall access rates in the U.S..

[27]In the appendix robustness checks (table 2.8), I also display the results of interacting MDU_s with a time trend. This is less flexible than the current specification and is therefore not used for the main empirical strategy. Results, however, are nearly identical to those presented below and the first stage is strong. I also show that a binary instrument which interacts an indicator for whether the state's MDU rate is above the median with an indicator for years prior 2003 is also highly predictive of Internet usage and produces nearly identical results.

stage is a linear probability model of HSI use on labor force participation:

$$y_{ist} = \widehat{HSI}_{ist}\beta_1 + X_i\beta_2 + S_{st}\beta_3 + \theta_t + \eta_s + \epsilon_{ist} \tag{2.3}$$

Where y_{ist} is initially labor labor force participation of individual i in state s and year t. The coefficient of interest is β_1, which measures the impact of HSI_{ist} on labor supply. X_i is a vector of individual controls and S_{st} is a vector of state-level controls, both of which are the same as described in section 2.4.1. θ_t are year fixed effects , η_s are state fixed effects and ϵ_{ist} is the error term. Standard errors are clustered at the state-year level.

2.5 Results

2.5.1 First Stage

Table 2.4 displays the first stage relationship estimated using equation 2.2, which indicates the sign and magnitude of the relationship between the MDU rates and home Internet usage follow the expected pattern over time. The year 2009 is used as the base year, which means all coefficients can be interpreted as the effect of state MDU rates on home Internet use relative to 2009. When coefficients are scaled by the mean of the dependent variable in each year (shown in italics below the coefficients and standard errors), there is a clear pattern of a positive and statistically significant, but declining, impact of MDU rates on home Internet use over time. In the most conservative, preferred specification with all control variables (column (5)), this implies that a one percent increase in the proportion of the state residing in a multiple dwelling unit increases home high-speed Internet use by 13.7 percent in 2000 relative to 2009. This magnitude falls each year thereafter and in 2007, a one percent increase in state MDU rates leads to only a 0.4 percent increase in usage relative to 2009.

In addition to displaying the predicted relationship between the instrument and Internet usage, tests for weak identification also indicate that the first stage relationship is strong.

To test for weak identification in the first stage, I cannot use the conventional F-Statistic because I have employed cluster-robust standard errors and the conventional F-Statistic will tend to over-reject the null of under-identification in the presence of non-i.i.d standard errors. Therefore, I alternatively employ a "robust" version of the F-Statistic which employs the Kleibergen and Paap (2006) rk statistic, as suggested by Baum, Schaffer, and Stillman (2007).[28] Henceforth, I will refer to this as simply the F-Statistic. Since no critical values have been developed for that statistic, Baum et al. (2007) suggest using either the Staiger and Stock (1997) "rule of thumb" of 10 or the Stock and Yogo (2005) critical values.[29] I will consider both. In the preferred specification, the F-Statistic is 27.55, which is above each of the aforementioned benchmarks and indicates the instrument is indeed powerful and relevant.

2.5.2 Labor Force Participation

Table 2.5 panel (b) shows the baseline instrumental variables results of estimating equations 2.2 and 2.3 for the impact of home Internet use on the labor force participation of married women. Column (1) displays the parsimonious specification, which includes only the state and year fixed effects. Columns (2)-(4) add the demographic control variables, with column (3) adds spouse's education and column (4) adds the controls for presence/age of children. As in the LPM specification, the model is initially estimated without presence/age of children since those variables may be endogenous to labor supply decisions. Their inclusion, however, does not fundamentally alter the result. Column (5) is the preferred specification

[28] The conventional Wald F-Statistic uses theCragg and Donald (1993) statistic for the rank of a matrix. This F-statistic instead uses the Kleibergen and Paap (2006) rk statistic, the robust analog of the theCragg and Donald (1993) statistic.

[29] Stock and Yogo (2005) provide two methods for evaluating the presence of weak instruments. The first considers the relative bias of IV as compared to OLS, where an instrument is considered weak at x percent if IV has a relative bias of more than x percent. Those critical values are as follows: 5% 16.85 10% 10.27 20% 6.71 30% 5.34. The second considers the Wald test, which rejects too often with weak instruments. An instrument is considered weak at x percent if the Wald Test has a rejection rate of x percent when it should have a rejection rate of 5%. Those critical values for my case are 10% 24.58 15% 13.96 20% 10.26 25% 8.31.

with a full set of state and demographic controls. The point estimates for the coefficient on *HSI* is 0.186 in the preferred specification, which is statistically significant at the one percent level. This indicates that being a home high-speed Internet user leads to a 18.6 percentage point increase in the probability of participating in the labor force for married women. Evaluated at the mean of the dependent variable, this corresponds to a 25 percent increase in the probability of participating.

Instrumental variable estimates are considerably larger than the LPM estimates. This suggests downward selection bias in the LPM estimate, which would occur if married women who do not work are more likely to take up the Internet than those who do work, perhaps because women who spend more time in the home place a higher value on the time-saving and entertainment uses of home Internet. For example, social uses of the Internet might be more valuable for a woman who spends the majority of her time in the home than for a women interacts with coworkers and has access to Internet in the workplace. The IV estimates are less precise than the baseline LPM estimates, but a Hausman specification test can reject the consistency of the LPM estimate at the ten percent level in the preferred specification.[30] The IV estimate of 18.6 percentage points is an admittedly large effect, particularly relative to the mean labor force participation rate of 73 percent. In section 2.6, I discuss the interpretation of the magnitude of these effects in much greater detail and show they can be explained by heterogeneity across the population and a decomposition of what the linear IV estimand identifies from the data.

2.5.3 Different Demographic Groups

Internet availability may have a relatively larger impact on women with more education and children, since women with more education are more likely to hold highly skilled oc-

[30] Since standard errors are clustered in both the OLS and 2SLS specification, OLS is not fully efficient and I compute the Hausman test statistic with bootstrapped variance estimates. I use 500 bootstrap replications and the p-value on the resulting test statistic is p=0.092.

cupations in which working at home online is feasible, and women with children may have more home production tasks to complete or face higher costs to working outside of the home Therefore, I split the sample by presence of children and education to see how results vary across groups. In addition to splitting the sample by whether or not a woman has any children, I look separately at women who have any children under 6 and those whose children are all school-aged. Although all women with children are expected to face psychic and non-psychic costs to working (notably, child care expenses), women with older children may find working from home an effective replacement to paid after-school child care, whereas small children may require too much supervision to make working in the home without paid child care a realistic option.

Panels (a) and (b) of table 2.6 display the results by the presence of children and education. Panel (a) column (3) displays the results for women with children, indicating home Internet usage is associated with a 29.3 percentage point increase in the probability of participating, which is statistically significant at the one percent level. Columns (4) and (5) split this group into women with any children under age 6 and women whose children are all age 6-18. Results for women with younger children are smaller than the overall average and insignificant, while the result for women with older children are significant at the 10 percent level and indicate Internet usage is associated with a 30 percentage point increase in participation for this group. Panel (b), columns (1)-(3) display the results for increasing amounts of schooling. Unfortunately, as the sample size falls first stage F-Statistics fall below the rule of thumb for the high school drop-out and college groups, so it is difficult to interpret those results as causal. Therefore, I split the sample by high school degree or less and some college or more in columns (4) and (5). In this case, the estimate on less educated women is small and insignificant while the estimate on more educated women indicates home Internet usage is associated with a statistically significant increase in participation of 38 percentage points for that group.

The results by demographic group have indicated that home Internet usage is associated with a 30 to 40 percentage point increase in participation for women with children and high levels of educations, and smaller, insignificant effects for other groups. The pattern of the magnitudes are generally opposite the observed pattern in the LPM estimates. This indicates that for women with children and more education, the downward bias in the LPM estimates is larger than the downward bias for other groups. This would be consistent with more educated women and those with children who do not work being more likely to take up the Internet than either their working counterparts or other types of women who do not work. This could occur if these women place more value on Internet services, for example, because non-working high skill women have more knowledge of the benefits of Internet technology than non-working low skill women.

2.5.4 Alternative Outcome Variables

Home Internet use may have a different impact on the intensive margin than it does on the extensive margin. On the one hand, telework and time saved in home production may allow individuals to increase their hours, while on the other hand, new entrants into the workforce may work fewer hours. Therefore, table 2.6, panel (c) column (1) displays the results of re-estimating equation 2.3, where the dependent variable is hours of work. The sample is limited to working individuals. The results indicate there is a positive and significant increase in usual hours worked by about 8.5 hours per week, indicating there is an intensive (as well as extensive) increase in labor supply associated with home Internet usage.

Next, I look at the choice between full-time and part-time work. Like hours worked, the predicted effect of Internet usage on part-time versus full-time work is ambiguous. On the one hand, Internet usage may facilitate engagement in the labor force, but only at the part-time level. On the other hand, it may be that the flexibility associated with Internet usage permits full-time participation when an individual would have otherwise only worked part-time. Columns (2) of table 2.6, panel (c) displays the results for the impact of home Internet use on an indicator for working full-time. This indicates that there is a 19.8 percentage point increase in the probability of working full-time (as opposed to part-time). This suggests that in fact, Internet usage encourages full-time engagement in the labor force.

Next, I examine changes in family income.[31] Table 2.6, panel (c), column (3) displays the estimated results.[32] The point estimate is positive and significant, and indicates a substantial increase in family income associated with a married women's high-speed Internet usage.[33] This point estimate is very large: almost $30,000. However, given the large predicted increases in participation it is fairly reasonable that family income would increase this substantially. Considering the average cost of a broadband subscription was about $30 a month in real dollars throughout the time period, this suggests high-speed Internet is potentially very valuable to families.

Finally, I examine the impact of Internet usage on employment status. In addition to the expansion of Internet services, the 2000-2009 time period was characterized by a recession and large changes in unemployment rates. It is therefore useful to see if the results hold alternatively estimating the effects of Internet usage on employment. If the results differ, it would suggest that Internet usage may encourage individuals to switch from non-participation to a state of unemployment, as opposed to working. This would change the interpretation of the results. Results of this specification are displayed in table 2.6, panel (c) column (4) and indicate that home Internet usage is associated with a statistically significant 17.5 percentage point increase in the probability of working. This result is extremely similar

[31] I have also alternatively estimated the effects on hourly wages. Unfortunately, however, weekly earnings information is only asked of outgoing CPS respondents, which means that I can only calculate hourly wages for one eighth of the original sample. The results indicate Internet usage is associated with a statistically insignificant $4.49 increase in hourly wages, however, the limited sample reduces the power of the first stage making causal inference difficult.

[32] Total family income is discretized, so to make this variable continuous I replace each value with the median value of the bin.

[33] I have also estimated the results for the dependent variable hourly wage (not shown here). Wage data is only recorded for outgoing rotation groups, which means the sample for whom wages are recorded is about one eighth the original sample. This reduces precision and the estimated effect is positive, but statistically insignificant.

to the participation result and suggests the participation results are in fact driven by women entering the workforce and becoming employed.

2.5.5 Men

Home Internet use should have an impact of individuals who are on the margin of entering the labor market *and* can benefit from working from home, job search, and/or time saving in home production. Since male labor force participation has historically been relatively inelastic and the home production and telework benefits of home Internet use are likely to be less relevant for men, there is less *a priori* reason to think that home Internet use should impact male labor supply. Column (5) of table 2.6 panel (c) displays results for married men and indicate that the impact of home Internet use on male participation is small and statistically insignificant. Although not displayed here, the specification was also estimated for single men and women and although first stage F-Statistics were too low for reliable inference for singles, the pattern on the point estimates was not similar to married women. Those results can be found in appendix table 2.8.

2.5.6 Falsification Test

The identification assumption for interpreting the instrumental variables estimate as causal is that state MDU rates would not have been correlated with subsequent trends in labor supply in the absence of broadband Internet diffusion. While there is no formal way to directly test this assumption, I can informally investigate whether this assumption seems reasonable by conducting a falsification test of the reduced form impact of the instrument on labor force participation *prior* to the availability of high-speed Internet. The intuition behind this exercise is that if a state's MDU rate affects labor force participation during a time period where high-speed Internet was not available, then it must affect labor supply through some other mechanism. If that were the case, there would be serious doubt that the exclusion restriction would be satisfied during the time period in which high-speed Internet was available.

The time period I use for this falsification test is 1990-1997, since the first residential broadband subscriptions became available in 1998 (Faulhaber, 2002). The convenient feature of studying this particular time frame for the falsification test is that both business broadband subscriptions and residential dial-up subscriptions were plentiful. Therefore, this analysis should help to shed light on whether or not a violation of the exclusion restriction occurs through one those channels. For example, one could imagine a scenario where participation increased in areas with more MDUs not because of residential high-speed Internet access, but because firms in those areas had differential trends in demand for Internet savvy workers. Since business broadband access and Internet knowledge were already established in this period, but home high-speed Internet was not, this strategy will be able to address this concern.

Figure 2.3 panel (a) displays the results of that exercise for the 1990-1997 period, and as a comparison, for the 2000-2009 sample used throughout the chapter in panel (b). Note that panel (a) uses 1990 MDU rates and panel (b) uses 2000 MDU rates. The data indicates that for the 1990-1997 period, there is no significant impact of an increase in state MDU rates on participation for any of the years shown (relative to the year 1997) and the point estimates display no clear trend or pattern, bouncing above and below, but remaining close, to zero. On the other hand, for the sample period used in this study, coefficients are consistently above zero and statistically different from zero for half of the years shown. Overall, this exercise has uncovered no evidence that state MDU rates have an effect on participation in the time period before high-speed Internet access was available, suggesting that state MDU rates do not affect participation except through their effect on Internet availability.

2.5.7 Robustness Checks

I implement a number of robustness checks on the empirical specification and data construction. This includes various tests on the exact definition of an MDU for the instrument, controlling for MSA level measures of wages, state time trends, and individual computer use. Those results are described in the appendix to this chapter (section 2.9). Although in some cases the sample sizes and/or the first stage F-Statistics become small and the results become imprecise, the point estimates are consistent across all specifications and well within a 95 percent confidence interval of the estimates obtained thus far.

2.6 Interpreting the Results as LATE

The results presented thus far have been estimated using a linear instrumental variables model, but it is well known that this strategy can only identify an average causal effect for those sub-populations whose treatment status is affected by the instrument. This parameter is called the local average treatment effect (LATE) and the population for whom it applies are often called the "compliers" (Imbens and Angrist, 1994; Angrist, Imbens, and Rubin, 1996).[34] In practice, complier status is unobservable and it is not possible extrapolate the estimated effects outside this unknown population without assuming homogeneous treatment effects. Given the observed heterogeneity in treatment effects by demographic subgroup, this assumption seems unrealistic and suggests an investigation of the external validity of the LATE is warranted.

In gauging the external validity of the LATE, I begin by considering the magnitude of the estimated LATE in the context of aggregate trends in female labor supply over the time period studied. If the LATE estimate of 0.186 represents an average effect that applies

[34] To interpret the IV estimate as a LATE, the monotonicity assumption must be satisfied. In this context, that assumption is that while some individuals may choose to use or not use the Internet regardless of their state of residence, residing in a state with more MDUs (and hence greater access) will only increase the probability that an individual uses the Internet at home.

to the entire population, then one would expect to see an increase in aggregate female labor supply. Figure 2.4 displays trends in aggregate labor supply over the 2000-2010 time period obtained from the American Communities Survey. Separate trends are shown for all married women, women with both a college education and children (for whom I estimate the largest effects), and as a comparison, men. There is a clear increase in female labor force participation beginning in about 2005, and this increase is larger for highly educated women with children. For men, on the other hand, labor force participation remains stable. While both the sign of these changes and the observed variation across groups is consistent with the results I have presented thus far, the magnitude of the aggregate changes in labor supply are much smaller than the estimated effects found in the IV specification. This is indicative that the magnitude of the LATE may not apply to the majority of the population. In what follows, I reconcile the difference in magnitude between the estimated effects found using linear IV and the aggregate trends in female labor supply.

2.6.1 Characterizing the Compliant Population

Without assuming homogenous treatment effects, the LATE found using linear IV is only interpretable as applying to the compliant population, which are those individuals whose treatment status is affected by the instrument. In the context of this chapter, that group is the individuals who use/do not use high-speed Internet because of the differential trends in access induced by their state's housing infrastructure. To garner some intuition on the type of individual that is expected to be a complier, it is useful to separate the first stage effect of the instrument on treatment status into two conceptual links: (1) the effect of the state MDU rates on home high-speed Internet access and (2) the effect of state access on take-up. Focusing on the first link, the compliant population will be states for which state MDU rates affected the timing of access, or states that would have received access later (or never) in the absence of this characteristic. I expect this to be states with relatively low latent demand for

Internet services, since ISPs would have otherwise had little incentive to enter those markets. For the second link, the compliers are individuals for whom access determines take-up. This rules out individuals with very low private demand for Internet services.[35] Since demand for Internet services is highly correlated with socioeconomic status, combining the two suggests that the compliant population should be relatively high SES individuals in relatively low SES states.

With some intuition on the characteristics of the compliers in hand, I now construct estimates of the proportion of the population which are compliers and tabulate their average characteristics. With a single binary instrument, this task is fairly straightforward. Unfortunately, I use multiple, continuous instruments, which complicates such an exercise. Fortunately, the IV strategy I use lends itself to a binary interpretation (since identification arises from differences across states and over time) and it turns out that that using an instrument based solely on comparing states with high/low MDU rates in the early/late periods of diffusion I am able to generate similar results to those using the more flexible specification used throughout the chapter.[36] Using those results, I can then easily calculate the fraction of the population which is compliant and compare their characteristics to the rest of the population.[37]

Summary statistics on the complier group's characteristics are displayed in table 2.7, panel (a). I find that 6 percent of the population is compliant (4 percent are home Internet users and 2 percent are non-users). This indicates that the LATE estimate applies to a small minority of the general population. In relation to the aggregate change in female labor force

[35]The non-compliant population includes the "always takers" and "never takers." Individuals who would always use high-speed Internet regardless of broadband access are those willing to pay for expensive satellite and business line subscriptions. In practice, 0.52 percent of Internet users in my data report having such an Internet connection 2000-2007, suggesting this group is quite small. The group that would never use high-speed internet, on the other hand, are individuals who are not interested or cannot afford an Internet subscription.

[36]The exact instrument used is $Z = \mathbf{1}[MDU \geq \overline{MDU}] * \mathbf{1}[Year \leq 2003]$. In this case, the estimated coefficient on HSI use is 0.201 with a standard error of 0.101

[37]The fraction is the population which is compliant is $P(Z = 1)((E(HSI|Z = 1) - E(HSI|Z = 0))/P(HSI = 1)$. The average of characteristic, x for the compliers (relative to the non-compliers) is $(E(HSI|Z = 1, x = 1) - E(HSI|Z = 0, x = 1))/(E(HSI|Z = 1) - E(HSI|Z = 0))$.

participation over the period (if we make the strong assumption of a zero treatment effect for the rest of the population), this would imply a 1.5 percentage point increase in aggregate labor supply, or approximately half of what was observed.

Table 2.7 also tabulates the characteristics of the compliant population, indicating that the compliers tend to have lower levels of education, are more likely to be a minority and are younger. This is suggests that the primary mechanism which restricts the compliant population size is the first channel described above: the compliers are those in states that would have otherwise not received access. To more formally investigate this hypothesis, I construct a composite measure of state socioeconomic characteristics that predict Internet usage.[38] I then calculate the weights linear IV puts on each quartile of that distribution in construction the IV estimate. The intuition behind this strategy is to investigate whether or not the compliers primarily come from states that ISPs would have otherwise been reluctant to provide access because latent demand was relatively low. The results of this exercise are displayed in the panel (b) of table 2.7 and indicate that indeed, the compliers are predominately from states whose socioeconomic characteristics would predict lower levels of Internet usage. In fact, individuals in states in the top quartile of the distribution of the state SES measure receive no weight in the IV estimate.

The results presented in this section are indicative that the compliant population is comprised of high SES individuals in states which would not have received Internet access in the absence of a favorable housing infrastructure which lowered the costs of access for the ISP. While this exercise has found that the LATE estimate applied to a minority of the population, it does indicate that the LATE may be of particular relevance for public policy. Currently, individuals without access are those individuals who reside in places that ISPs refuse to provide service. Thus, this estimate speaks to the effects of the provision of access to higher SES individuals in underserved locations.

[38]This measure as constructed using a linear projection of state-level demographic characteristics (fraction of the population in each race, education and age group used in the analysis) and economic characteristics (state-level variables used in the main specification) on individual high-speed Internet usage.

2.6.2 Marginal Treatment Effects and Linear IV Weights

Thus far, I have presented evidence that the LATE is applicable to a small population, but I have not explained why the magnitude is so large for that particular population or the magnitude of the effects for anyone outside that group. In this section, I will address these questions by decomposing the LATE estimate and estimating marginal treatment effects (MTEs) across the population.

In the presence of heterogeneous treatment effects the LATE estimate is a weighted average of subgroup-specific LATEs, where the weights placed on each subgroup will not in general be proportional to the population weight of the group. Instead, the weights are proportional to the average conditional variance of the first-stage fitted values for the group. Moreover, those weights are not always guaranteed to be non-negative. In practice, this means that not only is it possible for the LATE to over (under) weight some groups, but it is also possible that linear IV can produce a negative (positive) point estimate even when all subgroup specific LATEs are positive (negative) (Heckman et al., 2006).

The possibility of negative weights in the IV estimand occurs in the presence of unobserved heterogeneity (Heckman et al., 2006). The standard IV framework explicitly allows for sorting on *levels*: individuals may be unobservably more or less likely to take up treatment and those characteristics may be correlated with the outcome. It does not allow for sorting on *gains*. Sorting on gains occurs when individuals have some knowledge of their own expected gains from take-up and sort into treatment based on those expected gains. In the context studied here, this type of sorting would occur if individuals choose to take-up home Internet at least partly based on their own expected gains from home Internet usage in the labor market. In this context (and many others) this almost certainly seems likely to be true. For example, a family debating purchasing a costly Internet subscription will almost certainly be more likely to take-up the service if they expect it to bring in extra income for the family. Thus, it is clear that further investigation into the possibility of this type of sorting is warranted.

Heckman, Urzua, and Vytlacil (2006) demonstrate that it is possible to decompose the standard IV estimate into identifiable MTEs at different levels of unobservable factors that determine treatment status and calculate the weights placed on those MTEs in the standard IV estimand. This exercise allows me to investigate the labor supply response of individuals at different levels of unobservable factors that are correlated with home Internet take-up, and determine whether individuals sort on those unobservables.[39] To estimate those MTEs, I use a non-parametric local instrumental variables estimator (Heckman and Vytlacil, 2001, 2005).[40] Then, I calculate the weights on linear IV at different points in the distribution of unobservables from the data using the formulas provided by Heckman, Urzua, and Vytlacil (2006), where I evaluate the weights at the mean of the independent variables.

The resulting MTEs and standard errors (calculated using 500 bootstrap replications) are displayed in figure 2.5 (a). Note that the horizontal axis increases in the unobserved components that determine treatment status, so MTEs that increase in *P(z)* are consistent with positive sorting on gains. What emerges from this figure is a clear pattern of increasing treatment effects, which start out below zero and increase to very high levels at the extreme right tail of the distribution. This pattern is consistent with individuals sorting into Internet usage based on their expected labor market gains from usage. Moreover, figure 2.5 (b) displays population weights across the distribution of unobservables correlated with Internet usage. This makes it clear that the majority of the population actually resides in the left tail, where point estimates are smaller and sometimes negative.

Next, I examine the observable characteristics which are associated with high/low MTEs (and correspondingly, high/low expected gains from treatment). Table 2.7, panel (c) summa-

[39] Heckman, Urzua, and Vytlacil (2006) demonstrate that evaluating the treatment effect at different points in the distribution of the first stage fitted values(e.g., $P(z) = Pr(HSI = 1|Z, X)$) is equivalent to evaluating the treatment effect at different points in the distribution of unobserved determinants of treatment status. Thus, by calculating the treatment effect at each point in the distribution of P(z), I can investigate the possibility of sorting on gains.

[40] For the local instrumental variable estimator, I use a flexible polynomial in the propensity score to estimate the MTE at each point in the distribution of P(z). I estimate the propensity score using probit so that fitted values are ensure to lie on the interval [0,1]

rizes those characteristics for individuals in different parts of the distribution of unobservables correlated with Internet usage. Comparing the top two deciles of the distribution to the rest of the distribution indicates that the individuals with the very large treatment effects are significantly more likely to have a college education: 60 percent of individuals in the top two deciles have a college degree, as opposed to 30 percent for the middle 6 deciles and 22 percent for the bottom two deciles. Those individuals are also much more likely to have a college educated husband, have children (and particularly school-aged children) and are more likely to be white. Thus, it is clear that the very large point estimates found earlier for more educated women and women with children are a reflection of the fact that those groups tend to have higher treatment effects of home Internet use.

Finally, figure 2.5 (c) displays the weights that the instrumental variables estimator places on different parts of the distribution of unobservables correlated with Internet usage. To calculate the weights, I use the single instrument interaction of MDU rates with an annual time trend, which yields an IV point estimate and MTEs which are extremely similar to those used in the original specification.[41] The shape of the distribution of IV weights stands in sharp contrast to the shape of the distribution of population weights. Instead of placing relatively more weight on individuals towards the left tail, which is where the bulk of the population lies, the IV estimator places relatively more weight on individuals towards the right tail where the larger treatment effects lie. Moreover, the IV estimator actually places negative weights on some of the negative MTEs in the left tail, artificially inflating the IV estimate.

Based on the preceding analyses, it is clear that the large magnitude of the IV estimate is driven by a minority of individuals with very large treatment effects. From a policy standpoint, it is useful to uncover the magnitude of the effects that apply to the majority of the population. Unfortunately, the MTEs are fairly imprecise (particularly at the tails),

[41] The exact instrument used is $Z = MDU * t$. In this case, the estimated coefficient on HSI use is 0.184 with a standard error of 0.081. The distribution of MTEs using this instrument, as well as the IV estimate, can be found in the appendix to this chapter (section 2.9).

so I exercise caution in interpreting these effects as causal. However, using the weights and MTEs I have estimated here I am able to recover a predicted linear IV estimate of 0.196, a number which is extremely close to the actual linear IV estimate of 0.186 and provides some assurance that this exercise will be instructive.[42] Simply weighting the MTEs by population, I calculate an average treatment effect for the entire distribution of 0.041.[43] Yet, for the top quartile of the distribution of unobservables, the population weighted average is 0.356. This makes it is clear that the treatment effects that apply to the majority of the population are smaller than those estimated by linear IV, which applies to a smaller subset of the population and masks a significant amount of heterogeneity across the population. It also suggests that home high speed Internet usage can explain 90 percent of the observed increase in women's labor force participation over the period in question.[44]

2.7 Interpreting the Results in the Conceptual Framework

The analysis presented thus far indicates that home Internet use is associated with an increase in labor market participation, with larger effects found for women with children and more education. Since telework opportunities are expected to be particularly important for more educated women and those with older children, and time saved in home production is expected to be particularly important for women with children, the demonstrated heterogeneity in the estimated effects is suggestive evidence in favor of home production and telework as potential mechanisms explaining the results. Theory does not inform whether

[42]Note that the MTEs and weights are only calculated every 5 percentage points in the distribution of *P(z)* and are evaluated at the mean of the independent variables, so one would not expect to necessarily recover the exact point estimate. In order to recover the exact LATE, I would need to integrate out the independent variables, an exercise I do not conduct here.

[43]This average treatment effect differs from the parameter called the "average treatment effect" (e.g., the ATE). The average I describe here is really a local average treatment effect across the portion of the distribution of unobservables for which I have support of the propensity score. In practice, this is almost the entire distribution because the support of the propensity score is 0.1-0.97, but to calculate the ATE parameter would require full support. I could bound the ATE with the information I have, but I do not carry out that exercise here.

[44]Scaling by the increase in Internet usage over the period suggests Internet usage can explain 2.9 percentage points of the approximately 3.2 percentage point increase in participation.

or not job search is also a viable mechanism, nor does it identify if the estimated effects are attenuated by increases in time spent in leisure online. In this section, I explore the role of each mechanism in explaining the estimated effects. To do so, I examine data on the activities individuals in the sample report doing online. In 2000, 2001, and 2003, the Current Population Survey supplements record detailed information about activities individuals report engaging in online in the past year, which I use to tabulate summary statistics on individual's online behavior. The surveys vary in the specific questions asked, but I will focus on the following which were consistent across years: home production (with a separate category for shopping and paying bills), leisure (with separate categories for entertainment and checking news, weather, and sports scores), work-related and job search.[45]

Next, I draw a direct comparison between use of Internet for each activity and the magnitude of the estimated effects by comparing group level usage rates with the magnitude of the estimated effects of home Internet use on labor supply. To do so, I estimate group-level predicted changes in labor force participation when home high-speed Internet use increases from 0 to 1, $\widehat{LFP}_g$, which I compare to group-level mean rates of Internet use for each of the various tasks $\frac{1}{N_{sg}}\sum_{sg} HSI(task = t)$. $\widehat{LFP}_g$ is constructed by estimating group specific β_1's according to equation 2.3. Rates of use for each task are limited to users only, and groups are defined by census division, education and the presence of children. The goal of this analysis is to inform the extent to which Internet use for each activity contributes to the estimated effect. If there is no correlation between rates of use for a task and the predicted effects, this would suggest Internet use for that purpose is not an important driver of the results. On the other hand, a strong positive correlation suggests Internet use for that purpose may indeed play a role in explaining the results.

[45] The supplements ask about specific activities performed online in the past year. The data is only summarized for those who report using the Internet at home. Home production includes shopping and paying bills online, looking for government or health information, and looking for information about products and services. Leisure includes recreation (playing games, watching TV or movies, listening to music, recreation, entertainment or fun) and checking news, weather, or sports (NWS). Work related tasks, email and job search were consistently asked throughout the sample. See section 2.9 for more information on variable creation.

Figure 2.6 displays the results for work, job search, home production (including shopping/paying bills) and leisure (separately for entertainment and checking news, weather, and sport scores), respectively. Consistent with the predictions of the conceptual framework, each activity except "games/fun/recreation" is positively correlated with the predicted effects.[46] Overall, Internet use for work appears to be the leading explanation for increase participation. It has the largest coefficient on the line of best fit (0.19), and a comparison of individuals across the distribution of $\widehat{LFP}_g$ indicates that those in the top quartile have 13 percentage point higher rates of Internet use for work, or a 53 percent difference at the mean. This suggest Internet usage for telework is the primary driver of the results, although Internet use for job search and home production also play a role. In the third chapter of this dissertation, I further investigate each mechanism separately and provide further evidence confirming this result.

2.8 Conclusion

High-speed Internet has changed the way individuals live and work. Using an instrumental variables strategy that exploits supply-side constraints to high-speed Internet access, I find evidence that exogenously determined home Internet usage increases the labor market participation of married women. Estimates indicate that being a home high-speed Internet user is associated with a 18 percentage point increase in labor force participation. This estimate is large relative to aggregate changes in labor supply over the period, but a decomposition exercise indicates this estimate applies to a minority of the population who exhibit very large treatment effects, and the average effect among the overall population is 4 percentage points. Consistent with the proposed conceptual framework, those individuals with the largest estimated response to Internet usage are highly educated women with

[46] The negative correlation between "games/fun/recreation" and the predicted effects is consistent with Internet use for leisure mitigating the positive effects of telework, job search, and time saved in home production. This negative relationship is even more striking when one considers that Internet use for one task is highly predictive of Internet use for other tasks (correlations between tasks range from 0.15-0.3) and is reassuring for the validity of the strategy as a whole.

school-aged children. Supplemental analyses suggest that Internet use for telework is the primary explanation for the demonstrated increase in participation, although improved job search and time saved in home production also play a more modest role in explaining the estimated effects.

The provision of fast, affordable Internet access is a central policy objective. To the best of my knowledge, this is the first work to investigate the net impact of individual home Internet use on labor supply, recognizing that the Internet is truly a general purpose technology. This work therefore speaks directly to the potential labor market impact of extending high-speed Internet access, and for whom access is likely to be most important. More generally, this work can speak to the labor market impact of the diffusion of home technology. Unlike technology diffusion in the workplace, which may directly affect productivity, the link between home technologies and labor market outcomes is less clear. Similar to work on the diffusion of time-saving appliances in the twentieth century, I find that female labor supply is sensitive to technological progress in the home sector.

The conflicting demands of work and family force households to make tough decisions. I find that for highly skilled women with school aged children, Internet use facilitates entry into the labor market. This suggests that Internet use may allow women who "opted out" when their children were young to (re)enter the labor force. Moreover, the data indicates that employer-provided part time telework opportunities associated with home Internet use are the best explanation for this increase. This speaks to broader policy discussions about the potential benefits of telework and flexible scheduling policies. While it is generally accepted that flexibility in the workplace has the potential to benefit employers, employees and the economy as a whole, adoption is still low and there is little empirical evidence on the benefits/costs of these policies.[47] This chapter has demonstrated that Internet usage, via take-up of telework opportunities, has allowed a group of highly educated women to join the workforce, suggesting such policies may have the potential to encourage workforce entry

[47]See, for example, the report by the Council of Economic Advisers on Work-life Balance, March 2010.

by productive individuals and increase gender equity in the workplace.

2.9 Appendix

2.9.1 Data

Internet use is available in the October 1997, 2003, 2007 and 2009 school enrollment supplements and the December 1998, August 2000, and September 2001 Computer and Internet use supplements. Internet users are identified individually, while the type of connection is at the household level. In 1997 and 1998, household high speed, broadband connectivity is not identified, so those supplements are not used. The process for identifying Internet users changes slightly over time due to the nature of the questions asked in each supplement. In the 2000 and 2001 supplements, home Internet users were defined based on series of questions which asked about household computer and Internet use and finally, an individual level variable called "Internet use at home recode", which assigns individual home Internet use status by compiling household home Internet use information with a series of questions of the format "Does NAME/do you use the Internet at home for...?" From 2007-2009, identifying home Internet users is slightly trickier because this recoded variable was no longer available and dramatically fewer questions were asked. Therefore, home Internet users are defined as those who are deemed to be (1) in a household that uses Internet at home and (2) Internet users themselves (at any location). Next, the type of connection is identified at the household level. In each survey year (except 1997 and 1998), households who have home Internet were asked about the type of home Internet access they have. High speed access is defined as NOT using "regular, or 'dial-up' service." Note that in 2000, the survey specifically asked if users had "Higher speed Internet access service", whereas in following surveys users were asked if they had some combination of "Cable, DSL, fiber optics, satellite, wireless (such as Wi-Fi), mobile phone or PDA, or some other broadband Internet connection." Defining broadband as "not dial-up" is also how the CPS defines high speed access, which can be determined from the universe of respondents to a question in 2001, 2003, and 2009 surveys which asked "What is the main reason that you do not have high-speed (that is, faster than dial-up) Internet access at home?"

2.9.2 History of Broadband Deployment

In this section I briefly review broadband network architecture, the history of high speed Internet deployment and how it relates to my empirical strategy. For a more detailed description of broadband technology deployment see, for example Jackson (2002). There are two fundamental ways that high speed Internet services reaches the customer: cable-based services and DSL services over telephone lines. For both of these services, fiber-optic lines provide high speed Internet service up to a certain point and then traditional coaxial cable or copper telephone lines carry service the rest of the way. These fiber-optic lines may reach the ISPs central office, some remote terminal in the neighborhood, the curb, or reach all the way to the home. These connections are often called "fiber to the x" (Jackson, 2001). Once the line reaches residence, wiring inside the home is the property of the home owner (Ames, 2006). In general, those lines do not need to be upgraded, although in very recent years it is becoming more common to replace this wiring with fiber as well. For the sample period studied here, this would have been relatively uncommon.

There are a few differences between cable and DSL that are worth briefly highlighting. Cable companies began developing networks which expanded fiber to neighborhoods (and then coaxial wiring the rest of the way) beginning in the 1980s. This existing system was theoretically capable of simultaneously providing both cable and high speed Internet to some customers, but each additional customer on a single fiber line reduces the "downstream" capacity, so that if many individuals use Internet at the same time the capacity would be exhausted. Thus, in order to expand service cable companies needed to move the fiber lines closer to the home. For DSL, providers were able to take advantage of existing capacity in copper telephone wires that were not used for voice services. However, this was limited by distance of the wire and wire quality and in practice, providers had to upgrade their local systems to provide consumers with DSL access (Jackson, 2002).

In practice, what this means is that both cable and DSL companies had to make significant infrastructure investments to keep up with the rapid growth in demand for high speed Internet services. The cable companies had to invest in bringing fiber lines closer to the home and the the phone companies had to deal with wire quality issues. Both had to send installers to the home for the first installation, for which there was often significant wait times for consumers (Faulhaber, 2002). All of this meant that supply lagged demand, particularly in the early years of roll-out. It is clear therefore clear that ISPs would have had an incentive to choose which markets to enter carefully so as to reach as many potential consumers as possible quickly.

Simple intuition tells us that Internet service providers should have weighed the potential costs and benefits of entering each local market and chosen accordingly and there is evidence that this type of cherry picking did in fact occur. In the early years of diffusion, Internet service was not surprisingly offered to areas with high predicted rates of take up, e.g., high income areas. This source of variation in timing, however, is likely to be correlated with labor market outcomes and would therefore prove unsuitable for an instrumental variables strategy. My strategy will instead focus on the costs of providing service to local markets. Recall that both cable and DSL service required investment in moving/upgrading wiring closer to the dwelling. From there, the dwelling owner was responsible for wiring inside the home and in most cases this would not need to upgraded since most homes had cable and phone wiring by the 2000s. Note that this was the case for both a single family homes and multiple dwelling units (such as an apartment buildings), since the building owner would have owned the wiring that distributes Internet to each unit. This means that from the ISPs perspective, providing service to a dwelling with multiple potential customers was more cost effective than providing service to a single family home with just one potential customer.

Thus, all else equal, the ISP will prefer to enter markets with more multiple family dwellings. This is the source of exogenous variation I use to isolate the causal effect of Internet on labor supply.

2.9.3 Local IV Estimator

This section describes how I estimate the MTEs and weights described in section 3.6.3 using a non-parametric local IV estimator. For a much more detailed description of this estimator and its applications, see Heckman and Vytlacil (2001, 2005); Heckman et al. (2006). The goal of this exercise is to decompose the standard IV estimate into MTEs over the distribution of unobservable factors that determine treatment status and calculate the weights placed on those treatment effects in the linear IV estimate. Then, using those MTEs, I can calculate the average MTE over specific intervals, which is the LATE over that interval. Note that this "LATE" differs from the "LATE" described by Imbens and Angrist (1994), which I will henceforth refer to as the linear IV estimate. The linear IV estimate is a particular weighted average of MTEs, while the LATE I refer to here is simply the average MTE over some specified interval. I will also describe how I calculate those particular weights and reconstruct the linear IV.

To calculate the MTEs, the first step is to estimate the probability of treatment or the propensity score, $P(z) = Pr(HSI = 1|X, Z)$. I estimate *P(z)* using probit so that the predicted probabilities are guaranteed to lie within the interval [0,1]. Next, I determine the support of the propensity score, which is critical for determining the set over which the MTE can be identified. In practice, I find that I have close to full support and can identify the MTE over the interval [0.1, 0.97]. Next, I calculate the MTE at each point in the distribution *P(z)* using the non-parametric local IV estimator, which is a flexible polynomial in *P(z)*. Note that because *P(z)* decreases in *z*, the horizontal axis *increases* in the unobservables correlated with treatment status (in the notation of Heckman et al.

(2006), the horizontal axis is $P(z) = 1 - u_d$, as opposed to $P(z) = u_d$). The weights for the IV estimate are calculated for specific points in the distribution of *P(z)* at the mean of the independent variables using the formula for the IV weights provided by Heckman et al. (2006). Note that it is difficult to calculate these weights with multiple instruments so I use the single instrument interaction of MDU rates with a time trend. The MTEs for that specification are displayed in figure 2.7 of the appendix to this chapter. The population weights are a simple calculation of the fraction of the population in each interval of the distribution of *P(z)*.

2.9.4 Additional Results

Table 2.8 displays the results of various extensions and robustness checks. The first row tests the sensitivity of the estimated results to alternative IV strategies. One robustness check of particular interest is an investigation into the sensitivity of the results to using a different instrument. Column (1) presents the results using the instrumental variable strategy used by Stevenson (2009) to investigate the effects of Internet job search. The instrument is the interaction of automatic washing machine and telephone usage in 1960 (both constructed from the 1960 Decennial Census) with year fixed effects. The motivation for this strategy is the fact that state technology take up is highly correlated over time and states which were early adopters of previous technologies tend to be early adopters of new technologies (Skinner and Staiger, 2007). This instrument is less powerful than the one I use, but it is above the rule of thumb of 10. The point estimate of 0.130 is easily within a 95 percent confidence interval of the point estimate using the MDU instrument. A Hausman test cannot reject the null that the two estimators are both consistent.[48]

Next I consider alterations on the instrument used throughout the chapter. Column

[48] I compute a Hausman test with bootstrapped variance estimate, since both instrumental variables estimates are calculated with clustered standard errors and therefore are not fully efficient. The p-value on the Hausman test is p=0.255.

(2) presents the results from interacting MDU rates with a time trend as opposed to year fixed effects. While this specification is less flexible than the main specification, the results indicate a stronger first stage and nearly identical results to those using year fixed effects. Column (3) displays the results using a binary instrument based on whether MDU rates are above or below the median interacted with an indicator for whether the time period is prior to 2003. Again, the results persist with this less flexible specification. Column (4) from using both MDU rates and 1960s technology adoption as instruments (both interacted with year fixed effects). Note that the first stage is actually less powerful in this specification than that with only MDU rates, which is indicative that the technology adoption instrument has less predictive power than the MDU rate instrument.

Columns (5) and (6) display the results from altering the definition of a multiple dwelling unit used for the instrumental variable. It is not possible to perfectly map the FCC's definition of a MDU to what is available in the Census. For the main specification, I chose the one that matched most closest and had the largest first stage F statistic: dwellings with 3 or more units. Here, I alternatively try (5) two or more units, and (6) five or more units. In each case the results are similar to using the original definition.

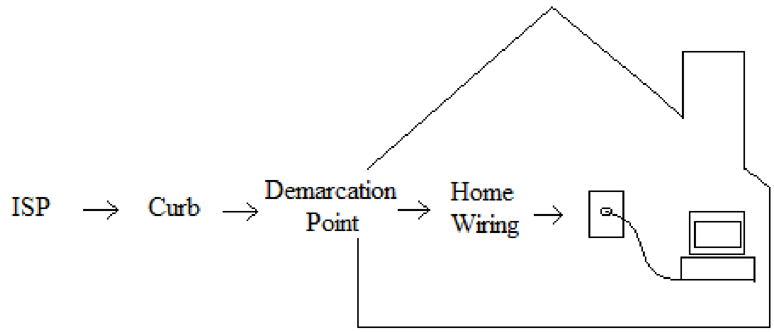
Next, I consider several checks on the data construction and specification in row 2 of table 2.8. In the first two columns, I omit from the sample those individuals whose Internet usage variable was allocated. In columns (3)-(4) I cluster standard errors at the state, as opposed to state-year level. This was not the main specification because there are 78 regressors and only 51 clusters, which means I cannot compute the variance-covariance matrix for the 2SLS specification without partialling out some control variables. In this case, I partial out the state fixed effects, demographic controls, and year fixed effects. Estimated standard errors increase slightly, but the point estimate remains statistically significant from zero. Finally, in columns (5)-(6) I include state-specific linear time trends, which allows for the possibility that women with plans to participate/not participate in the labor force might move into states with upward/downward trends in Internet usage. The inclusion of state trends is expected to be problematic for the instrumental variables estimates because the first stage is identified off of differences over time in the effect of state MDU rates on Internet use. Thus, state trends are likely to considerably reduce the power of the first stage. It is mechanically possible, however, to conduct such an estimate and although the power of the first stage is considerably weakened, the point estimate in 2SLS specifications remains quantitatively similar.

Next, I text the sensitivity of the results to the inclusion of several alternative control variables. One potential concern is that the included control variables do not accurately control for unobservables local labor market conditions. A state level measure of MDU living is used to deal with sorting bias in the instrument, but more local measures could be used for wages to capture more specific aspects of the labor market. Therefore, MSA level wages for both males and females are constructed from the Current Population Survey and included separately in the model. For individuals who do not live in an MSA, I use the rural portion of the state's average wage rates. This is shown in columns (1)-(4) of the third row of table 2.8 and the results are very similar to the original specification. Another concern is that Internet use is related to technology take-up more generally, which may also be correlated with labor market outcomes. One way to control for selection into technology usage that would lead to increased rates of Internet use is to specifically control for individual computer usage. Unfortunately, the sample will be severely limited by including this control variable, which is only available in 2000-2003 and is only asked at the household level. Results are displayed in the columns (5) and (6) of the third row of table 2.8. Because of the reduced sample size, the first stage is no longer strong enough to be confident about a causal interpretation of the point estimates. However, the point estimate is virtually unchanged in the linear IV estimation, suggesting that overall selection into technology use is not likely to be biasing results.

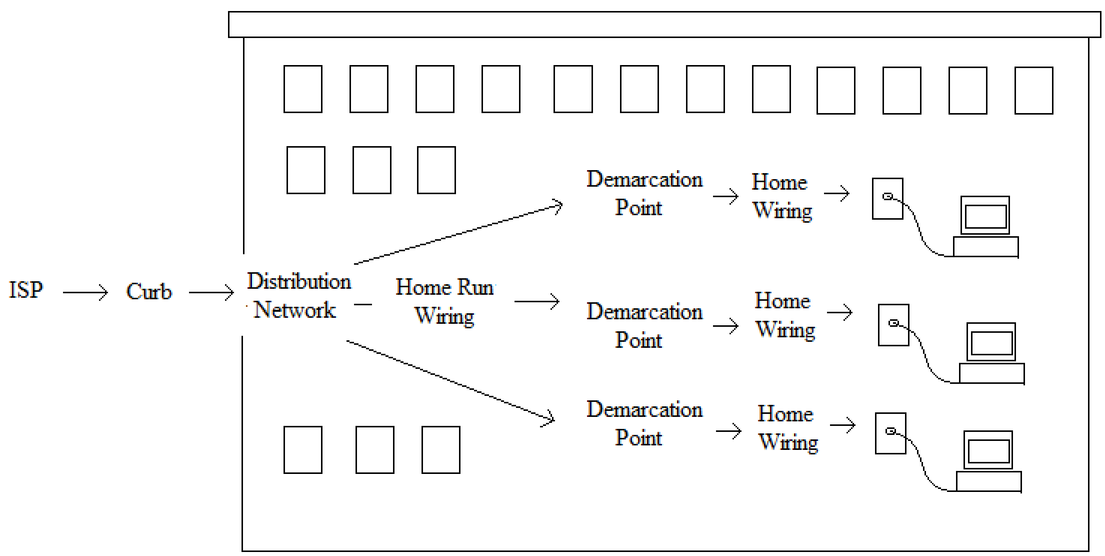
In addition to investigating the impact of Internet use on married women's participation, I have also investigated the impact for men and single women in the fourth row of appendix table 2.8. For both single women and men, the samples are small and first stage F statistics are too low to permit a causal interpretation. Overall, however, the magnitude and sign of the coefficients on *HSI* is not similar to that for married women. For men, the results are small and insignificant.

Figure 2.1: High-Speed Internet Installation Diagram

(a) Single Family Home

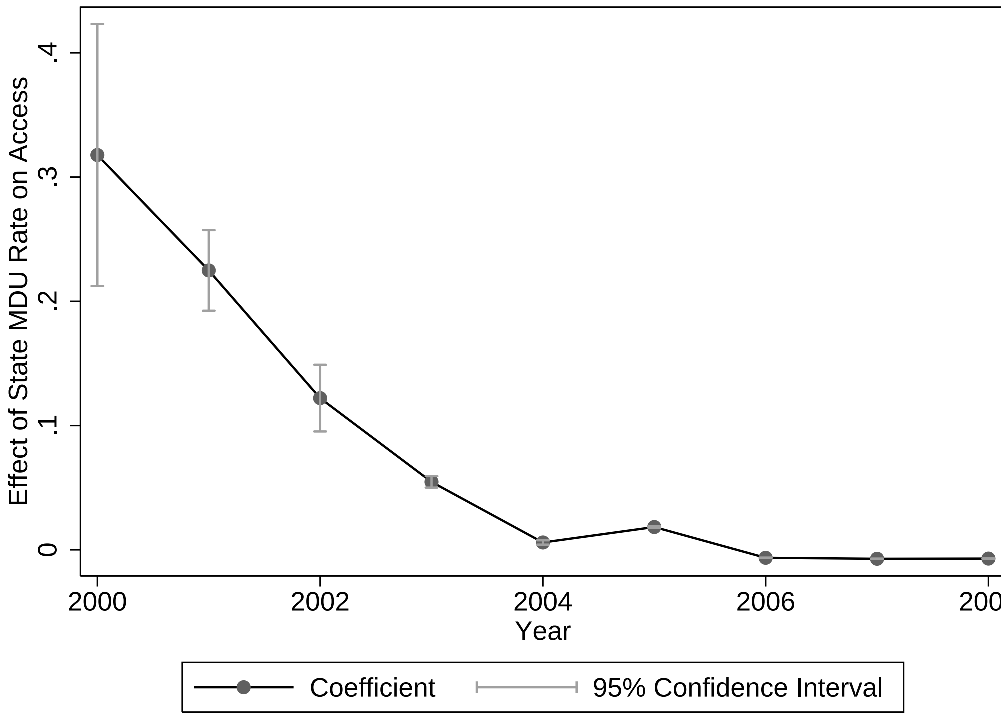


(b) Multiple Dwelling Unit (MDU)



Notes: Authors rendering based on Jackson (2002) and Ames (2006). "ISP" refers to the high speed Internet service provider.

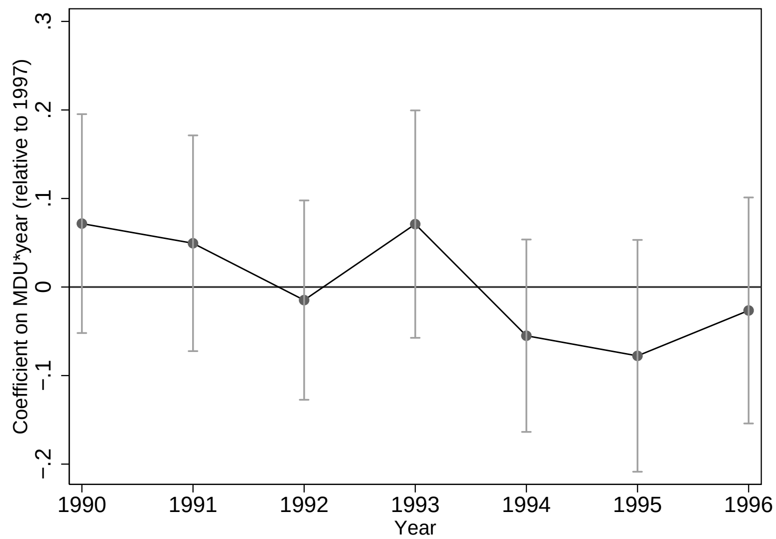
Figure 2.2: State MDU Rates and State Residential High-Speed Internet Access Rates



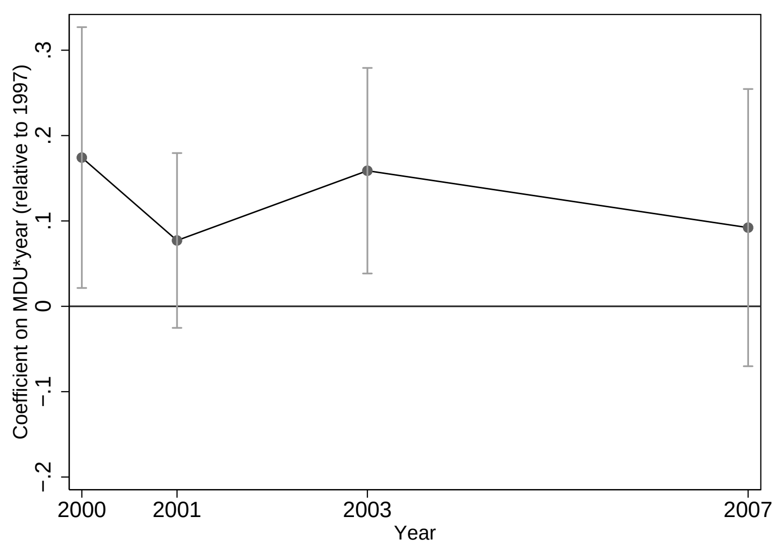
Notes: Displayed are the coefficients and 95 percent confidence intervals from a set of unconditional linear regressions relating yearly high-speed Internet access rates to 2000 state MDU rates. Data on state residential high-speed Internet access rates was obtained from the Federal Communications Commission Form 477 filing data for the years 2000-2008. Data on MDU rates and the number of households in each zip code was obtained from the 2000 Decennial Census. MDU rates are defined as the share of the 2000 state population residing in a structure with 3 or more units. Internet access rates are defined as the fraction of state's zip codes with at least one residential high-speed Internet customer in each year, weighted by the number of households in each zip code.

Figure 2.3: Reduced Form Relationship between MDU Rates and Labor Force Participation

(a) Years 1990-1997

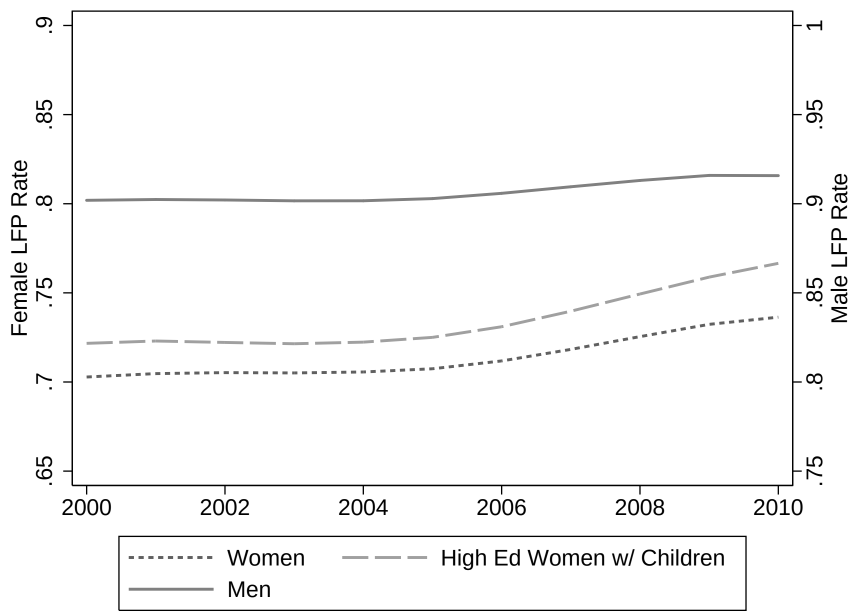


(b) Years 2000-2009



Notes: Displayed are the coefficients and 95% confidence intervals on the vector $MDU_s * \theta_t$ from estimation of the reduced from version of equation 2.2 and 2.3, which relates MDU rates to labor force participation. In panel (a) θ_t=1990-1997 and in panel (b) θ_t=2000, 2001, 2003, 2007 and 2009. Coefficients are relative to the base year of 1997 in panel (a) and 2009 in panel (b). 1990-1997 is the time period prior to the introduction of residential high-speed Internet access and 2000-2009 is the time period used throughout the chapter.

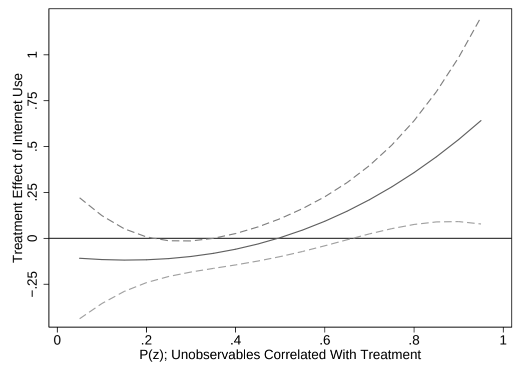
Figure 2.4: Aggregate Trends in Labor Force Participation 2000-2010



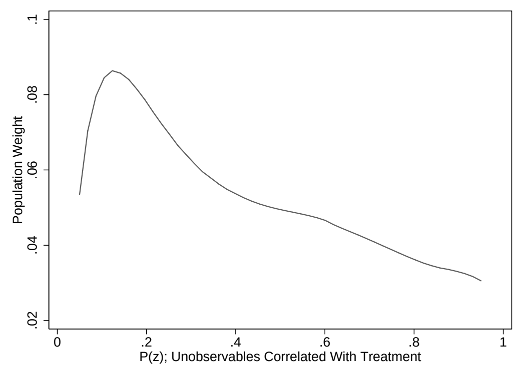
Notes: Displayed are aggregate labor force participation rates for married mean and women aged 18-59. Source is the American Communities Survey 2000-2010. Note that male and female rates are on different scales for ease of comparison.

Figure 2.5: Decomposition of Instrumental Variables Estimates

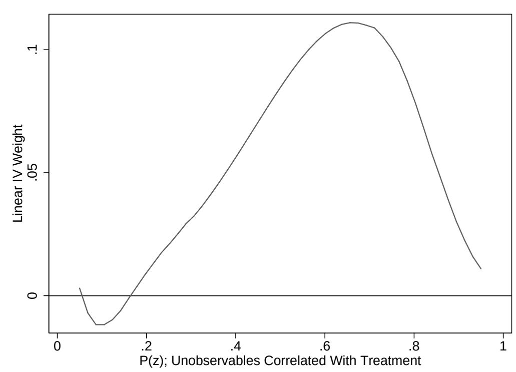
(a) Marginal Treatment Effects of Internet Usage



(b) Population Weights

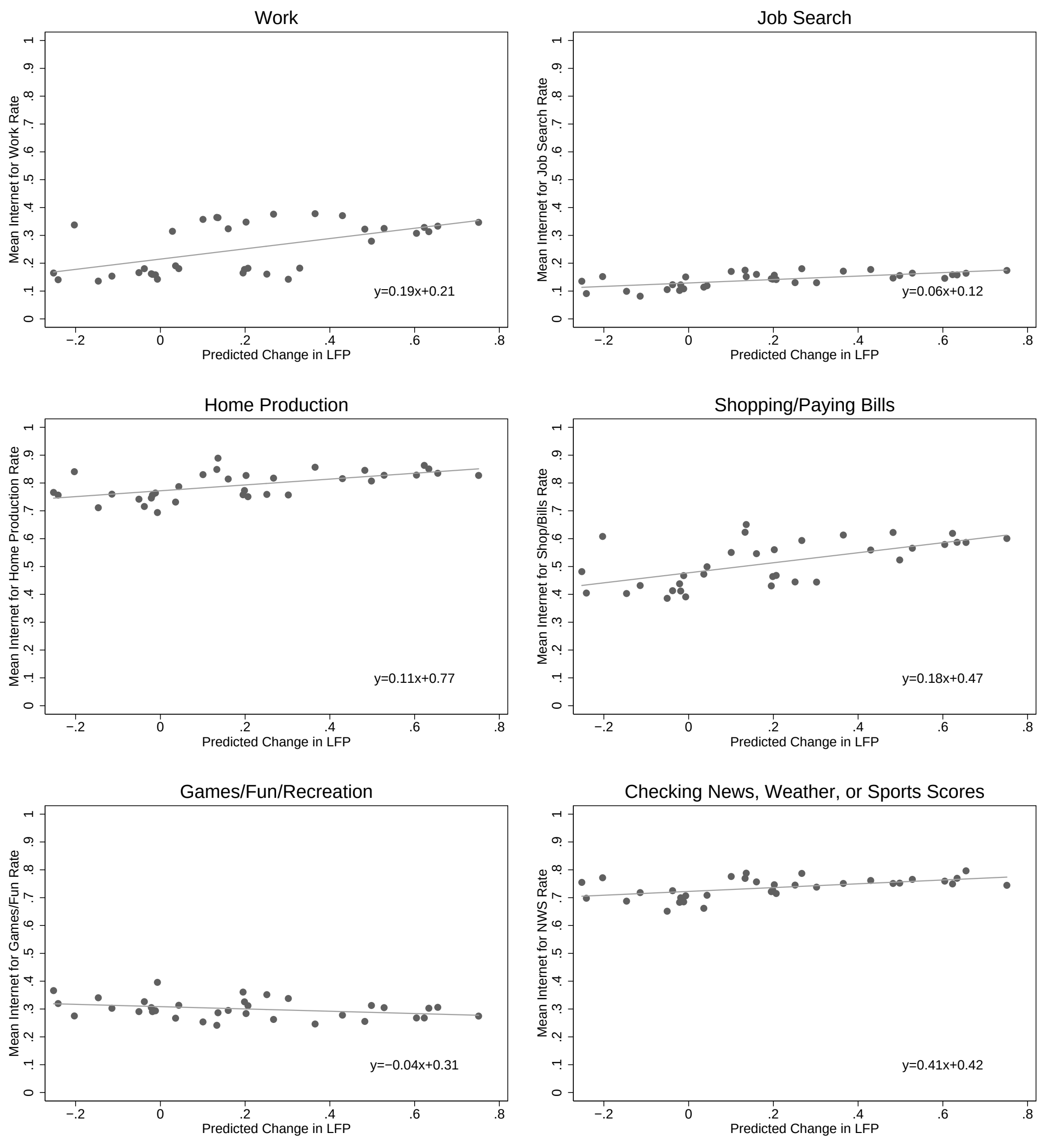


(c) Linear IV Weights



Notes: Panels (a) displays marginal treatment effects calculated using a local IV estimator. Panel (b) displays population weights and panel (c) displays the linear IV weights. In all cases the x-axis displays percentiles of the distribution of unobservables correlated with treatment status (the propensity score P(z)).

Figure 2.6: Group Mean Predicted Change in Participation and Rates of Internet Use for Different Tasks



Notes: Plotted is subgroup level mean rates of Internet use for each activity listed and subgroup-level predicted changes in labor force participation when high-speed Internet use increases from 0 to 1 ($\widehat{LFP}$), based on estimating equation 2.3 for each subgroup. The sample is limited to married women 18-59 and subgroups are defined by education, the presence of children and Census divisions.

Table 2.1: Summary Statistics on Married Women 2000-2009

Variable	Mean	Std. Deviation
high-speed Internet (HSI) User	0.356	0.479
.......in 2000	0.054	0.225
.......in 2001	0.112	0.315
.......in 2003	0.250	0.432
.......in 2007	0.662	0.473
.......in 2009	0.786	0.409
Labor Force Participation	0.731	0.444
Usual Hours Worked Per Week	37.51	11.19
Full Time Status	0.731	0.444
Less than High School	0.087	0.282
High School	0.303	0.460
Some College	0.293	0.455
College	0.317	0.465
Spouse-Less than High School	0.103	0.304
Spouse-High School	0.302	0.459
Spouse-Some College	0.265	0.442
Spouse-College	0.330	0.470
Lives in MSA	0.709	0.454
Lives in Central City	0.187	0.390
No Children Under 18	0.454	0.498
One Child Under 18	0.210	0.407
Two Children Under 18	0.218	0.413
3 or More Children Under 18	0.118	0.323
Any Children Under Age 6	0.242	0.428
White (NH)	0.766	0.424
Black (NH)	0.062	0.241
Hispanic	0.106	0.308
Other	0.066	0.249
Age	41.79	10.04

Notes: Displayed are means and standard deviations of the individual level dependent and independent variables from the 2000-2009 Current Population Survey supplements. The sample is limited to married women age 18-59. Full time status and hours worked are conditional upon participation in the labor market. The number of observations is 107,976.

Table 2.2: Summary of State-Level Variables 2000-2009

	Mean	Description	Source
Income PC	\$38,951 (\$5,701)	Income per capita	Bureau of Economic Analysis (BEA)
Avg Wage	\$43,866 (\$7,161)	Average annual wage income per job	Bureau of Economic Analysis (BEA)
Pop Density	135.1 (378.79)	Population Density	Census Land Area and Population Estimates
HPI	350.03 (110.57)	Housing Price Index	Federal Housing Finance Agency (FHFA)
Unemp Rate	5.48 (2.06)	Unemployment Rate	Bureau of Labor Statistics (BLS)
% Adopt	0.904 (0.066)	sum(Fraction Workers Industry i* Industry i Internet Adoption Rate)	Forman et al (2005) and BEA
% Enhance	0.129 (0.004)	sum(Fraction Workers Industry i* Industry i Enhancement Rate)	Forman et al. (2005) and BEA
% MDU	0.235 (0.059)	Percent Population in Multiple Dwelling Unit	2000 Census

Notes: Displayed are state-level control variables. Income per capita, average wage, and the house price index (HPI) are in 2009 dollars, inflated by CPI-U. Standard deviations displayed in parentheses.

Table 2.3: LPM Home Internet Use and Married Women's Labor Force Participation

	(1)	(2)	(3)	(4)	(5)	(6)
Dep Var: LFP						
HSI Use	0.0670***	0.0903***	0.0426***	0.0494***	0.0471***	0.0471***
	(0.00520)	(0.00699)	(0.00541)	(0.00531)	(0.00518)	(0.00518)
Year=2000		0.0471***	0.0174***	0.0240***	0.0185***	0.00818
		(0.00638)	(0.00578)	(0.00573)	(0.00565)	(0.0132)
Year=2001		0.0551***	0.0266***	0.0326***	0.0278***	0.0169
		(0.00583)	(0.00526)	(0.00522)	(0.00519)	(0.0122)
Year=2003		0.0390***	0.0168***	0.0217***	0.0188***	0.0101
		(0.00549)	(0.00499)	(0.00497)	(0.00481)	(0.00901)
Year=2007		0.00751	0.00254	0.00391	0.00280	-0.00834
		(0.00469)	(0.00463)	(0.00461)	(0.00452)	(0.00829)
Black (NH)			0.0550***	0.0465***	0.0488***	0.0488***
			(0.00597)	(0.00607)	(0.00593)	(0.00596)
Hispanic			-0.0299***	-0.0296***	-0.0185***	-0.0184***
			(0.00605)	(0.00601)	(0.00563)	(0.00564)
Other (NH)			-0.0517***	-0.0453***	-0.0486***	-0.0485***
			(0.00770)	(0.00753)	(0.00754)	(0.00756)
Age 20-24			0.126***	0.127***	0.149***	0.149***
			(0.0256)	(0.0254)	(0.0246)	(0.0246)
Age 25-29			0.162***	0.165***	0.199***	0.199***
			(0.0251)	(0.0249)	(0.0245)	(0.0244)
Age 30-34			0.158***	0.163***	0.212***	0.212***
			(0.0256)	(0.0254)	(0.0248)	(0.0248)
Age 35-39			0.186***	0.191***	0.218***	0.218***
			(0.0255)	(0.0253)	(0.0248)	(0.0248)
Age 40-44			0.221***	0.227***	0.207***	0.207***
			(0.0255)	(0.0252)	(0.0248)	(0.0248)
Age 45-49			0.235***	0.242***	0.188***	0.188***
			(0.0252)	(0.0249)	(0.0245)	(0.0245)
Age 50-54			0.203***	0.212***	0.145***	0.145***
			(0.0255)	(0.0252)	(0.0248)	(0.0248)
Age 55-59			0.106***	0.117***	0.0469*	0.0469*
			(0.0265)	(0.0262)	(0.0257)	(0.0257)
Central City			-0.0112**	-0.00873*	-0.0143***	-0.0141***
			(0.00486)	(0.00487)	(0.00495)	(0.00498)
MSA			-0.0175***	-0.0113***	-0.0104**	-0.0105***
			(0.00408)	(0.00401)	(0.00404)	(0.00403)

	(1)	(2)	(3)	(4)	(5)	(6)
Less than HS			-0.248***	-0.282***	-0.278***	-0.278***
			(0.00676)	(0.00734)	(0.00736)	(0.00735)
High School			-0.0783***	-0.124***	-0.128***	-0.128***
			(0.00382)	(0.00378)	(0.00373)	(0.00373)
Some College			-0.0297***	-0.0635***	-0.0656***	-0.0655***
			(0.00354)	(0.00352)	(0.00343)	(0.00343)
Spouse-Less than HS				0.0511***	0.0466***	0.0466***
				(0.00563)	(0.00570)	(0.00570)
Spouse-High School				0.0948***	0.0853***	0.0854***
				(0.00388)	(0.00395)	(0.00395)
Spouse-Some College				0.0852***	0.0781***	0.0781***
				(0.00346)	(0.00347)	(0.00348)
1 Child 0-5					-0.121***	-0.121***
					(0.00605)	(0.00605)
1 Child 6-18					0.0000731	0.0000932
					(0.00388)	(0.00388)
2 Children 0-5					-0.229***	-0.229***
					(0.00778)	(0.00778)
2 Children 6-18					-0.0430***	-0.0430***
					(0.00447)	(0.00447)
2 Children 0-18					-0.129***	-0.129***
					(0.00726)	(0.00727)
3+ Children 0-5					-0.354***	-0.354***
					(0.0179)	(0.0179)
3+ Children 6-18					-0.111***	-0.111***
					(0.00638)	(0.00639)
3+ Children 0-18					-0.264***	-0.264***
					(0.00721)	(0.00722)
State and Year FE		Yes	Yes	Yes	Yes	Yes
Basic Demographics			Yes	Yes	Yes	Yes
State-Level Controls						Yes
R^2	0.005	0.019	0.054	0.061	0.087	0.087
N	107976	107976	107976	107976	107976	107976

Notes: Basic individual demographic variables include fixed effects for age category, race, living in an MSA, and living in a central city. State-level controls include average wages, income per capita, unemployment rates, housing prices, population density, percent of state employees' industry internet adoption and enhancement, which are matched to the individual-level data at the state-year level. Standards errors clustered by state-year are in parentheses. The mean labor force participation rate is 0.73 and mean HSI use rate is 0.36. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 2.4: First Stage Relationship between Percent of State Residing in MDU and Home High-Speed Internet Use

	(1)	(2)	(3)	(4)	
%MDU*2000	0.565***	0.556***	0.566***	0.570***	0.736***
	(0.108)	(0.0996)	(0.0989)	(0.0989)	(0.0900)
	10.54%	*10.37%*	*10.56%*	*10.63%*	*13.73%*
%MDU*2001	0.583***	0.582***	0.585***	0.589***	0.716***
	(0.119)	(0.103)	(0.103)	(0.103)	(0.0816)
	5.19%	*5.18%*	*5.20%*	*5.24%*	*6.37%*
%MDU*2003	0.641***	0.631***	0.637***	0.638***	0.765***
	(0.116)	(0.105)	(0.104)	(0.103)	(0.0890)
	2.57%	*2.53%*	*2.55%*	*2.56%*	*3.07%*
%MDU*2007	0.280**	0.289***	0.290***	0.291***	0.252***
	(0.117)	(0.105)	(0.104)	(0.104)	(0.0917)
	0.42%	*0.44%*	*0.44%*	*0.44%*	*0.38%*
Control Variables:					
State and Year FE	Yes	Yes	Yes	Yes	Yes
Basic Demographics		Yes	Yes	Yes	Yes
Spouse's Education			Yes	Yes	Yes
Presence/Age of Children				Yes	Yes
State-level Controls					Yes
F Statistic	11.17	13.33	13.90	14.09	27.55
Over ID Stat	2.595	2.590	2.458	2.513	2.503
N	107976	107976	107976	107976	107976

Notes: Basic demographic variables include fixed effects for age category, race, number and ages of children, living in an MSA, and living in a central city. State-level variables include average wages, income per capita, unemployment rates, housing prices, population density, percent of state employees' industry internet adoption and enhancement, which are matched to the individual-level data at the state-year level. %MDU refers to the percent of state living in a multiple dwelling unit and is matched at the state-level. The left out category is %MDU*2009. Standards errors clustered by state-year are in parentheses. Percentages listed below each coefficient represent the percent change for a 1% increase in %MDU, scaled by the mean of the dependent variable in each year. The F Statistic refers to the cluster-robust F statistic described in the text. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 2.5: IV Home High-Speed Internet Use and Married Women's Labor Force Participation

	(1)	(2)	(3)	(4)	(5)
HSI Use	0.215**	0.182*	0.175*	0.204**	0.186***
	(0.0963)	(0.0954)	(0.0932)	(0.0945)	(0.0714)
Year=2000	0.138*	0.118*	0.115*	0.132*	0.114**
	(0.0705)	(0.0691)	(0.0675)	(0.0687)	(0.0534)
Year=2001	0.139**	0.120*	0.117*	0.133**	0.114**
	(0.0652)	(0.0638)	(0.0623)	(0.0634)	(0.0492)
Year=2003	0.106**	0.0908*	0.0885*	0.102**	0.0856**
	(0.0519)	(0.0508)	(0.0496)	(0.0504)	(0.0380)
Year=2007	0.0231*	0.0199	0.0197	0.0225*	0.0103
	(0.0126)	(0.0123)	(0.0120)	(0.0124)	(0.0113)
Black (NH)		0.0664***	0.0554***	0.0601***	0.0588***
		(0.00929)	(0.00844)	(0.00855)	(0.00764)
Hispanic		-0.0108	-0.0147	0.000287	-0.00182
		(0.0144)	(0.0125)	(0.0130)	(0.0103)
Other (NH)		-0.0444***	-0.0383***	-0.0398***	-0.0408***
		(0.00936)	(0.00934)	(0.00959)	(0.00867)
Age 20-24		0.131***	0.131***	0.155***	0.154***
		(0.0260)	(0.0258)	(0.0250)	(0.0247)
Age 25-29		0.165***	0.169***	0.204***	0.204***
		(0.0255)	(0.0253)	(0.0248)	(0.0246)
Age 30-34		0.161***	0.166***	0.217***	0.216***
		(0.0261)	(0.0258)	(0.0254)	(0.0251)
Age 35-39		0.188***	0.194***	0.224***	0.223***
		(0.0258)	(0.0255)	(0.0252)	(0.0250)
Age 40-44		0.224***	0.231***	0.214***	0.213***
		(0.0258)	(0.0256)	(0.0255)	(0.0251)
Age 45-49		0.239***	0.247***	0.196***	0.195***
		(0.0257)	(0.0255)	(0.0253)	(0.0249)

Age 50-54	0.210***	0.221***	0.156***	0.155***
	(0.0265)	(0.0264)	(0.0264)	(0.0256)
Age 55-59	0.118***	0.130***	0.0622**	0.0604**
	(0.0280)	(0.0280)	(0.0278)	(0.0267)
Central City	-0.0101**	-0.00769	-0.0131***	-0.0130***
	(0.00498)	(0.00500)	(0.00503)	(0.00491)
MSA	-0.0269***	-0.0186***	-0.0195***	-0.0186***
	(0.00797)	(0.00703)	(0.00705)	(0.00604)
Less than HS	-0.210***	-0.259***	-0.249***	-0.252***
	(0.0272)	(0.0187)	(0.0190)	(0.0154)
High School	-0.0562***	-0.111***	-0.112***	-0.114***
	(0.0162)	(0.0107)	(0.0108)	(0.00871)
Some College	-0.0197**	-0.0580***	-0.0586***	-0.0594***
	(0.00795)	(0.00545)	(0.00549)	(0.00491)
Spouse-Less than HS		0.0696***	0.0697***	0.0670***
		(0.0154)	(0.0158)	(0.0124)
Spouse-High School		0.107***	0.101***	0.0991***
		(0.0104)	(0.0107)	(0.00844)
Spouse-Some College		0.0909***	0.0853***	0.0844***
		(0.00556)	(0.00565)	(0.00476)
1 Child 0-5			-0.121***	-0.121***
			(0.00613)	(0.00612)
1 Child 6-18			-0.00520	-0.00456
			(0.00509)	(0.00464)
2 Children 0-5			-0.228***	-0.228***
			(0.00795)	(0.00788)
2 Children 6-18			-0.0481***	-0.0474***
			(0.00521)	(0.00502)

	(1)	(2)	(3)	(4)	(5)
2 Children 0-18				-0.129***	-0.129***
				(0.00717)	(0.00719)
3+Children 0-5				-0.355***	-0.355***
				(0.0179)	(0.0179)
3+ Children 6-18				-0.115***	-0.115***
				(0.00715)	(0.00691)
3+ Children 0-18				-0.263***	-0.263***
				(0.00738)	(0.00732)
State and Year FE	Yes	Yes	Yes	Yes	Yes
Basic Demographics		Yes	Yes	Yes	Yes
Spouse's Education			Yes	Yes	Yes
Presence/Age of Children				Yes	Yes
State-level Controls					Yes
LPM Estimates					
HSI Use	0.0903***	0.0426***	0.0494***	0.0471***	0.0471***
	(0.00699)	(0.00541)	(0.00531)	(0.00518)	(0.00518)
First Stage					
F Statistic	11.17	13.33	13.90	14.09	27.55
Over ID Stat	2.595	2.590	2.458	2.513	2.503
N	107976	107976	107976	107976	107976

Notes: Basic demographic variables include fixed effects for age category, race, number and ages of children, living in an MSA, and living in a central city. State-level variables include average wages, income per capita, unemployment rates, housing prices, population density, percent of state employees' industry internet adoption and enhancement, which are matched to the individual-level data at the state-year level. Instrument is percent of state living in a multiple dwelling unit (MDU) and is matched at the state-level. The F Statistic refers to the cluster-robust F statistic described in the text. Standards errors adjusted for clustering by state-year are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 2.6: Additional IV Estimates of Home High-Speed Internet Use and Participation

(A) Differences by Presence of Children					
	All	No Children	Children	Children Under 6	Children 6-18
OLS					
HSI Use	0.0471***	0.0606***	0.0359***	0.0492***	0.0247***
	(0.00518)	(0.00617)	(0.00631)	(0.00928)	(0.00726)
2SLS					
HSI Use	0.186***	0.0276	0.293***	0.176	0.302*
	(0.0714)	(0.0966)	(0.112)	(0.139)	(0.163)
F Stat	27.55	22.82	19.59	14.47	12.99
Mean LFP	0.73	0.75	0.71	0.62	0.76
Mean HSI Use	0.33	0.34	0.37	0.36	0.38
N	107976	48745	59231	26128	33103
(B) Differences by Education					
	LHS	HS/SC	College	HS or Less	Some or More College
OLS					
HSI Use	0.0925***	0.0523***	0.0413***	0.0681***	0.0335***
	(0.0205)	(0.00614)	(0.00739)	(0.00775)	(0.00533)
2SLS					
HSI Use	-0.136	0.229*	0.380*	0.0905	0.383**
	(0.397)	(0.125)	(0.204)	(0.0876)	(0.151)
F Stat	3.155	21.12	6.859	19.77	14.12
Mean LFP	0.50	0.73	0.79	0.66	0.78
Mean HSI Use	0.11	0.32	0.49	0.23	0.44
N	8999	64210	34767	41581	66395
(C) Alternative Outcomes and Extensions					
	Hours	Full Time	Family Income	Employment	Married Men
OLS					
HSI Use	0.270**	0.00216	16627.3***	0.0460***	0.0258***
	(0.118)	(0.00419)	(628.8)	(0.00524)	(0.00260)
IV					
HSI Use	8.513***	0.198*	27083.7**	0.175**	0.0772
	(2.791)	(0.107)	(13463.0)	(0.0711)	(0.0512)
F Stat	16.80	21.15	24.04	27.55	19.43
N	70018	76066	91290	107976	99592

Notes: All specifications include the following demographic variables: fixed effects for age category, race, number and ages of children, living in an MSA, and living in a central city; and state-level variables: average wages, income per capita, unemployment rates, housing prices, population density, percent of state employees' industry internet adoption and enhancement, which are matched to the individual-level data at the state-year level. Instrument is percent of state living in a multiple dwelling unit (MDU) and is matched at the state-level. The F Statistic refers to the cluster-robust F statistic described in the text. Standards errors adjusted for clustering by state-year are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 2.7: Decomposing the Instrumental Variables Estimates

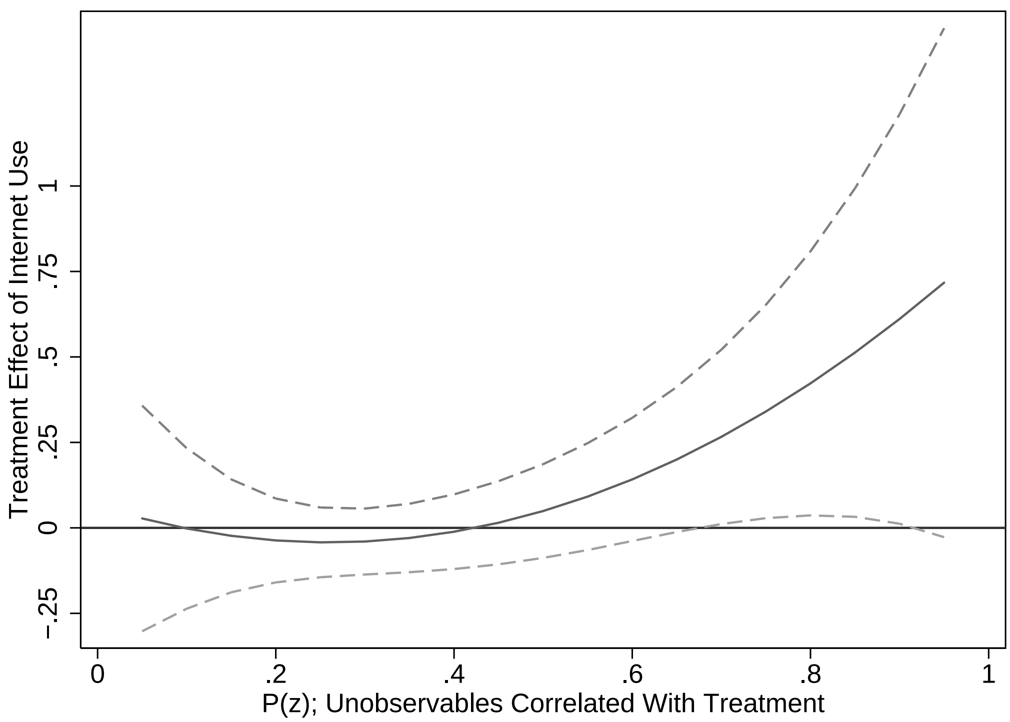
(a) Individual Characteristics of Compliers	Mean All	Mean Compliers	Relative Probability
Less than HS	0.09	0.10	1.09
High School	0.30	0.24	0.78
Some College	0.29	0.16	0.55
College	0.32	0.20	0.62
Spouse College Educated	0.33	0.27	0.81
Has Children	0.55	0.57	1.04
White (Non-Hispanic)	0.77	0.27	0.35
Black (Non-Hispanic)	0.06	0.07	1.16
Hispanic	0.11	0.06	0.55
Age 18-34	0.26	0.34	1.31
Age 35-59	0.74	0.65	0.89

(b) IV Weights on Composite State SES Measure	
Percentile of Distribution	Weight
0-25th Percentile	0.60
25th-50th Percentile	0.23
50th-75th Percentile	0.18
75th-100th Percentile	0.00

(c) Individual Characteristics across Distribution of Unobservables	0-0.2 Percentiles	0.2-0.8 Percentiles	0.8-1 Percentiles
Less than HS	0.07	0.05	0.00
High School	0.40	0.29	0.08
Some College	0.31	0.31	0.30
College	0.22	0.35	0.62
Spouse College Educated	0.24	0.36	0.62
No Children	0.47	0.46	0.38
Children Age <6	0.24	0.23	0.26
Children Age 6-18	0.29	0.31	0.36
White (Non-Hispanic)	0.77	0.78	0.87
Black (Non-Hispanic)	0.07	0.06	0.03
Hispanic	0.10	0.09	0.02
Age	41.65	42.30	41.58

Notes: In Panel (a), column (1) displays the mean of the characteristic indicated in each row, column (2) displays the mean of the characteristic for the compliant population, and column (3) dislplays the likelihood a complier has the characteristic indicated in each row. Column (2) was calculated from columns (1) and (3). Panel (b) displays the weights placed on each quartile of the distribution of the state SES measure by the IV estimate. The state SES measure is calculated based on a linear projection of state-level demographic and economic characteristics on high-speed Internet usage, as described in the text. Panel (c) displays the mean of the characteristic indicated in each row across different intervals in the distribution of unobservables correlated with Internet usage (P(z)).

Figure 2.7: Appendix - Marginal Treatment Effects for $z = MDU * t$



Notes: Displayed are marginal treatment effects calculated using a local IV estimator. The x-axis displays percentiles of the distribution of unobservables correlated with treatment status (the propensity score P(z)).

Table 2.8: Appendix - Extensions and Robustness Checks

	Alternative Instruments					
	Appliances	MDU*t	MDU*Pre	Both IVs	MDU=2+ Units	MDU=5+ Units
HSI Use	0.130*	0.184**	0.201**	0.173***	0.194**	0.222***
	(0.0676)	(0.0819)	(0.1013)	(0.0583)	(0.0842)	(0.0819)
F Stat	12.21	68.94	29.55	15.97	11.82	18.71
Over-ID Stat	-	-	-	12.06	0.241	5.382
N	107976	107976	107976	107976	107976	107976

	Drop Allocated		Cluster at State		State Trends	
	OLS	IV	OLS	IV	OLS	IV
HSI Use	0.0482***	0.231***	0.0471***	0.186*	0.0467***	0.219
	(0.00490)	(0.0820)	(0.00480)	(0.0964)	(0.00519)	(0.282)
F Stat		23.89		22.92		3.905
N	96770	96770	107976	107976	107976	107976

	MSA Male Wages		MSA Female Wages		Computer Use	
	OLS	IV	OLS	IV	OLS	IV
HSI Use	0.0498***	0.181**	0.0498***	0.191**	-0.00307	0.158
	(0.00603)	(0.0762)	(0.00604)	(0.0786)	(0.00517)	(0.389)
Male Wage	-0.000989*	-0.00137**				
	(0.000581)	(0.000606)				
Female Wage			-0.00121	-0.00172*		
			(0.000962)	(0.000991)		
Computer Use					0.0672***	0.0470
					(0.00483)	(0.0486)
F Stat		17.76		18.53		2.297
N	76683	76683	76683	76683	68273	68273

	Single Women		Married Men		Single Men	
	OLS	IV	OLS	IV	OLS	IV
HSI Use	0.0681***	-0.260*	0.0258***	0.0772	0.0438***	-0.461*
	(0.00604)	(0.151)	(0.00260)	(0.0512)	(0.00492)	(0.256)
F Stat		5.609		19.43		4.223
N	79679	79679	99592	99592	76497	76497

Notes:All specifications include fixed effects for age category, race, education, spouse's education, number and ages of children, living in an MSA, living in a central city, year and state (when appropriate). Also included is state income per capita, unemployment rates, housing prices, population density, percent of state employees' industry internet adoption and enhancement, which are matched at the state-year level. Instrument is the percent of state residing in a multiple dwelling unit and is matched at the state level (except where noted otherwise). Standard errors clustered by state-year are in parentheses (except where noted otherwise). * $p < .1$, ** $p < .05$, *** $p < .01$

Chapter 3

Home Internet and the Labor Market: Decomposing the Effects of Internet Usage on Labor Supply

3.1 Introduction

Home Internet technology has altered how, when and where individuals conduct numerous aspect of daily life. In this second chapter of this dissertation, I demonstrate that home Internet has a sizable impact on the labor supply of married women. This result perhaps raises more questions than it answers. Namely, if Internet usage had a profound effect on women's labor supply, what is the reason for that increase? The fact that Internet is used for many different tasks suggests numerous plausible candidate explanations – from the increased availability of flexible scheduling arrangements to reductions in the amount time that needs to be spent in home production. In this chapter, I disentangle the reasons for the demonstrated increase in female labor supply associated with Internet usage. First, I propose a theoretical model for investigating the impact of the Internet on labor supply via four different channels: work at home, home production, leisure and job search. I then provide empirical evidence on the extent to which each of those mechanisms contributes to observed change in labor supply for married women.

I begin the analysis of the mechanisms behind the demonstrated increase in labor supply by examining how the Internet is used, both across groups and over time. Using data from the Current Population Survey (CPS) and the PEW Internet and American Life Project I find that the Internet is used by most individuals for a diverse range of tasks. This includes activities that are best classified as home production, such as shopping, banking and paying bills, as well as activities that are best characterized as leisure, such as checking news, weather reports and sports scores and playing games. Individuals also report using the Internet at high rates for work and job search. Moreover, rates of use for most of these activities have increased between 2000 and 2010, both on the extensive and intensive margin. This is suggestive evidence that the introduction of high speed Internet – which increased in usage from about 5% of the population in 2000 to 74% in 2009 – may have facilitated use of Internet services for these activities.[49] This indicates that any analysis seeking to disentangle the reasons for the observed increase in labor supply associated with high speed Internet usage needs to consider numerous candidate explanations.

To capture the various ways Internet usage for different activities can affect labor supply decisions, I propose a theoretical framework based on a Becker (1965) model of time allocation. Key to this framework is the recognition that home Internet is not only an input in traditionally home-based activities such as leisure and home production, but can also be used for market work if individuals are able to telework or search for jobs online. In this framework, working from home lowers the costs of working, including the time and monetary costs of commuting and lost wages when children stay home from school or repairmen need to be met during work hours. Internet also reduces search frictions in the labor market, enabling individuals to more easily transition into market work. As with prior home technologies like washing machines and microwaves, the Internet can be used to save time in home production tasks, freeing up time for market work. On the other hand, the Internet offers a wide range of entertainment options, which can potentially offset any increase in labor supply. In sum, this framework implies that to understand the combined net effect of Internet usage on labor supply requires an understanding of the extent to which individuals use Internet for each of these tasks and the responsiveness of individual labor supply along each margin.

To empirically disentangle the net effect of Internet usage on labor supply, I separately analyze the effects of Internet usage for telework, job search, home production and leisure. To do so, I employ various sources of data on Internet usage, telework, employment histories and time use. An analysis of work at home rates confirms that those most affected by

[49] Author's calculation from chapter 2 of this dissertation.

Internet use are those in occupations with high rates of working from home: individuals in the top quartile in terms of predicted impact of Internet use on participation have work at home rates that are 45 percentage points (123 percent) higher than those in the bottom quartile. In addition, I examine different types of telework arrangements, including full time or part time, and for contract or through an employer. This analysis indicates that the impact of home Internet is through increased usage of employer-provided, part time work from home opportunities. Moreover, a back of the envelope calculation indicates that the rise in female participation induced by Internet can explain 27 percent of the rise in telework over the period studied.

To study job search, home production and leisure, I construct work histories for individuals who report using the Internet and examine their labor market outcomes for the following year. This analysis indicates that Internet use for job search increases the propensity for an individual to join the workforce in the following year and reduces the duration of the transition from non-participation to participation. Internet use for home production and leisure do not have significant impacts on participation or the speed with which transitions are made. However, an analysis of time use data indicates that Internet users spend approximately 2 fewer hours per week in home production than those who do not use the Internet, confirming that time is saved in home production online, although the effects are small. A back of the envelope calculation suggest that Internet usage for search and home production can explain 25 percent of the increase in labor supply induced by Internet usage found in the second chapter.

Based on the analyses conducted in this chapter, I conclude that telework, job search and time saved in home production are all plausible mechanisms explaining the results, and the ability to engage in part-time telework through an employer appears to be the greatest contributor to the estimated effects. This finding has important public policy implications, and speaks to broader policy discussions about the potential benefits of telework and flexible scheduling policies. While it is generally accepted that flexibility in the workplace has the potential to benefit employers, employees and the economy as a whole, adoption is still low and there is little empirical evidence on the benefits/costs of these policies.[50] This chapter has demonstrated that Internet usage, via take-up of telework opportunities, has allowed a group of highly educated women to join the workforce, suggesting such policies may have the potential to encourage workforce entry by productive individuals.

3.2 Background on Internet Usage Patterns

To understand the mechanisms via which Internet usage affects labor supply I begin by investigating how individuals use the Internet. To do so, I first examine data on the activities individuals report doing online, which is available Current Population Survey Internet and Computer Use supplements in 2000, 2001 and 2003. The surveys record detailed information about activities individuals report engaging in online in the past year, which I use to tabulate summary statistics on individual's online behavior. The surveys vary in the specific questions asked, so for consistency the questions are aggregated into categories: email, home production (with a separate category for shopping and paying bills), leisure (with separate categories for entertainment and checking news, weather, and sports scores), work-related and job search.[51] I show components of each category separately because some activities are more difficult to place into a single category than others. For example, while "entertainment" is almost certainly a leisure activity, "reading news" is more ambiguous.

Table 3.1 displays a summary statistics of this data and indicates that individuals use the Internet for a wide range of tasks that fall within each of the categories. Among married

[50]See, for example, the report by the Council of Economic Advisers on Work-life Balance in March 2010 (Council of Economic Advisers, 2010).

[51]The supplements ask about specific activities performed online in the past year. The data is only summarized for those who report using the Internet at home. Home production includes shopping and paying bills online, looking for government or health information, and looking for information about products and services. Leisure includes recreation (playing games, watching TV or movies, listening to music, recreation, entertainment or fun) and checking news, weather, or sports (NWS). Work related tasks, email and job search were consistently asked throughout the sample. See section 2.9 of this dissertation for more information on variable creation.

women, 36 percent use the Internet at home, and approximately 81 percent of those women use the internet for "home production", 75 percent for "leisure", 28 percent for work and 15 percent for job search. Within the home production and leisure categories, 54 percent use the Internet for "shopping or paying bills" and 29 percent use it for "entertainment." Among different groups, more educated women use the Internet more intensively for almost all tasks, including work and "shopping/paying bills", than either men or less educated women.

It is clear that the Internet is used across demographic groups for a variety of activities. It is also important to understand how these patterns have changed over time. This is particularly relevant for understanding the impact of broadband (as opposed dial-up), since that is the focus of the second chapter of the dissertation. Since the CPS data is only available until 2003 it is difficult to study trends using that data source. Fortunately, the PEW Internet and American Life Project collects trend data on Internet usage which is available from 2000 to 2010.[52] Although this data does not have the detailed demographic information the CPS has, it does have the advantage of having a measure of intensity of Internet usage: it asks individuals if they have "ever" done an activity online as well as their usage patterns "yesterday."

Figure 3.1 displays trends in Internet usage for email, work, job search, shopping, banking, and "just for fun." Figure 3.1 (a) displays trends in "ever use" and figure 3.1 (b) displays trends in use "yesterday." There is a clear pattern of increasing use of Internet shopping and banking, both extensively and intensively. This is consistent with the hypothesis that broadband might be particularly important facilitating those tasks. There is also a large increase in use of Internet "just for fun," suggesting usage for leisure has increased as well. While Internet usage for email and work "ever" remained fairly flat between 2000-2008 (data was not available post 2008 for this category), usage "yesterday" increases substantially for both uses over the period. This is also potentially consistent with increased usage for work

[52]PEW Internet and American Life Project "Usage Over Time" data can be found online at `http://pewinternet.org/Static-Pages/Trend-Data-(Adults)/Usage-Over-Time.aspx`. The PEW Internet and American Life Project bears no responsibility for the interpretations presented or conclusions reached in this dissertation.

associated with broadband deployment, although an important caveat is that the question does not distinguish between Internet usage in the home and at work.

Both the CPS and PEW Internet usage data indicate that individuals use the Internet for a diverse range of tasks, including home production tasks like shopping, leisure activities, work and job search. Moreover, trend data is indicative that usage of the Internet for those activities increased over the time period studied, and individuals are using Internet more intensively for most activities. In the next section, I propose a theoretical framework for incorporating these various uses of the Internet into labor supply decisions.

3.3 Theoretical Framework

To analyze the theoretical impact of Internet use on labor supply, I begin with a Becker (1965) model of time allocation. In this model, an agent has preferences over a set of commodities that are produced in home using a combination of inputs purchased in the market and the agent's own time. These commodities include home production goods like meals or a clean home, as well as leisure goods like watching television. As an example, purchased inputs may include raw ingredients and cooking utensils, which produce a meal in combination with time spent cooking and eating. The agent's time can be allocated between home and market work. Time in the market earns a wage, and this income can be used to purchase inputs for production.

To understand the impact of home Internet use and guide the empirical analysis, it is useful to consider the following formulation of Becker's (1965) model, which is similar to the framework outlined by Greenwood and Vandenbroucke (2005).[53] This framework will facilitate the illustration of the impact of Internet as a source of technological progress in the production of home goods and leisure, as well as an input in job search and telework. It should be noted that the goal of this exercise is not to specify a structural equation to

[53] Their model is a simpler, static version of the framework used in Greenwood, Sheshadri, and Yorukoglu (2005) to describe the impact of technological change in the home on female labor supply.

be estimated, but to serve as a framework for conceptualizing the various ways Internet use may affect labor supply in order to inform the empirical analysis that follows.

Consider an agent who has preferences over two goods: a market purchased consumption good c and a home produced good h. For the moment, the home produced good may be either a leisure good or a home production good, both of which are produced in the home using time and purchased inputs. Utility is described by the following additively separable function:

$$U(c) + V(h) \tag{3.1}$$

Where U_c, $V_h > 0$ and U_{cc}, $V_{hh} < 0$.[54] The home produced good is produced using the agents time t_h and market purchased inputs x according to the following constant returns to scale production function:

$$h = f(t_h, x) \tag{3.2}$$

Where f_t, $f_x > 0$, f_{tt}, $f_{xx} < 0$.[55] The price of x is q and the price of c is normalized to one. Time on the market is represented by t_m. When the agents spends t_m units of time in market work, he pays a time cost $w\rho t_m$ to working, where $0 < \rho < 1$. ρt_m represents time spent in commute, or in transit when an emergency or family obligation requires the agent to leave work (say, to meet a repairman).[56] The price of this time spent in commute is its opportunity cost, which is captured by the market wage w. The budget constraints can be written as:

$$c + qx = w(1 - \rho)t_m + B \tag{3.3}$$

[54] $U_c = \partial U/\partial C$.

[55] Diminishing marginal returns with time occur because agents get tired, begins to produce goods that have more market substitutes, and/or inputs start to wear out over time.

[56] This implies that commute time increases linearly in hours worked. While this assumption is oversimplified, it is not entirely unreasonable since commute time does tend to increase in *days* worked. Other authors have alternatively modeled commute as a fixed time cost. In this case, the budget constraint is non-convex which is problematic for optimization, but the resulting predictions in a static framework are the same (Black et al., 2008).

where B is unearned income. In a married couple (assuming no joint labor supply decisions) this might represent the husband's income, plus any household assets that earn income. The time endowment is T and the time constraint is:

$$t_h + t_m = T \tag{3.4}$$

Maximizing utility subject to the time and budget constraints leads to the following expression which governs market entry:[57]

$$\widetilde{w} \geq \frac{V_h f_t}{U_c} \tag{3.5}$$

Where $\widetilde{w} = w(1-\rho)$. This expression indicates that entry into the market depends on whether or not the net real wage is greater than or equal to the ratio of the marginal benefit of time spent in home production to the marginal utility of consumption. The ratio on the right hand side thus represents the reservation wage. Therefore, equation 3.5 indicates that the agent enters the market when the net market wage (the left hand side of 3.5) is greater than or equal to the reservation wage (the right hand side of 3.5). At an interior solution where the agent both works and spends time in home production, the agent is equating the marginal benefit of time spent in home production with its marginal cost (in terms of lost

[57]The Lagrangian optimization problem is represented by the following expression:

$$L = U(c) + V(f(t_h, x)) + \lambda((1-\rho)wt_m + B - qx - c) + \mu(T - t_m - t_h)$$

For simplicity, suppose expenditures on inputs $qx = \overline{qx}$ are fixed so that $\overline{qx} > 0$ (i.e., there is some indivisibility associated with x in production). Then, maximizing over c,t_h, and t_m implies that:

$$\begin{gathered} \lambda w(1-\rho) \leq \mu \\ V_h f_t \leq \mu \\ U_c \leq \lambda q \end{gathered}$$

Thus, for the case where $c > 0$:

$$\begin{array}{rr} \widetilde{w} = w(1-p) < \frac{V_h f_t}{U_c} & if\ t_m = 0 \\ \widetilde{w} = w(1-p) = \frac{V_h f_t}{U_c} & if\ 0 < t_m < T \\ \widetilde{w} = w(1-p) > \frac{V_h f_t}{U_c}. & if\ t_m = T \end{array}$$

Thus, market entry ($t_m > 0$) occurs when 3.5 is satisfied.

wages and foregone consumption utility).

It is then straightforward to show that in this framework to show that the ability to work from home can lead to an increase in labor force participation. Working from home is captured by a decrease in the fraction of time which much be spent in commute ρ, which implies an increase in the net wage $\widetilde{w}$. For some individuals, $\widetilde{w}$ will then become sufficiently large to satisfy 3.5 and these individuals will be induced to enter the market.

Internet may also affect the observed market wage w. Suppose the observed market wage is represented by the following function: $w = pw^{high} + (1-p)w^{low}$, so that with probability p the agent observes a high market wage and with probability *(1-p)* he observes a lower market wage. If the Internet reduces search frictions, this can be captured by an increase in p, and for some individuals $\widetilde{w}$ will become sufficiently large to induce entry. Similarly, if home Internet changes the agent's wage offer distribution by increasing his productivity at work, this can be expressed as an increase in w^{high} and/or w^{low}, which will also induce entry. This could be the case, for example, if telework increases the agent's productivity by decreasing absenteeism or if agent's are more productive working at home because there are fewer distractions.[58]

Next I consider the impact of Internet technology on the production of leisure and home production goods. In this case, home Internet can be viewed as technological progress in the purchased inputs to production, x. This is captured in the model by an increase in inputs x accompanied by a decrease in the price of inputs q. For simplicity, I assume that total expenditures $\overline{qx}$ remains unchanged.[59] Due to constant returns to scale in production, an increase in x leads to an increase in the amount of the home good h that can be produced given the same time input t_h.[60] The intuition behind this is fairly straightforward. Consider the example of purchasing clothing and other necessary household items. Prior to the Inter-

[58]To be clear, home Internet use (not workplace Internet use) is the mechanism considered that could alter productivity. There is some recent experimental evidence that telework can in fact increase productivity (Bloom, Liang, Roberts, and Ying, 2012).

[59]The assumption that qx remains unchanged is innocuous and eases interpretation.

[60]Suppose the new technology allows the production of αh goods given t_h, then constant returns to scale implies $\alpha h = \alpha f(t_h, x) = f(t_h, \alpha x)$.

net, to produce this home good the agent would have spent time traveling to multiple stores and waiting in lines. With Internet technology, the agent can shop with a click of a mouse, meaning that given the same time input, the agent can now produce much more "shopping" (alternatively, the agent can decrease the time cost associated with production of the old amount of "shopping").

For ease of composition, let the utility of home produced goods be represented by:

$$Z(t_h, x) = V(f(t_h, x)) \tag{3.6}$$

An increase in x affects the reservation wage, or the right hand side of 3.5. Substituting 3.6 and differentiating the right-hand side of 3.5 with respect to x leads to the following expression, which describes the effect of technological improvement in x on the reservation wage:

$$\frac{d}{dx}\left(\frac{Z_t(t_h, x)}{U_c(c)}\right) = \frac{Z_{tx}}{U_c} \tag{3.7}$$

Since $U_c > 0$, the sign of this expression depends on the sign of Z_{tx}, which describes the extent to which time and goods are substitutes or complements in utility. For a leisure good, time and inputs are expected to be complements in utility, which implies $Z_{tx} > 0$. This means that technological improvements in the production of leisure goods like watching television will increase the reservation wage, leading to a decrease in participation for those on the margin. Home production goods such as shopping and paying bills, on the other hand, are goods for which time and inputs are expected to be substitutes in utility. In this case, $Z_{tx} < 0$, which implies that for home production goods, technological progress in x leads to a decrease in the reservation wage, and an increase in labor force participation. Therefore, if home Internet is used exclusively for home produced goods such as shopping or paying bills, labor supply will increase. On the other hand, if home Internet is used exclusively for leisure, labor force participation will decrease. Since it is likely that Internet is an input in both types of goods (as well as telework and job search), the net predicted impact on participation is ambiguous.

The model also has an ambiguous prediction for hours of work conditional on participation in the market. When the agent works and spends time in home production, 3.5 holds with equality. Fully differentiating 3.5 leads to the following expressions for the impact of telework and job search (a change in $\widetilde{w}$ via a change in either w or ρ) or an improvement in home production or leisure technology (an increase in x) on time in the market t_m:

$$\frac{dt_m}{d\widetilde{w}} = \frac{U_c + U_{cc}\widetilde{w}t_m}{Z_{tt} + U_{cc}\widetilde{w}^2} \tag{3.8}$$

$$\frac{dt_m}{dx} = \frac{Z_{tx}}{Z_{tt} + U_{cc}\widetilde{w}^2} \tag{3.9}$$

The denominators of 3.8 and 3.9 are negative.[61] The numerator of 3.8 is ambiguous and depends on the relative magnitudes of U_c and $U_{cc}\widetilde{w}t_m$, since the former is positive and the latter is negative. The intuition behind this ambiguous prediction is offsetting income and substitution effects: the income effect of an increase in the wage leads the agent to demand more of all goods (including home produced goods), decreasing time on the market, while the substitution effect leads the agent to substitute away from home produced goods towards market purchased goods. As in the participation decision, the sign of 3.9 will depend on the sign of Z_{tx}, which is positive for leisure and negative for home production tasks (leading to an increase in hours for home production goods and a decrease in hours for leisure goods).

Next, I conduct an empirical investigation which considers separately the effects of Internet usage for each of the proposed mechanisms on labor supply. To do so, I will pool together various sources of data on each of mechanisms in question. The goal of these analyses will be to obtain a rich picture of how the Internet is used and how those uses contribute to affecting labor supply decisions, using the conceptual framework outlined above as a framework for

[61] Recall that $U_{cc} < 0$; $f_{tt} < 0$, and $V_{hh} < 0$; therefore $Z_{tt} = V_{hh}f_t + V_h f_{tt} < 0$.

the analyses.

3.4 Telework

Telework or "work from home," is growing in popularity and usage. A recent report by the Council of Economic Advisers finds that in 2008, 50 percent of employers allow some of their employees to work at home occasionally, and 23 percent allow some employees to work at home on a regular basis (2010). Moreover, the propensity to telework increased substantially. WorldatWork, which publishes statistics on telework behavior, finds that the propensity to work from home at least once per week increased 63 percent between 2002 and 2008 (WorldatWork, 2006, 2009). Moreover, that study also notes that teleworkers increasingly use broadband to work from home (WorldatWork, 2006). Oettinger (2011) finds that the wage penalty for working from home has decreased over time, which the author attributes in part to the rise of Information technology. Taking some liberty with that result would suggest that Internet usage may have made work from home easier and encouraged its take-up, both by employers and employees. To the best of my knowledge, however, there are no papers directly investigating the impact of Internet technology on telework take-up or telework patterns. In this section, I will investigate the link between Internet technology and telework usage.

3.4.1 Summary Statistics on Telework

In order to investigate the importance of telework in explaining the results, I begin by investigating data on telework trends across demographic groups, which can be found in the CPS Work Schedule Supplements from 2001 and 2004. Table 3.2 displays those data. Overall, about 24 percent of married women report working at least some hours in the home. Women have a slightly higher propensity to work from home than men, and they work slightly more days/hours in the home when they do so. The average woman who works from home does so approximately 3.5 days per week, 1.2 of which are spent working from home exclusively (i.e., not going into an office at all). Across groups, college educated women are much more likely to work from home than less educated women, although they do so less intensively. This appears to a product of the type of occupations each type of woman tends to hold. Column (6) of table 3.2 indicates that women with children are more likely than any other group to report working from home to "coordinate schedules with family or personal needs." This is direct evidence that working from home is a tool women use to balance the demands of work and family.

The information presented on telework patterns suggest that these may be tools women use to balance work and family needs, and therefore, this might be responsible for the increase in women's labor supply associated with Internet usage. An investigation into the effect of high speed Internet on work at home would be a more direct way of establishing this link, but unfortunately, the CPS Work Schedule data does not ask respondents about their own Internet usage so I cannot make this link using these data alone. As an alternative, I look across occupations in CPS Internet data and the CPS Work Schedule data to examine whether or not Internet users tend to have occupations where it is likely the individual teleworks. Figure 3.2 displays the correlation between occupation-specific work at home rates and occupation-specific Internet usage rates and indicates there is a strong, positive relationship.[62] Occupations in the top quartile of the distribution of mean Internet usage rates have work at home rates that are five times higher than those in the bottom quartile. A closer examination indicates that the top occupations in terms of work at home and Internet usage are computer scientists, lawyers and post-secondary teachers. This is suggestive evince that indeed, home Internet usage can facilitate telework.

[62] The CPS work schedule data is available in 2001 and 2004, while the CPS Internet data is available is 2000, 2001, 2003, 2007 and 2009. I construct averages using all available data for each occupation. Since occupation classifications changed dramatically between the 2000/2001 supplements and 2003-2009 supplements, I use the BLS CPS extracts to harmonize occupations over time `http://www.nber.org/data/cps_extract.html`.

Next, I investigate trends in telework to see if there have been any major changes in take up and/or types of telework arrangements in the period in which high speed Internet diffused. Figure 3.3 displays trends in telework between 2002 and 2010, which was compiled based on data presented in WorldatWork (2006, 2009, 2011). Figure 3.3 (a) displays trends in various types of telework defined by the type of contract, while Figure 3.3 (b) displays trends in types of telework defined by intensity of use. Both indicate that there have been large increases in telework, with telework "once per month" increasing from just over 20 million workers in 2002 to a peak of almost 35 million workers in 2008. Within the group of teleworkers, interesting trends also emerge in type of telework. Figure 3.3 (a) indicates "employee" telework, (working from home through an employer) has increased substantially while "contract/self-employed" telework (working from home on a for contract basis and/or through self-employment) has remained fairly stable. In terms of intensity, figure 3.3 (b) is suggestive that telework "once per year" and "once per week" have remained fairly stable while "every day" telework has declined. Telework "once per month" has grown the most substantially, suggesting that take up of occasional telework through an employer appears to be the major area of growth during the period in which high speed Internet was deployed. In the next section, I will more formally investigate the link between Internet usage, type of telework contract, and intensity of usage.

3.4.2 Self Employment and Full Time Telework

As described above, there are variations in the way in which an individual can engage in telework, from engaging full-time and exclusively working at home, to working at home just a few hours per month. Additionally, some individuals telework through an employer, while others do so on a contract basis or are self-employed. Figure 3.3 indicated that during the period in which high speed Internet diffused, occasional employer-provided telework was on the rise, and in this section I will more directly investigate this link. These distinctions are of particular interest for understanding the policy implications of the observed increase in labor supply associated with Internet usage. For example, if the observed change in labor supply can be attributed to an increase in employer-provided telework that would suggest an increase in the availability of those types of workplace policies could lead to further increases in female participation. On the other hand, if the increase in coming through increased rates of self-employment, that would suggest alternative policy levers could achieve those goals.

To formally investigate which types of telework arrangements are affected by Internet usage, I begin by investigating the link between individual high speed Internet usage and self-employment. Self-employment can be found in the CPS data that was summarized in this chapter, and used in the second chapter of this dissertation to study the link between home Internet usage and labor supply. To address the fact that Internet usage may be endogenous to the self-employment decision, I employ the instrumental variables strategy used in the second chapter of the dissertation. This strategy is motivated by the observation that diffusion of high speed Internet access was hampered by the presence of significant supply-side constraints. Namely, I exploit the fact that aspects of the underlying housing infrastructure affected the profitability of installation in particular areas, leading to different trends in access, and hence usage, of broadband Internet. To do so, I construct a measure of the fraction of the state population which resides in a multiple family dwelling – called a multiple dwelling unit (MDU) – and interact that measure with year fixed effects to capture the fact that this cross-sectional variation in the housing infrastructure created differential trends in access. The identification assumption is that the fraction of the population residing in a MDU would not have been systematically correlated with trends in labor supply in the absence of broadband Internet diffusion.

More formally, I estimate the following relationship between individual high speed Internet usage and self-employment using two stage least squares. The first stage is a linear probability model of the impact of the percent of the state residing in an MDU on HSI use:

$$HSI_{ist} = MDU_s\theta_t\gamma_1 + X_i\gamma_2 + S_{st}\gamma_3 + \theta_t + \eta_s + \nu_{ist} \tag{3.10}$$

HSI_{ist} is a dummy variable for whether or not individual i reports using the Internet in a household with broadband in state s and year t. MDU_s is the percent of the state's 2000 population which resides in a housing unit that is classified as an MDU, which was collected from the 2000 Census.[63] MDU_s is expected to affect home Internet use via the timing of Internet access, since places with a higher MDU_s are expected to receive Internet access earlier. Therefore, it is interacted with the vector of year fixed effects θ_t to allow the impact to vary separately by year. The instrument is the vector $MDU_s\theta_t$. The main effect of MDU_s only varies at the state level and is perfectly correlated with the state fixed effects, thus, it cannot be included in the model. The second stage is a linear probability model of HSI use on self-employment:

$$selfemployed_{ist} = \widehat{HSI}_{ist}\beta_1 + X_i\beta_2 + S_{st}\beta_3 + \theta_t + \eta_s + \epsilon_{ist} \tag{3.11}$$

Where $selfemployed_{ist}$ is self-reported self employment. The sample is limited to married, working women 18-59. The coefficient of interest is β_1, which measures the impact of HSI_{ist} on the propensity to be self-employed. X_i is a vector of individual controls and S_{st} is a vector of state-level controls, both of which are the same as described in detail in the second chapter of the dissertation. θ_t are year fixed effects , η_s are state fixed effects and ϵ_{ist} is the error term. Standard errors are clustered at the state-year level. Table 3.3 panel (a) present the results of that analysis. In columns (1)-(2) equations 3.10 and 3.11 are estimated for the entire sample of married women, so that the dependent variable is an indicator for whether an individual is self-employed or in any other employment/non-employment scenario. In columns (3)-(4), the sample is limited to employed individuals, so that dependent variable is instead an indicator for whether an individual is self-employed or working through

[63] Chapter 2 describes the data in more detail.

an employer. Columns (1) and (3) display the linear probability model (LPM) results, while columns (2) and (4) display the IV results. The baseline LPM results are indicative of a positive and statistically significant relationship between high speed Internet usage and self-employment in both samples, however, these are likely to be biased by reverse causality if self-employed individuals are more likely to take up home high speed Internet. The IV results on the other hand, are indicative of a negative, but statistically insignificant relationship. This suggests the naive LPM estimates are indeed biased upwards and that high speed Internet usage is not related to self employment in a statistically significant way.

In addition to differences in the type of contract, there are differences in the intensity with which individuals telework. This is also important in terms of policy, because if the results are explained by full-time telework, that is a very different workplace policy than allowing employees to work from home when necessary and engage in flexible scheduling. Unfortunately, as described above, information on telework behavior is not available in the CPS. Information on full-time work at home is, however, available in the American Communities Survey (ACS). This can be found in the survey question which asks "means of transportation to work", in which respondents may report "worked at home" as their means of transportation. The ACS does not, however, include information on Internet usage. Since geographic identifiers are available in both data sources over the same time frame, I can employ a two sample instrumental variables strategy (TSIV) (Angrist and Krueger, 1992). To do so, I use the CPS survey data to estimate the first stage relationship in equation 3.10 for the impact of MDU rates on home Internet use, and apply those parameter estimates to the ACS to estimate $\widehat{HSI}$ and the second stage relationship between home high-speed Internet use and full time telework.

Table 3.3 panel (b) display those results. In column (1) the first stage relationship for equation 3.10 is estimated. Column (2) displays the reduced form relationship between the instrument and full time work from home. This indicates that the instrument cannot predict full time work home. Column (3) displays the two sample IV estimate for the effect of $\widehat{HSI}$

on full time work from home. The standard error on this estimate was calculated using 500 bootstrap replications. The results indicate that home high speed Internet usage is associated with a statistically insignificant 2.5 percentage point decline in the probability of working from home full time. This is consistent with the summary data on trends in telework presented in figure 3.3, which indicated full time telework was falling slightly in the late 2000s. Overall, these analyses indicate that home high speed Internet usage does not increase full time telework or self employed telework, and trends in telework are consistent with high speed Internet usage increasing occasional telework through an employer.

3.5 Job Search

The idea that the Internet could have a profound effect on job search and the labor market was recognized early by economists (e.g., Autor, 2001). To date, however, the empirical evidence has been somewhat mixed on the effectiveness of Internet as a tool for job search. Kuhn and Skuterad (2004) study the effect of Internet search on unemployment durations using data from the CPS and find that Internet search *increases* unemployment durations. Subsequent work, however, has tended to find the opposite result (e.g.,Stevenson (2009), Kuhn and Mansour (2011)) or no relationship (e.g., Kroft and Pope (2010)). While previous literature has studied the relationship between Internet usage and unemployment durations, it has not looked at the effect of Internet usage for job search on labor force participation. In this section, I conduct such an analysis, using similar strategies to those used in the previous literature.

To study the effect of Internet usage for job search on transitions from non-participation to participation I will exploit the longitudinal nature of the CPS sampling frame to look at the post-survey labor market outcomes of respondents who were asked whether or not they used Internet for job search in the initial survey. In order to do so, I construct employment histories for the sample of married women who use Internet at home and do not participate in the labor force. Importantly, I need information on use of Internet for job search activities, which is included in the CPS supplements from 1998, 2000, 2001, and 2003. I no longer focus only on use of high speed Internet and look at all Internet users and compare those who used the Internet for job search to those who did not. Since the sample is limited to individuals with Internet at home, I do not need to be concerned about the endogeneity of Internet take-up to work patterns and I do not use the instrumental variables strategy used earlier. However, I do need to be concerned about take-up of job search practices in general, so I include as a control variable an indicator for whether or not the individual was "doing something to look for work in the past 4 weeks." I also include controls for the individual's age, race, ethnicity, age category, presence of children, year of the survey and state of residence. However, there is still a possibility that Internet take-up for job search is endogenous to work patterns and because I do not use an IV strategy, these results are best interpreted as descriptive and not causal.

In order to investigate the effect of Internet job search on labor force participation, I begin by estimating a linear probability model for the propensity for an individual to be a labor force participant one year after the initial survey according to the following specification:

$$LFP_{ist+12} = ijobsearch_{ist}\beta_1 + X_i\beta_2 + S_{st}\beta_3 + \theta_t + \eta_s + \epsilon_{ist} \tag{3.12}$$

Where LFP_{ist+12} is an indicator for whether or not an individual i in state s who is an Internet user in time period t was a participant in the labor force in time period *t+12*, or one year later. The coefficient of interest is β_1 which captures the effect of use of Internet for job search on future participation. Table 3.4 columns (1) and (2) display the results. Column (1) indicates that internet use for job search is associated with a positive and statistically significant increase in the probability of participating in the labor force. This relationship may be biased, however, if the indicator for Internet job search captures the effect of actively searching for work. Therefore, in column (2) I include as a control an indicator for whether or not an individual was "looking for work" in the past for 4 weeks. The effect of Internet use for job search is attenuated slightly but still indicates a positive and statistically significant 17.6 percentage point increase in the probability of participating. Not surprisingly, the coefficient on "looking for work" is also positive and statistically significant.

The preceding analysis indicates that Internet usage is associated with increased propensities to be a labor force participant in one year. However, this disregards information about how long the seeker searched. If home Internet eases transitions into participation, it may operate by allowing the individuals to become employed *sooner*. In order to incorporate all of the information available about length of time out of the labor force, I employ a duration model. I use the same method used by Kuhn and Skuterad (2004) in their analysis of Internet job search in the CPS, which addresses several unique features of the CPS sampling frame, including the fact that the CPS data is discrete, there are both left and right censored spells and eight month gaps in the data while respondents are out of the sample. The authors develop a discrete-time hazard model that accounts for those large windows but still uses a flexible baseline hazard. [64]

Table 3.4 columns (3)-(4) displays estimates of participation hazards, which include the same control variables described above. Again, column (3) omits the control for "looking for work" and column (4) includes it. The results indicate that, indeed, Internet job search is associated with a statistically significant increase in hazard rates (i.e., Internet search reduces non-participation durations). Thus, I conclude that Internet job search does in fact speed up transitions into participation among individuals who use Internet at home.

3.6 Home Production and Leisure

The final potential mechanisms through which home Internet usage may affect labor supply that I will consider is usage for home production and leisure. In this section, I

[64] See Kuhn and Skuterad (2004) for a detailed description of the model.

study the effect of Internet usage on these activities by conducting two separate analyses. First, I look at time use patterns for Internet users and non-users since Internet usage is expected to reduce/increase participation by changing the amount of time individuals spend in production of those goods (and hence, their valuation of time spent at home). Next, I use information from the CPS to look at the impact of Internet usage for leisure and home production on employment outcomes in the following year, using the same framework described above to study job search.

3.6.1 Evidence from Time Use Data

Internet usage for home production and leisure are expected to reduce/increase participation by changing the amount of time individuals spend in production of those goods. To discern whether or not those mechanisms are important I consider whether there are differences between home Internet users and non-users in how much time is spent in the production of those goods using the American Time Use Survey (ATUS), a survey which records time diaries of its respondents for 24 hours periods. The survey is administered to CPS respondents, so I am able to identify home Internet users in the ATUS data by linking it to the CPS Internet data.[65]

Table 3.5 summarizes time spent in work, home production, child care, and leisure among respondents in the linked ATUS and CPS data.[66] For each group, the top row is the difference between Internet users and non-users in the amount of time spent in each activity, while the bottom row is the group average hours per week spent in the activity.[67] The sample is

[65] ATUS data can downloaded from ATUS-X, which also provides the procedure for linking CPS and ATUS respondents. `http://www.atusdata.org/index.shtml` . The ATUS is a time-use survey which asks individual to record the number of minutes spent in various activities in a 24 hour period. I convert minutes per day into hours per week. Since there can be a a large time lag between when the supplement was administered and when the ATUS is administered, I limit the sample to only those individuals who respond to the ATUS within 6 months of the CPS supplement. This leaves a sample of approximately 2000 women.

[66] Child care is tabulated separately from home production as advised by Guryan, Hurst, and Kearney (2007).

[67] Data are collected in terms of minutes per day and converted into hours per week.

limited to adults 18-59 who are married. Note that these differences are merely correlations, although in panel (b) the sample is limited to individuals who are employed full time, so those results are conditional on differences across groups in the propensity to work.

The data suggest there are differences in time spent in each type of activity between Internet users and non-users. For women overall, Internet use is correlated with 1.88 *fewer* hours per week spent in home production and 1.5 *more* hours per week spent in market work. These estimates, however, are misleading due to differential rates of participation across demographic groups and between users and non-users. Panel (b) presents correlations conditional on employment, which indicate that employed women spend 0.63 *more* hours per in market work, 1.3 *fewer* hours per week in home production, and 0.5 *more* hours in leisure. In terms of heterogeneity, these gaps are largest for college educated women and women with children, while the relationships have the opposite sign for less educated women or men. This is consistent with the labor supply effects found in the second chapter of this dissertation, which indicated that Internet does not affect the labor supply decisions of less educated women or men, but has large effects on women with children and college educated women. Interestingly, I also find that Internet users spend more time in child care than non-users. This suggests some time saved in home production tasks or commuting may be used to spend time with children. I leave it to future work, however, to determine the impact of Internet use on children's well-being.

3.6.2 Evidence from the Current Population Survey

The preceding analysis suggested that Internet usage is associated with reductions in time spent in home production and increases in time spent in leisure for the groups whose labor supply was affected by Internet usage. These gaps, however, are small relative to the average amount of time an individual spends working and the approximately 8 hour change in hours worked per week induced by Internet usage that was found in the second chapter of this dissertation. Therefore, in this section I will directly estimate the effect of Internet usage for home production and leisure on participation to see if these factors can indeed explain the effects found. To do so, I conduct a similar exercise to the one used in section 3.5 to study the effect of Internet job search on participation.

Table 3.6 displays the results using the CPS data. In columns (1)-(2) I display the results of the LPM analysis, where the dependent variable is the probability an individual works in period $t+12$ if she used the Internet for each activity in period t (and was not in the labor force in period t). This indicates that women who use the Internet for home production are approximately 2.5 percentage points (with a standard error of 1.7 percentage points) *more* likely to both be a participant, while those who use Internet for leisure are approximately 1.11 percentage points (with a standard error of 1.8 percentage points) *less* likely to be a participant. While this is consistent with the time use data and predictions of the model, neither of these results is statistically significant at conventional levels. In columns (3) and (4) I present estimates from the duration model presented in section 3.5. Unlike Internet job search, it is not clear *ex ante* that use of Internet for home production should necessarily operate by reducing non-participation durations since there is no clear mechanism via which that might occur. Nonetheless, I conduct the analysis which indicates that use of the Internet for home production and leisure both increase non-participation durations, although neither of those results are statistically significant at conventional levels. Overall, these analysis suggest that Internet usage can save time in home production and may increase time spent in leisure, but neither is able to significantly explain changes in participation.

3.7 Decomposing the Net Effect

The evidence presented above has indicated that all of the potential mechanisms outlined in the conceptual framework have some explanatory power in explaining the estimated effects. Internet usage for occasional part-time telework has trended in a way that is similar to the diffusion of high speed Internet. Internet use for job search has been shown to increase the propensity for an individual join the labor force in a year's time and speed up transitions from non-participation to participation. Internet users do spend less time in home production and more time leisure, but Internet usage for those tasks cannot significantly predict changes in participation. In what follows, I will attempt to disentangle how much of the estimated effects on labor supply found in the second chapter of this dissertation can be attributed to each mechanism.

In order to decompose the extent to which each mechanism can explain the observed change in participation, information on the effects of each must be combined with information on usage rates for each activity. For example, even if Internet job search significantly predicts changes in employment, if only a small fraction of users use Internet for that activity than it cannot plausibly explain the bulk of the aggregate change in participation induced by high speed Internet usage. Therefore, I revisit the summary data present in table 3.1, which I combine with the point estimates from table 3.4 and 3.6 to construct "back of the envelope" estimates of the magnitude of each explanation in explaining changes in participation estimated in the second chapter of the dissertation. I take the point estimates in tables 3.4 and 3.6 and multiply each by the information in table 3.1 on the fraction of female Internet users who have used Internet for each activity. This exercise indicates that Internet job search can explain about 2.6 percentage points, or 14 percent of the estimated effect of Internet use on participation. Similarly, Internet use for home production can explain around 2 percentage points of the estimated effect, or 11 percent of the estimated increase in participation. In sum, use of Internet for search and home production can explain only 25 percent of the estimated positive effect Internet usage on participation combined.

Gauging the relative contribution of telework is less straight forward than it is for home production and search. Ideally, I'd use the same strategy to look at Internet use for work as I did above to look at job search and home production. Unfortunately, I cannot look at use of the Internet for work using this framework because individuals in the initial survey who do not work would not report using the Internet for work in that period. As an alternative, I use the labor supply results from the second chapter and compare them to changes in telework. I do so to gauge whether the observed aggregate changes in participation are plausibly related to changes in telework. If I make the strong assumption that all of the increase in labor force participation induced by Internet usage led to telework take-up, then diffusion of home broadband technology (via the effect on women's labor force participation) can explain 27 percent of the rise in telework.[68] Since Internet use for search and home production can only explain a combined 25 percent of the estimated effects, and telework take up changed so dramatically over this period, I conclude that Internet use for telework is leading explanation for the change in women's labor force participation induced by the diffusion of high speed Internet technology.

3.8 Conclusion

High speed Internet has affected many aspect of daily life, from facilitating telework to saving time shopping and paying bills. In this chapter, I provide theoretical and empirical evidence for the role of Internet in various aspects of daily life, and how those changes are related to the net effect of Internet usage on female labor supply. By investigating data on Internet usage for various activities, telework patterns, employment histories, and time use I find that use of Internet to engage in occasional telework through an employer is the primary explanation for the demonstrated increase in participation associated with high speed Internet use, although improved job search and time saved in home production also play a more modest role in explaining the estimated effects.

The role of telework in affecting women's labor supply has important public policy implications. While it is generally accepted that flexibility in the workplace has the potential to benefit employers, employees and the economy as a whole, adoption is still low and there

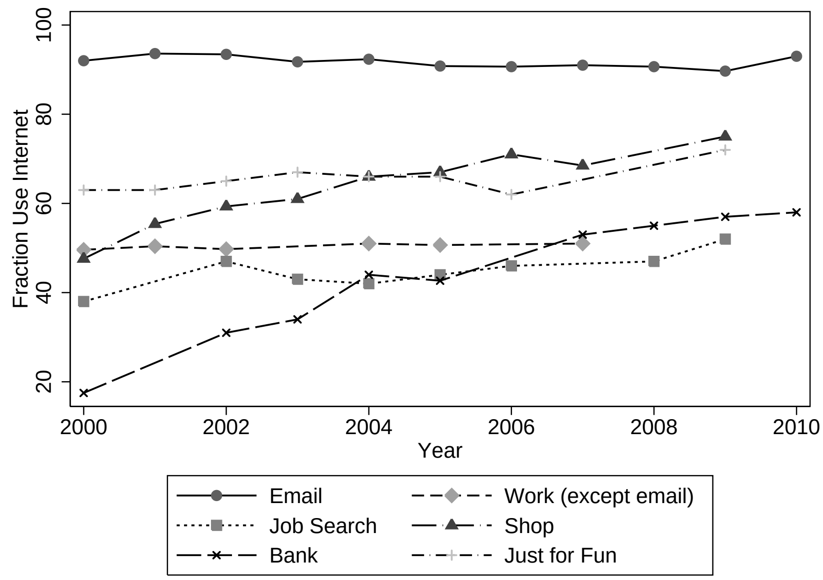
[68] Telework increased by 13 million 02-08. The results from chapter two imply 3.56 million more workers would join the workforce.

is little empirical evidence on the benefits/costs of these policies.[69] This chapter has demonstrated that Internet usage, via take-up of telework opportunities, has allowed a group of highly educated women to join the workforce, suggesting such policies may have the potential to encourage workforce entry by productive individuals.

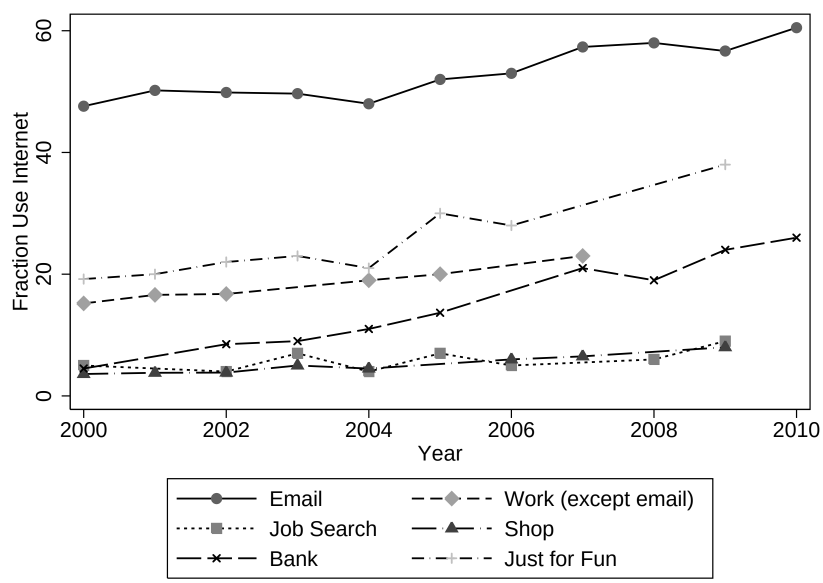
[69]See, for example, the report by the Council of Economic Advisers on Work-life Balance, March 2010 (Council of Economic Advisers, 2010).

Figure 3.1: Internet Usage Trends

(a) “Ever Use” the Internet for an Activity

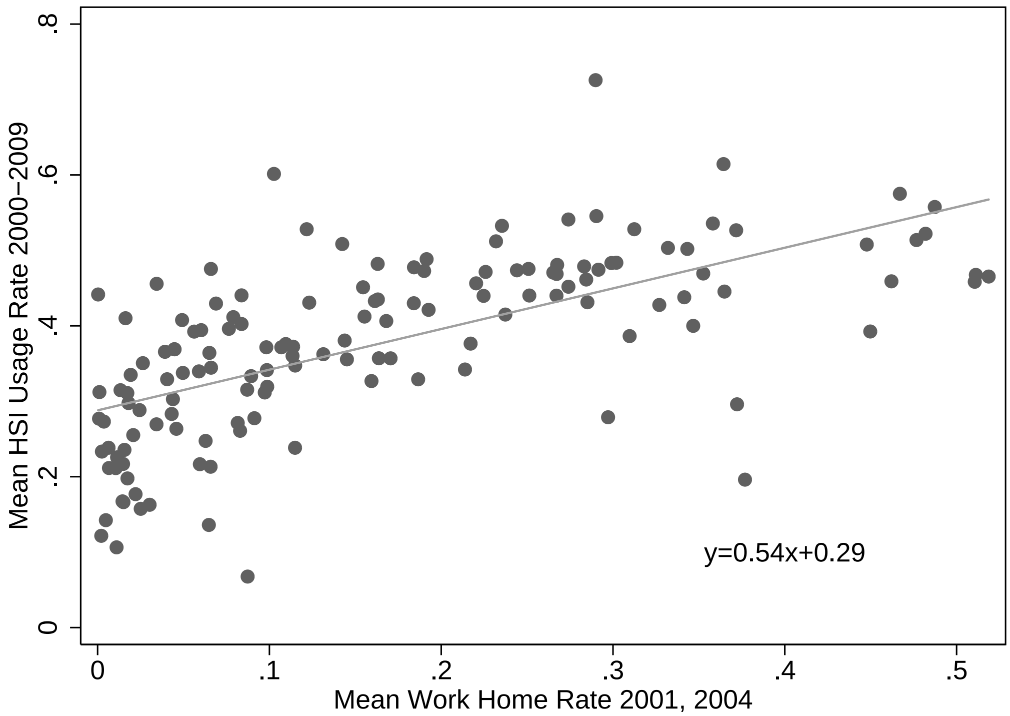


(b) Use the Internet “Yesterday” for an Activity



Notes: Displayed are trends in Internet usage for various activities. Data source is PEW Internet and American Life Project, “Usage Over Time” trend data (2012).

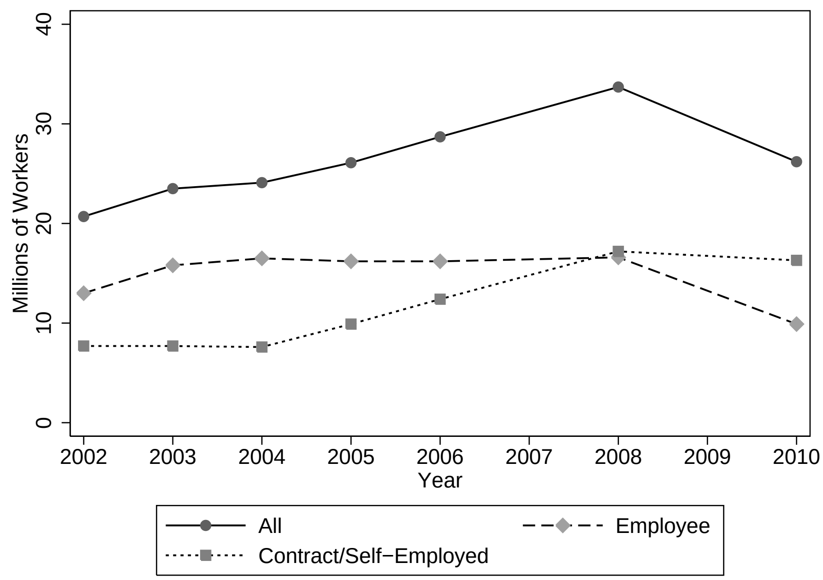
Figure 3.2: Occupation Mean Internet Usage and Work at Home Rates



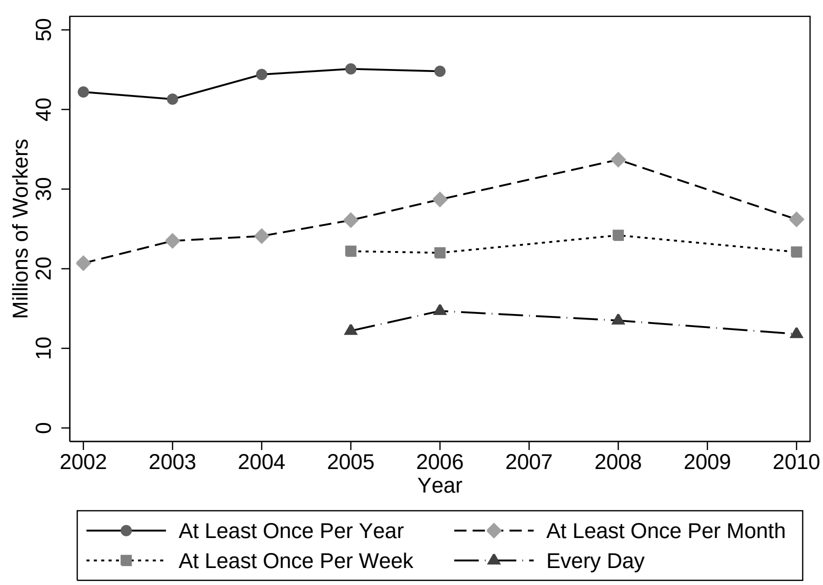
Notes: Plotted are occupation-specific work at home rates and occupation-specific high-speed Internet usage rates. Only occupations with at least 100 observations are shown and rates are calculated for working, married women 18-59. Internet usage rates were calculated in the 2000, 2001, 2003, 2007 and 2009 Current Population Survey supplements used in the main analysis and work at home rates were calculated in the 2001 and 2004 Current Population Survey work schedule supplements.

Figure 3.3: Trends in Telework 2002-2010

(a) Type of Contract



(b) Intensity of Telework



Notes: Displayed are trends in different types of telework. Data source is WorldatWork (2006, 2009) (2011).

Table 3.1: Activities Conducted Online: CPS 2000-2003

		Home Production:		Leisure:				
	All	All	Shop/Bills	All	Recreation	Check NWS	Work	Job Search
(a) All								
Married Women:	0.356	0.813	0.546	0.749	0.290	0.574	0.288	0.149
N=46690	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Less than HS	0.106	0.697	0.336	0.691	0.358	0.477	0.107	0.118
N=1338	(0.008)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.008)	(0.009)
HS/Some Coll	0.319	0.787	0.500	0.729	0.309	0.542	0.208	0.139
N=27262	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
College	0.492	0.860	0.632	0.783	0.256	0.628	0.422	0.167
N=18090	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
No Children	0.341	0.805	0.535	0.738	0.277	0.573	0.298	0.150
N=19998	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Children	0.369	0.818	0.555	0.757	0.300	0.574	0.282	0.149
N=26692	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Single Women	0.309	0.733	0.449	0.724	0.354	0.538	0.204	0.232
N=30383	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Married Men	0.345	0.806	0.532	0.814	0.338	0.703	0.373	0.172
N=41728	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)

Notes: Displayed are means and standard errors of for various reported uses of the Internet. The sample is limited to individuals 18-59. In columns (2)-(7) sample is only individuals who are reported to use the Internet at home. Source is 2000, 2001 and 2003 Current Population Survey Supplements.

Table 3.2: Summary Statistics on Telework

	Work Home	Work Home Once Per Week+	Days Work Home	Days Home Only	Hours Worked at Home	Work Home for Family Reasons
Married Women:						
All	0.239	0.811	3.480	1.231	12.998	0.069
(N=4871)	(0.006)	(0.006)	(0.029)	(0.030)	(0.231)	(0.004)
Less than HS	0.064	0.892	4.769	3.566	32.351	0.049
(N=82)	(0.027)	(0.035)	(0.189)	(0.283)	(2.390)	(0.024)
HS/Some College	0.146	0.758	3.585	2.006	17.415	0.094
(N=1780)	(0.008)	(0.010)	(0.051)	(0.059)	(0.469)	(0.007)
College	0.436	0.841	3.388	0.756	9.886	0.056
(N=3009)	(0.009)	(0.007)	(0.035)	(0.029)	(0.219)	(0.004)
No Children	0.233	0.817	3.592	1.212	12.634	0.033
(N=2353)	(0.009)	(0.008)	(0.043)	(0.043)	(0.322)	(0.004)
Children	0.247	0.791	3.241	1.351	13.846	0.144
(N=929)	(0.014)	(0.013)	(0.064)	(0.070)	(0.559)	(0.012)
Single Women	0.164	0.763	3.282	1.010	10.861	0.045
(N=2577)	(0.007)	(0.008)	(0.040)	(0.037)	(0.282)	(0.004)
Married Men	0.228	0.812	3.268	0.898	10.954	0.043
(N=7110)	(0.005)	(0.005)	(0.025)	(0.022)	(0.170)	(0.002)

Notes: Displayed are means and standard errors of various measures of working from home. The first column includes all full time working adults 18-59. The rest of the columns include only those individuals who report working from home. N also refers to number of individuals in cell who work from home. Source is 2001 and 2004 Current Population Survey.

Table 3.3: Home Internet and Telework Patterns

(a) Home Internet Use and Self-Employment: IV Using CPS

	All		Employed	
	OLS	IV	OLS	IV
HSI Use	0.0148***	-0.0525	0.0166***	-0.0987
	(0.00226)	(0.0433)	(0.00306)	(0.0748)
Mean Dep. Var.	0.074	0.074	0.102	0.102
F Statistic		27.55		21.15
N	107976	107976	76066	76066

(b) Home Internet Use and Full Time Telework: TSIV Using CPS and ACS

	First Stage : CPS	Reduced Form: ACS	TSIV
	HSI Use	Telework FT	Telework FT
HSI Use			-0.0251
			(0.0179)
MDU*2000	0.587***	-0.0124	
	(0.0833)	(0.0208)	
MDU*2001	0.600***	-0.0145	
	(0.0796)	(0.0135)	
MDU*2003	0.638***	-0.0150	
	(0.0831)	(0.0139)	
MDU*2007	0.251***	-0.0149	
	(0.0836)	(0.0102)	
Mean Dep. Var.	0.378	0.051	0.051
F Statistic	21.15		
N	76066	962005	962005

Notes: Panel (a) displays results when the dependent variable is self-employment status. In columns (1) and (2) equation 3.10 is estimated for the whole sample, and in columns (3) and (4) for employed individuals only. The instrument described in more detail in the text. Panel (b) displays results from the two sample IV analysis, where column (1) shows the first stage results estimating equation 3.10 in the CPS, column (2) displays the reduced form results estimating equation 3.10 (except the outcome variable is full time telework in the ACS and column (3) displays the two sample IV estimate of the effect of HSI Use on full time telework. The standard error on the two sample IV estimate was calculated using 500 bootstrap replications. Demographic variables included in all specifications include fixed effects for age category, race, number and ages of children, living in an MSA, and living in a central city. state-level variables include average wages, income per capita, unemployment rates, housing prices, population density, percent of state employees' industry internet adoption and enhancement are matched at the state-year level. %MDU refers to the percent of state living in a multiple dwelling unit and is matched at the state-level. The left out category is % MDU*2009. Standards errors clustered by state-year are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.4: Internet Usage for Job Search and Participation

	LPM: LFP One year later		Duration Model: Participation Hazard	
	(1)	(2)	(3)	(4)
Internet Job Search	0.192***	0.176***	0.489***	0.435***
	(0.0218)	(0.0219)	(0.0640)	(0.0646)
Looking For Work		0.267***		0.977***
		(0.0528)		(0.188)
N	4292	4292	4111	4111

Notes: Columns (1)-(3) display the results for estimation of a linear probability model for participation in the labor force 12 months after the initial survey. Columns (4)-(6) display the results estimating the duration model described in the text, where the dependent variable is the participation hazard for the year following the initial survey. Includes controls for age, race, education, number and age of children, state of residence, and year of survey. The estimates for Internet job search also include a control for participation in active job search. The sample is limited to the married women 18-59 who use the Internet at home and were not participating in the labor force in the initial survey. Standards errors are in parentheses. Data source is Current Population Survey 1998, 2000, 2001 and 2003. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 3.5: Means and Differences Between Users and Non-Users In Hours Spent per Week Spent In Different Time Use Categories

		Home Production	Child Care	Work	Leisure
(a) All					
Married Women	Difference	-1.88	2.87	1.54	-0.30
	Mean	25.81	9.51	19.12	30.53
Less than High School	Difference	1.14	3.37	-5.14	1.07
	Mean	31.79	6.01	13.62	32.20
High School/Some College	Difference	-0.88	1.02	0.53	0.58
	Mean	26.33	7.69	18.49	32.24
College	Difference	-0.69	2.06	0.60	1.20
	Mean	23.98	12.64	20.99	27.87
No Children	Difference	-1.55	0.88	6.84	-2.81
	Mean	24.70	1.08	21.41	35.20
Children	Difference	-2.33	1.75	-0.24	1.94
	Mean	26.34	13.55	18.02	28.28
Single Women	Difference	-1.94	-0.34	6.99	-1.40
	Mean	19.17	4.91	24.47	34.54
Married Men	Difference	1.06	0.76	1.62	-2.24
	Mean	16.15	5.16	33.50	33.88
(b) Employed					
Married Women	Difference	-1.30	2.55	0.63	0.05
	Mean	24.34	7.93	25.44	28.44
Less than High School	Difference	6.08	3.57	-11.71	6.11
	Mean	30.07	5.16	24.33	27.49
High School/Some College	Difference	-0.12	1.02	-0.91	0.22
	Mean	25.12	6.46	24.63	29.93
College	Difference	-1.13	2.48	2.08	0.83
	Mean	22.69	10.12	26.61	26.66
No Children	Difference	-1.28	1.32	4.72	-3.11
	Mean	23.42	1.07	27.23	31.90
Children	Difference	-1.53	1.74	-1.02	2.36
	Mean	24.81	11.44	24.53	26.68
Single Women	Difference	-2.45	-0.77	5.18	0.42
	Mean	18.66	4.48	29.34	32.77
Married Men	Difference	1.06	0.76	1.62	-2.24
	Mean	16.03	5.21	35.44	32.47

Notes: The top row displays the difference between high speed Internet users and non-users in hours per week spent in each activity. The bottom row displays the group mean hours per week spent in each activity. Sample includes adults 18-59, and in the bottem panel only those who are currently employed full time. Data source is American Time Use Survey and CPS 2000, 2001 and 2003.

Table 3.6: Internet Usage for Home Production or Leisure and Participation

	LPM: LFP One year later		Duration Model: Participation Hazard	
	(1)	(2)	(3)	(4)
Internet for Home Production	0.0248		-0.0213	
	(0.0172)		(0.0493)	
Internet for Leisure		-0.0112		-0.0789
		(0.0172)		(0.0493)
N	4292	4292	4111	4111

Notes: Columns (1)-(2) display the results for estimation of a linear probability model for participation in the labor force 12 months after the initial survey. Columns (3)-(4) display the results estimating the duration model described in the text, where the dependent variable is the participation hazard for the year following the initial survey. Includes controls for age, race, education, number and age of children, state of residence, and year of survey. The estimates for The sample is limited to the married women 18-59 who use the Internet at home and were not participating in the labor force in the initial survey. Standards errors are in parentheses. Data source is Current Population Survey 1998, 2000, 2001 and 2003. * $p < .1$, ** $p < .05$, *** $p < .01$

Chapter 4

House Prices and Birth Rates: The Impact of the Real Estate Market on the Decision to Have a Baby

4.1 Introduction

This project investigates how changes in Metropolitan Statistical Area (MSA)-level house prices affect household fertility decisions. The conceptual approach is based on an economic model of fertility that recognizes that changes in house prices potentially have offsetting effects on fertility. Assuming that children are normal goods, and recognizing that housing is a major cost associated with (additional) children, an increase in the price of housing will have a negative substitution effect on the demand for children in the current period, ceteris paribus. This is true for both potential first-time homeowners (i.e., current renters who would buy a house with the addition of a child) and current homeowners who might buy a larger house with the addition of a child. On the other hand, for a homeowner, an increase in MSA-level house prices increases home equity. This could lead to an increase in birth rates among homeowners through two channels – a traditional wealth effect and/or an equity extraction effect. In either case, when home prices increase, homeowners might use some of their new housing equity to fund their childbearing goals. The net effect of house prices on aggregate birth rates will depend on individual's responsiveness along these margins and rates of home ownership.

We are interested in identifying the causal relationship between movements in local area house prices and current period fertility rates. Conceptually, we are examining how short-term fluctuations in house prices affect current period fertility rates, separately for owners and non-owners, all else equal. Our main analyses focus on the housing price cycle of 1997 to 2006, a period of general housing price growth. We additionally separately consider the adjacent housing market cycles characterized by falling house prices. We begin our empirical investigation with a set of ordinary least square (OLS) regressions of MSA-demographic group level fertility rates on MSA level house prices interacted with a baseline measure of MSA-group level home ownership rates, controlling for conditional variable main effects, time-varying MSA conditions, and MSA fixed effects. Fertility rates are constructed from vital statistics natality files and groups are defined by age and race/ethnicity. Our main source of house price data is the Federal Housing Finance Agency House Price Index (FHFA - HPI), which we use in combination with 2000 MSA-level median home values to generate house price levels over time. Our regression specification controls for MSA fixed effects so that the estimated relationship between house prices and birth rates is not confounded by time-invariant differences in preferences for children across MSAs. If couples with lower preferences for children sort into areas with higher costs of living – driven by other amenities – there will be a negative correlation between house prices and fertility. Our estimated relationship of interest will be net of any such sorting patterns.

The key to our empirical approach is to separately identify effects for owners and non-owners. As a preview of our findings, the data suggests that as the non-owning share of an MSA increases, rising house prices exert a negative effect on current period fertility rates; as the owning share increases, rising house prices exert a net positive effect on current period fertility rates. Any alternative explanation (other than a causal effect of house prices) for these observed relationships must differentially impact owners and non-owners. The underlying assumption of our empirical strategy is that conditional on all of the controls in our regression model, there are not omitted factors that drive both local area house prices and affect the fertility behavior of owners and renters differently.

To address the possibility that other local factors are biasing our OLS estimates we additionally implement an instrumental variables (IV) strategy that exploits exogenous variation in house price movements induced by variation across MSAs in their housing supply elasticity. We predict annual MSA-level house prices using the interaction between the national version of the FHFA house price index with the cross-sectional MSA housing supply elasticity, as measured by Saiz (2012). This instrument was first proposed by Chetty and Szeidl (2012) who define it at the state level in their analysis of the relationship between housing equity and debt and portfolio choice. The conceptual motivation for this instrumental variables strategy is that it isolates movements in local-area housing prices coming from the local area response to national-level aggregate demand shocks, as determined by the degree of local-area housing supply elasticity. The intuition is that MSA-level house prices will tend to covary more with national prices in places where the housing supply is more constrained.

Both our OLS and IV results indicate that as the proportion of individuals in a demographic cell who are home owners increases, an increase in house prices is conditionally associated with an increase in current period fertility rates. This is consistent with a positive "home equity effect" that dominates any negative price effect. The data also indicate that as the proportion of homeowners approaches zero, an increase in MSA-level house prices leads to a decrease in current period fertility rates, which is consistent with a negative price effect among non-owners. In general, the main results hold across race/ethnic groups and are equally driven by first, second, and higher-parity births.

These main results are statistically significant and economically meaningful. Employing our regression estimates in a straightforward simulation exercise, we find that a \$10,000 increase in home prices is associated with a 5 percent increase in fertility rates in MSA cells with 100 percent ownership rates. For MSA cells with zero percent home ownership rates, we estimate a corresponding decrease in fertility rates of 2.4 percent. For an MSA-group, as the home ownership rates increase from 30 to 40 percent, the net effect of a \$10,000 increase in house prices becomes positive. Under the assumption of linear effects, these estimates suggest that all else held constant, the roughly \$108,000 average increase in house prices during the housing boom of 1997 to 2006 would have led to a 9 percent increase in births over that time.[70]

We implement a number of robustness checks on the model specification and sample

[70] The population weighted average home price change for the 154 MSAs in our sample from 1997 to 2006 was \$108,038

construction. We also examine longer averages of house prices; compare the effects across different demographic subgroups; and examine, first, second, and higher parity births. We also turn to individual-level data from the Current Population Survey (CPS) to confirm that the pattern of effects we see in the aggregate data is found at the individual-level in the ways expected. In addition, we estimate our model on data from two housing bust periods, to see how the estimated relationships compare to those estimated during the 1997-2006 housing boom period. And finally, we tabulate data from the American Housing Survey (AHS) to see if home equity extraction - via mortgage refinancing or home equity loans/lines of credit - is a viable mechanism contributing to the positive effect of house price increases on fertility for home owners.

The main contribution of the chapter is to provide an empirical examination of how aggregate movements in house prices affect aggregate level birth rates. First, as an issue of economic demography, it is informative to understand how movements in the real estate market affect current period birth rates, overall and for various demographic subgroups. Second, within the research literature on the nature of the demand for children, an examination of the effect of house prices on the fertility outcomes of homeowners constitutes a useful test of wealth effects. Third, our work highlights the importance of including housing markets in any model of how economic conditions affect fertility outcomes. In fact, as an empirical matter, we find that changes in house prices exert a larger effect on current period birth rates than do changes in unemployment rates. Fourth, our results potentially speak to the role of credit constraints, and imperfect capital markets, in affecting the timing of fertility decisions. This is an issue that features prominently in the literature on the cyclicality of fertility timing, as reviewed in Hotz, Klerman, and Willis (1997). Our finding of a positive effect among home owners suggests that some individuals may consume out of home equity to fund their childbearing goals. And finally, there is a literature, described below, on the tendency of individuals to consume out of housing wealth. To our knowledge, that literature has not previously considered children as a potential "consumption" good in this regard.

Our results provide clear empirical support for the idea that house prices impact birth rates in a statistically significant and economically meaningful way.

4.2 Conceptual Framework and Related Literature

There is a large literature in neoclassical economics investigating the nature and determinants of fertility in developed countries. In the most simple static approach to this question, parents are viewed as consumers who choose the quantity of children that maximizes their lifetime utility subject to the price of children and the budget constraint that they face. Children are conventionally thought to be normal goods, but an empirical puzzle presents itself in both time series and cross-sectional data, which tend to show a negative correlation between income and number of children.

There are two leading explanations for this observed correlation that maintain the basic premise of children as normal goods: (1) the quantity/quality trade-off (Becker, 1960) and (2) the cost of time hypothesis (Mincer (1963); Becker (1965)). The first refers to the observation that parents have preferences for both the quantity and quality of children. If the income elasticity of demand for quality exceeds the income elasticity of demand for number of children, then as income rises, parents will substitute away from the number of children, toward quality per child. The second hypothesis attributes the observed negative relationship between income and fertility to the higher cost of female time experienced by higher income families, either because of increased female wage rates or because higher household income raises the value of female time in non-market activities. There is a long and active literature that attempts to estimate the effect of changes in family income and of own-prices on fertility.[71]

[71]The key empirical challenge in this literature is to find variation that is exogenous to women's (or couple's) preferences and that alter price or income without affecting the opportunity cost of women's time. Many of these papers are reduced-form in nature, and include examinations, for example, of the effect of direct pro-natalist government payments (e.g., Milligan (2005); Cohen, Dehejia, and Romanov (2007)) and of exogenous changes in income (Lindo (2010); Black et al. (2011)).

There exists a closely related literature investigating the cyclicality of fertility, which is a literature about fertility timing (e.g., Galbraith and Thomas (1941); Becker (1960); Silver (1965); Ben-Porath (1973)). Changes in the unemployment rate are typically thought to affect the wages of women and their husbands. Under the standard assumption that women bear the primary responsibility for child rearing, it becomes optimal for woman to select into childbearing at times when their opportunity cost is lowest, that is, when economic conditions are least favorable. Another consideration affecting optimal timing with regard to unemployment rates is skill depreciation (Happel, Hill, and Low, 1984).

In a world with imperfect capital markets and credit constraints, women might not be able to optimally time fertility with regard to opportunity cost and skill depreciation considerations. In particular, though some women might optimally choose to select into childbearing during economic downturns, they might not be able to afford to do this, if husbands' income is also negatively affected. Schaller (2011) provides a recent examination of this issue and explicitly considers the role of gender-specific labor market conditions. Her results confirm previous empirical findings that increases in overall unemployment rates are associated with decreases in birth rates. In other words, her empirical work confirms that births are pro-cyclical. In support of the predictions of Becker's time cost model, she further finds that improved labor market conditions for men are associated with increases in fertility, while improved labor market conditions for women have the opposite-signed effect.[72]

In many respects, the context of real estate markets is more straightforward to consider conceptually because changes in house prices do not affect the cost of parental time. Our conceptual framework is thus not encumbered by considerations of skill depreciation or opportunity cost of time. We motivate our empirical model and interpret our estimated effects simply in terms of housing costs (which affect the price of childbearing) and housing income

[72] Dehejia and Lleras-Muney (2004) suggest that relatively more white women opt into childbearing during economic downturns than black women; they attribute this difference to credit constraints facing blacks. Neither Schaller (2011) nor we find evidence in the data consistent with this idea. In particular, we find a statistically significant negative relationship between unemployment rates and birth rates among whites and a statistically insignificant relationship among blacks.

effects (which affect ability to consume in the current period).[73] Our focus on current period prices and contemporaneous fertility allows us to look separately for price and "income" effects. Changes in the real estate market are expected to generate price effects because housing costs are estimated as the greatest portion of the annual cost of raising a child: greater than food, child care, or education (Lino, 2007).

We qualify the term "income" because an increase in house prices does not necessarily imply increased wealth or income for home owners. If price increases are viewed to be permanent and homeowners see their home as a store of wealth, an increase in house prices can be thought of as an increase in (perceived) wealth for existing homeowners. This could lead to an increase in the demand for children in the current period, as well as in a completed lifetime setting. But, if homeowners do not intend to "cash out" and move to a lower-priced real estate market during their lifetime, or if they view the increase in house prices as transitory and expect it to be undone at a later period, there is no change in actual wealth or permanent income. However, if homeowners are otherwise credit constrained but can liquefy increases in home equity, there can be an increase in current period accessible income and this could lead to an increase in current period birth rates. To the extent that equity extraction is driving our results, our work potentially speaks to the role that credit constraints play in affecting the timing of childbearing. For the sake of convenience of exposition, we refer to this general class of explanations as a "home equity effect".

One could reasonably argue that in contrast to unemployment rates – which are generally understood to be cyclical – movements in the housing market over the period we analyze were likely to have been perceived at least in part as permanent. This would follow from the observation that the national trend in housing prices between 1997 and 2006 was steadily increasing. This suggests our results may be indicative of a change in completed fertility, as

[73]There exists a class of dynamic or life-cycle models of fertility decisions, which recognize that changes in prices and income over the life cycle may result in changes in the timing of childbearing, even if they do not cause completed lifetime fertility to change. The Handbook chapter by Hotz et al. (1997) provides an overview of these theoretical models. Heckman and Walker (1990) provides an empirical examination of the effect of income and wages on life-cycle fertility using data from Sweden.

opposed to simply a story about timing or cyclicality. We give a cursory treatment of this possibility in our empirical analyses below - in particular by looking at higher-order births - but we leave it to future research to thoroughly examine this possibility.

Finally, we acknowledge that we talk about fertility throughout the chapter as though it is a simple decision. Of course, fertility is a stochastic outcome, albeit one that is to a large extent controllable by individual's actions with regard to sexual activity, contraceptive use, fertility treatments, and abortion. We recognize, however, that latent demand for fertility timing will not be perfectly realized. Thus, any response we see of fertility to house prices will be a muted reflection of a couple's desired fertility response.

4.3 Data and Empirical Approach

The main empirical approach of this chapter is to empirically relate MSA level fertility rates to demeaned MSA-level house prices, interacting house prices with a baseline measure of group-level home ownership rates and controlling for time-varying MSA level characteristics. The three main data requirements are (1) MSA-level fertility rates, (2) MSA-level house prices, and (3) group-level home ownership rates. In this section we describe our main data sources and briefly describe how we construct the relevant variables. Table 4.1 provides details on explanatory variables and associated data sources.

4.3.1 Data

Data on births come from the Vital Statistics Natality Files, years 1990 to 2007. Vital statistics data contain birth certificate information for virtually every live birth that takes place in the United States. Vital statistics data identifies the race/ethnicity, marital status, age, and education of the mother, as well as some limited information about the baby's pregnancy conditions, and the baby's health status at time of birth. (These data do not include information about home ownership status of the parent.) For the purposes of matching births to our explanatory variables, we create a file of conceptions for the years 1990 to 2006, using information on the date of birth and length of gestation to identify year of conception. We do this because in terms of the decision-making process, the most relevant decision is the decision to get pregnant in a given time period. It is thus the economic conditions that exist at the conception decision point that are relevant, as opposed to the economic conditions in place at the time when the birth actually occurs (typically 40 weeks later.) To be precise, our analysis sample is a sample of conceptions that result in live births in year t.

We construct MSA-year-group level fertility rates by aggregating births and female population counts to the MSA-year-group cell, where groups are defined by the interaction of race/ethnicity and age category. We define three mutually-exclusive race/ethnic groups: Non-Hispanic White, Non-Hispanic Black, and Hispanic. We exclude other race/ethnicities from the analysis. We define two age categories, 20-29 and 30-44. We obtained annual female population counts (by age, race, ethnicity, and county) from the National Center for Health Statistics (National Center for Health Statistics, 2003 2010). We use these data to construct MSA-group-level fertility rates, defined as the total number of births to women in the MSA-year-group cell divided by the MSA-year-group population. We obtained access to confidential natality files that identify the mother's state and county of residence. We use the county-level identifiers in the confidential Vital Statistics natality files to construct MSA level fertility rates, using the MSA definitions that are used in the federal housing data sets: 5-digit MSAs and Divisions as defined by the Office of Management and Budget in December 2009 (Bulletin 10-02).

We identify a total of 384 MSAs in the birth records. We restrict our sample to MSAs that have at least five births in every year-group cell, which leaves us with a sample of 222 MSAs.[74] When we further restrict the sample to those MSAs for which all explanatory

[74]Other empirical papers that have used aggregate level MSA data have used the following rules: Blau, Kahn, and Waldfogel (2000) look at MSA level marriage rates and MSA level indicators of labor and marriage market conditions. They use a rule of 20 observations per race-education group. Blau et al. (2004) look at MSA level single motherhood and headship rates and welfare benefits. They use a rule of 10 observations

variables used in the baseline specification are available, we are left with a sample of 154 MSAs.[75]

The main data source used to construct MSA-level house prices is the Federal Housing Finance Agency (FHFA) housing price index (HPI), previously known as the OFHEO housing price index. The FHFA index is available for nearly all metropolitan areas in the United States.[76] It measures the movement of single family home prices by looking at repeat mortgage transactions on homes with conforming, conventional mortgages purchased or securitized through Fannie Mae or Freddie Mac since 1975.[77] Since the index looks at repeat mortgages of the same home, it is continually revised to reflect current MSA boundaries. This is the reason we must use the most current definitions of MSAs in constructing the birth data. We annualize the index (which is available quarterly) by taking the mean value of the index over the four quarters of a year.

We use the FHFA index to construct real house prices for each MSA-year by combining it with information on median home values obtained from the 2000 census. The 2000 Census records median home values for each county in the U.S. We use the same county crosswalk used to construct MSAs in the birth data to construct MSA-level median 2000 house values, which are the population-weighted average across all counties in each MSA. Home values are scaled by the relevant change in the FHFA index over time and are adjusted to 2006 dollars using the CPI-U "All items less shelter" series. This measure serves as a proxy for real house

per race-education group.

[75]While this process eliminates 60 percent of MSAs, it only eliminates about 15 percent of births.

[76]FHFA requires a metro area to have at least 1,000 transactions before it is published.

[77]Conventional mortgages are those that are neither insured nor guaranteed by the FHA, VA, or other federal government entities. Mortgages on properties financed by government-insured loans, such as FHA or VA mortgages, are excluded from the HPI, as are properties with mortgages whose principal amount exceeds the conforming loan limit. Mortgage transactions on condominiums, cooperatives, multi-unit properties, and planned unit developments are also excluded. This contrasts to the alternative Case-Shiller index, which includes all homes, but is only available for 37 states and a more limited set of MSAs. Additional differences between the two indices are that the Case-Shiller index puts more weight on more expensive homes and the Case-Shiller index uses purchases only, whereas the FHFA index also includes refinance appraisals. As a robustness check, we have re-estimated our results using the Case-Shiller index. Note that we use the "all transactions" version of the FHFA index, which includes both sales and refinancings of existing mortgages. We do not use the "sales only" version of the index because it is available for only a small subset of MSAs.

price movements of median value homes in each MSA.[78]

The third main variable we need to construct is a measure of mean group-level home ownership rates at the MSA level. This is key to our analysis because conceptually, we expect there to be heterogeneous responses of birth rates to home prices across groups with different rates of home ownership. Recall that Vital Statistics data do not include information about home ownership status, so we can not separately tabulate current period births (or conceptions) separately for home owners and non owners. Furthermore, we ideally do not want to use an individual-level measure of realized home ownership rate, as that is potentially endogenously determined with childbearing outcomes. Our implemented solution is to use MSA-group level home ownership rates calculated from the 1990 five percent sample of the decennial census. As above, groups are defined by race/ethnicity and age category. We match the MSA definitions provided in the Census to the 2009 MSA definitions used for the birth and housing price data according to the crosswalk procedure described in the appendix (section 4.7). To be clear, our group-level measure of home ownership is taken at baseline and is time invariant.

4.3.2 Descriptive Statistics and Trends

Figure 4.1 displays trends in mean (CPI adjusted) house prices, constructed as described above, in our sample, both in levels (panel (a)) and yearly percentage changes between year $t-1$ and t (panel (b)). Figure 4.1 also displays house prices alternatively constructed using the Case-Shiller Index to scale 2000 median home prices. The three housing cycles that fall within our period of study are highlighted: the 1990-96 period of price decline, the 1997-2006 housing boom, and the subsequent 2007-2010 housing bust. Appendix table 4.11 lists the 154 MSAs included in our analysis sample, ranked according to the percentage increase in housing prices between 1997 and 2006. The table also lists the computed median home price

[78]This is the same procedure used by Glaeser et al. (2008).

in 2006 and the fertility rate in 2006. The top seven ranking MSAs/MSADs in terms of the percent change in house prices during the boom cycle are all in California. Among the top 25 ranked, fifteen are in California, nine are in Florida, and the remaining one is the Washington, DC Metropolitan area. The most expensive housing market in the nation is the San Francisco MSA, with a 2006 median home price of $781,891. The MSAs with the least house price growth during this boom cycle are Dayton, OH (3.92 percent), and Fort Wayne, IN (3.27 percent).[79]

Figure 4.2 displays the time-series correlation between fertility rates and house prices and then between fertility rates and unemployment rates, for the period 1990-2006, averaged across the MSAs in our sample. These plots suggest that movements in fertility rates track movements in house prices fairly closely, particularly in more recent periods. In fact, a comparison of the graphs reveals that the time-series correlation between aggregate fertility rates and housing prices is much greater than it is between aggregate fertility rates and unemployment rates, .85 versus -.04 This provides a prima facie case for the importance of considering housing prices when investigating how economic conditions affect current period birth rates.

Table 4.2 provides summary statistics from the 1997-2006 Vital Statistics natality files and the 1990/2000 Census. These data are used collectively in various analyses presented below. All measures are female-population weighted. The first three columns summarize the main dependent variable of interest: fertility rates (group-level births per 1000 women age 20-44), overall and for first and higher parity births. The overall fertility rate in our sample is 70 births per 1000 women aged 20-44. The highest fertility rates are found among

[79] There is an active literature exploring various explanations for the boom and bust in house prices experienced in recent decades. As summarized by Sinai (2012) – who offers citations for the various factors – these potential explanations include "changing interest rates, sub-prime lending, irrational exuberance on the part of home buyers, a shift to speculative investment in housing, contagion and fads, and international capital flows." Sinai's 2012 paper presents a set of empirical facts about the recent housing cycle, including information about how the amplitude and timing of house price appreciation and depreciation varied across MSAs. One of the observations he makes that is particularly relevant to our current empirical approach is that "demand fundamentals" do not have the same amplitude as price cycles nor does the time pattern of the growth in fundamentals match the timing of the growth in house prices across MSAs.

Hispanics age 20-29: 154 births per 1000 women. The lowest rate is among Black mothers age 30-44: 38 births per 1000 women.

The next column summarizes data from the 1990 census on MSA-group level home ownership rates. The overall home ownership rate among our sample of women age 20-44 is 44 percent. The highest home-ownership rates are found among older (age 30-44) white women, who have an ownership rate of 67 percent. The lowest rates are found among younger (age 20-29) Black women, whose ownership rate is on average 8 percent. This indicates there is substantial variation across groups in rates of ownership. For the sake of comparison, the next column shows the rates as calculated from the 2000 census. Comparing the group-level ownership rates in 1990 and 2000 we see that home ownership rates are extremely stable over this time period. The final column displays the range of the 1990 ownership rate across MSAs, for each group. These numbers indicate that in addition to the substantial variation across groups in rates of home ownership, there is also substantial variation within groups across MSAs.

4.3.3 Empirical Specification

Our initial empirical analysis consists of ordinary-least squares regressions (OLS) at the MSA-group-year level. For our baseline analysis, we restrict our attention to the housing cycle of 1997-2006. This facilitates interpretation as the period was one of nearly uniform house price growth, and is recognized by the real estate literature as a housing boom period. We will subsequently consider two housing bust periods: the early 1990s bust period (1990-1996) and the post-2006 housing bust (2007-2010).

We estimate regression models of the following form:

$$\begin{aligned} ln(FertRate_{mtg}) \; &= \beta_0 + \beta_1(HousePrices_{mt-1} * OwnRate_{mg}) + \beta_2 HousePrices_{mt-1} \\ &+\beta_3 OwnRate_{mg} + \beta_4 \mathbf{X}_{mt-1} + FracColl_{mgt} + \gamma_m + \gamma_t + \gamma_g \\ &+\gamma_m * (t-1) + \gamma_m * OwnCat_{mg} * (t-1) + \epsilon_{mgt} \end{aligned} \tag{4.1}$$

The level of analysis is an MSA-year-group cell. In the above equation, the subscript m denotes MSA division, t denotes year of the birth (where *t-1* refers to the year of conception), and g denotes group.[80] There are six groups, defined by the interaction of our three race/ethnic groups (non-Hispanic white, non-Hispanic black, and Hispanic) and two age categories (age 20-29 and age30-44). Our final analysis sample consists of 9,240 observations (10 years * 6 groups * 154 MSAs). All regression are weighted by the total number of births in each cell.[81]

The coefficients of primary interest are β_1 and β_2 , which capture the conditional effect, respectively, of MSA-year house price index (HPI) interacted with a baseline measure of MSA-group-level ownership rates and the conditional main effect of the MSA-year house prices ($HousePrice_{mt-1}$) on fertility rates. The former indicates how an increase in home ownership rates affects the relationship between de-meaned (and sometimes de-trended) MSA house prices and births. The conditional main effect of $HousePrice_{mt-1}$ indicates how movements in house prices affect fertility rates net of ownership interactions, all else held constant. We interpret this to be the conditional relationship between $HousePrice_{mt-1}$ and log fertility rates among a non-home-owning population of households.

The variable $OwnRate_{mg}$ is the MSA-group level home ownership rate measured in the 1990 five percent sample of the decennial census. This measure is taken at baseline to minimize concerns about the endogeneity of year-specific MSA-home ownership rates and year-specific MSA-fertility rates. Taking a baseline measure of home ownership rates for a

[80]For the sake of convenience, we write *t-1*, but our empirical analysis is precise in dating the year of conception by taking the date of birth and subtracting off the reported weeks of gestation.

[81]Results alternatively weighting by total female population in each cell are similar and available from the authors upon request.

group is therefore preferable. As reported in Table 4.2, there is considerable heterogeneity across groups in home ownership rates, as well as heterogeneity within groups across MSAs. It is also true that home ownership rates are quite stable over time within groups, which means the baseline measure is highly predictive of current period home ownership rates. Therefore, this approach does not entirely eliminate any concern about endogenously determined current period births and our measure of home ownership rates. We control for this conditional main effect to facilitate a causal interpretation of β_1, but we are careful not to assign a causal interpretation to the coefficient on ownership rates.

We are interested in identifying the causal relationship between lagged house prices and fertility rates. It is thus important to control for other time-varying MSA-level economic conditions that potentially covary with real estate markets and also fertility timing decisions. Our regression specification includes controls for MSA-year unemployment rate, MSA-year male wages included in the vector X_{mt} in equation 4.1. The specification also controls for $FracColl_{mgt}$, the fraction college educated in each MSA-group-year. This is calculated as a three year moving average using data from the Current Population Survey. Data on MSA-year level unemployment rates come from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics. Our measure of MSA-year level male wages is the 25th, 50th, and 75th percentile male wage, which was calculated by MSA and year in the Current Population Survey. Percentiles of the wage distribution were constructed based on hourly earnings for full-time, full-year male workers.[82] Unemployment rates were collected at the county level and aggregated to MSAs using the crosswalk procedure described in the appendix (section 4.7). The wage and fraction college measures were calculated using the MSA definitions available in the CPS and translated to 2009 MSAD definitions using the crosswalk procedure.

The regression also includes controls for MSA fixed effects (γ_m), year fixed effects (γ_t), group fixed effects (γ_g), and in some specifications, MSA-specific time trends ($\gamma_m * (t-1)$)

[82]We construct wages as in Autor, Katz, and Kearney (2008). We define full time as 35 or more hours per work, and full year as 40 or more weeks worked in the past year. We drop individuals who make less than one half the 2006 minimum wage (in 2006 dollars). Top-coded observations are multiplied by 1.5.

and MSA-ownership-cell-specific time trends ($\gamma_m * OwnCat_{mg} * (t-1)$). It is imperative that the regression specification control for MSA fixed effects so that the estimated relationship between house prices and birth rates is not confounded by time-invariant differences in preferences for children across MSAs. If couples with lower preferences for children sort into areas with higher costs of living – driven by other amenities – there will be a negative correlation between house prices and fertility.[83] Given our goals in this chapter, we want to isolate the effect of house prices on current period fertility net of these sorting patterns. It is thus important that our regressions control for mean MSA-level differences in birth rates. The resulting regression estimate of the relationship between house prices and birth rates is identified off within-MSA changes in house prices.

We additional include MSA specific time trends in some specifications to allow for the possibility that individuals with plans to increase or decrease their fertility move into MSAs with upward or downward trending house prices. If there exist trends of this kind that are distinct for groups with high and low ownership rates, our estimated β_1 might be a biased estimate of the conditional causal effect of interest. We thus additionally allow for separate MSA specific time trends based on whether a group's level of ownership (in a particular MSA) is above or below the median home ownership rate (of 30.5 percent in our sample of MSA*group cells), yielding two values of $OwnCat_{mg}$. So, for example, this allows for white women age 30-44 in the Boston metro area to be on a different trend then black women age 20-29 in the Boston metro area. Finally, we directly address the possibility that reverse causality or correlated unobservables are biasing the OLS estimates of β_1 and β_2 by implementing an instrumental variables approach that exploits this exogenous variation in

[83]For example, consider the hypothetical case of two couples, in which one moves to San Francisco, where household expenses are high, because they expect to have few children and spend their time and money instead indulging in city-type amenities. The other couple moves to Wichita, in expectation of buying a big house at a much lower cost per square foot, and filling it with kids. If these couples are typical, then high-latent-fertility couples will sort into lower priced real estate markets and low-latent-fertility couples will sort into lower priced real estate.Simon and Tamura (2008) examine the cross-sectional relationship between fertility and the price of living space across U.S. metropolitan areas, as captured by the average rent per room in an urban area (calculated among renting households.) Their baseline specification, which controls for region effects and demographic composition, suggests that a one percent increase in rent is associated with 0.16 fewer children per household.

house price movements induced by the interaction between national house price movements and MSA housing supply elasticities. We describe these approaches in more detail below.

As noted above, our empirical analysis is designed to capture current period fertility responses to movements in local house prices. Certainly it would be interesting to know whether any short term responses observed translate into changes in completed fertility. To the extent that we observe a change in higher-order births, we can speculate that those changes reflect changes in total completed fertility. But, we leave it to future research to carefully consider the lifetime implications of any short term changes that we find. Such an analysis requires a different empirical framework.

4.4 Estimation Results

4.4.1 Ordinary Least Squares (OLS) Specifications

Table 4.3 presents the results of estimating equation 4.1. Column (1) reports the results with all fixed effects included, but without MSA-specific controls for labor market conditions. This sparse specification yields a point estimate of β_1 of 0.0468 and a point estimate on β_2 of -0.0124, both statistically significant at the one percent level. The positive and statistically significant point estimate on the interaction term $HousePrice_{mt-1} * OwnRate_{mg}$ indicates that as home ownership rates increase, higher house prices lead to an increase in current period births, all else held constant. This implies that a positive home wealth effect dominates any negative price effect among current home owners. The negative and statistically significant point estimate on $HousePrice_{mt-1}$ is consistent with a negative price effect of house prices on current period fertility for non-home owners. Column (2) adds the unemployment rate and wage measures. The main point estimates of interest are qualitatively unchanged. Looking at other explanatory variables, we see that the estimated coefficient on the mean ownership rate is positive, but statistically insignificant. As noted above, we do not propose a causal interpretation to this relationship. The unemployment rate is found to be negatively related to the fertility rate in all specifications, but it does not enter with statistical significance.

We next check the sensitivity of the results presented above to the inclusion of MSA-specific linear trends, MSA-specific quadratic trends, and MSA-ownership category specific linear trends. Columns (3) and (4) report the results with MSA specific linear trends and MSA specific quadratic trends, respectively. The pattern remains the same – a positive coefficient on $HousePrice_{mt-1} * OwnRate_{mg}$ and a negative coefficient on $HousePrice_{mt-1}$– and the magnitudes of the coefficients are similar to the specification without any trend terms included, giving us no reason to suspect that individuals with plans to increase or decrease their fertility systematically move into MSAs with upward or downward trending house prices.

In Column (5), we include distinct MSA-specific trends for cells with ownership rates above and below the median home ownership rate. These trends also do not alter the pattern of estimates, but the point estimates are attenuated toward zero. In this model, the estimated coefficient on the $HousePrice_{mt-1} * OwnRate_{mg}$ interaction is 0.0276 (with a standard error of .00375) and the estimated coefficient on $HousePrice_{mt-1}$ is -0.00509 (standard error of .00184).[84]This is arguably the most conservative of the OLS specifications. These estimates suggest that if house prices increase by $10,000, as we move from an MSA-group with an ownership rate of 0 to a cell with an ownership rate of 1, there would be a relative increase of 2.5 percent in fertility rates. More usefully, if house prices increase by $10,000, comparing MSA-groups with ownership rates of 0.25 to those with ownership rates of 0.75, we would see a relative increase of 1.25 percent in fertility rates.

[84]As described earlier, these estimates are weighted by the total number of births in each cell. If we alternatively weight by total female population in each cell we obtain an estimate of 0.0316 with a standard error of 0.00429 on the interaction term and an estimate of -0.0087 with a standard error of 0.0021 on the main effect.

4.4.2 Instrumental Variables (IV) specifications: Exploiting local housing supply elasticity

The main threat to assigning a causal interpretation to the estimated β_1 and β_2 is the possibility of reverse causality or some correlated unobservable to house prices that affect fertility rates. If it were simply the case that in MSAs where people demanded more children house prices were driven up in equilibrium, ceteris paribus, then both $\beta_1$and β_2 would be estimated to be positive. For the finding of separating effects to be explained by the alternative reverse causality story, it must be the case that fertility-related demand pressures occur disproportionately in areas with relatively higher rates of home ownership (as measured in a pre-period baseline year). This confounding story is one of fertility-preference demand driven price changes.

In order to address the possibility that reverse causality or correlated unobservables are biasing our estimates we make use of the Saiz (2012) measure of housing supply elasticity. The measure is based on non-linear combinations of both the Saiz (2008) geographic limitations measure and the Wharton Residential Urban Land Regulation Index created by Gyourko et al. (2008). We propose that concerns about fertility-preference demand driven price changes are less likely to be a concern in places with lower housing supply constraints, or higher housing market supply elasticities. We thus estimate our regression models separately for MSAs with higher and lower levels of supply elasticity, as captured by the Saiz (2012) measure. If the estimated relationship is maintained in less supply constrained places, we are more confident in the assumption that our estimated effect is not driven by homeowners with babies (or fertility intentions) bidding up the prices of inelastically supplied houses.

Table 4.4 reports the results using the baseline specification displayed in column (2) of table 4.3, which includes both MSA and year fixed effects. We choose this to be our baseline specification because it is the one we will estimate with the IV specification, as described below. Column (1) reports the results for the sample of MSAs with supply elasticities below the median and column (2) above the median. In fact, moving from column (1) to (2), the estimated positive coefficient on $HousePrice_{mt-1} * OwnRate_{mg}$ increases in magnitude, as does the estimated conditional negative main effect of $HousePrice_{mt-1}$. This is opposite of what would be expected under the reverse causality scenario.

Next, we more formally incorporate the supply elasticity measure by employing an instrumental variables strategy similar to that employed by Chetty and Szeidl (2012); we instrument for MSA-level house prices with the interaction of a baseline supply elasticity measure with the national trend in housing prices. The intuition here is that aggregate demand shocks that affect the national housing market are expected to exert a relatively larger influence on local housing prices in MSAs which are more supply constrained. The identification assumption is that the interaction between baseline MSA housing supply elasticity and national house price trends would not have been systematically correlated with trends in fertility rates in the absence of MSA house price changes.

In order to implement this strategy we interact the measure of supply elasticity with the national version of FHFA housing price index. Since we have two potentially endogenous variables –the level measure of house prices and the interaction term with MSA-group level ownership rates–we use the triple interaction of the MSA supply elasticity with the national house price index and MSA-group level ownership rate as a second instrumental variable. Since nationally, house prices grew almost linearly in this time period, we do not include MSA-specific time trends as our baseline IV specification; MSA-specific time trends are highly collinear with the instrument, which substantively reduces the statistical power of the first stage. [85]We also do not include the conditional main effect of housing supply elasticity because it is measured at the baseline and does not vary across groups, so it is absorbed by the MSA fixed effects. Table 4.4 presents the first-stage estimation results for the interaction term $Houseprice_{mt-1} * OwnRate_{mg}$ (column (3)) and the level term $Houseprice_{mt-1}$ (column

[85]If we include MSA time trends, the results are very similar but the first stage F Statistics fall below conventional levels. The coefficients on $\beta_1$and β_2 are 0.07148 and -0.01439 (with standard errors of 0.0098 and 0.0039), respectively. However, the first stage F Statistic is the equation for $Houseprice_{mt-1}$ is reduced to 8.71, which is below the conventional rule of thumb of 10.

(4)). As expected, increasing supply elasticity is associated with reductions in MSA level house price growth as national house prices increase.

Table 4.4, column (5) presents the IV results. The estimated effects of interest maintain their signs of direction, but increase in magnitude. The point estimate of β_1 of 0.0723 and a point estimate on β_2 of -0.0239, both statistically significant at the one percent level. These estimates suggest that if house prices increase by \$10,000, as we move from an MSA-group with an ownership rate of 0 to a cell with an ownership rate of 1, there would be a relative increase of 7.2 percent in fertility rates. More usefully, if house prices increase by \$10,000, comparing MSA-groups with ownership rates of 0.25 to those with ownership rates of 0.75, we would see a relative increase of 3.6 percent in fertility rates. These estimates are larger in magnitude than the OLS results and a Hausman specification test can reject the consistency of the OLS estimate at the 1 percent level.[86] However, the net effect at the mean is almost identical between the OLS and IV specification: both specifications indicate that at the mean U.S. home ownership rate, the net effect of a \$10,000 increase in house prices is a 0.8 percent increase in fertility rates.We put these numbers into context below with the use of simulation exercises.

4.4.3 Robustness Checks

In this section we implement various robustness checks on the model specification and sample construction. We begin by considering how our estimates change if we replace the house price in the year of conception with alternative measures of house prices. We do not have a strong a priori reason to believe that house prices in the year of conception is the most relevant measure, as opposed to, say, house prices averaged over the three years prior. It may be the case that couple's fertility decisions are based on a longer time horizon or on

[86]Since standard errors are clustered in both the OLS and 2SLS specification, OLS is not fully efficient and I compute the Hausman test statistic with bootstrapped variance estimates. The variance was calculated using 500 bootstrap replications.

longer terms averages. Table 4.5 reports the results of estimating alternative models of this sort, using both the OLS and IV strategies. Columns (1)-(4) use house prices in the years 1, 2, 3, and 4, respectively, prior to conception. Columns (5), (6), and (7) use the 3-year moving average of house prices over the two, three, and four years, respectively, prior to conception. In all of these cases, we instrument for house prices using the same instrumental variables strategy described above. In all of these seven alternative models, the familiar pattern emerges of a positive coefficient on the interaction between $HousePrice_{mt-1}$ and $OwnRate_{mg}$and a negative coefficient on $HousePrice_{mt-1}$. for both the OLS and IV results. Results are within a reasonable range of magnitudes – with the point estimates of β_1 ranging from 0.0468 to 0.0718 for the OLS and 0.0723 to 0.0995 for the IV results, giving us no reason to prefer one of these specifications over our baseline specification.

Next, we estimate various alternative specifications to equation 4.1 above, providing some robustness checks on the main MSA-group level analysis. Table 4.6 reports these results. Column (1) reproduces the main OLS and IV results from Tables 4.3 and 4.4 for the sake of comparison. First we consider alternative measures of the labor market conditions. In column (2) we replace male wages with separate measures of male and female wages. In column (3) we replace the wage distribution measures with the mean wage. In column (4) we replace the wage distribution measures with a measure of income per capita collected from the Bureau of Economic Analysis (BEA) regional economic accounts. To create this variable at the MSA-year level, we employ our crosswalk procedure described in the appendix (section 4.7). Income included in this measure includes all wage and salary income as well as supplements to wages and salaries, proprietor's income with inventory valuation and capital consumption adjustments, rental income of persons with capital consumption adjustment, personal dividend income, personal interest income, and personal current transfer receipts, less contributions for government social insurance. In each case, the coefficients are virtually unchanged.

Next, we consider that owners and non-owners might be differentially affected by general economic conditions in a way that is not captured by simply including a measure of wages. If this were the case, the coefficient on $HousePrice_{mt-1} * OwnRate_{mg}$ might capture this difference, leading to a biased estimate of the causal effect of interest. To do this, we interact the home-ownership rate with the wage measures. Column (5) displays the results of this exercise. The coefficient on $75thWage_{mt-1} * OwnRate_{mg}$ is positive and statistically significant, indicating owner's fertility decisions are positively effected by increases in male wages at the top of the distribution. However, the coefficients on $HousePrice_{mt-1} * OwnRate_{mg}$ and $HousePrices_{mt-1}$ remain unaffected.

Next, we add a control variable for average rental prices in the MSA-year. Average rental prices and house prices tend to covary, but there are years during which the two series are more or less closely aligned. Our measure of average rental prices comes from the Department of Housing and Urban Development (HUD) Fair Market Rents program, used for the purpose of calculating rent for the Section 8 housing assistance payment program.[87] We adjust the nominal values of rental prices to 2006 dollars using the CPI-U "all items less shelter" series (as we do for the HPI). As shown in the table, the inclusion of this control variable does not appreciably change the estimated coefficients on our two explanatory variables of interest: $HousePrice_{mt-1}$ and $HousPrice_{mt-1} * OwnRate_{mg}$.

Finally, in column (7) of Table 4.6 we consider an alternative sample of MSAs to check on whether the changing boundaries of MSAs over the sample period is influencing our estimates. We re-estimate the specifications reported in Table 4.4 using a restricted set of MSAs. In particular, we limit the sample to MSAs whose boundaries did not change between 1990 and 2009. This is done as a check on the sensitivity of our estimates to the crosswalk procedure we have used to link current MSAs (2009 OMB definitions) to vintage MSAs (1983 OMB definitions) which we use to match the group level home ownership rates to the rest of the data. This procedure effectively ignores boundary changes that have occurred over our

[87] Calculated rent is inclusive of utilities and is typically calculated at the 40th percentile of the rent distribution by number of bedrooms. Prior to 1995, rent was calculated at the 45th percentile. Some cities are calculated at the 50th percentile. We take the unweighted average of the reported fair market rental value for zero to four bedroom units as the average rental price in a given city.

sample period. Though the sample size is reduced, the range of point estimates on the two coefficients of primary interest is not qualitatively altered.[88]

4.4.4 Different Demographic Groups

In Table 4.7, we report the results of estimating equation 4.1 for various demographic subgroups and for first and higher order births using both the OLS and IV strategies. Column (2) reports the results for non-Hispanic whites, column (3) reports the results for non-Hispanic blacks and column (4) reports the results for Hispanic whites. The point estimate on the interaction term $HousePrice_{mt-1} * OwnRate_{mg}$ is always positive while the coefficient on $HousePrice_{mt-1}$ is negative, implying a net positive effect of house price increases among home owners and a negative effect among non-owners across all groups

We next consider whether the effects of house prices on current period births are driven by first births, second births, or higher parity births. It is not clear a priori which would be more price or income elastic. On the one hand, the optimal timing of first births might be less constrained, since mothers tend to be younger and might consider that a deliberate delay will be less consequential, as they have more childbearing years ahead of them. Also, if couples have specific ideas about optimal spacing, they might be more flexible about the timing of their first birth. On the other hand, subsequent births might be more "marginal" and thus might exhibit a great degree of elasticity with respect to price or a wealth shock. An additional motivation for this analysis is that an effect on higher order births might be indicative of a change in completed fertility.

Table 4.7 columns (5)-(7) report the results. For both first, second, and higher parity births, the estimated coefficient on the interaction between $HousePrice_{mt-1}$ and ownership rate is positive and statistically significant, with similar magnitudes: 0.0538, 0.0474

[88] An additional specification we estimated uses only variation across MSA/age cat cells to define average ownership rates. This approach purges that ownership variable of variation coming from race/ethnic groups. The results are very similar to those obtained when ownership rates are defined by MSA/age cat/race-ethnicity. We do not report those results here for the sake of space.

and 0.0434 in the OLS specification and 0.0947, 0.0754 and 0.0574 in the IV specification, respectively. The point estimate for the coefficient on $HousePrice_{mt-1}$ is negative and statistically significant for first, second, and higher-order births. The finding of an effect on both first and higher-order births is potentially informative about the nature of the effects we are estimating. Increases in first births might reasonably be interpreted as a change in timing, while changes in higher order births might reasonably be interpreted as an increase in the total number of children, particularly for third and higher parity births. These interpretations are merely speculative, and warrant further investigation.

Given that a previous literature exists on the relationship between unemployment rates and contemporaneous fertility rates, it is interesting to consider the estimated coefficients on the unemployment rate. Our regression models yield statistically significant negative estimates of the relationship between unemployment rates and fertility rates among whites, but not among blacks or Hispanics. When house prices are not included in the model (not shown in the table), the estimated relationship is largely unchanged for whites (a statistically insignificant -0.0036), but it becomes positive and statistically significant for blacks and Hispanics. It is also interesting to note that in terms of separate effects by birth parity, the unemployment rate is negatively related to first and second births, but not discernibly related to higher-order births. This would be consistent with the unemployment rate having an effect on the timing of childbearing initiation, but potentially not with completed fertility. To the extent that this interpretation is warranted, this is an interesting contrast to the potentially more permanent effect of house prices. Again, we think these considerations deserve further examination, although it is outside the scope of this work.

4.4.5 Individual level estimation – using Current Population Survey (CPS) 1996-2006

The empirical results presented above suggest that an increase in MSA-level house prices exert a negative price effect on births among non-owners and a net positive effect on births among owners, all else equal. These estimates are generated by an aggregated cell-level analysis, but the underlying conceptual framework is at the individual level. We thus turn to individual-level Current Population Survey (CPS) data to check that the story told by aggregate level data is confirmed with individual level data. We map the older MSA designations provided in the CPS (as in the Census) to the 2009 MSA designations provided in the FHFA house price data using the crosswalk procedure described in the appendix (section 4.7). In the CPS we do not see the full population of births, as we do with an analysis of Vital Statistics birth data. However, as a supplementary data source, the CPS offers the distinct advantage of directly identifying home-owners.

In this individual level analysis, we define Own_i as an indicator for whether the individual in the CPS is the household head or head's spouse and the household is reported to own their home. In the aggregate analysis above, ownership was defined at the group level in the baseline year of 1990. Here it is defined at the individual level in the current year, as we have no measure of lagged home ownership available. Caution should thus be exercised in assigning a causal interpretation to the $HousePrice_{mt-1} * Own_i$ interaction term in this specification, since individuals who intend to have a baby this year might decide to buy a house in anticipation of that event. This is another reason we consider this analysis supplementary to the main analysis above.

We define the dependent variable $Pr(Birth)_i$ to be one if there is a child under the age of one in the household. All the other variables are defined at the MSA level as defined in equation 4.1 above. Explanatory variables, including the house price index, are matched to observations by the year prior to the survey year in order to capture the effect of conditions in the year of the baby's conception. (We do not have perfect birth-date or gestation information, as we do in the Vital Statistics natality files, and so here we use year minus one as an approximation.) Recall that our goal is to obtain an estimate of the causal relationship of house prices on current period birth rates. Even if we had individual level house prices, we would not use them because individuals are likely to sort into houses at least in part based on their expectations of number of children. For example, individuals intending to have more children will likely seek larger houses or houses in better school districts, which tend to be more expensive. We thus use MSA-level house prices conditional on MSA fixed effects (to control for endogenous sorting into higher or lower priced MSAs) in all our analyses.

Table 4.8 reports the results estimated using a linear probability model.[89] In the pooled sample regression reported in column (1), we see the familiar pattern of point estimates – a negative point estimate on $HousePrice_{mt-1}$and a positive point estimate on the interaction of $HousePrice_{mt-1} * Own_i$ (significant at the 1 percent level). Columns (2)-(3) report the results including the additional MSA and year fixed effects and the time-varying MSA-level controls, unemployment rates and wages. Column (4) reports the results from the IV specification described above. Although the IV results are not precisely estimated, the magnitudes of the coefficients of primary interest are similar to the OLS estimates. We have additionally estimated these models restricting the CPS sample to individuals who do not report moving over the past year. The results are not qualitatively altered. Overall, this set of individual-level results give us confidence that our interpretation of the results from the aggregate level analyses is appropriate. In particular, we see that the positive effect is being driven by individuals that are self-reported to be home owners.

[89]Probit marginal effects are similar and available upon request.

4.4.6 Housing Bust Periods

Our analysis has thus far been limited to a period of history characterized by rising house prices. It is interesting to consider explicitly the relationship between housing price decreases and birth rates. There might be asymmetric effects, whereby an increase in housing wealth might lead people move up their period of childbearing to a greater extent than a decrease in housing wealth will lead people to delay. One possible reason for such an asymmetry is that there is a biological timing constraint that individuals are reluctant to push against. It becomes an empirical question as to whether there are differential responses to house price rises and declines. To consider this explicitly, we want use data from two periods of house price decline: 1990-1996 and 2007-2010. Figure 4.1 shows these two periods: between 1990-1996 prices declined gradually and between 2007-2010 there is a dramatic decline in prices. Unfortunately Vital Statistics birth data is not available for conception years past 2006, so we can only look at the 1990-1996 housing bust period using the approach of the main analysis. We therefore turn to individual level data sources for 2007-2010 period.

We begin by examining the 1990-1996 bust period using the approach used in the main analysis. Table 4.9, columns (1) and (2) display the results using OLS and the IV strategy, respectively. The pattern on the coefficients remains similar to the 1997-2006 period – a positive coefficient on $HousePrice_{mt-1} * Own_{mg}$ and a negative coefficient $HousePrice_{mt-1}$. In columns (3) and (4) we report results from an individual level analysis from the CPS for this time period. Again, results are very similar to those found in the 1997-2006 period: a positive and statistically coefficient on $HousePrice_{mt-1} * Own_{mg}$ and a negative, but not statistically significant effect on $HousePrice_{mt-1}$. As in the analysis for 1997-2006 period, the the IV estimates are similar but less precise.

Next, we move on to the 2007-2010 housing bust period, which is characterized by a steep decline in prices. First, we repeat the individual-level CPS analysis for this period. Table 4.9, columns (5) and (6) displays results. The pattern on the coefficients is extremely similar to both the earlier bust period (1990-1996) and the housing boom period (1997-2006). Since Vital Statistics birth data is not available for this period, we supplement the analysis by examining data from the American Communities Survey (ACS), conducted annually by the U.S. Census Bureau, beginning in 2000. We obtained this data from IPUMS. The data is available with the equivalent of MSA identifiers starting in 2005.[90] We construct the indicator variables "Pr(Birth)" and "own home" in the same manner as described above for the CPS data. Again, the coefficient on $HousePrice_{mt-1} * Own_i$ is positive and statistically significant at the 1% level. The coefficient on $HousePrice_{mt-1}$ is positive, although it is not statistically significant.

These findings give us some confidence that it is appropriate to use our preferred aggregate results above – generated from data for the years 1997-2006, a period characterized by house price increases – to make out-of-sample predictions to more recent years, characterized by house price declines. Between 2006 and 2010, housing prices fell $63,000 among the MSAs in our sample. At the mean rate of home ownership, our estimates imply that this would lead to a 7.5 percent decline in subsequent year births. We can also simulate the effect of the rise in unemployment rates over the period. [91]Between 2006 and 2010, unemployment rates rose 5.14 percentage points. Holding housing prices fixed, our estimates imply that this corresponds to a 6.8 percent decline in births. A simulation of the two changes in tandem implies that in the great recession period 2007-2011 the decline in housing prices and increase in unemployment rates is associated with a 13.8 percent decline in current-period birth rates. According to the National Center for Health Statistics, the national fertility rate dropped from 69.3 in 2007 to 63.8 in 2011, a 7.9 percent decline.

[90]The ACS identifies PUMAs (Public Use Microdata Areas), which IPUMS has matched to MSAs. We then use the crosswalk procedure described in the data appendix (section 4.7) to match to the housing data (Ruggles et al., 2010). PUMAs are also identified in 2003, but we do not use this data.

[91]Both the average fall in home prices and the average increase in unemployment rates are population weighted average changes for the 163 MSAs in our sample between 2006 and 2010.

4.5 Interpreting the Results

4.5.1 Interpreting the Magnitudes of the Estimated Effects: Simulation Exercise

Our analysis of Vital Statistics birth data coupled with MSA-level house prices shows that an increase in MSA-level house prices, all else held constant, is associated with fewer births among non-owners and a net increase in births among owners. We interpret this pair of findings as indicative of a negative price effect among non-owners and a dominant housing wealth/equity effect among owners. These patterns hold for whites, blacks, and Hispanics and appear for both first and higher parity births.

In order to facilitate an understanding of whether these results are economically large or small, we have conducted a simple simulation exercise. Figure 4.3 presents the predicted effect of a $10,000 increase in house prices on births for each race/ethnic group as well as first and higher parity births.[92] The x-axis represents group home ownership rates and the y-axis represents the net predicted percentage change in births from of a $10,000 increase in house prices, conditional on each level of home ownership. The prediction is indicated by the solid line and a 95% confidence interval is indicated by the dashed lines.[93] The predictions are calculated based on IV point estimates displayed in Table 4.7, which include all of the main demographic group and MSA level control variables, and MSA and year fixed effects.

In all cases, the exercise suggests a positive, linear relationship between home ownership rates and the change in births due to a $10,000 increase in house prices. The net effect for

[92] Since the effects are similar for second and third or higher parity births we combine the two for these simulation exercises.

[93] We predict the percentage change in fertility rates from a $10,000 increase in mean housing prices for ownership rate o: $(FertRate|HousePrice = h + 10k, OwnRate = o) - (FertRate|HousePrice = h, OwnRate = o)/(FertRate|HousePrice = h, OwnRate = o)$. For each group, we calculate the standard error of the prediction at the mean of the independent variables using 100 bootstrap replications and apply that standard error to calculate the confidence interval at each level of o.The solid line represents the predicted effect and the dashed line represents a 95% confidence interval, both of which were smoothed using a locally weighted linear regression.

all demographic groups implies that as the ownership rate increases from 30 percent to 40 percent, the net effect become positive. This implies that in MSAs with sizable rates of home ownership, the positive home equity effect among owners is large enough to outweigh the negative price effect, leading to increases in MSA-level birth rates. Among whites, the impact switches from negative to positive between 40 and 50 percent ownership. For blacks, the impact becomes positive between 20 and 30 percent, and for Hispanics, it switches between 40 and 50 percent.

We also consider what changes in home prices imply for group specific fertility rates, since there is heterogeneity in the magnitude of the price and home equity effects as well as in rates of home ownership. Overall in our data the population weighted mean home ownership rate is 44 percent. At this rate, the net effect of a $10,000 increase in prices is a 0.8 percent increase in births. Among whites, the mean home ownership rate is 53 percent, which is associated with a 0.7 percent increase in births for that group. Among blacks, the mean home ownership rate is 24 percent, which is associated with a net increase of 0.2 percent in births. And among Hispanics, the mean home ownership rate is 29 percent, which is associated with a net decrease in births of 0.3 percent. This indicates that although Hispanic home ownership rates are higher than Black home ownership rates, the net effect is smaller for Hispanics because the estimated home equity effect is smaller for that group.

Finally, it is useful to consider an out-of-sample prediction assuming extreme values of ownership rates, to have an estimate of the effect among owners and non-owners. Assuming a 100 percent ownership rate, the net impact of a $10,000 house price increase is a 5 percent increase overall. Separately by race/ethnicity, our simulations suggest a 5.9 percent increase for whites, a 7.9 percent increase for Blacks, and a 1.9 percent increase for Hispanics. These figures imply that among owners, the increase in house prices during the recent housing boom led to a sizable impact on the likelihood of giving birth in a given year.[94]

[94] These results are comparable to those found in a contemporaneous working paper by Lovenheim and Mumford (2011), which investigates the relationship between changes in home value and current period fertility using individual-level data from the Panel Study of Income Dynamics (PSID) from 1990-2007. The authors estimate linear probability models of the probability that a woman gives birth in a given year as a

An interesting empirical exercise is to consider the relative impact of unemployment rates versus housing prices. Using the same simulation procedure described above, we estimate the relative impacts of a one standard deviation increase in housing prices and decrease in unemployment rates. We find that at the mean rate of ownership (44 percent), a one standard deviation increase in housing prices leads to a 8.3 percent increase in births while a one standard deviation increase in unemployment rates leads to only a 2.1 percent decrease across all rates of ownership (note that this estimate is based on the point estimate in table 4.4, column (5). Even among renters, the negative price effect an increase in housing prices is 21 percent, 10 times as large as the effect of unemployment rates. This highlights the importance of considering housing markets in any empirical analysis of how economic conditions affect fertility outcomes.[95]

function of two and four year changes in the reported market value of her home. The authors find that a $10,000 increase in an individual's real housing wealth is associated with a 0.07 percentage point (1.3 percent at the mean) increase in the probability of having a child. It is also useful to compare our estimates to those found by Black et al. (2011), in their analysis of the effect of earnings on current period birth rates. Those authors use the experience of the coal boom and bust during the 1970s and 1980s in the Appalachian region of the U.S. to examine the effect of an exogenous increase in male's earnings (because females almost never work in the coal industry) on fertility rates. They estimate changes in county-level birth rates as a function of differences in lagged county level log earnings, conditional on state and year fixed effects. They find that a ten percent increase in county-level earnings is associated with a one percent increase in the subsequent year's birth rates. For earnings increases coming specifically from the coal boom, they estimate that a ten percent increase in coal-related earnings is associated with a seven percent increase in the subsequent year's birth rates. To put our findings in comparable terms, recall that we simulated that a $10,000 increase in house prices is associated with a 0.8 percent net increase in birth rates. Consider that a $10,000 increase in house prices is about a five percent increase off the 2006 median house price in the U.S. of roughly $260,000. So our estimates would suggest that a 10 percent increase in house prices would be associated with a 1.6 percent net increase in birth rates.

[95]Schaller (2011) provides the most up to date and arguably compelling empirical analysis of the relationship because local area unemployment rates and current period birth rates. Her analysis finds that a one percentage-point increase in unemployment rates is associated with a 0.7 to 2.5 percent decrease in birth rates, depending on specification. Our specification finds that a one percentage-point increase in unemployment rates is associated with a 1.37 percent decrease in birth rates, which is well within her range of estimates. It is this estimated coefficient that we translate into a standard deviation measurement to compare to the effect of house prices. Therefore, our conclusion that house prices exert a larger effect on birth rates than do unemployment rates would apply even if we took an estimate of the cyclicality of birth rates from outside our own analysis.

4.5.2 The Home Equity Effect

We have interpreted the positive effect of house price increases for owners – inferred from the estimated coefficient on the $HousePrice_{mt-1} * Own_{mg}$ interaction in the MSA-group level analyses and the individual-level CPS analysis – as a net positive housing wealth effect. As noted above, we use this term to encompass two potential mechanisms. First, there could be a traditional wealth effect that increases the demand for children. Second, there could be an increase in liquifiable housing wealth that otherwise credit-constrained consumers use to fund current period consumption, including child-related expenses. (Note that the first effect likely implies an increase in the number of children ever born, while the second effect could simply reflect a change in timing.) We are agnostic about the extent to which each of these two mechanisms is separately contributing to the empirical effect we observe in the data. However, it is interesting to consider the feasibility of the home equity extraction explanation.

Before examining data on home equity extraction, it is useful to place our finding in the context of the existing literature on consumption and housing wealth. There is a large body of research on the propensity for households to fund current consumption out of housing wealth (See for example, Case, Quigley, and Shiller (2005); Benjamin, Chinloy, and Jud (2004); Bostic, Gabrial, and Painter (2009); Haurin and Rosenthal (2005)). Most work in this literature finds that the propensity to consume out of housing wealth is substantially higher than the propensity to consume out of financial wealth, although there is disagreement about the magnitudes of these distinct marginal propensities to consume (Greenspan and Kennedy, 2007). It is also understood that there are two distinct effects of housing values on consumption – the traditional wealth effect and a home equity extraction effect. According to the Federal Reserve Survey of Consumer Finances, among families in the 40-60th percentile of the income distribution in 2004, housing represents an average of 48 percent of a household's total assets (Bucks et al., 2009). This indicates an increase in home prices could lead to a substantial increase in wealth for many homeowners. But, if households do not intend to realize these gains over their lifetime by selling their current house and moving to a lower-priced real estate market, there is not necessarily an increase in permanent wealth. However, there might still be an equity extraction effect. That effect is practically realized by refinancing one's mortgage, or obtaining a second mortgage, home equity loan, or home equity line of credit.

Greenspan and Kennedy (2007) empirically investigate the use of home equity extractions to fund consumption during the period 1991 to 2005, a similar time period to the one that we study in the present chapter. They report that during this period, free cash resulting from the three types of equity extraction averaged about $530 billion annually. Equity extracted through sales of existing homes accounted for about two-thirds of total free cash; home equity loans accounted for close to 20 percent, and cash-out refinancings about 13 percent. The extracted cash was used to finance consumer spending, outlays for home improvements, debt repayment, acquisition of assets, and other uses. In general, the use of housing equity to fund consumption is a relatively expensive approach. Hurst and Stafford (2004) propose that extracting home equity to fund consumption, as opposed to tapping into more liquid assets, should be relatively more common among individuals with lower amounts of liquid assets. They empirically confirm this pattern using individual-level data on households during the period 1991 to 1996 (Hurst and Stafford, 2004). Mian and Sufi (2009) estimate that the average homeowner extracted 25 to 30 cents for every dollar increase in home equity during the 1997 to 2009 period. They further find that money extracted from increased home equity was not used to purchase new real estate or pay down high credit card balances, which they interpret as suggesting that borrowed funds were used for consumption or home improvement expenses. In addition, they find that home equity-based borrowing was strongest among younger households. We view these patterns as being potentially consistent with the suggestion of our work that individuals used increased housing equity to pay for child-related expenses.

We use data from the 1997-2009 files of the American Housing Survey (AHS) to tabulate rates of home equity borrowing and refinancing, and the extent to which this is related to rising house prices. Our main goal is to simply observe the extent to which households are accessing their housing wealth. This speaks to the possibility that part of our documented "home equity effect" might be driven by the use of extracted home equity to fund current childbearing-related expenses. The AHS includes a survey of about 60,000 housing units across all 50 states and the District of Columbia. It is conducted every two years, in odd-numbered years.[96]

Table 4.10 reports the results from estimating regressions of the likelihood of having an equity loan/line or a refinanced mortgage on $HousePrice_{mt}$, as well as MSA-year means for the relevant variables. The top panel of the table reports rates of housing equity loans and lines of credit and mortgage refinancing. These questions are asked of home owners only. We see in the data that 20 percent of owners report having an equity loan or line of credit and as prices increase homeowners are more likely to tap into housing wealth in this matter. In column (3) we can see that this is being driven by increases in home equity lines of credit.

The AHS also gives information about rates of refinancing. The mean rate of having refinanced a first mortgage is 35 and the mean rate of having refinanced a second mortgage is 7 percent. The survey asks homeowners who report having refinanced why they chose to refinance, which is reported in the bottom panel of table 10. Eighty-six percent respond to obtain a lower interest rate. Lower interest rates leave people with lower monthly payments, which would make more disposable income available to fund current consumption. Interestingly, the only motivating factor in the decision to refinance which is increasing with house prices is explicitly "to get cash." This speaks directly to the use of housing equity to fund current consumption and its relationship to rising home prices.

[96]The AHS also includes a metropolitan area survey that has included varying numbers of areas and has been conducted at varying intervals over the past 30 years.

4.6 Conclusion

This chapter has investigated how current house prices affect current period fertility. Our results suggest that house prices are a relevant factor in a couple's decision to have a baby at the present time. House prices lead to a negative price effect that conditionally reduces birth rates in the current period, and an offsetting positive home equity effect that leads to a net increase in births among homeowners. We use the estimated coefficients from our regression analyses to simulate the effect of a \$10,000 increase in house prices on current year births. This exercise indicates that when home ownership rates reach 30 percent, the net effect becomes positive. At the mean U.S. home ownership in our sample period, the net effect of a \$10,000 increase in prices is a 0.8 percent increase in births. Given underlying differences in home ownership rates and heterogeneity in the point estimates, the predicted net effect of house price changes varies across race/ethnic groups. We simulate that a \$10,000 increase in MSA-level house prices leads to a 0.7 percent increase in current year births among whites, a 0.2 percent increase in births among blacks, and a 0.2 percent decrease in births among white Hispanics. Interestingly, these effects are substantially larger than the effects of changes in the unemployment rate. Moreover, using our estimates to make an out-of-sample prediction of the the impact of the "Great Recession", we find that the fall in housing prices between 2006 and 2010 was associated with a 7.5 percent decline in births.

Our work is written within the paradigm of the empirical literature on the cyclicality of fertility and as such, it is about the timing of fertility decisions. The evidence presented in this chapter suggests that couples use some of their increased housing wealth to "fund" their childbearing goals. This chapter potentially demonstrates empirically that (imperfect) credit markets affect fertility timing. We have discussed our results in terms of the decision couples make with regard to whether or not to have a baby in the current period. We leave it to future research to investigate how house prices affect completed fertility or the demand for children more generally. In addition, it might also be true that when house prices increase or decrease, parents increase (or decrease) quality investments in children, where quality of children is meant in the Beckerian sense. For example, perhaps some home-owning parents use their increased home equity to purchase, say, private education for their children. Once we allow for this possibility, it becomes clear that our empirical analysis is not designed to capture the full range of how real estate markets might affect childbearing and child rearing decisions.

4.7 Appendix

4.7.1 Metropolitan Statistical Areas

Metropolitan statistical areas are defined by the Office of Management and Budget. Their geographic definitions are based on core urban areas with a population of 50,000 or more and adjacent counties with a “high degree of social and economic integration (as measured by commuting to work) with the urban core” (Census Bureau documentation). Current metropolitan area definitions include both metropolitan areas (MSA) and divisions (MSADs), which are smaller units within this metropolitan area. Current definitions also include an alternative to the MSA/MSAD for metropolitan areas in the New England states, which are called New England City Town Areas (NECTAs). The boundaries of MSAs change over time as city populations change. The Office of Management and Budget releases revised definitions based on the decennial census and yearly census population estimates, and in addition to changing MSA compositions, sometimes changes the labels associated with each type of unit. A major change was done in 2003, at which point the coding system changed from a 4-digit coding system to a 5-digit coding system. Prior to 2003, instead of MSADs Primary Metropolitan Statistical Areas (PMSAs) were used and instead of NECTAs, New England Metropolitan County Areas (NECMAs) were used.

The housing price index is available at the level of MSA/MSAD, based on the November

2008 definitions (released in December 2009) , Since the index is based on repeat sales of the same home, the 2009 definitions apply throughout the data. For example, suppose a home sells once in 1980, 1990, and 2005. Suppose that in 1980 and 1990 it was not in an MSA, but in 2005 it was. Then, the home is considered part of the MSA and the housing price indices for 1980 and 1990 are revised to reflect the current boundaries.The rest of this appendix explains how we harmonize all other data sources to match this level of aggregation. Table 4.1 lists the level of geographic detail available for each of our control variables.

Whenever county level data is available, it is the preferred level of disaggregation because we can use it to construct MSA/MSADs which will exactly match the housing price index data. Data available at this level of disaggregation includes the Vital Statistics Natality Data (confidential files), Vital Statistics population data, Census median home value data, Bureau of Labor Statistics Unemployment data, Bureau of Economic analysis income per capita data, and the Rappaport and Sachs (2003) coastal measure. To construct MSAs from the county level data, we use the 2009 metropolitan area definition files available from the Census Bureau at: `http://www.census.gov/population/metro/files/lists/2009/List1.txt`. These files map entire counties to a 2009 OMB MSA/MSAD definitions, thus, we can construct MSAs/MSADs that are exactly equivalent to those used in the housing price data.

It is worth noting a few technical points about linking counties to MSAs. First, Miami-Dade County, FL was renamed between the 1990 and 2000 census; so in all cases we have assigned the post-2000 FIPS code to this county.[97] Another issue concerns BLS Local Area Unemployment (LAU) Statistics, which are calculated at the county level, but use a coding system based on what are called "areas". For the most part, the area codes are simply county FIPS codes. However, for counties which had large populations (50,000-100,000 and 100,000 plus) in 1970; a different coding system is applied.[98] We construct a crosswalk between

[97] See, for example,`http://www.census.gov/popest/archives/files/90s-fips.txt`
[98] See `http://www.bls.gov/lau/laucodes.htm`

the two using state FIPS codes and county names using vintage 2009 county FIPS codes.[99] Finally, in the BEA personal income data, BEA combines some counties/county equivalents in Virginia and assigns new county codes. We re-assign those counties which are contained within an MSA to one of the combined counties' FIPS code. In all cases these combinations were wholly contained within one MSA/MSAD.[100]

For the case when counties are not available, but vintage metropolitan area definitions are available, we use those. By vintage metropolitan area definitions, we are referring to metropolitan areas based on historical definitions which may differ in composition from the 2009 definitions. Data that is available in this manner includes the 1990 and 2000 Census microdata (used to construct home ownership rates), the Current Population Survey data (used to construct wages and fraction college educated), and the Saiz (2012) elasticity measure. The vintage definitions used in these data include the 1983 MSA/PMSA, 1993 MSA/PMSA, 1999 MSA/NECMA, and 2003 MSA/NECTA codes, as described in table 4.1.

To match the vintage definitions to the 2009 definitions, we begin by creating a crosswalk that links the counties that make up the different metropolitan areas over time. Unlike the current 2009 MSA/MSAD definitions (and vintage 2003 MSA/MSAD definitions) which directly map entire counties to MSAs, the earlier metropolitan area (and NECTA/NECMA) definitions allow for a single county to be in multiple metropolitan areas. For the case when a single county is in multiple MSAs/PMSAs/NECTAs/NECMAs, we use 1990 population counts of the minor civil divisions (a smaller unit within the metropolitan area) to assign the county to whichever MSAs/PMSAs/NECTAs/NECMAs the majority of the population resides.

From this county-msa crosswalk, we construct vintage MSA-to-2009 MSA/MSAD crosswalks. In most cases, there is a one to one match between the vintage MSA definitions and the 2009 definitions. In some cases, however, its possible for a vintage metropolitan area to have split into two or combined to form a single metropolitan area by 2009. For metropolitan

[99] `http://www.census.gov/popest/geographic/codes02.html`

[100] See `http://www.bea.gov/regional/docs/msalist.cfm`

areas that have combined to form one metropolitan area by 2009, we use 1990 population weights to create a population weighted average of the data. For metropolitan areas that have split , we apply the single data point to all the split-off areas.

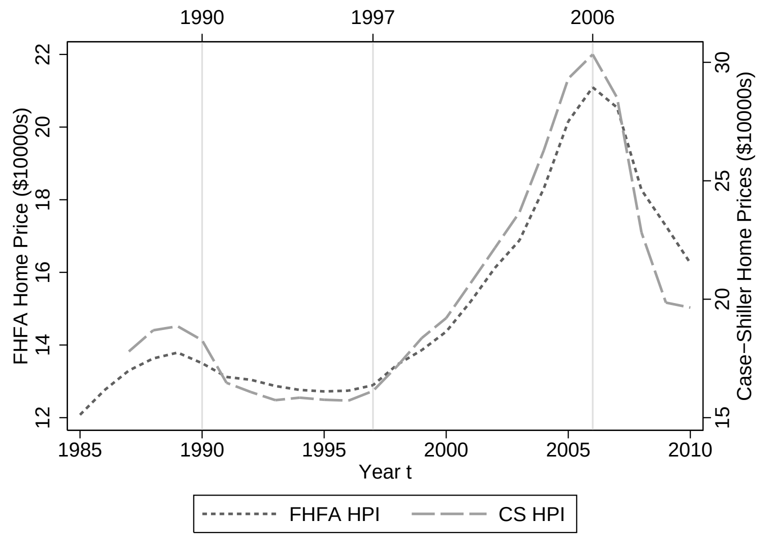
In the individual level data, we are given the vintage metropolitan area codes. In this case, we need to construct the housing price, wage, and unemployment data according to those definitions. In the individual CPS we are provided with 1983, 1993 and 2003 MSA/2003 NECTA codes and in the AHS we are provided with 1980 MSA codes. For the ACS, only PUMAs (Public Use Microdata Areas) are provided, however, IPUMS has created a cross-walk procedure and attached 1993 MSA codes, which we will use Ruggles et al. (2010). Recall the unemployment data is at the county level. In this case, we use county-to-vintage MSA cross walk described in the section above. For the wage data, linking to the CPS is trivial since it was constructed in the CPS and therefore uses the same MSA definitions. For linking the wage data to the ACS and for linking the housing data to the CPS, ACS and AHS , we again use the county-to-vintage MSA crosswalk described above. In this case, if multiple 1980/1983/1993/2003 MSA/2003 NECTA combine to form a single MSA in 2009, we assign the housing price data to each vintage MSAs. For the case when a single 1980/1983/1993/2003 MSA/2003 NECTA splits to form multiple MSAs in 2009, we we use 1990 population weights to assign a weighted average of home prices to the vintage metropolitan areas codes. Finally, since CPS uses different MSA codes over time which are not consistent, we use the linked 2009 MSA definition for the fixed effects. In the case where the vintage MSA split into multiple 2009 MSADs, we use assign the code of the MSAD with the largest population share.

4.7.2 Construction of House Prices

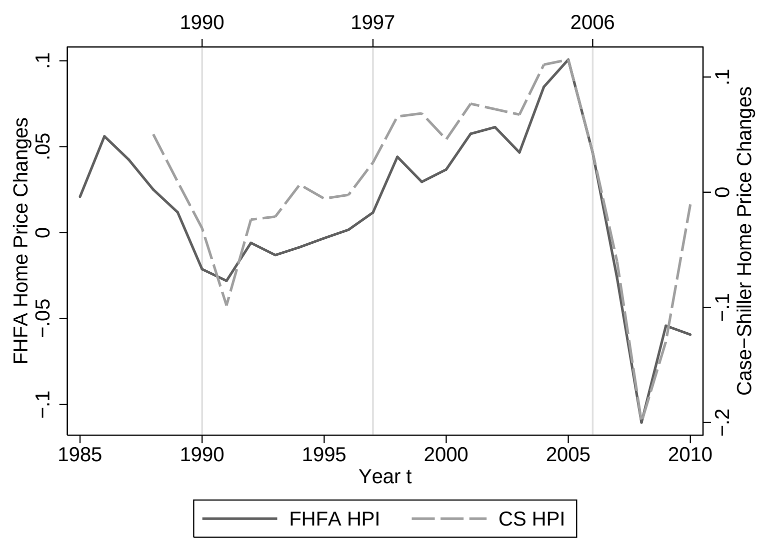
We use the same procedure used by Glaeser et al. (2008) to construct house prices. First, we construct a 2000 median home value from county-level census data, using the crosswalk procedure outlined above to create a population-weighted median home value. We inflate this value to 2006 dollars using the CPI-U "All Items-Less Shelter Series." We then take this value and scale it by the percent change in the housing price index from 2000 to the year of interest, which is calculated: $(hpi_t - hpi_{2000})/hpi_{2000}$. The housing price index is also inflated to 2006 dollars using the the CPI-U "All Items-Less Shelter Series" prior to scaling. This gives us a value that proxies for the price growth of a median value home in each MSA over time.

Figure 4.1: Housing Price Index (FHFA and Case-Shiller)

(a) House Prices

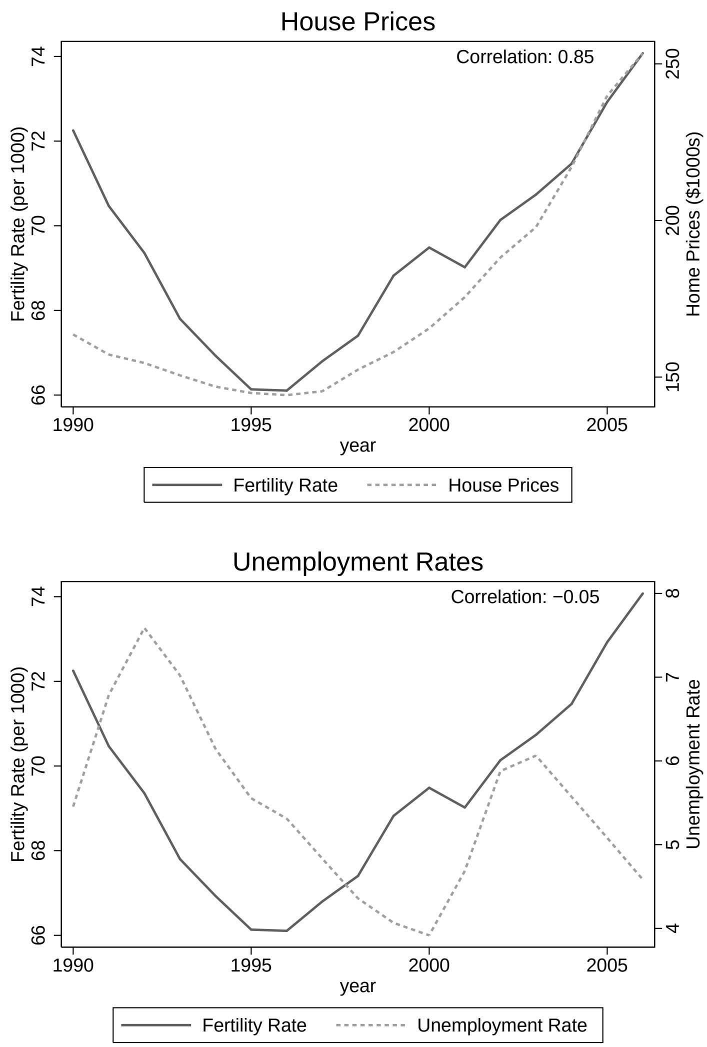


(b) Percentage Change House Prices Year t-1 to Year t



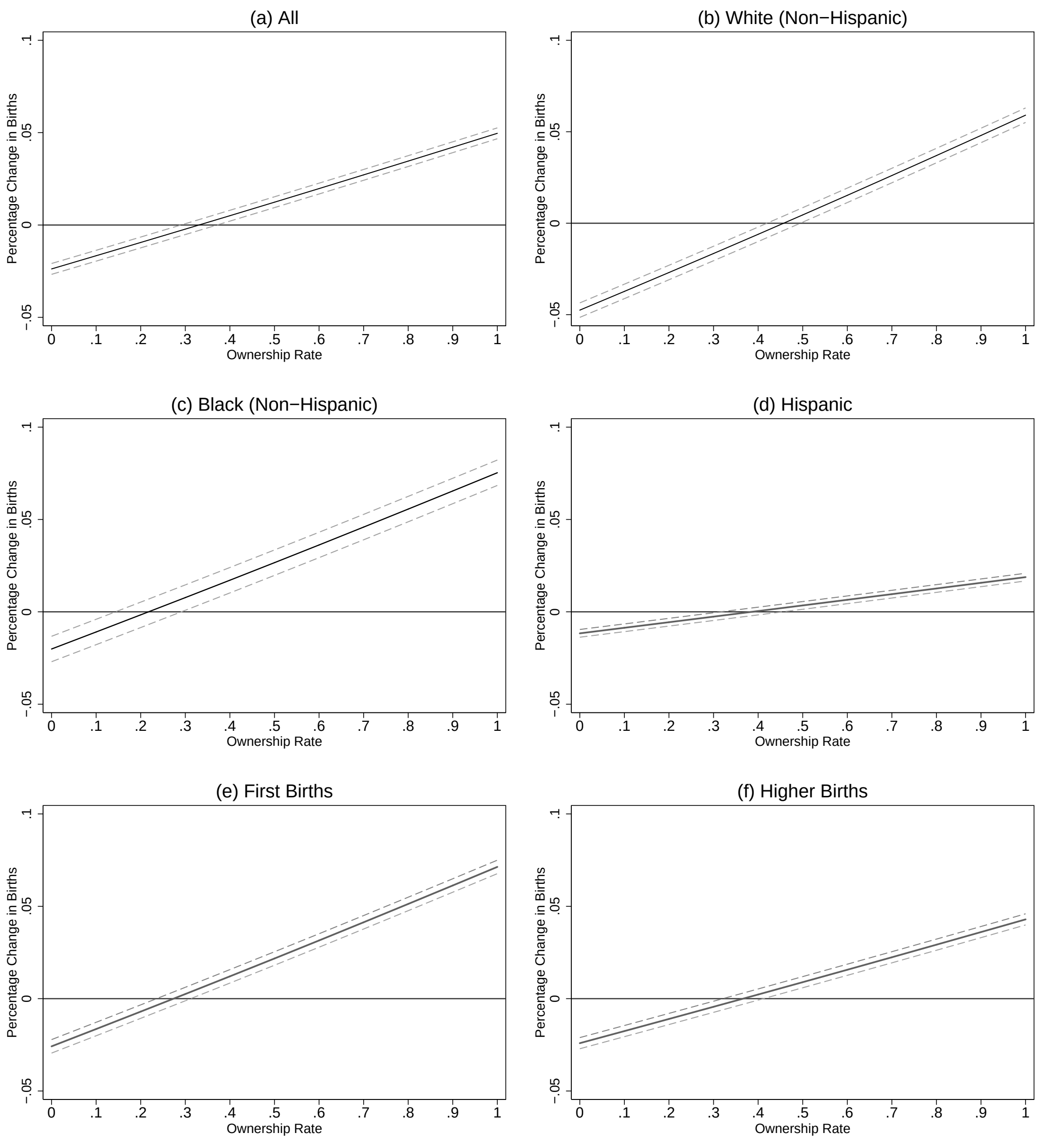
Notes: House prices are calculated using 2000 MSA median home values, which are scaled by either the FHFA house price Index or the Case-Shiller house price Index to create MSA-year median home values, which are then averaged over the 154 MSAs (27 MSAs for the Case-Shiller Index) in our sample each year 1984-2010. Both are adjusted to 2006 dollars using CPI-U "all items less shelter" series. Percentage change in home prices is calculated as $(HousePrice_t - HousePrice_{t-1})/HousePrice_{t-1}$. In both figures, the left y-axis is represents the mean value of the FHFA-constructed prices and the right y-axis represents the mean value of the Case-Shiller constructed prices.

Figure 4.2: Fertility Rates and Macro Indicators



Notes: Displayed are trends in fertility rates, housing prices, and unemployment rates. Annual fertility rates (births per 1000 women) are calculated using yearly totals of MSA-level births to women age 20-44 divided by total female population age 20-44, both obtained from the National Center for Health Statistics, National Vital Statistics System. House Prices are 2000 median home values scaled by the Federal Housing Finance Agency (FHFA) housing price Index, and are displayed in 2006 dollars. Unemployment rate is the annual mean unemployment are taken from Bureau of Labor Statistics local area unemployment statistics. All three measures are yearly mean values calculated based on the 154 MSAs in our sample.

Figure 4.3: Predicted Percentage Change in Births for a \$10,000 Increase in MSA Housing Prices



Notes: These figures display the results of simulation exercises using estimates from the group specific IV regression specifications displayed in Table 7. We predict the percentage change in fertility rates from a \$10,000 increase in mean housing prices for each ownership rate o displayed on the x axis: $(FertRate|HousePrice = h + 10k, OwnRate = o) - (FertRate|HousePrice = h, OwnRate = o)/(FertRate|HousePrice = h, OwnRate = o)$. For each group, we calculate the standard error of the prediction at the mean of the independent variables using 100 bootstrap replications and apply that standard error to calculate the confidence interval at each level of o.The solid line represents the predicted effect and the dashed line represents a 95% confidence interval, both of which were smoothed using a locally weighted linear regression.

Table 4.1: Aggregate Variables

Variable	Mean	Std Dev.	Source	Description	Geographic Detail
HPI	163.64	38.14	Federal Housing Finance Agency	House Price Index (All Transactions)	MSA Divisions (2009)
Home Price	\$162,356	\$89,700	Census (2000) and HPI	Average MSA home price in 2000 scaled by Housing Price Index (HPI) to create yearly series	County
Male Wages:			Current Population Survey	Individual wage and salary	Primary MSAs
25th Percentile Wage	\$12.94	\$2.33		income divided by the product	(1983, 1993, 2003)
50th Percential Wage	\$19.10	\$3.62		of weeks and hours worked	
75th Percentile	\$28.09	\$6.52		for full time working adult men	
Mean	\$23.17	\$4.38			
All Wages:				Individual wage and salary	
25th Percentile Wage	\$11.58	\$1.79		income divided by the product	
50th Percentile Wage	\$17.00	\$2.71		of weeks and hours worked	
75th Percentile	\$24.91	\$4.27		for all full time adult workers	
Unemployment Rate	4.84	1.74	Bureau of Labor Statistics Local Area Unemployment Statistics	Number of unemployed divided by the total labor force	County
Income Per Capita	\$39,740	\$7,149	Bureau of Economic Analysis Regional Economic Accounts	Sum of income from all sources divided by the total population	County
Average Rent	\$785	\$214	Department of Housing and Urban Development	Mean fair market rent for 0-4 bedroom residences	County
Housing Supply Elasticity	1.96	0.97	Saiz (2011)	Measure of elasticity of housing supply	Primary MSAs and NECMAs (1999)
Fraction College	0.19	0.19	Current Population Survey	Fraction of MSA-Group with a college degree	Primary MSAs (1983, 1993, 2003)

Notes: Listed are aggregate level variables and their means for the 154 MSAs used in the baseline specification. All variables are aggregated up the the MSA level from the level of geographic detail (column 6) available using the crosswalk procedure described in the text and data appendix. All nominal values are CPI adjusted to 2006 dollars.

Table 4.2: Summary Statistics

		Vital Statistics			Census	
	Fertility Rate (1000)	First Birth Fertility Rate	Higher Birth Fertility Rate	Home Ownership Rate 1990	Home Ownership Rate 2000	Min/Max Ownership Rate
All	70.09	24.60	45.49	0.44	0.44	[0.00, 0.80]
	(36.40)	(16.55)	(22.68)	(0.23)	(0.23)	
White 20-29	88.48	42.02	46.46	0.27	0.25	[0.10,0.39]
	(20.76)	(6.96)	(14.43)	(0.06)	(0.07)	
Black 20-29	118.00	40.19	77.81	0.08	0.10	[0.00,0.29]
	(17.18)	(5.11)	(16.78)	(0.03)	(0.03)	
Hispanic 20-29	153.78	54.70	99.09	0.14	0.15	[0.00,0.50]
	(31.03)	(9.82)	(24.71)	(0.06)	(0.06)	
White 30-44	48.28	14.84	33.44	0.67	0.68	[0.47,0.80]
	(8.30)	(4.62)	(4.74)	(0.07)	(0.08)	
Black 30-44	37.60	7.90	29.70	0.34	0.35	[0.06,0.60]
	(7.93)	(2.75)	(5.80)	(0.08)	(0.08)	
Hispanic 30-44	59.69	10.81	48.88	0.40	0.41	[0.00,0.80]
	(11.19)	(3.07)	(10.69)	(0.13)	(0.13)	

Notes: Fertility rates are total births over the total female population in each MSA, year of conception, age category and race/ethnicity cell for women age 20-44. Mean home ownership rates are calculated in 1990 Census by year, msa, age category, and race/ethnicity. Min/Max home ownership rates are based on 1990 data.Sources for aggregate birth data and population data is Vital Statistics birth certificate data (1997-2007) and population data (1996-2006), and for home ownership data is the decennial Census (1990 and 2000). All means displayed are population weighted. Standard deviations are in parentheses.

Table 4.3: Housing Prices and Fertility Rates: 1997-2006

	(1)	(2)	(3)	(4)	(5)
Dep. Var. $log(FertRate)_{mgt}$					
$HousePrice_{mt-1} * OwnRate_{mg}$	0.0468***	0.0468***	0.0481***	0.0485***	0.0276***
	(0.00443)	(0.00448)	(0.00481)	(0.00488)	(0.00375)
$HousePrice_{mt-1}$	-0.0124***	-0.0128***	-0.0111***	-0.0160***	-0.00509***
	(0.00103)	(0.00115)	(0.00243)	(0.00195)	(0.00184)
$OwnRate_{mg}$	0.0545	0.0542	0.0460	0.0445	0.0852
	(0.318)	(0.318)	(0.320)	(0.323)	(0.267)
$WhiteAge20-29$	0.948***	0.949***	0.953***	0.953***	0.902***
	(0.128)	(0.128)	(0.127)	(0.128)	(0.120)
$BlackAge20-29$	1.335***	1.334***	1.337***	1.335***	1.255***
	(0.185)	(0.185)	(0.184)	(0.186)	(0.174)
$HispanicAge20-29$	1.569***	1.569***	1.570***	1.567***	1.500***
	(0.164)	(0.164)	(0.163)	(0.164)	(0.164)
$BlackAge30-44$	0.0112	0.0111	0.0121	0.00995	-0.0484
	(0.0749)	(0.0748)	(0.0743)	(0.0749)	(0.0680)
$HispanicAge30-44$	0.371***	0.371***	0.369***	0.365***	0.328***
	(0.0790)	(0.0790)	(0.0784)	(0.0791)	(0.0599)
$FracColl_{mgt-1}$	-0.349***	-0.352***	-0.375***	-0.395***	-0.358***
	(0.0641)	(0.0648)	(0.0682)	(0.0710)	(0.0679)
$UnempRate_{mt-1}$		-0.00255	-0.00116	-0.00156	-0.00155
		(0.00335)	(0.00211)	(0.00204)	(0.00200)
$25thWage_{mt-1}$		0.00138	0.000674	0.000237	0.000744
		(0.00106)	(0.000660)	(0.000550)	(0.000621)
$50thWage_{mt-1}$		0.000166	0.000978*	0.000475	0.000839*
		(0.000904)	(0.000534)	(0.000424)	(0.000480)
$75thWage_{mt-1}$		0.000623	0.000241	0.000182	0.000199
		(0.000552)	(0.000236)	(0.000216)	(0.000230)
MSA Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
MSA Trends	No	No	Yes	Yes	Yes
MSA Quadratic	No	No	No	Yes	No
MSA-Own Category Trends	No	No	No	No	Yes
R^2	0.908	0.908	0.910	0.910	0.937
Number of MSAs	154	154	154	154	154
N	9240	9240	9240	9240	9240

Notes: Fertility rates are total births over the total female population in each MSA, year of conception, age category and race/ethnicity cell for women age 20-44. Mean home ownership rates are calculated in 1990 Census by year, msa, age category, and race/ethnicity. Fraction of cell that is a college graduate is matched by msa, year, age category and race. House prices (10,000s), unemployment rates, and male wages are matched by msa and year of conception. Data sources are described in the text. All specifications are weighted by the total number of births in the cell. Standard errors adjusted for clustering at the msa level are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 4.4: Housing Prices and Fertility Rates by MSA Supply Elasticity: 1997-2006

	(1) Low Elasticity	(2) High Elasticity	(3) First Stage 1: Hprice*Own	(4) First Stage 2: Hprice	(5) IV
$HousePrice_{mt-1} * OwnRate_{mg}$	0.0428***	0.0609***			0.0723***
	(0.00448)	(0.0128)			(0.00997)
$HousePrice_{mt-1}$	-0.0104***	-0.0170***			-0.0239***
	(0.00117)	(0.00498)			(0.00328)
$Elasticity_m * HPI_t * OwnRate_{mg}$			-0.0137***	-0.000647**	
			(0.00200)	(0.000287)	
$Elasticity_m * HPI_t$			-0.0111***	-0.0571***	
			(0.00186)	(0.00956)	
$OwnRate_{mg}$	0.143	-0.394	17.93***	0.193	-0.171
	(0.433)	(0.363)	(2.332)	(0.353)	(0.291)
$FracColl_{mgt-1}$	-0.419***	-0.111	4.995***	0.674	-0.492***
	(0.0856)	(0.0762)	(0.804)	(0.673)	(0.102)
$UnempRate_{mt-1}$	-0.00644	0.0000946	-0.111	-0.994***	-0.0137**
	(0.00459)	(0.00270)	(0.0887)	(0.321)	(0.00563)
$25thWage_{mt-1}$	0.00253	-0.00106	-0.0147	0.0338	0.00226
	(0.00152)	(0.000871)	(0.0319)	(0.101)	(0.00166)
$50thWage_{mt-1}$	0.00134	-0.00117*	0.0154	0.116	0.00153
	(0.00141)	(0.000674)	(0.0268)	(0.0786)	(0.00115)
$75thWage_{mt-1}$	0.00111	0.000417	-0.0000601	0.0304	0.00120*
	(0.000790)	(0.000314)	(0.0188)	(0.0434)	(0.000704)
MSA Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
R^2	0.900	0.939	0.876	0.932	0.897
F Statistic			29.60	18.32	
Number of MSAs	77	77	154	154	154
N	4620	4620	9240	9240	9240

Notes: Fertility rates are total births over the total female population in each MSA, year of conception, age category and race/ethnicity cell for women age 20-44. Mean home ownership rates are calculated in 1990 Census by year, msa, age category, and race/ethnicity. Fraction of cell that is a college graduate is matched by msa, year, age category and race. House prices (10,000s) are matched by msa and year of conception (or years prior to conception where noted). Elasticity refers to the Saiz (2011) supply elasticity measure and HPI refers to the national version of the FHFA house price index. First stage 1 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1} * OwnRate_{mg}$ and first stage 2 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1}$. All regression include group, MSA and year fixed effects, as well as MSA-year unemployment rates and male wages. Data sources are described in the text. All specification are weighted by the total number of births in the cell. Robust standard errors are in parentheses and clustered at the msa level. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 4.5: Alternative House Price Measures

	(1) $Price_{mt-1}$	(2) $Price_{mt-2}$	(3) $Price_{mt-3}$	(4) $Price_{mt-4}$	(5) $AvgPrice_{mt-2}$	(6) $AvgPrice_{mt-3}$	(7) $AvgPrice_{mt-4}$
Dep. Var. $log(FertRate)_{mgt}$							
OLS							
$HousePrice_{mt-1} * OwnRate_{mg}$	0.0468***	0.0534***	0.0630***	0.0718***	0.0501***	0.0541***	0.0581***
	(0.0045)	(0.0052)	(0.0063)	(0.0077)	(0.0048)	(0.0052)	(0.0057)
$HousePrice_{mt-1}$	-0.0128***	-0.0145***	-0.0172***	-0.0199***	-0.0136***	-0.0146***	-0.0157***
	(0.0012)	(0.0013)	(0.0017)	(0.0023)	(0.0012)	(0.0014)	(0.0015)
IV							
$HousePrice_{mt-1} * OwnRate_{mg}$	0.0723***	0.0823***	0.0919***	0.0995***	0.0770***	0.0814***	0.0853***
	(0.0100)	(0.0111)	(0.0120)	(0.0127)	(0.0105)	(0.0109)	(0.0113)
$HousePrice_{mt-1}$	-0.0239***	-0.0274***	-0.0313***	-0.0353***	-0.0255***	-0.0272***	-0.0288***
	(0.0033)	(0.0036)	(0.0040)	(0.0045)	(0.0034)	(0.0036)	(0.0038)
$OwnRate_{mg}$	-0.171	-0.301	-0.423	-0.509*	-0.232	-0.288	-0.336
	(0.291)	(0.282)	(0.275)	(0.271)	(0.287)	(0.283)	(0.280)
$UnempRate_{mt-1}$	-0.0137**	-0.0137**	-0.0120**	-0.00893*	-0.0138**	-0.0134**	-0.0126**
	(0.00563)	(0.00573)	(0.00547)	(0.00494)	(0.00567)	(0.00559)	(0.00542)
$25thWage_{mt-1}$	0.00226	0.00222	0.00211	0.00202	0.00224	0.00220	0.00216
	(0.00166)	(0.00164)	(0.00155)	(0.00146)	(0.00165)	(0.00162)	(0.00158)
$50thWage_{mt-1}$	0.00153	0.00145	0.00134	0.00137	0.00149	0.00144	0.00142
	(0.00115)	(0.00115)	(0.00109)	(0.00103)	(0.00115)	(0.00112)	(0.00109)
$75thWage_{mt-1}$	0.00120*	0.00121*	0.00120*	0.00113**	0.00120*	0.00120*	0.00118*
	(0.000704)	(0.000701)	(0.000649)	(0.000561)	(0.000702)	(0.000685)	(0.000653)
$FracColl_{mgt-1}$	-0.492***	-0.505***	-0.502***	-0.491***	-0.498***	-0.499***	-0.497***
	(0.102)	(0.102)	(0.0991)	(0.0951)	(0.102)	(0.101)	(0.0997)
Group Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F Stat 1	29.60	27.83	25.94	23.78	28.87	28.24	27.67
F Stat 2	18.32	18.69	18.84	18.23	18.61	18.91	19.27
Number of MSAs	163	163	163	163	163	163	163
N	9780	9780	9780	9780	9780	9780	9780

Notes: Fertility rates are total births over the total female population in each MSA, year of conception, age category and race/ethnicity cell for women age 20-44. Mean home ownership rates are calculated in 1990 Census by year, msa, age category, and race/ethnicity. Fraction of cell that is a college graduate is matched by msa, year, age category and race. House prices (10,000s) are matched by msa and year of conception (or years prior to conception where noted). Average house price refers to the average home price from the year indicated up to the year of conception. The instrumental variable is the interaction between Saiz (2011) supply elasticity measure and the national version of the FHFA house price index. First stage 1 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1} * OwnRate_{mg}$ and first stage 2 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1}$. F Stat 1 refers to the first stage F statistic where the dependent variable is $HousePrice_{mt-1} * OwnRate_{mg}$, and F Stat 2 for $HousePrice_{mt-1}$. All regression include group, MSA and year fixed effects, as well as MSA-year unemployment rates and male wages. Data sources are described in the text. All specification are weighted by the total number of births in the cell. Robust standard errors are in parentheses and clustered at the msa level. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 4.6: Alternative Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Var. $log(FertRate)_{mgt}$							
OLS							
$HousePrice_{mt-1} * OwnRate_{mg}$	0.0468***	0.0468***	0.0468***	0.0468***	0.0389***	0.0469***	0.0432***
	(0.00448)	(0.00447)	(0.00447)	(0.00446)	(0.00497)	(0.00449)	(0.00437)
$HousePrice_{mt-1}$	-0.0128***	-0.0128***	-0.0127***	-0.0128***	-0.0101***	-0.0126***	-0.0122***
	(0.00115)	(0.00115)	(0.00116)	(0.00119)	(0.00123)	(0.00109)	(0.00136)
IV							
$HousePrice_{mt-1} * OwnRate_{mg}$	0.0723***	0.0723***	0.0723***	0.0722***	0.0739***	0.0724***	0.0599***
	(0.00997)	(0.00997)	(0.00996)	(0.00995)	(0.0128)	(0.00999)	(0.0102)
$HousePrice_{mt-1}$	-0.0239***	-0.0239***	-0.0239***	-0.0240***	-0.0244***	-0.0255***	-0.0196***
	(0.00328)	(0.00325)	(0.00328)	(0.00334)	(0.00412)	(0.00343)	(0.00327)
$OwnRate_{mg}$	-0.171	-0.171	-0.171	-0.171	-0.0915	-0.171	-0.0546
	(0.291)	(0.292)	(0.291)	(0.291)	(0.349)	(0.292)	(0.223)
$UnempRate_{mt-1}$	-0.0137**	-0.0139**	-0.0138**	-0.0122**	-0.0138**	-0.0172***	-0.00481
	(0.00563)	(0.00568)	(0.00573)	(0.00530)	(0.00570)	(0.00546)	(0.00414)
$FracColl_{mgt-1}$	-0.492***	-0.491***	-0.490***	-0.486***	-0.493***	-0.497***	-0.366***
	(0.102)	(0.102)	(0.101)	(0.101)	(0.0977)	(0.103)	(0.109)
$25thWage_{mt-1}$	0.00226				-0.00341	0.00257	0.000622
	(0.00166)				(0.00769)	(0.00172)	(0.00105)
$50thWage_{mt-1}$	0.00153				0.00752	0.00144	0.00104
	(0.00115)				(0.00550)	(0.00117)	(0.00108)
$75thWage_{mt-1}$	0.00120*				0.000790	0.00103	0.000670
	(0.000704)				(0.00501)	(0.000738)	(0.000669)
$25thWageAll_{mt-1}$		0.00468**					
		(0.00231)					
$50thWageAll_{mt-1}$		-0.000306					
		(0.00185)					

$75thWage_{mt-1}$		0.00262**					
		(0.00121)					
$MeanWage_{mt-1}$			0.00329***				
			(0.00120)				
$IncomePC_{mt-1}$				0.00603*			
				(0.00328)			
$25thWage_{mt-1} * OwnRate_{mg}$					0.0147		
					(0.0189)		
$50thWage_{mt-1} * OwnRate_{mg}$					-0.0156		
					(0.0148)		
$75thWage_{mt-1} * OwnRate_{mg}$					0.00106		
					(0.0124)		
$AvgRent_{mt-1}$						0.000158*	
						(0.0000831)	
Group Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA Boundaries Constant	No	No	No	No	No	No	Yes
F Stat 1	29.60	29.76	29.81	28.09	28.27	28.33	32.52
F Stat 2	18.32	18.26	18.44	16.09	18.25	17.95	22.41
No. of MSAs	154	154	154	154	154	154	108
N	9240	9240	9240	9240	9240	9240	6480

Notes: Fertility rates are total births over the total female population in each MSA, year of conception, age category and race/ethnicity cell for women age 20-44. Mean home ownership rates are calculated in 1990 Census by year, msa, education category, age category, and race. Fraction of cell that is a college graduate is matched by msa, year, age category and race. House prices (10,000s), Income per capita, Male wages, All wages, and average rent are matched by MSA and year of conception. All regressions include group, MSA and year fixed effects. In column (7) the sample is limited to MSAs with boundaries that do not change 1990-2006. The instrumental variable is the interaction between Saiz (2011) supply elasticity measure and the national version of the FHFA house price index. First stage 1 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1} * OwnRate_{mg}$ and first stage 2 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1}$. Data sources are described in the text. All specifications are weighted by the total number of births in the cell. Standard errors adjusted for clustering at the msa level are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 4.7: Different Groups

	(1) All	(2) White	(3) Black	(4) Hispanic	(5) First Births	(6) Second Births	(7) Higher Births
Dep. Var. $log(FertRate)_{mgt}$							
OLS							
$HousePrice_{mt-1} * OwnRate_{mg}$	0.0468***	0.0733***	0.0788***	0.0220***	0.0538***	0.0474***	0.0434***
	(0.00448)	(0.00762)	(0.00922)	(0.00302)	(0.00522)	(0.00409)	(0.00551)
$HousePrice_{mt-1}$	-0.0128***	-0.0306***	-0.0142***	-0.00690***	-0.0122***	-0.0118***	-0.0152***
	(0.00115)	(0.00286)	(0.00141)	(0.00196)	(0.00121)	(0.00117)	(0.00169)
IV							
$HousePrice_{mt-1} * OwnRate_{mg}$	0.0723***	0.106***	0.0926*	0.0305***	0.0947***	0.0754***	0.0574***
	(0.00997)	(0.0133)	(0.0483)	(0.00999)	(0.0129)	(0.0114)	(0.00988)
$HousePrice_{mt-1}$	-0.0239***	-0.0485***	-0.0201***	-0.0116***	-0.0260***	-0.0230***	-0.0256***
	(0.00328)	(0.00598)	(0.00694)	(0.00339)	(0.00386)	(0.00372)	(0.00382)
$OwnRate_{mg}$	-0.171	1.746***	-2.332***	-0.746***	-0.149	-0.0477	-0.423
	(0.291)	(0.613)	(0.486)	(0.141)	(0.311)	(0.415)	(0.299)
$UnempRate_{mt-1}$	-0.0137**	-0.0158**	0.000361	0.00269	-0.0237***	-0.0111**	-0.00947
	(0.00563)	(0.00703)	(0.00810)	(0.00496)	(0.00631)	(0.00513)	(0.00757)
$25thWage_{mt-1}$	0.00226	0.00131	0.00350*	0.000918	0.00294	0.00185	0.00254
	(0.00166)	(0.00157)	(0.00186)	(0.00302)	(0.00204)	(0.00155)	(0.00223)
$50thWage_{mt-1}$	0.00153	0.00118	0.000151	0.00185	0.00177	0.00126	0.00168
	(0.00115)	(0.00131)	(0.00145)	(0.00297)	(0.00122)	(0.00112)	(0.00185)
$75thWage_{mt-1}$	0.00120*	0.00111	0.00114**	0.000876	0.000955	0.00109**	0.00164
	(0.000704)	(0.000944)	(0.000499)	(0.00158)	(0.000734)	(0.000553)	(0.00121)
Group fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F Stat 1	29.60	24.57	15.81	22.57 29.54	30.59	27.60	
F Stat 2	18.32	18.58	17.19	20.71 19.26	18.86	16.32	
N	9240	3080	3080	3080	9234	9240	9240

Notes: Fertility rates are total births over the total female population in each MSA, year of conception, age category and race/ethnicity cell for women age 20-44. Mean home ownership rates are calculated in 1990 Census by year, msa, age category, and race/ethnicity.House prices (10,000s), unemployment rates, and male wages are matched by msa and year of conception. The instrumental variable is the interaction between Saiz (2011) supply elasticity measure and the national version of the FHFA house price index. First stage 1 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1} * OwnRate_{mg}$ and first stage 2 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1}$. All regression include group, MSA and year fixed effects, as well as MSA-year unemployment rates and male wages. Data sources are described in the text. All specifications are weighted by the total number of births in the cell. Standard errors adjusted for clustering at the msa level are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 4.8: Individual Level Analysis Using CPS: 1997-2006

	(1) No MSA/Year FE	(2) MSA/Year FE	(3) MSA/Year FE	(4) IV
Dep. Var. $Pr(Birth)_i$				
$HousePrice_{mt-1} * Own_i$	0.000631***	0.000569***	0.000567***	0.000427
	(0.000197)	(0.000197)	(0.000197)	(0.000409)
$HousePrice_{mt-1}$	-0.000134	-0.000307**	-0.000331**	-0.0000331
	(0.0000985)	(0.000131)	(0.000138)	(0.000337)
Own_i	0.0367***	0.0377***	0.0378***	0.0403***
	(0.00385)	(0.00393)	(0.00393)	(0.00782)
$UnempRate_{mt-1}$			0.000156	0.000541
			(0.000650)	(0.000735)
$25thWage_{mt-1}$			-0.000382	-0.000350
			(0.000639)	(0.000637)
$50thWage_{mt-1}$			0.000361	0.000338
			(0.000551)	(0.000551)
$75thWage_{mt-1}$			0.000229	0.000197
			(0.000274)	(0.000280)
Demographics	Yes	Yes	Yes	Yes
MSA fixed Effects	No	Yes	Yes	Yes
Year fixed Effects	No	Yes	Yes	Yes
IV	No	No	No	Yes
Mean Had Baby	0.063	0.063	0.063	0.063
Mean Own	0.498	0.498	0.498	0.498
F Stat 1				25.64
F Stat 2				17.08
N	192788	192788	192788	192788

Notes: Sample is women age 20-44 in March Current Population Survey 1998-2007. Dependent variable is an indicator for having a child under one. House prices (10,000s), unemployment rates, and male wages are matched by msa and year. Ownership is the household's home ownership status, which is assigned as a 1 when the household owns a home and the respondent is the household head or spouse of the household head. All regressions include fixed effects for education, year, age category, race, Hispanicity, and msa. The instrumental variable is the interaction between Saiz (2011) supply elasticity measure and the national version of the FHFA house price index. First stage 1 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1} * OwnRate_{mg}$ and first stage 2 refers to the first stage regression where the dependent variable is $HousePrice_{mt-1}$. Data sources are described in the text. Standard errors adjusted for clustering at the msa level are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 4.9: Real Estate Bust Periods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Dep. Var., Source	log(FertRate), Aggregate-90-96		Pr(Birth), CPS-90-96		Pr(Birth), CPS-07-09		Pr(Birth,) ACS-07-09	
$HousePrice_{mt-1} * Own$	0.0604***	0.0722***	0.000707*	0.000741	0.000703***	0.000978***	0.000445***	0.000463**
	(0.00765)	(0.00962)	(0.000402)	(0.000587)	(0.000224)	(0.000353)	(0.000115)	(0.000182)
$HousePrice_{mt-1}$	-0.0117**	-0.0206***	-0.000560	-0.000379	-0.000689	-0.00110	0.0000239	0.000410
	(0.00526)	(0.00732)	(0.000501)	(0.00199)	(0.000480)	(0.00122)	(0.000242)	(0.000344)
Own	0.103	0.0170	0.0362***	0.0357***	0.0318***	0.0258***	0.0374***	0.0370***
	(0.365)	(0.343)	(0.00511)	(0.00770)	(0.00496)	(0.00711)	(0.00213)	(0.00330)
$UnempRate_{mt-1}$	-0.00761**	-0.0127*	-0.000774	-0.000649	-0.000322	-0.000763	0.000573	0.00113*
	(0.00306)	(0.00698)	(0.000971)	(0.00155)	(0.00145)	(0.00220)	(0.000487)	(0.000603)
$25thWage_{mt-1}$	-0.000971	-0.00130	-0.000730	-0.000725	0.000577	0.000571	-0.0000345	-0.0000391
	(0.00125)	(0.00151)	(0.000753)	(0.000752)	(0.00128)	(0.00128)	(0.000356)	(0.000348)
$50thWage_{mt-1}$	0.00203*	0.00269*	0.000514	0.000502	-0.000342	-0.000383	-0.000393	-0.000329
	(0.00121)	(0.00138)	(0.000726)	(0.000733)	(0.00118)	(0.00117)	(0.000349)	(0.000350)
$75thWage_{mt-1}$	0.000598	0.000663	0.000464	0.000464	-0.000485	-0.000458	0.000243	0.000225
	(0.000796)	(0.000745)	(0.000487)	(0.000485)	(0.000602)	(0.000605)	(0.000173)	(0.000178)
Group fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
MSA fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Fert Rate	0.069	0.069	0.064	0.064	0.066	0.066	0.058	0.058
Mean Own	0.45	0.45	0.43	0.43	0.46	0.46	0.45	0.45
F Stat 1		22.55		15.88		36.93		45.46
F Stat 2		8.48		3.94		13.88		27.55
N	6348	6348	126799	126799	71317	71317	801798	801798

Notes: In columns (1)-(2) the sample is all births to women age 20-44 for the bust period of 1990-1996 according to the specification in table 3, column (6) and the dependent variable is the fertility rate. In column (3)-(4), the sample is women 20-44 in the March Current Population Survey for the bust period 1990-1996 and in column (5)-(6) for the bust period 2007-2009. In column (7)-(8) the sample is women 20-44 in the American Communities Survey for the bust period 2007-2009. The dependent variable in columns (2)-(8) is an indicator for having a child under one. In column (1)-(2) ownership rates are matched by MSA, age category and race/ethnicity. In columns (3)-(8), ownership is the household's home ownership status, which is assigned as a 1 when the household owns a home and the respondent is the household head or spouse of the household head. House prices (10,000s), unemployment rates, and male wages are matched by msa and year. Column (1)-(2) includes group, MSA and year fixed effects. Columns (3)-(8) include fixed effects for MSA, year, race/ethnicity, age category and education. The instrumental variable is the Saiz (2011) supply elasticity measure interacted with the national version of the FHFA house price index. F Stat 1 refers to the first stage F statistic where the dependent variable is $HousePrice_{mt-1} * OwnRate_{mg}$, and F Stat 2 for $HousePrice_{mt-1}$. Data sources are described in the text. Standard errors adjusted for clustering at the msa level are in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$

Table 4.10: House Prices and Home Equity Withdrawal Behavior

	Dependent Variable: Does the Home Owner Currently Have a....?				
	Equity Loan or Line of Credit	Equity Line of Credit	Equity Loan	First Mortgage Refinanced	Second Mortgage Refinanced
$HousePrice_{mt}$	0.00182**	0.00299***	-0.000478	0.00269*	-0.00171
	(0.000752)	(0.000890)	(0.000509)	(0.00145)	(0.00219)
Mean Dep Var.	0.196	0.185	0.0835	0.349	0.0755
	(0.00272)	(0.00329)	(0.00193)	(0.00449)	(0.00642)
R^2	0.055	0.262	0.026	0.104	0.072
N	21209	13948	20527	11289	1696
	Dependent Variable: Why Did You Refinance Your First Mortgage?				
	Lower Interest Rate	To Get Cash	Renew or EYestend	Increase Payments	Reduce Payments
$HousePrice_{mt}$	-0.00319	0.00491***	-0.000503	0.000822	0.0000489
	(0.00199)	(0.00187)	(0.000527)	(0.000830)	(0.00147)
Mean Dep. Var	0.857	0.131	0.00965	0.0221	0.104
	(0.00558)	(0.00537)	(0.00156)	(0.00234)	(0.00487)
R^2	0.107	0.053	0.043	0.043	0.068
N	3937	3937	3937	3937	3937

Notes: Displayed is the coefficient of MSA-year house prices on the probability of making different types of home equity withdrawals, as well as the mean of each dependent variable. All regressions include control for msa, year, race, ethnicity, and age. House prices (10,000s) are matched by MSA and year. Source is American Housing Survey, National Version, Every other year 1997-2009. Refinancing is only available in 2001-2009. Questions only asked of home owners. Data is in panel form and respondents are asked if they have ever done the specific activity. For instance, the question will ask, "is the respondents first mortgage a refinancing of a previous mortgage?" "Why refinance?" is only asked for those who have refinanced. The categories are not mutually exclusive and respondents may respond yes to multiple categories.

Table 4.11: Characteristics of Metropolitan Areas in the Sample

Metropolitan Area Name (2009 MSAD)	Percent Change Prices 97-06	Home Price 2006	Fertility Rate 2006	Elasticity of Supply	Unemp Rate 2006	Median Wage 2006
Salinas, CA	171.4%	$ 588,736	103.5	1.10	6.92	$ 15.83
Santa Barbara-Santa Maria-Goleta, CA	165.5%	$ 628,696	86.3	0.89	4.04	$ 12.02
Riverside-San Bernardino-Ontario, CA	162.7%	$ 338,547	90.6	0.94	4.92	$ 16.83
Los Angeles-Long Beach-Glendale, CA	162.6%	$ 515,061	76.1	0.63	4.78	$ 16.83
Vallejo-Fairfield, CA	151.9%	$ 398,870	77.8	1.14	4.87	$ 18.27
San Diego-Carlsbad-San Marcos, CA	151.4%	$ 474,242	81.1	0.67	3.96	$ 19.23
Oxnard-Thousand Oaks-Ventura, CA	148.7%	$ 556,284	86.3	0.75	4.30	$ 19.23
Fort Lauderdale-Pompano Beach-Deerfield Beach, FL	146.5%	$ 267,370	67.9	0.65	3.07	$ 15.11
Stockton, CA	146.4%	$ 331,468	92.6	2.07	7.42	$ 21.31
Modesto, CA	145.8%	$ 314,814	90.7	2.17	7.96	$ 18.69
Oakland-Fremont-Hayward, CA	144.1%	$ 564,108	74.7	0.70	4.37	$ 24.04
Miami-Miami Beach-Kendall, FL	142.6%	$ 295,201	71.4	0.60	4.08	$ 15.11
West Palm Beach-Boca Raton-Boynton Beach, FL	140.7%	$ 290,506	76.3	0.83	3.64	$ 15.11
Santa Rosa-Petaluma, CA	131.4%	$ 510,412	73.5	1.00	3.99	$ 28.00
Cape Coral-Fort Myers, FL	131.3%	$ 241,947	86.4	1.28	2.88	$ 17.79
Fresno, CA	126.2%	$ 264,992	97.8	1.84	8.01	$ 16.17
North Port-Bradenton-Sarasota, FL	125.7%	$ 247,504	75.6	0.92	3.06	$ 19.40
Port St. Lucie, FL	125.4%	$ 238,298	79.9	1.19	3.89	$ 17.55
Bakersfield-Delano, CA	123.9%	$ 230,694	105.8	1.64	7.54	$ 18.33
San Francisco-San Mateo-Redwood City, CA	122.9%	$ 781,891	64.0	0.66	3.89	$ 24.04
Deltona-Daytona Beach-Ormond Beach, FL	117.7%	$ 194,512	65.8	1.07	3.24	$ 16.83
Washington-Arlington-Alexandria, DC-VA-MD-WV	117.2%	$ 401,637	74.7	1.61	3.13	$ 23.08
Palm Bay-Melbourne-Titusville, FL	117.0%	$ 210,691	67.1	1.04	3.23	$ 17.55
San Jose-Sunnyvale-Santa Clara, CA	116.9%	$ 698,468	82.6	0.76	4.55	$ 24.04
Tampa-St. Petersburg-Clearwater, FL	112.7%	$ 186,334	69.3	1.00	3.43	$ 19.23
Orlando-Kissimmee-Sanford, FL	107.0%	$ 221,037	72.3	1.12	3.12	$ 15.42
Bethesda-Rockville-Frederick, MD	106.4%	$ 445,442	80.1	1.61	2.87	$ 23.08
Phoenix-Mesa-Glendale, AZ	106.4%	$ 252,298	87.2	1.61	3.62	$ 16.35
Atlantic City-Hammonton, NJ	105.1%	$ 258,536	71.7	1.12	5.68	$ 23.32
New York-White Plains-Wayne, NY-NJ	104.7%	$ 472,658	69.3	0.80	4.79	$ 21.63
Providence-New Bedford-Fall River, RI-MA	98.4%	$ 272,732	61.2	1.34	5.37	$ 13.10
Visalia-Porterville, CA	94.6%	$ 224,105	108.8	1.97	8.50	$ 16.83
Boston-Quincy, MA	94.2%	$ 349,966	61.4	0.86	4.66	$ 25.56
Poughkeepsie-Newburgh-Middletown, NY	94.0%	$ 290,782	73.6	1.79	4.12	$ 23.61
Las Vegas-Paradise, NV	92.6%	$ 288,093	85.8	1.39	4.28	$ 17.55
Jacksonville, FL	92.6%	$ 181,653	73.2	1.06	3.26	$ 19.62
Baltimore-Towson, MD	89.8%	$ 258,936	66.9	1.23	4.00	$ 23.45
Ocala, FL	89.2%	$ 147,002	74.9	1.73	3.39	$ 12.02
Newark-Union, NJ-PA	88.6%	$ 381,183	73.0	1.17	4.65	$ 21.63
Charleston-North Charleston-Summerville, SC	85.6%	$ 176,563	74.0	1.20	5.10	$ 19.23
Virginia Beach-Norfolk-Newport News, VA-NC	85.5%	$ 221,630	71.6	0.82	3.32	$ 16.68
Reno-Sparks, NV	84.7%	$ 316,406	80.0	1.39	4.12	$ 16.83
Lakeland-Winter Haven, FL	82.9%	$ 141,875	84.9	1.56	3.60	$ 15.68
Peabody, MA	81.8%	$ 335,451	68.9	0.86	5.09	$ 25.56
Bridgeport-Stamford-Norwalk, CT	81.0%	$ 468,745	74.9	0.98	3.90	$ 29.91
Trenton-Ewing, NJ	80.7%	$ 276,183	67.0	1.88	4.24	$ 23.02
Worcester, MA	80.6%	$ 245,583	66.0	0.86	5.10	$ 20.00
Seattle-Bellevue-Everett, WA	77.4%	$ 369,177	66.6	0.88	4.28	$ 23.94
Tucson, AZ	76.5%	$ 198,430	73.7	1.42	4.01	$ 14.42
Cambridge-Newton-Framingham, MA	76.0%	$ 380,977	63.6	0.86	3.94	$ 25.56
Gainesville, FL	75.6%	$ 169,875	52.6	2.48	2.66	$ 16.67
New Haven-Milford, CT	71.3%	$ 261,064	65.8	0.98	4.85	$ 23.62
Tacoma, WA	71.0%	$ 261,712	73.2	1.21	5.05	$ 23.94
Camden, NJ	70.5%	$ 232,665	69.7	1.65	4.68	$ 21.45
Norwich-New London, CT	68.2%	$ 252,431	61.6	1.46	4.17	$ 17.31

Metropolitan Area Name (2009 MSAD)	Percent Change Prices 97-06	Home Price 2006	Fertility Rate 2006	Elasticity of Supply	Unemp Rate 2006	Median Wage 2006
Philadelphia, PA	67.6%	$ 211,020	67.9	1.65	4.50	$ 21.45
Minneapolis-St. Paul-Bloomington, MN-WI	66.5%	$ 221,112	74.0	1.45	3.83	$ 22.12
Wilmington, DE-MD-NJ	63.9%	$ 232,676	68.8	1.95	3.92	$ 21.45
Pensacola-Ferry Pass-Brent, FL	61.2%	$ 154,225	73.7	1.48	3.07	$ 15.38
Vineland-Millville-Bridgeton, NJ	61.0%	$ 166,452	89.8	1.85	6.92	$ 19.23
Springfield, MA	59.4%	$ 209,209	57.4	1.52	5.20	$ 20.60
Richmond, VA	58.0%	$ 185,521	69.2	2.60	3.19	$ 19.23
Olympia, WA	57.3%	$ 248,438	66.9	1.75	4.56	$ 21.06
Hartford-West Hartford-East Hartford, CT	54.2%	$ 238,239	63.0	1.50	4.54	$ 24.04
Portland-Vancouver-Hillsboro, OR-WA	52.9%	$ 283,856	72.1	1.07	5.02	$ 20.41
Allentown-Bethlehem-Easton, PA-NJ	52.8%	$ 207,168	67.6	1.77	4.55	$ 19.23
Albany-Schenectady-Troy, NY	52.8%	$ 187,483	60.3	1.70	3.96	$ 18.63
Asheville, NC	51.8%	$ 159,848	69.3	1.55	3.75	$ 17.79
Chicago-Joliet-Naperville, IL	49.7%	$ 251,216	74.0	0.81	4.46	$ 20.66
Denver-Aurora-Broomfield, CO	40.1%	$ 216,292	77.5	1.53	4.46	$ 21.37
Milwaukee-Waukesha-West Allis, WI	38.2%	$ 180,640	71.3	1.03	4.85	$ 20.14
Spokane, WA	37.4%	$ 182,067	71.9	1.64	4.94	$ 15.87
York-Hanover, PA	36.8%	$ 172,685	68.4	1.99	3.97	$ 21.15
St. Louis, MO-IL	35.9%	$ 136,093	69.3	2.36	5.09	$ 19.23
Lake County-Kenosha County, IL-WI	35.4%	$ 258,459	75.9	1.00	4.65	$ 20.66
Racine, WI	35.1%	$ 161,554	73.0	1.77	5.61	$ 20.19
Madison, WI	34.0%	$ 201,667	63.0	2.25	3.41	$ 19.23
Reading, PA	33.9%	$ 166,269	70.5	2.03	4.34	$ 20.61
Fayetteville-Springdale-Rogers, AR-MO	32.7%	$ 132,146	82.9	2.06	3.55	$ 16.33
Austin-Round Rock-San Marcos, TX	32.7%	$ 157,418	75.6	3.00	4.15	$ 19.79
Lancaster, PA	31.8%	$ 180,168	84.7	2.24	3.47	$ 19.23
Binghamton, NY	31.7%	$ 107,227	65.2	2.26	4.65	$ 20.43
Houston-Sugar Land-Baytown, TX	30.9%	$ 113,243	85.0	2.23	5.00	$ 15.87
Utica-Rome, NY	30.1%	$ 106,208	67.2	2.79	4.57	$ 9.68
Colorado Springs, CO	30.0%	$ 198,587	78.2	1.67	4.69	$ 17.09
Atlanta-Sandy Springs-Marietta, GA	29.5%	$ 179,457	76.0	2.55	4.63	$ 18.13
Albuquerque, NM	28.2%	$ 183,194	76.2	2.11	3.92	$ 16.33
Mobile, AL	27.9%	$ 107,656	74.1	2.04	3.60	$ 22.24
Niles-Benton Harbor, MI	27.8%	$ 127,136	73.9	2.06	6.89	$ 23.56
Kansas City, MO-KS	27.6%	$ 132,024	77.6	3.19	5.03	$ 20.31
Salt Lake City, UT	27.4%	$ 225,663	97.0	0.75	2.96	$ 17.09
Lafayette, LA	27.1%	$ 120,141	72.9	4.84	2.89	$ 18.03
Baton Rouge, LA	27.0%	$ 122,634	72.3	1.74	3.93	$ 20.03
Ann Arbor, MI	26.6%	$ 212,388	55.7	2.29	4.56	$ 35.26
Chattanooga, TN-GA	26.4%	$ 119,291	66.3	2.11	4.40	$ 16.35
El Paso, TX	25.9%	$ 99,742	95.5	2.35	6.71	$ 13.46
Knoxville, TN	25.8%	$ 131,259	64.7	1.42	4.15	$ 19.51
Lexington-Fayette, KY	25.8%	$ 140,213	66.8	2.63	4.63	$ 19.90
Birmingham-Hoover, AL	24.7%	$ 121,813	72.7	2.14	3.22	$ 19.23
San Antonio-New Braunfels, TX	24.6%	$ 102,182	81.0	2.98	4.61	$ 14.05
Syracuse, NY	24.5%	$ 116,321	63.5	2.21	4.69	$ 19.44
Corpus Christi, TX	24.1%	$ 94,566	76.4	1.65	4.95	$ 14.79
Harrisburg-Carlisle, PA	23.6%	$ 150,613	68.4	1.63	3.67	$ 18.38
Nashville-Davidson–Murfreesboro–Franklin, TN	23.2%	$ 161,298	72.7	2.24	4.23	$ 18.49
Augusta-Richmond County, GA-SC	22.9%	$ 109,720	74.0	3.57	5.85	$ 20.09
Columbia, SC	22.8%	$ 119,656	68.6	2.64	5.53	$ 14.42
Lansing-East Lansing, MI	22.4%	$ 133,123	58.0	2.58	5.79	$ 19.71
Scranton–Wilkes-Barre, PA	22.2%	$ 123,566	63.9	1.62	5.15	$ 15.82
Detroit-Livonia-Dearborn, MI	21.4%	$ 113,685	65.5	1.24	8.39	$ 24.04
Des Moines-West Des Moines, IA	20.0%	$ 129,175	80.5	3.66	3.38	$ 19.23
Davenport-Moline-Rock Island, IA-IL	19.4%	$ 105,763	73.4	4.11	4.28	$ 16.44
Durham-Chapel Hill, NC	19.4%	$ 171,002	67.9	2.11	3.91	$ 16.49

Metropolitan Area Name (2009 MSAD)	Percent Change Prices 97-06	Home Price 2006	Fertility Rate 2006	Elasticity of Supply	Unemp Rate 2006	Median Wage 2006
Dallas-Plano-Irving, TX	19.1%	$ 124,055	80.2	2.18	4.82	$ 18.73
Little Rock-North Little Rock-Conway, AR	19.0%	$ 108,394	74.1	2.79	4.68	$ 14.74
Louisville/Jefferson County, KY-IN	18.6%	$ 127,502	70.8	2.34	5.67	$ 16.02
Pittsburgh, PA	18.3%	$ 110,200	60.9	1.20	4.68	$ 19.15
Warren-Troy-Farmington Hills, MI	18.2%	$ 186,579	64.0	1.30	6.42	$ 24.04
Beaumont-Port Arthur, TX	18.1%	$ 76,356	79.5	2.49	5.89	$ 14.62
Grand Rapids-Wyoming, MI	16.5%	$ 131,715	77.5	2.39	5.80	$ 18.18
Charlotte-Gastonia-Rock Hill, NC-SC	16.3%	$ 151,494	77.2	3.09	4.76	$ 17.20
Omaha-Council Bluffs, NE-IA	16.2%	$ 124,335	85.0	3.47	3.45	$ 19.23
Kalamazoo-Portage, MI	15.7%	$ 127,799	67.0	2.48	5.45	$ 23.08
Fort Worth-Arlington, TX	15.7%	$ 108,505	81.7	2.80	4.72	$ 18.73
Hickory-Lenoir-Morganton, NC	15.6%	$ 105,663	64.7	2.41	5.85	$ 16.15
Cincinnati-Middletown, OH-KY-IN	15.5%	$ 138,971	71.8	2.51	5.14	$ 21.18
Lubbock, TX	15.0%	$ 83,529	72.7	4.33	3.98	$ 14.66
Peoria, IL	14.4%	$ 110,385	76.4	3.23	4.17	$ 16.08
Columbus, OH	13.8%	$ 146,041	73.0	2.71	4.65	$ 19.23
Rockford, IL	13.8%	$ 123,510	75.5	3.68	5.63	$ 16.83
Toledo, OH	13.8%	$ 114,567	65.4	2.21	5.99	$ 24.55
Raleigh-Cary, NC	13.7%	$ 178,667	76.2	2.11	3.70	$ 21.45
Flint, MI	13.6%	$ 108,328	70.2	2.75	8.04	$ 19.23
Greenville-Mauldin-Easley, SC	13.5%	$ 117,495	72.7	2.71	5.64	$ 16.83
Gary, IN	13.2%	$ 131,597	72.5	1.74	5.35	$ 20.66
Springfield, MO	13.1%	$ 114,971	71.1	3.60	3.87	$ 16.24
South Bend-Mishawaka, IN-MI	12.8%	$ 107,023	74.1	4.36	5.14	$ 17.63
Winston-Salem, NC	12.4%	$ 128,671	72.5	3.10	4.27	$ 13.22
Memphis, TN-MS-AR	11.8%	$ 108,828	75.7	1.76	5.68	$ 19.71
Wichita, KS	11.6%	$ 96,868	88.7	5.45	4.57	$ 16.35
Montgomery, AL	11.1%	$ 108,348	70.5	3.58	3.45	$ 14.90
Ogden-Clearfield, UT	10.9%	$ 179,163	109.0	0.75	3.14	$ 22.84
Saginaw-Saginaw Township North, MI	10.9%	$ 100,272	66.1	2.23	7.35	$ 14.42
Buffalo-Niagara Falls, NY	9.9%	$ 115,101	60.5	1.83	5.11	$ 19.23
Greensboro-High Point, NC	9.8%	$ 122,924	66.1	3.10	4.79	$ 16.83
Akron, OH	9.0%	$ 132,057	62.6	2.59	5.16	$ 18.03
Cleveland-Elyria-Mentor, OH	8.7%	$ 140,618	66.6	1.02	5.52	$ 19.47
Canton-Massillon, OH	8.6%	$ 117,865	68.9	3.03	5.70	$ 15.87
Fayetteville, NC	7.8%	$ 107,010	83.9	2.71	5.38	$ 14.42
Youngstown-Warren-Boardman, OH-PA	7.5%	$ 97,192	64.3	2.59	6.07	$ 18.96
Rochester, NY	7.3%	$ 117,940	63.9	1.40	4.56	$ 19.23
Indianapolis-Carmel, IN	7.0%	$ 130,473	78.3	4.00	4.36	$ 19.23
Spartanburg, SC	6.8%	$ 100,563	70.5	2.71	6.59	$ 16.75
Dayton, OH	3.9%	$ 119,761	68.0	3.71	5.65	$ 17.07
Fort Wayne, IN	3.3%	$ 101,869	81.1	5.36	4.89	$ 19.23

Bibliography

Aguiar, M. and E. Hurst (2007). Measuring trends in leisure: The allocation of time over five decades. *The Quarterly Journal of Economics 122(3)*, 996–1006.

Ames, M. C. (2006). A review of the FCC's inside wiring rules. *Mimeo*.

Anderson, P. M. and P. B. Levine (2000). Child care and mothers' employment decisions. In D. Card and R. Blank (Eds.), *Finding Jobs: Work and Welfare Reform*. Russel Sage.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91(434)*, 444–455.

Angrist, J. D. and A. B. Krueger (1992). The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. *Journal of the American Statistical Association 87(418)*, 328–336.

Autor, D. (2001). Wiring the labor market. *Journal of Economic Perspectives 15(1)*, 25–40.

Autor, D. H., L. E. Katz, and M. S. Kearney (2008). Trends in U.S. wage inequality: Revising the revisionists. *Review of Economics and Statistics 90(2)*, 300–323.

Baker, M., J. Gruber, and K. Milligan (2008). Universal child care, maternal labor supply and family well-being. *Journal of Political Economy 116(4)*, 709–745.

Baker, M. and K. Milligan (2008). How does job protected maternity leave affect mothers' employment and infant health. *Journal of Labor Economics 26(4)*, 655–692.

Baum, C. F., M. E. Schaffer, and S. Stillman (2007). Enhanced routines for instrumental variables/GMM estimation and testing. *Boston College Economics Working Paper 667*.

Becker, G. (1960). An economic analysysis of fertility. In N. . National Bureau of Economic Research Series (Ed.), *Demographic and Economic Change in Developed Countries*, pp. 209–231. Princeton University Press.

Becker, G. (1965). A theory of the allocation of time. *Economic Journal 75(299)*, 493–517.

Belkin, L. (2003). The opt-out revolution. *New York Times Magazine October 26*.

Ben-Porath, Y. (1973). Short-term fluctuations in fertility and economic activity in Isreal. *Demography 10(2)*.

Benjamin, J. D., P. Chinloy, and G. D. Jud (2004). Why do households concentrate their wealth in housing? *Journal of Real Estate Research 26(4)*, 329–343.

Bertrand, M., C. Goldin, and L. F. Katz (2009). Dynamics of the gender gap for young professionals in the corporate and financial sectors. *NBER Working Paper 14681*.

Bertrand, M. and K. Hallock (2001). The gender gap in top corporate jobs. *Industrial and Labor Relations Review 55*, 3–21.

Bhuller, M., T. Havnes, E. Leuven, and M. Mogstad (2011). Broadband internet: An information suprehighway to sex crime? *IZA Discussion Paper Series No. 5675*.

Bianchi, S. M., M. Milkie, L. Sayer, and J. Robinson (2000). Is anyone doing the housework? Trends in the gender division of household labor. *Social Forces 79*, 191–228.

Bittlingmayer, G. and T. W. Hazlett (2002). The financial effects of broadband regulation. In R. H. Crandall and J. H. Alleman (Eds.), *Broadband: Should We Regulate High-Speed Internet Access*, pp. 245–277. AEI-Brookings Joint Center for Regulatory Studies.

Black, D., N. Koleskinova, S. G. Sanders, and L. J. Taylor (2011). Are children 'normal'? *Research Division Federal Reserve Bank of St. Louis Working Paper 2008-40E*.

Black, D., N. Kolesnikova, and L. J. Taylor (2008). Why do so few women work in New York (and so many in Minneapolis)? *Working Paper*.

Black, D. A., A. M. Haviland, S. G. Sanders, and L. J. Taylor (2008). Gender wage disparities among the highly educated. *Journal of Human Resources 43*, 630–659.

Black, S. E. and A. Spitz-Oener (2010). Explaining women's success: Technological change and the skill content of women's work. *The Review of Economics and Statistics 92(1)*, 187–194.

Blau, F., L. M. Kahn, and J. Waldfogel (2000). Young women's marriage rates in the 1980s: The role of labor and marriage market conditions. *Industrial and Labor Relations Review 53(4)*, 624–647.

Blau, F., L. M. Kahn, and J. Waldfogel (2004). The impact of welfare benefits on single motherhood and headship of young women: Evidence from the census. *Journal of Human Resources*.

Bloom, N., J. Liang, J. Roberts, and Z. J. Ying (2012). Does working from home work? Evidence from a chinese experiment. *Working Paper*.

Bostic, R., S. Gabrial, and G. Painter (2009). Housing wealth, financial wealth, and consumption: New evidence from micro data. *Regional Science and Urban Economics 39(1)*, 79–89.

Boushey, H. (2005). Are women opting out? Debunking the myth. *CEPR Briefing Paper*.

Brencic, V. (2012). Search online: Evidence from online job boards and resume banks. *Working Paper*.

Bucks, B. K., A. B. Kennickell, T. L. Mach, and K. B. Moore (2009). Changes in u.s. family finances from 2004 to 2007: Evidence from the survey of consumer finances. *Federal Reserve Bulletin*.

Cardia, E. (2010). Household technology: Was it the engine of liberation? *Working Paper*.

Case, K. E., J. Quigley, and R. Shiller (2005). Comparing wealth effects: The stock market versus the housing market. In *Advances in Macroeconomics: 5(1)*. The Berkeley Electronic Press.

Cavalcanti, T. V. and J. Tavares (2008). Assessing the "engines of liberation": Home appliances and female labor force participation. *Review of Economics and Statistics 90(1)*, 81–88.

C.E.A. (2010). Work-life balance and the economics of workplace flexibility. *Council of Economic Advisors Report March 2010*.

Chetty, R. and A. Szeidl (2012). The effect of housing on portfolio choice. *NBER Working Paper No.15998*.

Coen-Pirani, D., A. Leon, and S. Lugauer (2008). The effect of household appliances on female labor force participation: Evidence from micro data. *Tepper School of Business Working Paper 54*.

Cohen, A., R. Dehejia, and D. Romanov (2007). Do financial incentives affect fertility? *NBER Working Paper No. W13700*.

Cortes, P. and J. Tessada (2008). Cheap maids and nannies: How low-skilled immigration is changing the labor supply of high-skilled american women. *Working Paper*.

Cragg, J. G. and S. G. Donald (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory 9(2)*, 222–240.

Dehejia, R. and A. Lleras-Muney (2004). Booms, busts, and babies' health. *The Quarterly Journal of Economics August*.

Falck, O., R. Gold, and S. Heblich (2012). E-lections: Voting behavior and the internet. *IZA Discussion Paper Series No. 6545*.

Faulhaber, G. R. (2002). Broadband deployment: Is policy in the way? In R. H. Crandall and J. H. Alleman (Eds.), *Broadband: Should We Regulate High-Speed Internet Access*, pp. 223–244. AEI-Brookings Joint Center for Regulatory Studies.

F.C.C. (2007). Exclusive service contracts for provision of video services in multiple dwelling units and other real estate developments, MB Docket No. 07-51, Report and order and further notice of proposed rulemaking, F.C.C. 07-189.

F.C.C. (2008). 47 C.F.R. § 76.2000(b). Exclusive access to multiple dwelling units generally.

F.C.C. (2010). Federal communications commission's broadband internet access page, Accessed february 16 2010.

Forman, C., A. Goldfarb, and S. Greenstein (2003). Digital dispersion: A geographic census of commercial internet use. In S. Craner and Lorrie (Eds.), *Rethinking Rights and Regulations: Institutional Responses to New Communication Technologies*. MIT Press.

Galbraith, V. L. . and D. S. Thomas (1941). Birth rates and the interwar business cycles. *Journal of the American Statistical Association 26*, 465–476.

Gelbach, J. (2002). Public schooling for young children and maternal labor supply. *American Economic Review 92(1)*, 307–322.

Gibson, C. (2013). The american demographic history chartbook: 1790-2010.

Glaeser, E., J. Gyourko, and A. Saiz (2008). Housing supply and housing bubbles. *Journal of Urban Economics 64(2)*, 198–217.

Greenspan, A. and J. Kennedy (2007). Sources and uses of equity extracted from homes. *Finance and Economics Discussion Series 2007-20, Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board, Washington DC*.

Greenstein, S. and J. Prince (2007). The diffusion of the internet and the geography of the digital divide in the united states. In D. Q. Mansell, Robin and R. Silverstone (Eds.), *Oxford Handbook on ICTs*. Oxford University Press.

Greenwood, J., A. Sheshadri, and M. Yorukoglu (2005). Engines of liberation. *Review of Economic Studies 72*, 109–133.

Greenwood, J. and G. Vandenbroucke (2005). Hours worked (long run trends). In L. E. Blume and S. N. Durlauf (Eds.), *The New Palgrave Dictionary of Economics*, pp. 75–81. London: Palgrave Macmillan.

Grubesic, T. H. and A. T. Murray (2002). Constructing the divide: Spatial disparities in broadband access. *Papers in Regional Science 81*, 197–221.

Guryan, J., E. Hurst, and M. S. Kearney (2007). Parental education and parental time with children. *Journal of Economic Perspectives 22(3).*

Gyourko, J. E., A. Saiz, and A. A. Summers (2008). A new measure of the local regulatory environment for housing markets: Wharton residential land use regulatory index. *Urban Studies 45(3)*, 693–729.

Happel, S. K., J. Hill, and S. A. Low (1984). An economic analysis of the timing of childbirth. *Population Studies XXXVIII*, 299–311.

Haurin, D. R. and S. S. Rosenthal (2005). The impact of house price appreciation on portfolio composition and savings. *Office of Policy Development & Research, U.S. Department of Housing and Urban Development*.

Hausman, J. (2002). Internet-related services: The results of asymmetric regulation. In R. W. Crandall and J. H. Alleman (Eds.), *Broadband*, pp. 129–156. AEI-Brookings Joint Center for Regulatory Studies.

Hausman, J., J. G. Sidak, and H. J. Singer (2001). Residential demand for broadband telecommunications and consumer access to unaffiliated Internet content providers. *Yale Journal on Regulation*.

Heckman, J. J. (1993). What has been learned about labor supply in the past twenty years. *American Economic Review 83(2)*, 116–121.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics 88(3)*, 389–432.

Heckman, J. J. and E. Vytlacil (2001). Local instrumental variables. In C. Hsiao, K. Mormimune, and J. Powell (Eds.), *Nonlinear Statistical Modeling: Proceedings of the Thirteenth Annual Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya*, Cambridge, U.K., pp. 1–46. Cambridge University Press.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econmetrica 73(3)*, 669–738.

Heckman, J. J. and J. R. Walker (1990). The relationship between wages and income and the timing and spacing of births: Evidence from swedish longitudinal data. *Econometrica 58(6)*, 1411–1441.

Herr, J. L. and C. Wolfram (2009). Work environment and 'opt-out' rates at motherhood across high-education career paths. *NBER Working Paper 14717*.

Hotz, J. V., J. A. Klerman, and R. J. Willis (1997). The economics of fertility in developed countries. In *Handbook of Population and Family Economics*. North-Holland.

Hurst, E. and F. Stafford (2004). Home is where the equity is: Mortgage refinancing and household consumption. *Journal of Money, Credit, and Banking 36(6)*, 985–1014.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62(2)*, 467–475.

Jackson, C. L. (2002). Wired high-speed access. In R. W. Crandall and J. H. Alleman (Eds.), *Broadband*, pp. 129–156. AEI-Brookings Joint Center for Regulatory Studies.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*, 97–126.

Kroft, F. and D. Pope (2010). Does online search crowd-out traditional search and improve matching efficiency? Evidence from craigslist. *Working Paper*.

Kuhn, P. and H. Mansour (2011). Is internet job search still ineffective? *Working Paper*.

Kuhn, P. and M. Skuterad (2004). Internet job search and unemployment durations. *American Economic Review 94(1)*, 218–232.

Lindo, J. M. (2010). Are children really inferior goods? Evidence from displacement driven income shocks. *Journal of Human Resources 45(2)*, 301–327.

Lino, M. (2007). Expenditures on children by familiies, 2006. *U.S. Department of Agriculture, Center for Nutrition Policy and Promotion Misc Publication No. 1528-2006.*

Lister, K. and T. Harnish (2011). The state of telework in the u.s.: How individuals business, and government benefit. *Telework Research Network*.

Lovenheim, M. and K. J. Mumford (Mimeo 2011, Mimeo). Do family wealth shocks affect fertility choices? Evidence from the housing market boom and bust.

Macunovich, D. J. (2010). Reversals in the patterns of women's labor supply in the united states, 1977-2009. *Monthly Labor Review*, 16–36.

Mian, A. and A. Sufi (2009). The consequences of mortgage credit expansion: Evidence from the U.S. mortgage default crisis. *Quarterly Journal of Economics 124(4)*.

Milligan, K. (2005). Subsidizing the stork: New evidence on tax incentives and fertility. *Review of Economics and Statistics 87 (3)*, 539–555.

Mincer, J. (1963). Market prices, opportunity costs, and income effects. In *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld*. Stanford University Press.

National Center for Health Statistics (2003). Bridged-race intercensal estimates of the july 1, 1990-july 1, 1999, united states resident population by state, county, age group, sex, race, and hispanic origin. prepared by the u.s. census bureau with support from the national cancer institute.

National Center for Health Statistics (2010). Postcensal estimates of the resident population of the united states for july 1, 2000-july 1, 2009, by year, county, age, bridged race, hispanic origin, and sex (vintage 2009). prepared under a collaborative arrangement with the u.s. census bureau.

Oettinger, G. S. (2011). The incidence and wage consequences of home-based work in the united states: 1980-2000. *Journal of Human Resources 46(2)*.

Rappaport, J. and J. D. Sachs (2003). The united states as a coastal nation. *Journal of Economic Growth 8*, 5–46.

Rosenbush, S., D. Carney, A. Borrus, T. Lowry, P. Elstrom, C. Haddad, and R. Grover (2001). Broadband: What happened? *Businessweek June 11, 2001*.

Ruggles, S., K. G. J. Trent Alexander, R. Goeken, M. B. Schroeder, and M. Sobek. (2010). *Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota*. Minnesota: University of Minnesota.

Ruhm, C. (2004). How well do parents wtih young children combine work and family life? *NBER Working Paper 10247*.

Saiz, A. (2008). On local housing supply elasticity. *Working Paper SSRN No.1193422*.

Saiz, A. (2012). The geographic determinants of housing supply. *Quarterly Journal of Economics*.

Sasser, A. (2005). Gender differences in physician pay: Tradeoffs between career and family. *Journal of Human Resources 40*, 477–504.

Schaller, J. (2011). Booms, busts, and fertility: Testing the becker model using gender-specific labor demand. *University of California-Davis Mimeo*.

Shang, Q. and B. A. Weinberg (2009). Opting for families: Recent trends in the fertility of highly educated women. *NBER Working Paper 15074*.

Silver, M. (1965). Births, marriages, and the business cycles in the united states. *Journal of Political Economy 73(3)*, 237–255.

Simon, C. J. and R. Tamura (2008). Do higher rents discourage fertility? Evidence from u.s. cities, 1940-2000. *Regional Science and Urban Economics 39(1)*, 33–42.

Skinner, J. and D. Staiger (2007). Technology apoption from hybrid corn to beta blockers. *Hard to Measure Goods: Essays in Honor of Zvi Grilliches*, 545–570.

Stevenson, B. (2009). The internet and job search. In D. Autor (Ed.), *Labor Market Intermediation*. University of Chicago Press.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear iv regression. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models*, pp. 80–108. Cambridge University Press.

Story, L. (2005). Many women at elite colleges set career path to motherhood. *New York Times September 20*.

Stross, R. (2011). Cheap, ultrafast broadband? Hong Kong has it. *The New York Times March 5, 2011*.

Violante, G. L. (2008). Skill-biased technical change. In S. N. Durlauf and L. E. Blume (Eds.), *The New Palgrave Dictionary of Economics*. Palgrave Macmillan.

Wallis, C. (2004). The case for stayng home. *TIME May 10*, 44.

Weinberg, B. (2000). Computer use and the demand for women workers. *Industrial and Labor Relations Review 53(2)*, 290–308.

WorldatWork (2006). WorldatWork 2006 telework trendlines; commissioned from The Dieringer Research Group Inc.

WorldatWork (2009). WorldatWork telework trendlines 2009; Data from The Dieringer Research Group Inc.

WorldatWork (2011). Telework 2011: A WorldatWork special report.