Title
Essays in Applied Microeconomics

Author
Eun, Dong Ook

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Essays in Applied Microeconomics

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Economics

by

Dong Ook Eun

2021

ABSTRACT OF THE DISSERTATION

Essays in Applied Microeconomics

by

Dong Ook Eun

Doctor of Philosophy in Economics

University of California, Los Angeles, 2021

Professor Moshe Buchinsky, Chair

These essays contribute towards our understanding of applied microeconomics. This dissertation is composed of three chapters. Chapter 1 investigated whether granting legal status to undocumented immigrants improved infant health outcomes in the United States. To answer this question, I've used the Immigration Reform and Control Act of 1986 that legalized about 2.7 million undocumented immigrants in the US as my historical setting. My research design is essentially an analysis on the county level using varying treatment intensities that come in the form of geographical variation in the proportion of legalized immigrants. In or-

der to deal with endogeneity concerns that arise from immigrants' location choice, I employ the "ethnic enclave instrument" and the "distance to the US-Mexico border" instrument. As a summary my results, I found that for a 1% increase in the proportion of the legalized immigrants in a county, I found a corresponding 2-3% decrease in the infant mortality rate.

For Chapter 2, we turn our attention to the self-reported health measure. Self-reported health is widely used in economic models to measure general health status. Most major surveys include some form of a question, in which respondents are typically asked to rate their health on a five-point scale from excellent to poor. Despite its widespread usage, we understand little about the process individuals use to position themselves on the scale. Furthermore, the process itself may have changed over time as knowledge and perceptions about particular health conditions and their medical treatments have evolved. Using the National Health Interview Study, we show that use of the scale has changed substantially over the past 22 years. We find the change is due not only to changes in underlying health, but also to changes in the way individuals regard their health in relation to the scale.

For Chapter 3, we explored how within classroom ordinal height or income rank of elementary school students affected their future academic performance. Adolescents' perceived lower social hierarchy is associated with adverse outcomes, and we focused on these two factors that are known to shape one's self-view. We used panel data from Seoul Education Longitudinal Study 2010. We exploit the feature that two students with identical height or family income can be ranked differently based on which classroom they are in. Our results indicate that height, not ordinal height rank, is associated with better academic performance in later years, while ordinal income rank is consistently shown to have an impact on future academic performance. We look for potential mechanisms and arrive at the conclusion that income rank affects future academic performance through increased parental investment.

The dissertation of Dong Ook Eun is approved.

Sarah J Reber

Youssef Benzarti

Bernardo Santos Da Silveira

Rodrigo Ribeiro Antunes Pinto

Moshe Buchinsky, Committee Chair

University of California, Los Angeles

2021

DEDICATION

To,

my parents Sung Soo Eun & Hei Won Han,

my church friends and mentors,

and my soon-to-be wife Emily Li

for their support, encouragement, and prayers over the years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

As I look back the past six years, many names come to mind who helped make the completion of my graduate study possible. First and foremost, I must thank my advisor Professor Moshe Buchinsky. With a deadline looming and stuck in my progress in my second year of the program, he opened his schedule and office so that I can ask him for help anytime. Since then, he's given me much needed direction and guidance so that I can continue in my program. I would not have been able to get to this point without his help. Special thanks to him also for helping me graduate with a job!

I thank also members of my committee who gave me valuable feedback on my research. They are Professors Youssef Benzarti, Bernardo Silveira, and Rodrigo Pinto. They took time out of their busy schedules to provide much appreciated help with my research. They also helped prepare me for the job market, which I am very grateful for. I would also like to thank Professor Sarah Reber who helped me as the external committee member and was involved with my advancement to candidacy.

Special thanks also goes to my co-authors for Chapters Two and Three without whom I would not have been able to get to this point. For Chapter Two, my co-authors are Professors Moshe Buchinsky, Nicole Maestas, and Illiana Reggio. I'm grateful that they allowed me to join in their project. Chapter Two is an updated work from Buchinsky, Moshe and Maestas, Nicole and Reggio, Iliana. "The Evolution of Self-Reported Health," 2014. I updated the analysis to cover 16 extra years of data, while exploring reasons for the observed changes we observe in our analysis. All errors are mine. The work is in preparation for publication.

I'm grateful to Changsu Ko who worked with me on Chapter Three. He reached out to me with this interested project that he was working on, and I'm thankful that I got to be part of the work. Changsu Ko worked on the main analysis while I extended it to various specifications and exploration of potential mechanisms. All errors are my own.

Special shoutout goes to Jeong-Hwan Kim and Suk-Ho Bang. I don't think I would've been able to get through first year of graduate school without you guys!

I want to thank my parents, without whom I would not be here. I do not take it for granted that I grew up in a loving and supportive family. My dad worked hard to support our family and provide for my education. He showed me by his own example how to work hard and persevere through things. Mom, day in and day out you were there to support us. Thank you for loving me concretely with your acts of service even when I gave you such a hard time.

There were many who carried me through the ups and downs of life through prayers, encouragement, and, of course, good food. Alex Lee, Jake Myung, and Brian Chan, the "Grad Haven" days will be missed. Together, we've grown a lot (physically?) and I'm thankful for how you guys have shown me a lot of grace over the years. I'm truly thankful we went through graduate school together.

Special thanks also to my church mentors and friends. Especially when things were tough and I needed a lot of guidance, I'm grateful to you all for being there for me. Proverbs 11:14 says *"Where there is no guidance, a people falls, but in an abundance of counselors there is safety."* Amen. For my friends (older, younger, and my peers), my life is rich because of you all. I am thankful that you guys are my friends who will go through life together and pursue our common mission. Also a special shoutout to Camden Bros!

Emily (Li soon to be Eun), I am truly thankful for you and I look forward to our next chapter in life! Thank you for supporting me along the way!

Last and most important of all, I thank Jesus Christ my Lord and Savior! He has been my good Shepherd. He proved to be faithful to me and I trust that He will continue to be faithful. I end with the words of Psalm 23: *"Surely goodness and mercy shall follow me all the days of my life, and I shall dwell in the house of the LORD forever."* Amen!

VITA

2015	B.A., Economics University of California, Berkeley Berkeley, California
2016	M.A., Economics University of California, Los Angeles Los Angeles, California
2017	C.Phil., Economics University of California, Los Angeles Los Angeles, California

1 The Effect of Immigration Amnesty on Infant Health Outcomes: An Analysis Through the Immigration Reform and Control Act of 1986

1.1 Introduction

The question on how to address undocumented immigrants has become more urgent as their number reached close to 12 million as of 2015 according to the Department of Homeland Security (DHS). There is the matter of border security as well as whether the government should deport or grant their stay. One recurring proposal in dealing with undocumented immigrants has been to grant amnesty to those who meet certain requirements. Examples include Deferred Action for Childhood Arrivals (DACA) which started as President Obama's 2012 executive order and is continuing under the Trump administration, in addition to the Development, Relief, and Education for Alien Minors (DREAM) Act which sought for a wider amnesty program through legislative means. As expected, it did not go unchallenged. Concerns were raised regarding increased competition for jobs and resources as well as fear of immigrant crime as we hear today.

Much of the existing research on immigration amnesty has focused on labor market outcomes on its recipients and their "competitors" (i.e. native US workers or immigrants who remain undocumented). However, the benefits are not limited to the labor market alone. Legalization can be thought of as a packaged deal with benefits that range from tangible to intangible ones. Establishing proper identity being a key requirement for many of the services in the US, granting legal status would presumably have wider impact than it being limited to a better labor market outcome. For example, many studies document the plight of undocumented immigrants having difficulty in accessing services such as medical care, banks, credit, and insurance as well as fearing deportation (Gonzales and Chavez 2012; Paulson et al. 2006; Cavazos-Rehg et al 2007). Furthermore, the National Children's Study Symposium in 2011 pointed out that the lack of legal status is the single largest barrier for "Hispanic SES mobility, better health, and integration." This highlights a need for us to understand the wider impact of an immigration amnesty program on its recipients.

One important question to ask then is amnesty's effect on immigrant children, and in particular on infant health. One obvious reason is the availability of near-universal infant mortality data from the US. On top of that, given the sensitivity of infant health to even in-utero conditions that Almond and Currie (2011) point out, the effect of amnesty can be significant. And the impact is not limited to infant health alone but later life outcomes, as studies that point to its lifelong consequences demonstrate (Currie and Hyson 1999; Behrman and Rosenzweig 2004; Almond and Currie 2011). With immigration amnesty likely removing the adverse circumstances tied to undocumented status, it would be pertinent to know whether infant health improvements are realized through such programs.

Specifically, this paper focuses on whether granting legal status had any positive effect on infant mortality rates and birthweight of undocumented immigrants' infants. However, this is not an easy question to answer. First, undocumented immigrants are, as their name indicates, undocumented. Also, immigration amnesty is not a common occurrence, and simply looking at those who obtain a green card or become naturalized would not be a fair comparison to the 12 million undocumented immigrants, the majority of whom do not have such options. Therefore, even a simple correlation between an outcome of interest and legalized status is hard to obtain since we have no representative data.

The Immigration Reform and Control Act of 1986 (IRCA) passed under the Reagan administration gives us a chance to address some of the questions raised above. It legalized about 2.7 million undocumented immigrants and created a wide geographic distribution of legalized immigrants across the United States. I use this county level variation in the proportion of legalized immigrants as "treatment intensity" to conduct a county-level analysis. In my main explanatory variable, I also incorporate the randomness in the timing in which legalizations were granted. This is evidenced by a low R^2 value between geographic location and the time it took for the applications to be processed.

Given that location choices for undocumented immigrants are not random, there is a concern for endogeneity that stem from whether counties with high level of undocumented immigrants had factors that contributed to better infant health outcomes. I address this concern in this paper by using the "ethnic enclave" and "distance to the border" instrumental variables which are used in the immigration literature (Card 2001; Ottaviano and Peri 2005; Peri 2012; Baker 2013).

The 1992 Legalization Summary Tapes provides data on all IRCA applicants. From this I can construct a "treatment intensity" variable, which is the proportion of legalized undocumented immigrants to the total population residing in a county. For county level infant births and deaths data I utilize the U.S. County-Level Natality and Mortality Data. This is made available through the Inter-university Consortium for Political and Social Research (ICPSR 36603). This dataset is constructed from all records of birth certificates and death certificates in the United States during the study period.

Results consistently point to improvements in infant mortality rates among counties with higher proportion of IRCA recipients. I also find evidence of lower incidence of low birthweight. For example, I find that for a one percent increase in proportion of legalized population in a county led to decrease of about 1.5% to 3% in infant mortality rates as well as a decrease of about 0.8% to 1.3% in incidence of low birthweight births. These results confirm that legalization of undocumented immigrants have welfare benefits beyond the labor market and in particular on infant health.

The remainder of the paper is as follows. I begin with a survey of relevant literature. In Section 1.3 I give an overview of the IRCA. Section 1.4 covers data and descriptive statistics. In Section 1.5 I discuss this paper's empirical strategy and in Section 1.6 I report results. In Section 1.7 I conclude with an exploration of various mechanisms and what we learn from this paper.

1.2 Review of Relevant Literature

While a significant portion of the immigration literature focused on the changes in labor supply induced by immigration, a study of the impact of legalization differs in that we are examining what is essentially a status change of those already present in the country. Recent interests in immigration issues produced a body of literature that I will examine here.

The connection between legal status and health has been highlighted in some recent studies. As noted earlier, the National Children's Study Symposium in 2011 identified the lack of legal status as the barrier to better health for Hispanic undocumented immigrants. Many studies document the plight of undocumented immigrants having difficulty accessing services such as medical care, banks, credit, and insurance, as well as fearing deportation (Gonzales and Chavez 2012; Paulson et al. 2006; Cavazos-Rehg et al 2007).

Some studies looked into the connection between a restrictive immigration law and health. When Arizona's legislature in 2010 signed into law the Arizona Senate Bill 1070 (SB1070), the body of the law stated that it was "intended to ... discourage and deter the unlawful entry and presence of aliens." One study noted that it contributed to decreases in use of preventive health care and public assistance among Mexican-origin adolescent mothers (Toomey RB, Umaña-Taylor AJ, Williams DR, et al. 2014). Furthermore, Torche and Sirois (2019) found evidence of lower birthweight among Latina immigrant women as a result of the threat of SB1070.

This is of importance given what we know of how infant health can have lasting impact on later life outcomes. Almond and Currie (2011) surveyed the epidemiological and economics literature to conclude that even in-utero conditions are considered to be important determinants of infant health and later-life outcomes. Infant birthweight specifically is often tied to increased schooling and improved labor market outcomes (Behrman and Rosenzweig 2004; Currie and Moretti 2007; Oreopoulos, Stabile, Walld, and Roos 2008; Black Devereux, and Salvanes 2007).

Another way that legalization can impact health is through improved labor market outcomes, which in turn improve health through increases in wage. Most papers that studied IRCA recipients agree that legal status led to significant improvements in labor market outcomes. Pan (2012) uses a regression discontinuity method to examine the labor market outcome of recipients by comparing 1975-1981 arrivals and 1982-1986 arrivals. The former cohort outperforms the latter in "males wages, female employment probability, and male English-speaking ability." Kassoudji and Cobb-Clark (2002), Amuedo-Dorantes et al. (2007), and Amuedo-Dorantes et al. (2011) estimate the program's impact on the recipients of the amnesty by comparing those surveyed in the Legalized Population Survey (LPS) and NLSY79. In particular, the wage penalty found by Kassoudii et al. (2002) ranges from 14% to 24% for having undocumented status, and 6% wage benefit under IRCA. The median annual wage of IRCA applicants was $10500, and 6% increase in wage comes out to an annual wage increase of $1440 in 2021 dollars. To put this number into perspective, Gelber, Moore, and Strand (2018) estimate that a $1000 increase in Social Security disability insurance payments decreased annual mortality rate by 0.1 to 0.25 percentage points.

On top of this, the association between parents' income and children's health is well documented. For example, Case, Lubotsky, and Paxson (2002) find a strong relationship between the two using several large nationally available datasets, the National Health Interview Survey (NHIS), Panel Study of Income Dynamics (PSID), and National Health and Nutrition Examination Survey (NHANES) to name a few. Wilkinson and Marmot (2003) also present a survey of the vast literature that confirms this observation.

Other related studies also examine the impact of legalization on crime rates. For example, an interesting study was done by Bell, Machin, and Fasani (2013) who looked at two immigration waves in 1990s and 2000s to the United Kingdom. They find property crimes increased with the first group which was composed mostly of asylum seekers legally barred from working pending their application process. The other group consisted mostly of legal workers from recently admitted EU countries. Another study by Mastrobuoni and Pinotti (2015) find that "legal status leads to 50% reduction in recidivism, and explains half to two-thirds of the observed differences in crime rates between legal and illegal immigrants." These studies suggest that legal status indeed plays a significant role.

More specifically looking at crime rates after the implementation of IRCA, Baker (2013) finds that for a one percent increase in the legalized population, it lowered crimes by 2% to 6%, primarily through the channel of property crimes. This amounts to 80,000 and 240,000 fewer "violent and property crimes" each year which is quite significant. The empirical strategy employed by Baker (2013) is the most relevant for this paper.

There are some studies that also sought to investigate concerns regarding negative effects of immigration amnesty to the native population. For the labor market outcomes on natives, Bailey (2002) finds minimal effect of IRCA. In terms of government resources, Cascio and Lewis (2017) find sizable effect of increased Earned Income Tax Credit (EITC) transfers as a result of IRCA. They estimate that permanent residency raised the annual EITC transfer to the average applicant by about $800 by 1996.

Lastly, Cortes (2013) uses difference-in-differences method to look at the effect of IRCA on immigrant youth postsecondary educational access. She compares 1975-1981 arrivals and 1982-1986 arrivals and further distinguishes between economic immigrants (i.e. IRCA treated immigrants) and refugees who already had legal status. She finds that those who benefited from IRCA were more likely to pursue postsecondary education.

1.3 Immigration Reform and Control Act of 1986

The Migration Policy Institute (2005) provides a clear and concise overview of the Immigration Reform and Control Act of 1986 (IRCA). The IRCA passed under President Ronald Reagan with the main goal of stopping illegal immigration. It had three main components to the program, which were employer sanction for hiring undocumented immigrants, stricter border control, and legalization for undocumented immigrants already present in the country. It was a result of a prolonged debate on immigration reform that sought to satisfy multiple interests, with the first draft of the bill proposed in 1982. Eventually, out of the 3 million applicants who applied for amnesty through the IRCA, about 2.7 million went on to receive permanent residency.

The scale of legal amnesty to undocumented immigrants was unprecedented and has not been observed since as you can note in Figure 1.1. Legalization happened through two main programs: Section 245A and the Special Agricultural Worker (SAW) programs. Section 245A was the general legalization program which granted path to legal status and towards citizenship to those who were present in the US since 1982. Applicants had to pay a $185 application fee and could not have any criminal record. Those who qualified under this program would receive temporary residency which would last 18 months after which they could apply to become permanent residents after meeting further requirements such as English competency. About 1.6 million were eventually approved for temporary resident status through this program.

The SAW program was added to the IRCA in response to the demands from the agricultural sector which depended heavily on immigrants for their labor supply. Qualifications for this program were much more lax, where one only had to demonstrate that one had 60 days of agricultural work experience in certain crops between May 1985 to May 1986. SAW applicants did not have to pay any application fees and to qualify for permanent residency, they did not have to meet requirements such as the English competency requirement under Section 245A. About 1.3 million were granted legal status through this program.

Application period for the IRCA were between 1987 and 1988. In order to process over 3 million applications, the Immigration and Naturalization Service (henceforth, INS) set up temporary local offices. Once local offices reviewed the applications, the ones approved were sent to the Regional Processing Facilities for review for a final decision.

There were notable reports of fraud under the IRCA, especially under the SAW program. Baker (2013) notes that the number of SAW applicants in California were far greater than previously available estimate of the entire agricultural population. He also noted that there were interviewers who found applicants under the SAW program showing little knowledge of agricultural products they purported to work with. But this had little bearing as there was political pressure to grant as many amnesty as possible. Drawing from an estimate of 3.2 million undocumented immigrants present before the enactment of IRCA from Woodrow and Passel (1990), this suggests a near universal amnesty program for undocumented immigrants who were in the country.

Baker (2013) notes that the INS intended the application process "to be standardized across the country, with no strong regional variation in the decisions regarding similar legalization applications as had been seen in decades past with naturalization decisions." Running a regression to see how much of the application decision dates are explained by where these applicants resided, I find an R^2 value of 0.035. A similar regression between the length of time it took for a decision to be made (time between application file date and application decision date) and location also yields an R^2 value of 0.035. These R^2 values suggest little ties between the decision process and locations where the applicants lived. This provides greater confidence in the estimation process that I introduce later.

Limitations were in place for government assistance programs for those who were to receive legal status. Most public benefits were barred from this population until they became permanent residents. Baker (2013) notes that this included food stamps, Medicaid, and most other programs based on financial need. However, exceptions were in place for programs that served the disabled, pregnant women, and children. The government was charged with the responsibility of raising awareness to the newly available services to the IRCA recipients.

According to the IRCA legislation itself:

> "Attorney General, in cooperation with qualified designated entities, shall broadly disseminate information respecting the benefits which aliens may receive under this section and the requirements to obtain such benefits."

1.4 Data and Descriptive Statistics

1.4.1 Number of Legalized Immigrants Through IRCA

1992 Legalization Summary Tapes is a comprehensive record prepared by the INS which included all the applications that it received and their decision status as of 1992. Given that all applications were received by 1988, Legalization Summary Tapes captures most if not all applications sent to the INS. This data is available from the National Archives[1] and provides information on each applicant such as age, race, marital status, county of residence, and time of legalization.

1.4.2 Applicant Information

Based on the legalization data, we are able to describe some of the characteristics of the relatively less known undocumented immigrant population. As you can see in Figure 1.2 and Table 1.1, the undocumented immigrant population is mostly between 18 to 35 of age and the overwhelming majority are of Hispanic origin. In fact, most of them are from Mexico. About half of the applicants were married and they had a median income of $12,000 annually

[1] I would like to thank Linda Bailey and Scott Baker for generously sharing the Legalization Summary Tapes data

(for income distribution, see Figure 1.3). In 1987, the median annual household income was about $26,000 for the whole nation.

Referring to Tables 1.4 and 1.5, location-wise most of the IRCA applicants resided in California, Texas, New York, Illinois, and Florida. In terms of the proportion of undocumented immigrant population in relation to the total population of a county, the counties that have greater proportions are mostly in California, Arizona, and Texas. Imperial, Yuma, and Tulare counties, which are the top 3 counties in terms of proportion, are predominantly agricultural economies reflecting how many of the undocumented immigrants were agricultural workers. Focusing our attention on those qualifying under the general legalization program, most undocumented immigrants worked as laborers and in the service occupations as we can see in Table 1.3. The rest, even with the fraud claims under the SAW program, would have been employed in some form in the agricultural sector.

Once we turn our attention to the counties that have the greatest number of undocumented immigrants, top at the list is Los Angeles county. In light of the fact that Los Angeles is one of the largest counties in the U.S. in terms of population, 8.6% of that population being legalized is a sizable impact.

1.4.3 Infant Mortality

For infant mortality, U.S. County-Level Natality and Mortality Data will be used which is available through Inter-university Consortium for Political and Social Research (ICPSR 36603). The advantage of this data set is that while it is not possible to identify infant mortality specifically for those of hispanic origin, it is a product of meticulous data work that many of the data cleaning concerns are already addressed.

It has information on live births, infant deaths, and general mortality on the county level. For the purposes of this study, the time frame of interest is between the years 1984 to 1995. It was chosen to allow sufficient time before and after the enactment of the IRCA to see pre-trends and post-legalization behavior.

The original data source for birth in the ICPSR data set for the time frame of interest is from the Natality Detail File Series (ICPSR 36) which in turn is from Vital Statistics Cooperative Program (VSCP). It included all records of birth certificates to participating states. By 1984, 46 states participated and by 1985 all states would participate. Infant mortality data is originally from Multiple Cause of Death Data File from the National Bureau of Economic Research which in turn is from the death certificate filed through the National Vital Statistics System of the National Center for Health Statistics. All deaths that occurred in the United States are included.

This gives us greater confidence for the use of this data in analyzing health outcomes realized for the undocumented immigrant population. At the least, using a near-complete account of births and deaths is a significant improvement from using other available data sources that suffer from larger under-representation of the undocumented population.

Certain things are worth pointing out about this data. One notable aspect of the mortality data is that only counties with population over 100,000 or more can be identified beginning 1989. This will be taken into account in the empirical analysis where I will restrict the analysis to counties with population greater than 100,000 (for a visual representation, see Figure 1.4). A key decision to make is in the process of calculating crude infant mortality rates which will be used as the main outcome of interest. Births and infant deaths are reported in two forms: by residence and occurrence. The main distinction is that data reported by residence are restricted to births and death that occurred to residents of a given county, whereas data reported by occurrence will include all births and deaths that occur in a given county. Further distinction is that only data reported by occurrence includes births and deaths of non-residents. For this reason, as our population of interest will likely be captured in the non-resident portion, the main analysis will be done with data reported by occurrence. However, as a check, results using data reported by residence will be available as well.

In dealing with the issue of missing data, if data reported by residence and occurrence are both missing, it is taken out of the analysis. However, for a given county and year, if one of them is available while the other missing, then I made the decision of simply imputing the missing data with the available data.

Looking at infant mortality as a health outcome variable has some potential advantages though it cannot readily be tested. One is that the problem of under-representation of undocumented immigrants in the data is less so of a problem. It is presumably difficult to hide an infant's death so that it would not be on the record, and also there is incentive to report an infant's birth for medical care for the infant as well as to obtain citizenship of for being born within the U.S.

1.4.4 Summary Statistics

In order to see if there were any systematic differences between counties that had a large presence of undocumented immigrants versus ones with a smaller presence, a two-sided t-test of differences in means is conducted. In Table 1.6, comparison is made between two groups of counties: counties that had above median values in the proportion of county population legalized by year 1992 and counties that were below the median. What seems evident is that there are significant differences between these two groups of counties. As expected, undocumented immigrants go to counties with higher per-capita income and with significantly higher population (see Figure 1.5 for a more detailed information about population). These variables will be controlled for in our later analysis.

Though location choice is endogenous and could be problematic, it does seem that undocumented immigrants do not choose based on infant health prospects. In fact, the difference in infant mortality rates between these two groups are negligible as evidenced by the p-value computed (0.65). This holds true for the rate of incidence of low weight births with p-value of (0.72). These two outcome variables being balanced is worth highlighting for the purpose of our analysis, as it gives us greater confidence that in terms of infant health, these two groups are comparable. Finally, this is in line with what we know from immigration literature that immigrants mainly select based on presence of similar immigrant communities (Bartel 1989; Zavodny 1999).

Lastly, in order to provide some level of control for county characteristics, I find data on county population from the National Bureau of Economic Research and per-capita income level on counties from the Census Bureau.

1.5 Empirical Strategy

The ideal data would be a randomized control trial with individual level data on undocumented immigrants. However, without such data, this paper will take advantage of the wide geographic variation in the level of undocumented immigrants granted legal status as well as the aforementioned quasi-randomness in the timing of the legalization process.

Earlier, it was noted that a regression to see how much of the application decision date and processing length are explained by geographic factors yielded a low R^2 value of 0.035. To provide additional qualitative evidence, Baker (2013) notes that:

> "The practice of 'remoting' and the unfamiliarity of the INS with the massive undertaking, as well as the underestimate of the number of applicants, meant the INS was overwhelmed and the application approval process took much longer for some applicants than others. In essence, it meant that if two identical IRCA applicants both applied in mid-1987 in the same county, one might be legalized by the end of 1987 and the other remaining without legal status for up to 3 additional years."

1.5.1 Regional Variation

If we want to use regional variation in the "treatment intensity" of legalization, it would need to be a pretty sizable effect especially given that our main outcome variable is also at the county level. If only a few people in each county are legalized, then it would be difficult to argue that the this is evidence of an impact of legalization. However, 2.7 million undocumented immigrants were legalized through the IRCA. Therefore, it is sizable enough to identify potential effects of the program using geographical variation using a regional level analysis.

Figures 1.7 and 1.8 serve to show a wide geographic variation in the United States of the legalization treatment intensity, and that there are sufficient number of counties that have sizable treatment amount. I also present Figure 1.8 to show the interquartile range. To address concerns about using all the counties across the US, I also present analysis to only those counties with IRCA recipients. Among counties with IRCA recipients, the 75th percentile treatment intensity is at 0.9% of the 1986 county population legalized while the median value is at 0.16%. The treatment intensity ranges from 0.009% to 20%.

1.5.2 Specification

The main empirical strategy will follow the approach of Baker (2013) in a panel OLS method using regional variation in treatment intensity. In order to address endogeneity concerns, instrumental variable strategies will also be employed later on as well. The following is the first main specification.

$$HealthOutcome_{ct} = \beta_0 + \beta_1 CumIRCA\%_{ct} + \gamma' X_{ct} + Year_{FE} + County_{FE} + \epsilon_{ct}$$

This is a group-level analysis and would need to be weighted by the population size of each county, with standard errors correlated at the state level following the suggestion by Angrist and Pischke (2009) in dealing with serial correlation in group level time series analysis. Two $HealthOutcome_{ct}$ variables will be used: the infant mortality rate and the rate of incidence of low weight infant births. Infant mortality rate is defined as infant deaths per 1000 births in a given county and year. Rate of incidence of low weight births is defined as number of infant births weighing less than 2500 grams per 1000 births in a given county and year. X_{ct}, though limited in our analysis, is county characteristics such as population size and per-capita income. As for the main explanatory variable it is a cumulative measure with the following construction:

$$CumIRCA\%_{ct} = \frac{Legalized_{ct} + Legalized_{ct-1} + ... + Legalized_{c1984}}{Population_{ct}} \times 100$$

It has the advantage of incorporating the quasi-random timing of legalizations granted across counties, even as the total numbers of legalizations of undocumented immigrants accumulate over the years. The resulting variable takes on the form as in Figure 1.9, where prior to the IRCA the cumulative measure remains at 0 then begins to increase with some degree of randomness until by 1992 most legalization applications are processed. The decrease that happens after 1992 in the measure is due to the population size increasing while the number of legalized immigrants remain constant.

1.5.3 Limitations

Key identifying assumption in using regional variation in treatment intensity over years is that the legalized population, even if they moved, largely remained in the same county that they were present when filing their applications. Baker (2013) provides a limited test using the Legalized Population Survey which was conducted in 1989 and 1992 on the newly legalized population.

Those who did not move or moved within the same zip code would still be in the same county. In the 1992 survey, though about 33% moved within their state of residence, Baker (2013) notes that there are about 40,000 zip codes but only 3,000 counties, and that there are 522 zip codes in LA where most IRCA recipients resided. He uses this fact to argue that most IRCA recipients would have stayed in the same county in which they were living at the time of the application. Of course, this is a less tenable position to maintain as we extend the time frame of our analysis.

An additional concern is that the locations where immigrants chose to live is not random but endogenously determined due to some characteristics of those counties. This is a valid concern, though as previously noted, undocumented immigrants' location choice does not seem to be influenced by health related factors. However, to deal with this, distance measure from the US-Mexico border as well as the immigrant enclave instrument will be used in instrumental variable estimation.

1.5.4 Concerns over Underreporting

There are valid concerns over underreporting of birth and death data from the undocumented immigrants, but less so for births since there are obvious incentives to report them for citizenship purposes. This is a problem if the effect of legalization is to make undocumented immigrants more likely to become "documented." In that case, it would be challenging to distinguish between the effect of legalization versus the effect of better reporting on infant health outcomes. One thing to note, however, is that the direction of the bias as a result of increased representation in data will be toward higher infant mortality rates. William et al. (1986) show US-born infants with Spanish surnames have slightly higher mortality rates. The introduction of a greater number of this population to the birth data should raise, not lower, mortality rates. Further, as Powell-Griner et al. (1982) notes, evidence seems to suggest that there is underreporting of infant deaths among Spanish surname population. As long as legalization of undocumented parents do not cause them to increase in their underreporting behavior, the increase in representation from the legalization should point the bias upwards. The infant health benefit then should be greater than what has been estimated through our analysis.

1.6 Results

1.6.1 OLS Results

We come to the results. The OLS results are shown in Table 1.7 and analysis using both reported by occurrence and residence data are shown. Though both are statistically significant, we find a greater impact of IRCA when using the occurrence data that includes births to non-residents. Focusing our attention on column (1), we find that for a 1% legalization of a county population, there is a 0.28 decrease in infant deaths per 1000 births. Given that mean infant mortality rates for all counties is around 10 deaths per 1000 births, that amounts to about a reduction of 2.8% in infant mortality rates for a 1% legalization of a county population.

In order to examine whether other measures of infant health were affected from the legalization, Table 1.8 shows the effect of legalization on rate of incidence of births with low infant weights. Focusing our attention on column (1), we find that a 1% legalization of a county's population is associated with a decline of 0.72 incidence of low weight births per 1000 births. We noted earlier in Table 1.6 that the mean of low birthweight births per 1000 births was 63. This amounts to a 1.14% reduction in low birthweight births for a 1% legalization of a county's population. OLS results consistently point to an improved infant health outcome through the legalization program.

1.6.2 Instrumental Variable Estimation (2SLS)

We have seen earlier in Table 1.6 causes for concern: there are systematic differences between counties with higher proportion of undocumented immigrant population versus those that have lower proportion. While we didn't notice any cause for concern in health-related factors we examined, if there are any omitted variables that affect infant health and are correlated with higher number of undocumented immigrants, that would introduce bias to our results.

In order to deal with this concern, an emphasis was made about the quasi-random nature of the timing of legalization. However, to address this issue further, I implement an instrumental variables estimation. For instruments I use distance to the US-Mexico border and the predicted number of immigrants in each state based on their location in 1970. These are instruments previously used in the literature to deal with such endogeneity concerns.

The distance instrument is constructed using ArcGIS and NHGIS data. Centroid of each county is provided through the NHGIS data and the shape file for Mexico and US border is provided through the ArcGIS website which is turn is from the International Boundary and Water Commission. Then I use ArcGIS to calculate the length of the straight line from the centroid of each county to the nearest point on the border. Using distance as instrument is for the purpose of isolating variations that arise only from moving costs faced by immigrants that would induce them to stay nearer to the border. Looking at Figure 1.10, we have greater confidence that distance does not affect infant health directly.

The second instrument, commonly called "ethnic enclave" instrument, is constructed from the Census data where I use the number of immigrants in existing ethnic enclaves to compute the predicted the number of immigrants in each state in 1990. This is so that one could isolate the variations that come only from the supply side shift while using the common behavior of immigrants to settle where other immigrant communities are. One thing to note is that this instrument is based on the state level, not the county level.

The more detailed construction of this "ethnic enclave" instrument is as follows:

$$PredictedImmig_s^{1990} = \sum_c FB_{s,c,1970} \times \frac{FB_{c,1990}}{FB_{c,1970}}$$

where $FB_{s,c,1970}$ is the number of foreign born from country (enclave) c and state s in 1970. Then we compute the predicted number of immigrants from this county (enclave) c by using the observed nation wide growth rate of that immigrant group population. We can then sum up all the predicted immigrant numbers and find the 1990 predicted immigrants in state s: $PredictedImmig_s^{1990}$. This is then interacted with year dummies to capture time the dynamic:

$$PredictedImmig_s^{1990} \times 1(year = t) = PredictedImmig_{st}$$

The first stage of the estimation is

$$\begin{aligned} CumIRCA\%_{ct} =& \gamma_0 + \gamma_1 Distance_{ct} + \gamma_2 PredictedImmig_{st} + \gamma_3 X_{ct} + YearDummies \\ &+ CountyDummies + \epsilon_{ct} \end{aligned}$$

where $Distance_{ct}$ also is a variable created by interacting year dummies to static distance to capture time dynamic. The second stage will be:

$$HealthOutcome_{ct} = \beta_0 + \beta_1 Cum\widehat{IRC}A\%_{ct} + \gamma' X_{ct} + Year_{FE} + County_{FE} + \epsilon_{ct}$$

1.6.3 IV (2SLS) Result

IV results show even greater estimates of the impact of legalization on improving infant mortality rates compared to OLS estimates. Focusing our attention on column (1), we find that for 1% legalization of a county population, there is a 0.35 decrease in infant deaths per 1000 births. This amounts to a 3.5% reduction in infant mortality rates given a 1% increase in the proportion of legalized immigrants in a given county. This is a slightly larger estimate than the OLS estimate.

IV results show greater estimates of the impact of legalization on improving infant birthweight outcomes compared to OLS estimates. Focusing our attention on column (1), we find that for 1% legalization of a county population, there is a 0.76 decrease in low weight infant births per 1000 births. This amounts to 1.2% less births with low birthweight relative to the mean. This is a similar estimate compared to the OLS estimate.

1.6.4 Evidence of Causality

In order to check that the impact of legalization that we've estimated is actually realized *after* the implementation of the legalization program, we plotted Figure 1.11. The plotted points are coefficient estimates of a regression between infant mortality rate and total percentage of IRCA recipients in a county interacted with year dummies. We can confirm from this figure that the observed decreases in the infant mortality rate happens *post*-IRCA. This gives us greater confidence that we are not wrongly attributing the infant health improvements on the legalization program.

1.6.5 Focusing on Female Recipients

In our main specification, the way we constructed our treatment variable was using the total number of legalized immigrants. The main specification is preferred as it remains agnostic as to through what channels the legalization benefits are realized. However, one could consider an alternative way to define the main treatment variable, this time focusing on the female population.

$$CumIRCA\%_{ct}^{females15-44} = \frac{Legalized_{ct}^{f15-44} + Legalized_{ct-1}^{f15-44} + ... + Legalized_{c1984}^{f15-44}}{Population_{ct}^{females15-44}} \times 100$$

Here we restrict our attention to females of childbearing age from ages 15 to 44. $Legalized_{ct}^{f15-44}$ refers to the number of female undocumented immigrants between the ages of 15 to 44 who were legalized in year t from county c. $Population_{ct}^{females15-44}$ is the total female population between the ages of 15 to 44 from county c in year t. This construction is equivalent to our main treatment variable but restricted in focus. The 2SLS results using this definition are provided in Table 1.11 and 1.12

One major thing to note from this result compared to our main results is that the estimated impact using the female recipients of legal amnesty is smaller in size. A straightforward interpretation of the results show that for a 1% increase in the proportion of female undocumented immigrants granted legal status in a given county (ages 15-44), I estimate a corresponding 0.29 decrease in infant deaths per 1000 births, and 0.6 decrease in the incidence of low weight births per 1000 births. These results correspond to 2.9% decrease in infant mortality rate and 0.95% decrease in the incidence of low weight births as a result of legalizing 1% of female undocumented immigrants in childbearing age of a given county. These results are not too far from the main results that we presented earlier.

1.7 Conclusion

This paper evaluated whether granting legal status to undocumented immigrants led to improvements in infant health. In order to answer this question, we looked at the Immigration Reform and Control Act of 1986 as our historical setting. While we were limited by the lack of available data on the undocumented population, this setting gave rise to an indirect approach using the geographic variation in the number of legalized immigrants. We found that for a 1% increase in the proportion of legalized immigrants in a county, there was a corresponding 2.8% to 3.5% decrease in the infant mortality rate. We also noted a corresponding 1.2% decrease in the incidence of low weight births. A rudimentary back-of-the-envelope calculation suggests that this amounted to 800 to 1200 less infant deaths nationwide as a result of the legal amnesty.

Given that the literature consistently point to long-term benefits of childhood investment and health, it is yet to be known what are the full ramifications of the legalization program and how improved infant health outcomes would have contributed to their long-term outcomes. At the least, it is once again clear that one's immigration legal status plays an important factor in one's overall well-being, and I hope this paper contributes to the immigration debate that we have today.

Table 1.1: Race of IRCA Applicants

Race	Freq.	Percent
Asian	132,647	4.40
Black	132,839	4.40
Hispanic	2,618,342	86.79
White, Nonhispanic	73,839	2.45
Unknown	59,216	1.96
Total	3,016,883	100.00

Table 1.2: Marital Status of IRCA Applicants

Marital	Freq.	Percent
Divorced	74,575	2.47
Married	1,250,077	41.44
Not Married	1,534,114	50.85
Separated	92,082	3.05
Unknown	31,845	1.06
Widowed	34,190	1.13
Total	3,016,883	100.00

Table 1.3: Occupation of IRCA Applicants

Rank	Occupation	Number	Percent
1	Laborers	427,770	24.26%
2	Service Occupations	376,776	21.37%
3	Students and/or Children Under 16	219,685	12.46%
4	Precision Production / Repair	192,494	10.92%

Table 1.4: List of Counties Ranked By Percent Legalized

Rank:	State and County	% Legalized	Total Legalized
1	California Imperial County	20.06%	19,845
2	Arizona Yuma County	10.91%	9,880
3	California Tulare County	9.15%	26,167
4	California Los Angeles	8.66%	727,213
5	California Merced	8.21%	13,057
6	California Fresno	7.77 %	46,186
7	California Santa Cruz	7.72 %	16,734
8	Texas Hidalgo County	7.33 %	25,711

% legalized is calculated by dividing the total number of legalized immigrants by county population in 1986.

Table 1.5: List of Counties Ranked By Total Legalized

Rank	State and County	% Legalized	Total Legalized
1	California Los Angeles County	8.66%	727,213
2	California Orange County	5.59%	124,667
3	Illinois Cook County	2.3%	119,084
4	Texas Harris County	4.21%	117,215
5	California San Diego County	4.06%	89,221
6	Texas Dallas County	3.38%	61,046
7	Florida Miami-Dade County	3.27%	58,970
10	New York Queens County	2.29 %	44,756

% legalized is calculated by dividing the total number of legalized immigrants by county population in 1986. Rank 8 and 9 omitted to show a New York county.

Table 1.6: Counties With Population Above 100,000

	Above Median $CumIRCA\%_{c1992}$		Below Median $CumIRCA\%_{c1992}$		P_{val}
	1984	1995	1984	1995	Δ 1984
Per-Capita Income	14,700 (240)	23,800 (460)	13,300 (120)	22,200 (210)	0
Pop	553,000 (70000)	638,000 (80000)	281,000 (17,000)	287,000 (16,000)	0
Infant Deaths Per 1000 (All)	9.43 (.36)	7.27 (.29)	9.67 (.38)	9.78 (.31)	0.65
Infant Deaths Per 1000 (Residents)	10.4 (.19)	7.13 (.16)	10.6 (.19)	8.39 (.15)	0.49
% Births At Hospital (All)	98.25 (.23)	98.75 (.11)	99.14 (.07)	98.19 (.6)	0
% Births At Hospital (Residents)	98.46 (.17)	98.84 (.08)	99.17 (.07)	99.05 (.1)	0
Low Birth Weight Per 1000 (All)	62.57 (1.4)	66.46 (1.74)	63.27 (1.44)	71.64 (1.81)	0.72
Low Birth Weight Per 1000 (Residents)	65.53 (1.06)	69.4 (1.05)	65.22 (.97)	74.76 (1.08)	0.82
Number of Counties	[c = 174]	[c = 174]	[c = 242]	[c = 267]	

Mean and standard errors are reported here. Comparison is made between two groups of counties: counties that had above median values in the proportion of county population legalized by year 1992 and counties that were below the median. Summary statistics from 1984 (before IRCA) and 1995 (after IRCA) are reported here. Counties with infant mortality rates below 1-percentile or above 99-percentile levels are excluded, as well as counties with population below 100,000. P-value is computed from a two-sided t-test of difference in means

Table 1.7: OLS Results For Infant Mortality

	(1) Occurrence	(2) Occurrence	(3) Residence	(4) Residence
Cumulative % IRCA	-0.2835***	-0.2065***	-0.0811***	-0.0750***
	(0.0973)	(0.0670)	(0.0243)	(0.0222)
log(Population)	3.1160**	2.9412**	-1.0668	-0.9020
	(1.3918)	(1.1704)	(0.8482)	(0.8745)
log(Per Capita Income)	-3.1263	-3.2553*	-1.6760	-1.9143*
	(1.9001)	(1.7572)	(1.0713)	(1.1042)
Constant	-0.7905	2.7690	41.1406**	41.4129**
	(26.1477)	(26.1507)	(15.7776)	(16.1452)
Only IRCA	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	5,114	4,156	5,219	4,212
Counties	489	369	490	369
States	49	45	49	45
R-squared	0.8438	0.8728	0.7916	0.8177

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Restricted to counties with population greater than 100,000. The result is a weighted least squares with county population as weights. Standard errors are clustered at the state level. For results using occurrence data, infant mortality rates below and above the 1st and 99th percentiles are excluded. Column (1) includes all counties and occurrence data, whereas Column (2) is excludes counties that did not have any IRCA recipients. Columns (3) and (4) use residence data

Table 1.8: OLS Results for Incidence of Low Birthweight

	(1) Occurrence	(2) Occurrence	(3) Residence	(4) Residence
Cumulative % IRCA	-0.7231***	-0.8236***	-0.8552***	-0.8616***
	(0.1012)	(0.1369)	(0.1116)	(0.1172)
log(Population)	-16.7704***	-16.9166***	-15.9163***	-16.2882***
	(3.5858)	(3.7809)	(3.5726)	(3.7613)
log(Per Capita Income)	4.2282	3.6588	1.9831	2.9276
	(4.1463)	(4.3371)	(4.0570)	(4.0370)
Constant	252.5134***	262.4742***	262.8311***	260.9155***
	(67.6943)	(70.0084)	(69.8201)	(70.7276)
Only IRCA	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	4,942	4,059	4,942	4,059
Counties	457	365	457	365
States	48	43	48	43
R-squared	0.9532	0.9577	0.9440	0.9491

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Restricted to counties with population greater than 100,000. Low birth weight rate is incidence per 1000 births. The result is a weighted least squares with county population as weights. Standard errors are clustered at the state level. For results using occurrence data, low birth weight rates below and above the 1st and 99th percentiles are excluded. Column (1) includes all counties and occurrence data, whereas Column (2) is excludes counties that did not have any IRCA recipients. Column (3) and (4) use residence data

Table 1.9: IV (2SLS) Results for Infant Mortality

	(1) Occurrence	(2) Occurrence	(3) Residence	(4) Residence
$\widehat{Cumulative}\%IRCA$	-0.3532***	-0.2543***	-0.0627**	-0.0602**
	(0.1078)	(0.0847)	(0.0299)	(0.0303)
log(Population)	3.2821**	3.0471***	-1.1397	-0.9073
	(1.3607)	(1.1444)	(0.8240)	(0.8367)
log(Per-Capita Income)	-3.7317**	-3.6893**	-1.5724	-1.8208*
	(1.7696)	(1.7430)	(1.0488)	(1.1029)
Constant	12.6016	2.9725	40.3118***	36.4568**
	(22.7805)	(24.3065)	(13.5438)	(14.6671)
Only IRCA	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	5,114	4,156	5,114	4,156
Counties	489	369	489	369
States	49	45	49	45
1st Stage F	181	166	181	166
R-squared	0.8436	0.8727	0.7925	0.8172

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Restricted to counties with population greater than 100,000. The result is a weighted least squares with county population as weights. Standard errors are clustered at the state level. For results using occurrence data, infant mortality rates below and above the 1st and 99th percentiles are excluded. First stage F-Statistic is reported in the table. Column (1) includes all counties and occurrence data, whereas Column (2) is excludes counties that did not have any IRCA recipients. Column (3) and (4) use residence data.

Table 1.10: IV (2SLS) Results for Incidence of Low Birthweight

	(1) Occurrence	(2) Occurrence	(3) Residence	(4) Residence
$\widehat{Cumulative\%IRCA}$	-0.7591***	-0.8545***	-0.9381***	-0.9411***
	(0.1867)	(0.2152)	(0.1550)	(0.1665)
log(Population)	-16.6935***	-17.5056***	-14.8550***	-15.1779***
	(3.6465)	(3.8163)	(3.5620)	(3.7055)
log(Per-Capita Income)	3.8907	3.2105	0.6056	1.4631
	(4.4205)	(4.4791)	(3.6627)	(3.6940)
Constant	234.1023***	215.8560***	216.5221***	216.7056***
	(59.0293)	(64.8728)	(58.7189)	(61.7379)
Only IRCA	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	4,942	3,962	4,746	3,944
Counties	453	363	459	366
States	49	44	48	44
1st Stage F	197	210	201	205
R-squared	0.9548	0.9576	0.9378	0.9431

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Restricted to counties with population greater than 100,000. The result is a weighted least squares with county population as weights. Standard errors are clustered at the state level. For results using occurrence data, counties with infant mortality rates below and above the 1st and 99th percentiles are excluded. First stage F-Statistic is reported in the table. Column (1) includes all counties and occurrence data, whereas Column (2) is excludes counties that did not have any IRCA recipients. Column (3) and (4) use residence data.

Table 1.11: IV Results (Infant Mortality) For Females

	(1) Occurrence	(2) Occurrence	(3) Residence	(4) Residence
$\widehat{Cumulative}\%IRCA$	-0.2858***	-0.2074***	-0.0484**	-0.0464*
	(0.0891)	(0.0697)	(0.0233)	(0.0237)
log(Population)	3.0760**	2.8930**	-1.1967	-0.9627
	(1.3612)	(1.1388)	(0.8144)	(0.8258)
log(Per-Capita Income)	-4.1460**	-4.0724**	-1.5993	-1.8401*
	(1.7766)	(1.7528)	(1.0499)	(1.1084)
Constant	18.8580	8.6353	41.2246***	37.3488**
	(22.8691)	(24.2771)	(13.5966)	(14.7293)
Only IRCA	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	5,073	4,115	5,073	4,115
Counties	482	362	482	362
States	48	44	48	44
1st Stage F	230	238	230	238
R^2	0.8440	0.8734	0.7935	0.8184

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Restricted to counties with population greater than 100,000. The result is a weighted least squares with county population as weights. Standard errors are clustered at the state level. For results using occurrence data, infant mortality rates below and above the 1st and 99th percentiles are excluded. First stage F-Statistic is reported in the table. Column (1) includes all counties and occurrence data, whereas Column (2) is excludes counties that did not have any IRCA recipients. Column (3) and (4) use residence data.

Table 1.12: IV Results (Incidence of Low Birthweight) For Females

	(1) Occurrence	(2) Occurrence	(3) Residence	(4) Residence
$Cumulative\widehat{\%IRCA}^{females15-44}$	-0.5978***	-0.6761***	-0.7434***	-0.7469***
	(0.1457)	(0.1686)	(0.1209)	(0.1294)
log(Population)	-17.2087***	-18.0928***	-15.4543***	-15.7943***
	(3.5236)	(3.6932)	(3.4586)	(3.5931)
log(Per-Capita Income)	3.2069	2.3966	-0.1420	0.6662
	(4.5264)	(4.5945)	(3.7603)	(3.7973)
Constant	246.7058***	231.2233***	230.7270***	232.2887***
	(61.0677)	(67.3059)	(60.8919)	(64.2546)
Only IRCA	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Observations	4,727	3,921	4,705	3,903
Counties	446	356	452	359
States	48	43	47	43
1st Stage F	250	357	297	228
R^2	0.9549	0.9577	0.9377	0.9431

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Restricted to counties with population greater than 100,000. The result is a weighted least squares with county population as weights. Standard errors are clustered at the state level. For results using occurrence data, counties with infant mortality rates below and above the 1st and 99th percentiles are excluded. First stage F-Statistic is reported in the table. Column (1) includes all counties and occurrence data, whereas Column (2) is excludes counties that did not have any IRCA recipients. Column (3) and (4) use residence data.

Figure 1.1: Number of Immigrants Granted Legal Status 1970-2018

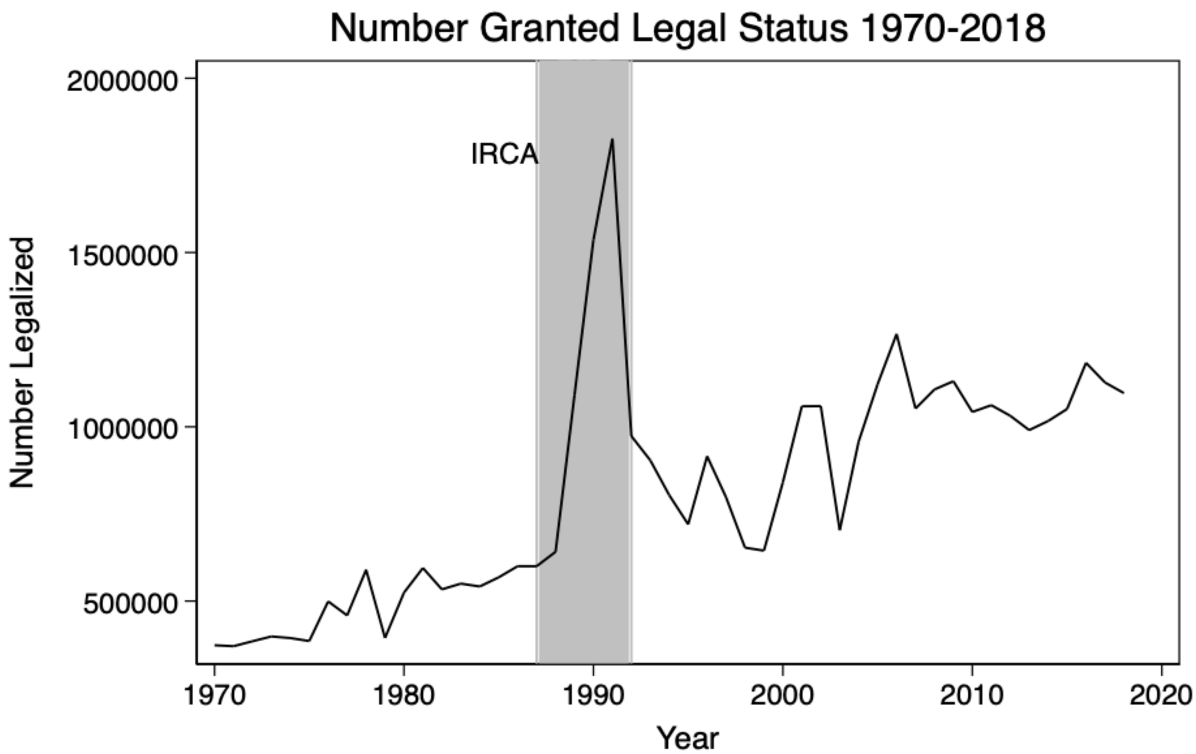


Figure 1.2: IRCA Recipients By Gender

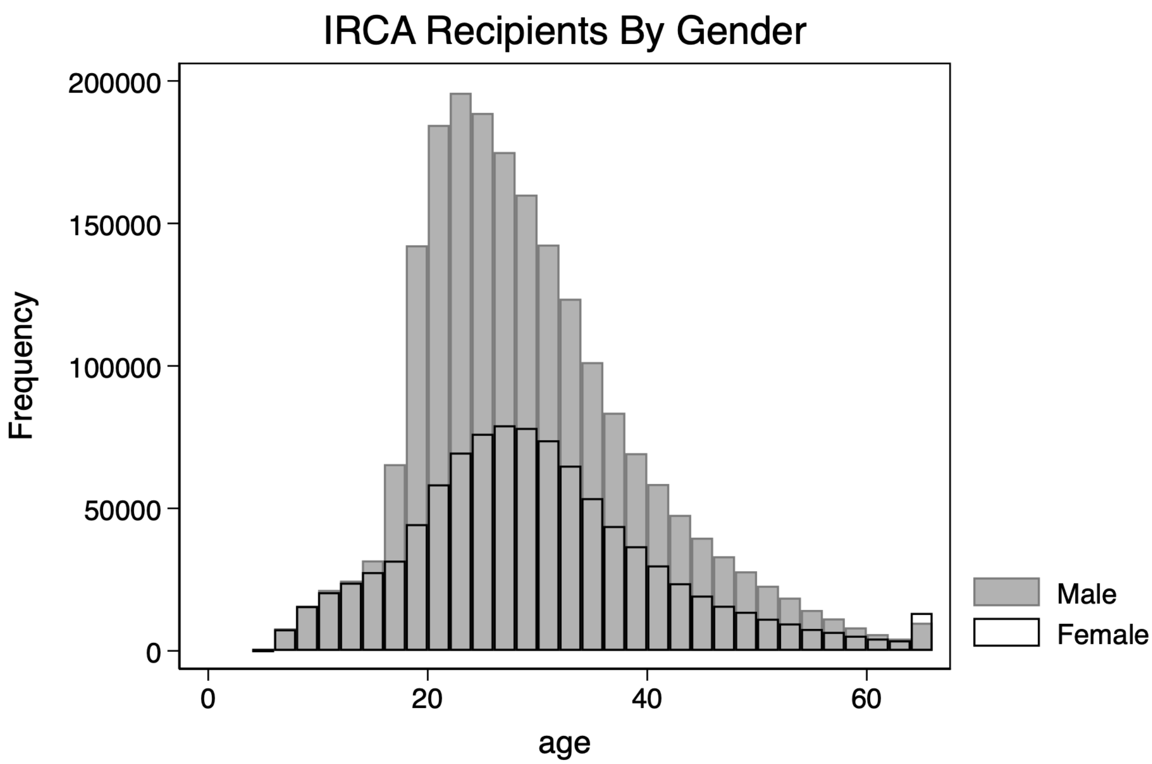
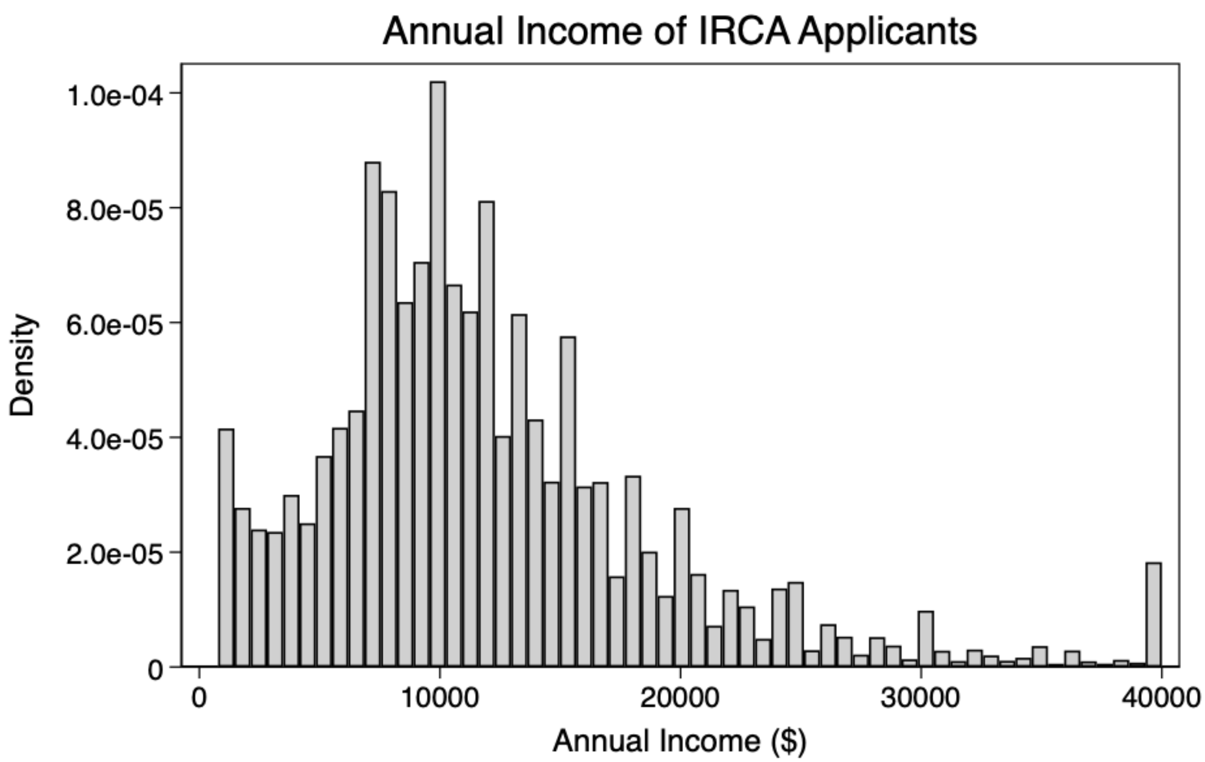


Figure 1.3: Annual Income of IRCA Applicants



$10,000 in 1987 is about $21,443 in 2017, $20,000 is $42,887, $30,000 is $64,331, and $40,000 is $85,775

Figure 1.4: Counties That Had More Than 100,000 In Population in 1992

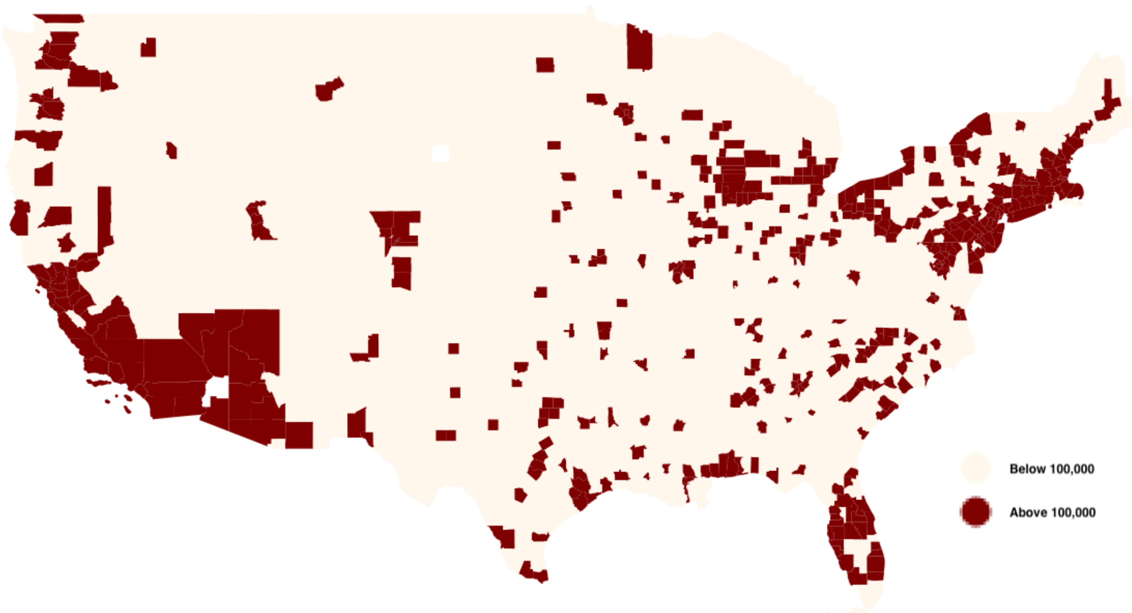


Figure 1.5: Kernel Density Plot of County Population in 1984

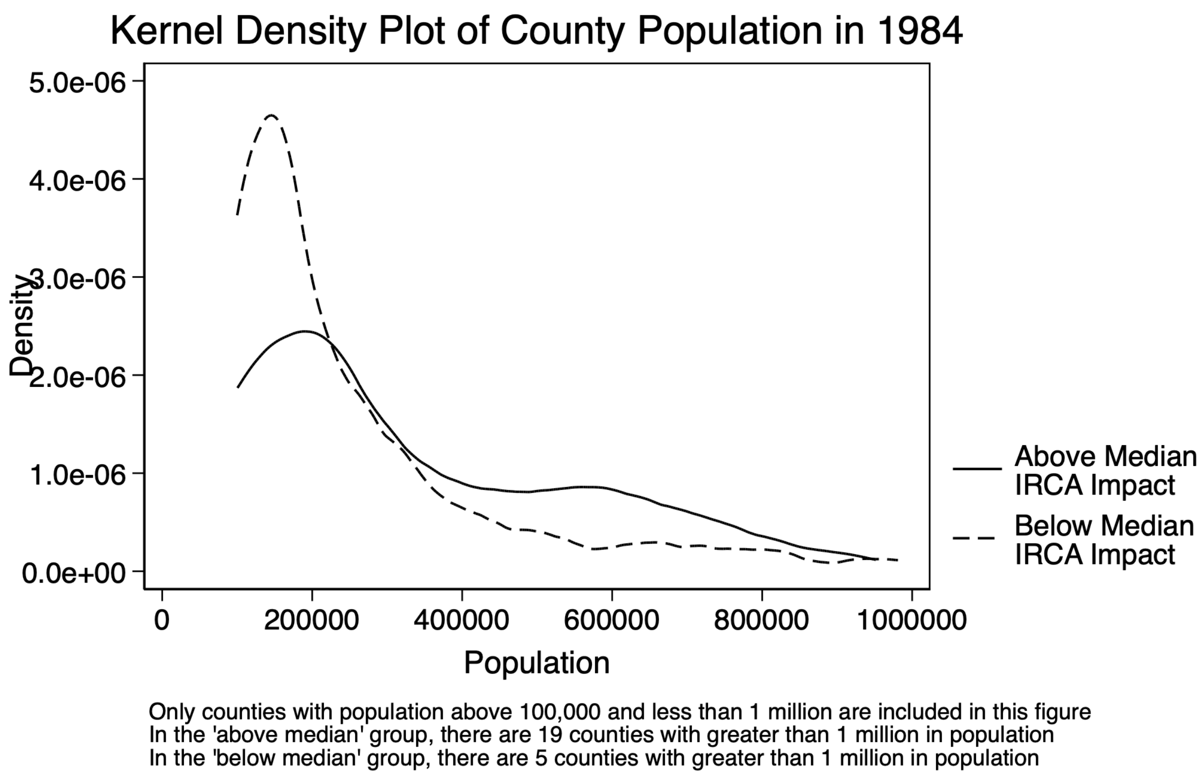


Figure 1.6: Percent of IRCA Recipients in 1992

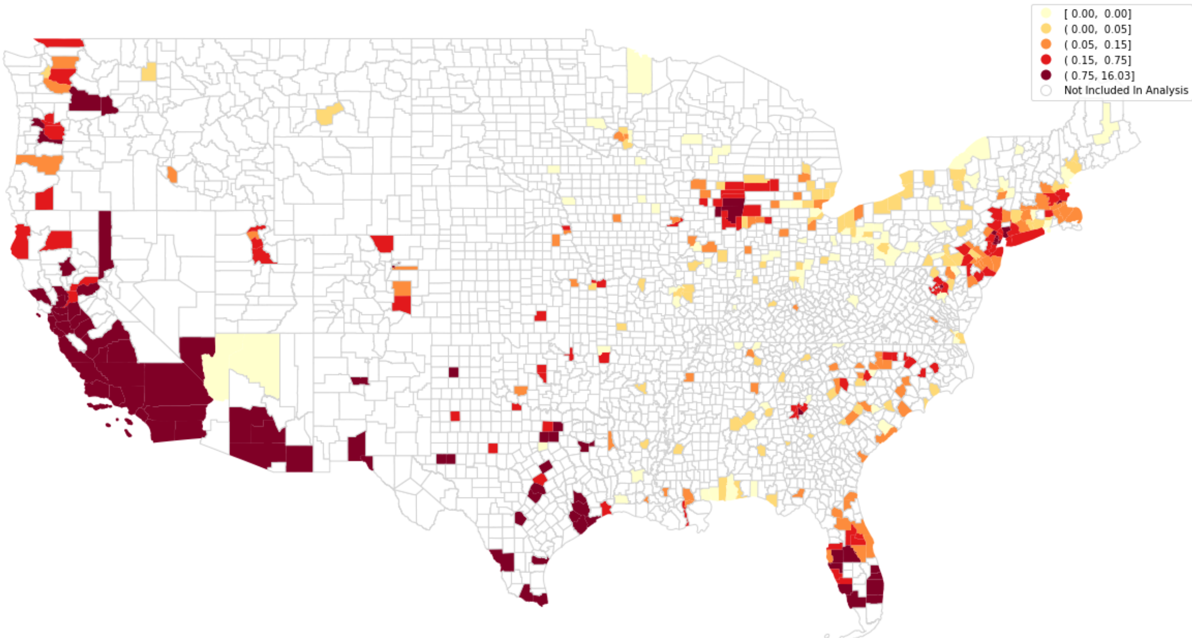


Figure 1.7: Percent of 1986 County Population Legalized

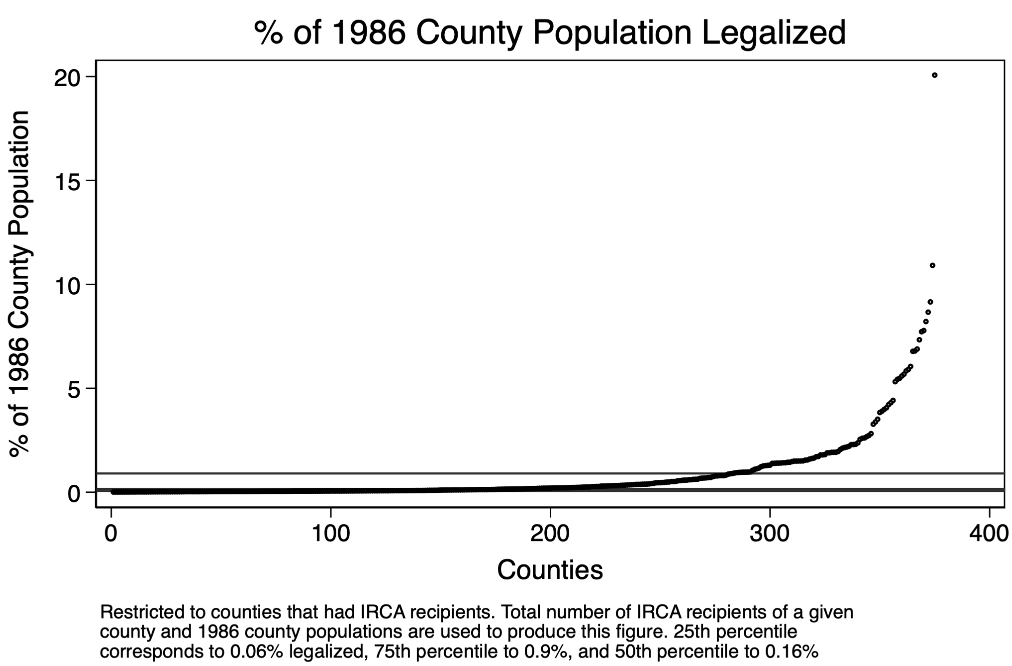
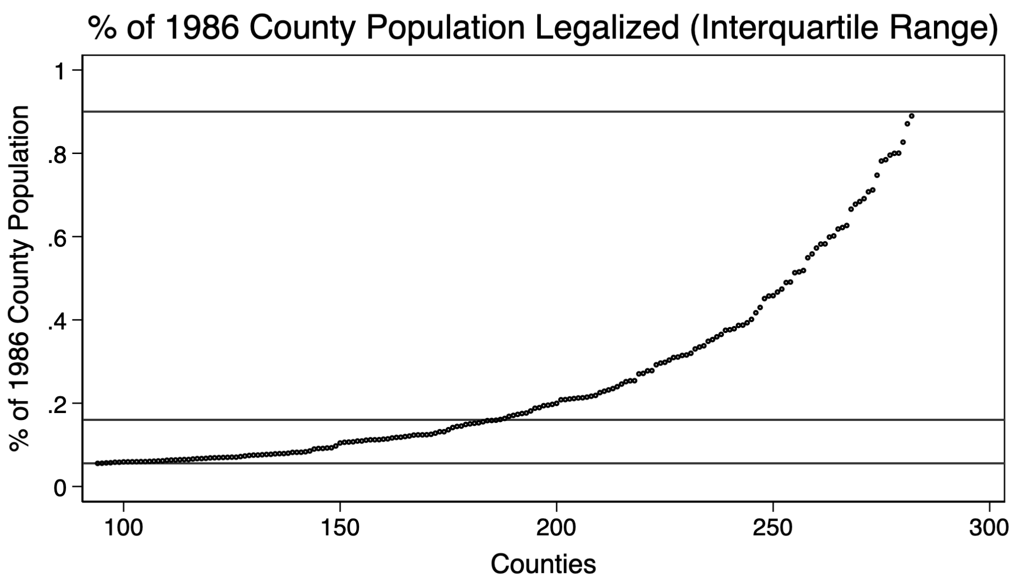


Figure 1.8: Percent of 1986 County Population Legalized (Interquartile Range)



Restricted to counties that had IRCA recipients. Total number of IRCA recipients of a given county and 1986 county populations are used to produce this figure. 25th percentile corresponds to 0.06% legalized, 75th percentile to 0.9%, and 50th percentile to 0.16%

Figure 1.9: Select Main Explanatory Variable

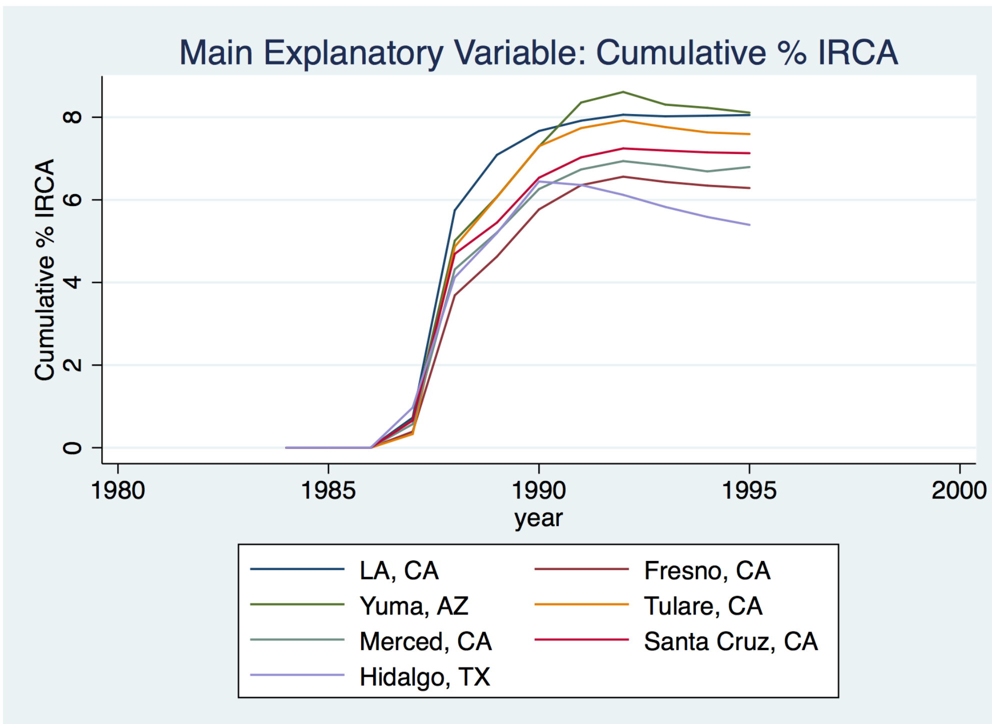
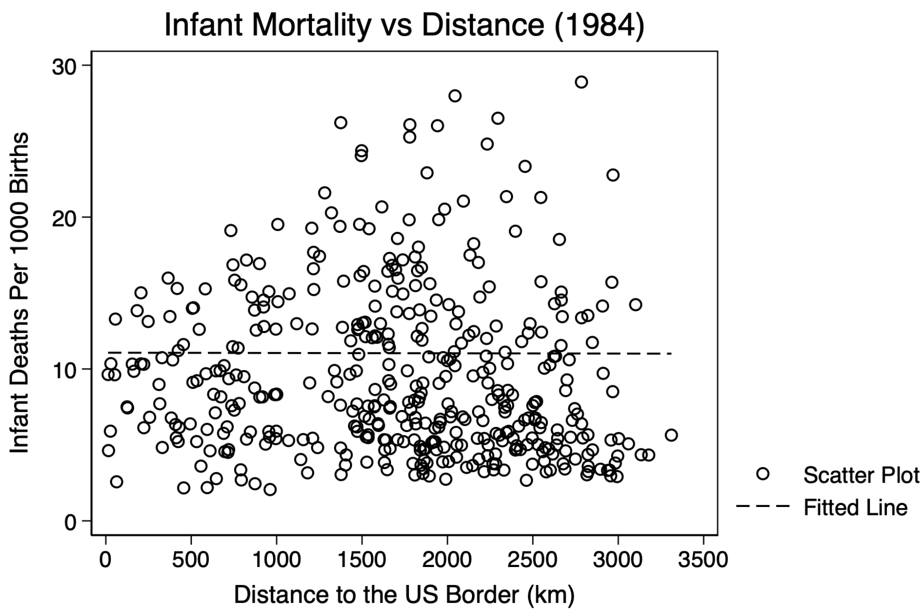
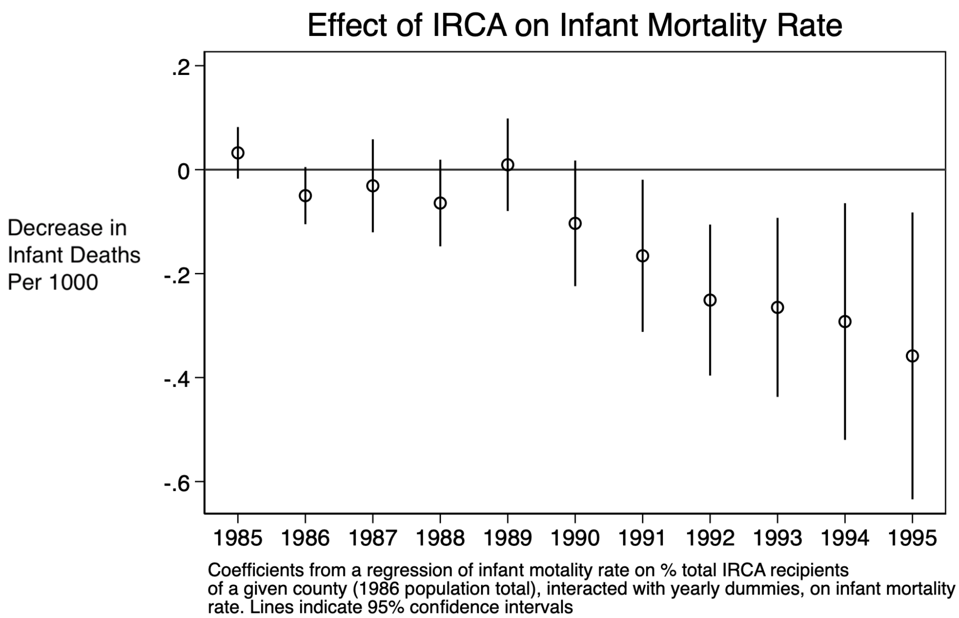


Figure 1.10: Relationship Between Distance and Infant Mortality



A scatter plot between infant mortality rate and distance to the US-Mexico Border. Only counties that are included in the analysis are depicted here (Counties that had infant mortality rates below and above 1 and 99 percentiles are removed). A regression between infant mortality rates in 1984 and distance to the border, weighted by county population, reveals no relationship between the two: t-statistics of -0.07 with p-value of 0.95

Figure 1.11: Effect of IRCA on Infant Mortality Rate



1.8 References

"DACA Statistics, FY2016 Q3." USCIS. USCIS, 2016. Web. 28 Nov. 2016.

"Department of Justice." USCIS. US Department of Justice, INS, 30 May 2001. Web. 28 Nov. 2016.

"Legalization Public Use Microdata Files, 5/5/1987 - 8/15/1992." National Archives and Records Administration. National Archives and Records Administration, Web. 26 June 2017.

"LPS: Legalized Population Survey." Princeton University. Web. 26 June 2017.

Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." Journal of Economic Perspectives, 25 (3): 153-72.

Altonji, Joseph G., and David Card. "The effects of immigration on the labor market outcomes of less-skilled natives." Immigration, trade, and the labor market. University of Chicago Press, 1991. 201-234.

Amuedo-Dorantes, Catalina & Bansak, Cynthia, 2011. "The Impact of Amnesty on Labor Market Outcomes: A Panel Study Using the Legalized Population Survey," IZA Discussion Papers 5576, Institute for the Study of Labor (IZA).

Amuedo-Dorantes, Catalina, Cynthia Bansak and Steven Raphael.2007. "Gender Differences in the Labor Market: Impact of IRCA." American Economic Review, 97(2): 412-416.

Angrist, Joshua D., and Jorn-Steffen Pischke. Mostly harmless econometrics: An empiricist's companion. Princeton university press, 2008.

Bailey, Martha, Karen Clay, Price Fishback, Michael Haines, Shawn Kantor, Edson Severnini, and Anna Wentz. U.S. County-Level Natality and Mortality Data, 1915-2007. ICPSR36603-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2016-10-27. https://doi.org/10.3886/ICPSR36603.v1

Bailey, Martha, Karen Clay, Price Fishback, Michael Haines, Shawn Kantor, Edson Severnini, and Anna Wentz. U.S. County-Level Natality and Mortality Data, 1915-2007. ICPSR36603-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2016-10-27. https://doi.org/10.3886/ICPSR36603.v1

Baker, Bryan, and Nancy Rytina. "Estimates of the Unauthorized Immigrant Population Residing in the United States: January 2012." Department of Homeland Security. Department of Homeland Security, Mar. 2013. Web. 28 Nov. 2016.

Baker, Scott R. "Effects Of Immigrant Legalization on Crime: The 1986 Immigration Reform And Control Act." (2013).

Bartel, Ann P. "Where Do the New U.S. Immigrants Live?" Journal of Labor Economics 7, no. 4 (1989): 371-91. Accessed February 12, 2021. http://www.jstor.org/stable/2535134.

Behrman, Jere & Rosenzweig, Mark. (2004). "Returns to Birth Weight. The Review of Economics and Statistics." 86. 586-601. 10.1162/003465304323031139.

Bell, Brian, Francesco Fasani, and Stephen Machin. "Crime and immigration: Evidence from large immigrant waves." Review of Economics and statistics 21.3 (2013): 1278-1290.

Black, Sandra & Devereux, Paul & Devereux, Michael & Salvanes, Kjell. (2007). "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." The

Quarterly Journal of Economics. 122. 409-439. 10.1162/qjec.122.1.409.

Borjas, George J. "The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market." The quarterly journal of economics 118.4 (2003): 1335-1374.

Card, David. "Using regional variation in wages to measure the effects of the federal minimum wage." ILR Review 46.1 (1992): 22-37.

Card, David. Industrial and Labor Relations Review, Vol. 43, No. 2. (Jan., 1990), pp. 245-257.

Card, David. Journal of Labor Economics, Vol. 19, No. 1. (Jan., 2001), pp. 22-64.

Cascio, Elizabeth and Lewis Ethan. "How Much Does Amnesty Strengthen the Safety Net? Evidence from the Immigration Reform and Control Act of 1986" NBER. 2017

Case, Anne, Darren Lubotsky, and Christina Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient ." American Economic Review, 92 (5): 1308-1334.

Cavazos-Rehg, Patricia A., Luis H. Zayas, and Edward L. Spitznagel. "Legal Status, Emotional Well-Being and Subjective Health Status of Latino Immigrants." Journal of the National Medical Association 99.10 (2007): 1126 - 1131. Print.

Cooper, Betsy, and Kevin O'Neil. "Lessons From The Immigration Reform and Control Act of 1986." Migration Policy Institute. Migration Policy Institute, Aug. 2005. Web. 26 June 2017.

Cortes, Kalena E. 2013. "Achieving the DREAM: The Effect of IRCA on Immigrant Youth

Postsecondary Educational Access." American Economic Review, 103(3): 428-32.

Currie, Janet, and Douglas Almond. "Human capital development before age five." Handbook of labor economics. Vol. 4. Elsevier, 2011. 1315-1486.

Currie, Janet, and Jonathan Gruber. "Saving babies: the efficacy and cost of recent changes in the Medicaid eligibility of pregnant women." Journal of political Economy 104.6 (1996): 1263-1296.

Currie, Janet, and Enrico Moretti. "Biology as destiny? Short-and long-run determinants of intergenerational transmission of birth weight." Journal of Labor economics 25, no. 2 (2007): 231-264.

Finch, Brian Karl, and William A. Vega. "Acculturation stress, social support, and self-rated health among Latinos in California." Journal of immigrant health 5.3 (2003): 109-117.

Friedberg, Rachel M., and Jennifer Hunt. "The impact of immigrants on host country wages, employment and growth." The Journal of Economic Perspectives 9.2 (1995): 23-44.

Gelber, Alexander, Timothy Moore, and Alexander Strand. "Disability insurance income saves lives." NBER Working Paper (2018).

Hasanali, Stephanie. "Understanding the Role of Immigration Policy in Health Care of Immigrants." Thesis. Penn State University, 2015. Penn State. Web. 17 Mar. 2017.

Heckman, James J. (2008) : The Case for Investing in Disadvantaged Young Children, CESifo DICE Report, ISSN 1613-6373, ifo Institut für Wirtschaftsforschung ander Universität München, München, Vol. 06, Iss. 2, pp. 3-8

Mastrobuoni, Giovanni, and Paolo Pinotti. 2015. "Legal Status and the Criminal Activity of Immigrants." American Economic Journal: Applied Economics, 7(2): 175-206.

"Mexico and US Border." Arcgis.com. ArcGIS, May 2017. Web. 26 June 2017. https://www.arcgis.com/home/item.html?id=d33dc009aacf496aae0e2fc050a978d5

Minnesota Population Center. National Historical Geographic Information System: Version 11.0. Minneapolis: University of Minnesota. 2016. http://doi.org/10.18128/D050.V11.0. http://www.nhgis.org

Oreopoulos, Philip, Mark Stabile, Randy Walld, and Leslie L. Roos. "Short-, medium-, and long-term consequences of poor infant health an analysis using siblings and twins." Journal of human Resources 43, no. 1 (2008): 88-138.

Ottaviano, Gianmarco IP, and Giovanni Peri. Rethinking the Gains from Immigration: Theory and Evidence from the US. No. w11672. National Bureau of Economic Research, 2005.

Pan, Y. J Labor Res (2012) 33: 119. doi:10.1007/s12122-012-9134-0

Passel, Jeffrey S., and Mark Hugo Lopez. "Up to 1.7 Million Unauthorized Immigrant Youth May Benefit from New Deportation Rules." Pew Research Center's Hispanic Trends Project. Pew Research Center, 14 Aug. 2012. Web. 28 Nov. 2016.

Paulson, Anna et al. "Financial Access for Immigrants." Financial Access for Immigrants - Federal Reserve Bank of Chicago. Federal Reserve Bank of Chicago, May 2006. Web. 26 June 2017.

Peri, Giovanni. "The effect of immigration on productivity: Evidence from US states."

Review of Economics and Statistics 94, no. 1 (2012): 348-358.

Powell-Griner, Eve & Streck, D. (1982). A closer examination of neonatal mortality rates among the Texas Spanish surname population. American journal of public health. 72. 993-9. 10.2105/AJPH.72.9.993.

Roberto G. Gonzales and Leo R. Chavez, "Awakening to a Nightmare: Abjectivity and Illegality in the Lives of Undocumented 1.5-Generation Latino Immigrants in the United States," Current Anthropology 53, no. 3 (June 2012): 255-281.

R L Williams, N J Binkin, and E J Clingman "Pregnancy outcomes among Spanish-surname women in California.", American Journal of Public Health 76, no. 4 (April 1, 1986): pp. 387-391.

SENATE BILL 1070. 2010. State of Arizona Senate.

Toomey RB, Umaña-Taylor AJ, Williams DR, Harvey-Mendoza E, Jahromi LB, Updegraff KA. Impact of Arizona's SB 1070 immigration law on utilization of health care and public assistance among Mexican-origin adolescent mothers and their mother figures. Am J Public Health. 2014 Feb;104 Suppl 1(Suppl 1):S28-34. doi: 10.2105/AJPH.2013.301655. Epub 2013 Dec 19. PMID: 24354823; PMCID: PMC3924594.

Torche F, Sirois C. Restrictive Immigration Law and Birth Outcomes of Immigrant Women. Am J Epidemiol. 2019 Jan 1;188(1):24-33. doi: 10.1093/aje/kwy218. PMID: 30358825.

The National Children's Study Symposium. 2011. Health Disparities Among Children of Immigrants. [online] U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, Chapter 6. Available: **here** [Accessed 6 January 2021].

United States Census Bureau, USA Counties Data.

United States. Immigration and Naturalization Service (1978). Statistical yearbook of the Immigration and Naturalization Service. Washington, D.C.: The Service.

Wilkinson, Richard G., and Michael Marmot, eds. Social determinants of health: the solid facts. World Health Organization, 2003.

Woodrow, Karen and Jeffrey Passel. "Post-IRCA Undocumented Immigration to the United States: Assessment Based on the June 1988 CPS," Program for Research on Immigration Policy, 1989.

Zavodny, Madeline. "Determinants of Recent Immigrants' Locational Choices." The International Migration Review 33, no. 4 (1999): 1014-030. Accessed February 11, 2021. doi:10.2307/2547361.

2 Evolution of Self-Reported Health[2]

2.1 Introduction

Few measures of health status are as ubiquitous in social science surveys around the world as the global self-reported health question. A typical example, from the U.S. National Health Interview Survey (NHIS), reads: "*Would you say your health in general is excellent, very good, good, fair, or poor?*"[3] The question is remarkably simple in content, offering no particular guidance as to which aspects of health should be considered or how they should be combined. Respondents are free to interpret the question in any way they like, prioritize those health domains they deem most relevant, or evaluate their health relative to a particular point in time or group of peers. Indeed, the cognitive processes underlying response formation to abstract self-assessments are not well understood. Moreover, it has been pointed out that a response to the above survey question may also be related to the sequence of questions that precede this particular question. Consequently, the literature has been struggling to interpret the responses, and posed the question: What does self-reported health really measure?

The most striking result from twenty-five years of epidemiological research on this question has been the consistent finding that self-reported health is highly predictive of subsequent

[2]As noted in the ACKNOWLEDGEMENTS section, this work is an updated version of Buchinsky, Moshe and Maestas, Nicole and Reggio, Iliana. "The Evolution of Self-Reported Health," 2014. The work is in preparation for publication.

[3]The precise wording varies from survey to survey and four response categories are sometimes used instead of five. For example, a common alternative is "*In general, would you say your health is: excellent, good, fair, or poor?*" In some European countries the convention is to label the five response categories "*very good, good, fair, bad, and very bad.*" Some variants ask individuals to rate their health over a particular time period, relative to a point in the past, or compared to a group of peers.

mortality, a highly objective measure of "true" health by any standard (Idler and Benyamini 1997). The relationship between self-reported health and mortality holds with respect to both long-term and short-term mortality, and persists in the presence of extensive controls for known health risk factors and socioeconomic status (Idler and Benyamini 1997). Even the dose-response relationship is robust across studies, with mortality risk rising monotonically as health ratings progress from excellent to poor (Idler and Benyamini 1997). Moreover, its predictive ability is robust to variations in question wording and translation, national origin, the number and labeling of response categories, whether a time frame is specified, or whether a particular reference group is indicated (Idler and Benyamini 1997).

Although the mortality relationship has been the focus of most studies, self-reported health also predicts future decline in physical functioning (Idler and Kasl 1995; Lee 2000). Largely, the conclusion of this literature is that self-reported health is an accurate measure of true health, perhaps more accurate and more inclusive than other measures on account of its apparent ability to tap domains of health that are beyond the reach of other measures (Idler and Benyamini 1997).

While the vast majority of studies have related baseline levels in self-reported health to subsequent mortality, a related question that has received much less scrutiny is, what do changes in self-reported health measure? Measurement theory would indicate that even if the level of self-reported health is correlated with true health, unless that correlation is perfect, there is also an error component that may not be random (Bound 1991; Currie and Madrian 1999). Understanding the nature of this error component and its dynamic properties are crucial for interpreting changes in self-reported health over time. This is a salient issue for policy making on account of the fact that the time profile in self-assessments like self-reported health are often used to measure trends in population health, health inequality, and the decline of health over the life-cycle (see e.g., Crimmins 1990; Kunst et al. 2004; Contoyannis, Jones and Rice 2004; Case and Deaton 2005). The conceptual difficulties with trend analysis of health statistics were noted by Drury and Wilson (1984), and more recently, by Waidmann, Bound and Schoenbaum (1995). Both studies argue that purported health declines in the 1970s may have in fact been due to changes in other factors, such as disease detection and awareness.

One particularly useful framework for interpreting change over time in self-assessments is that of Golembiewski, Billingsley and Yeager (1979). They posit three types of changes. The first, called *alpha* change, refers to an absolute quantitative change in the construct of interest. In the context of health status, an alpha change would be a change in true health status (typically what we are trying to measure). The second type, called *beta* change, describes change resulting from recalibration of the measuring instrument. For example, a change in the threshold value of latent true health that distinguishes fair from poor health would be a beta change. Beta change is also known as "response shift," whereby the internal standards of measurement or the response criteria change (Howard and Dailey 1979). Such effects are hypothesized to arise, for example, if individuals beset with illness learn to cope or change their expectations (Bjorner, Fayers and Idler 2005).

The third type, called *gamma* change, describes a major change in "the perspective or frame of reference within which phenomena are perceived and classified" (Golembiewski, Billingsley and Yeager, 1979, p. 135). In the context of self-reported health status, gamma change would result from a change in the state variables guiding the decision process. Such state variables could include social phenomena, reference groups, public health information, scientific advancements, technological innovations, and economic conditions. Policy interventions themselves can cause gamma change if, for example, individuals change their outlook about a particular issue as a result of an intervention.[4]

The Golembiewski, Billingsley and Yeager framework reveals a fundamental identification challenge. Does an observed change in self-reported health represent alpha change, or is it confounded by beta or gamma changes? The identification problem exists whether we consider individual-level change or population-level change, and confronts self-assessments generally, not just measures of self-reported health. In many settings, it is not possible to separately identify these three alternative factors. Hence, it makes it almost impossible to directly test for the presence of confounding changes such as these. In health research, only a small group of studies has attempted to test for beta or gamma change in individual-level responses, although many more have documented seemingly inconsistent response patterns that point to the likely presence of either beta or gamma change (see Sprangers and Schwartz (1999) for a short review of the epidemiological literature, and Currie and Madrian (1999) for a review of the economics literature).[5]

Among studies that have tested directly for changes of this nature are Allison, Locker and Feine (1997), who document evidence of reference group change in the measurement of health-related quality of life. Lindeboom and van Doorslaer (2004), find inconclusive evidence of "cut-point shift" (beta change) in self-reported health responses in the Canadian National Population Health Survey. In this paper, we present an analysis of the extent to which changes in self-reported health reflect beta and gamma changes in addition to true changes in health status. Using 21 years of data from the U.S. National Health Interview Survey (NHIS), we examine the time trend in self-reported health as it relates to changes in body

[4]Sprangers and Schwartz (1999) offer a revised typology in which the term "response shift" encompasses gamma changes as well as beta changes.

[5]A related strand of research has investigated differential use of response scales across countries in order to facilitate cross-country health comparisons (e.g., Jurges 2007; Kapteyn et al. 2007).

mass index (BMI), a well-defined and relatively objective measure of health (or health risk). First we show that virtually every segment of the population (e.g., stratified by gender, age, socioeconomic status, region) experienced a notable rise in BMI during the period from 1982 to 2002. We interpret this rise in BMI as evidence of a decline in the true underlying health of the U.S. population. During the same period, mean self-reported health declined slightly and there was a reduction in the variance in self-reported health as individuals became less likely to classify their health as excellent and more likely to choose very good; however, this was not all that occurred. We show that the mapping between BMI and self-reported health changed in two important ways. First, the cut points in an ordered response model of self-reported health have changed over time, and second we show that the coefficient defining the relationship between self-reported health and BMI also changed. We interpret the former as evidence of beta change and the latter as evidence of gamma change. We offer support for our interpretation of gamma change by presenting evidence that the variation over time in self-reported health is highly responsive to a number of macroeconomic variables. In particular, we find evidence of a strong pro-cyclical relationship–when the economy improves, so does health. This pattern is contrary to previous work that showed a countercyclical relationship between the number of acute and chronic health conditions (self-reported) and the unemployment rate during the 1970s (Ruhm 2003).

2.2 The Data

The data for this study comes from the National Health Interview Survey (NHIS). This data set is the principal source of information on the health of the civilian non-institutionalized population living in the United States at the time of the interview.[6] The NHIS was originated by the National Health Survey Act of 1956. The main objective of the NHIS is to monitor the health of the United States population through the collection and analysis of data on a broad range of health topics. The NHIS provides detailed information about health characteristics broken down by many demographic and socioeconomic characteristics. The NHIS is used extensively by the Department of Health and Human Services (DHHS) to monitor trends in illness and disability and to track progress toward achieving national health objectives. It also provides the primary source of information for studies in various disciplines, including public health, epidemiology, sociology, economics, and many others.

The NHIS is a cross-sectional household interview survey that has been conducted continuously since 1957. The sampling plan, which was redesigned in 1995, follows a multistage area probability design that permits the representative sampling of households. Overall the NHIS data are collected annually from approximately 45,000 households, including about 93,000 individuals. While survey participation is voluntary, the annual response rate of the NHIS is greater than 90% of the eligible households.

However, the content of the survey has been updated from time to time. In particular, in 1997 there was a substantial revision. One of the major revision implemented was for the purpose of reducing proxy-reporting bias: per household a sample adult and a sample child were interviewed and others were not allowed to answer for those sampled. While the changes introduced in the NHIS dramatically improved the ability of the survey to provide important health information, it also introduced some problems including the inconsistency of some of the variables (including some of those we use in this study) over time.

[6]Because of logistical problems, several segments of the population are not included in the sample. These include individuals living in long-term care facilities; persons on active duty with the Armed Forces; and U.S. nationals residing abroad.

Since 1997 the NHIS has three parts or modules: a Basic module, a Periodic module, and a Topical module. The Basic module replaced the core questionnaire and remain largely unchanged. It contains three components: the Family Core, the Sample Adult Core, and the Sample Child Core. The Family Core gathers information on everyone in the family. This includes: household composition and sociodemographic characteristics; tracking information; information for matches to administrative databases; and basic indicators of health status and utilization of health care services.

From each family in the NHIS, one sample adult and one sample child, if applicable, are randomly selected and information on each individual is collected with the Sample Adult Core and the Sample Child Core questionnaires. While the two questionnaires differ somewhat, they both collect basic information on health status, health care services, and behavior. The Periodic Module collected more detailed information on topics related to the Basic Module, and the Topical Module asked questions on current health topics.

In this study we mainly use information about the individual and family characteristics and a variety of responses, objective and subjective, regarding the health status of each individual. Table 2.1 provides summary statistics on the extract used in this study. The average age of the individual in the NHIS was 44 in 1997 and has increased to about 47 by 2018. This largely comes from the fact that the U.S. population is aging. About 51% of the sample is women, especially for the sample years from 1997 onward.

An average individual in the sample reports that his/her health is somewhat better than good (where good is the middle point on a scale from 1 to 5). One particularly interesting development is with respect to the *body mass index* (BMI). This index increased substantially over the sample years (from 26 in 1997 to 28 by 2018). Also there are continuous declines over the sample years in the fraction of people having education level of high school or lower and a dramatic increase in the fraction of individuals who are college graduates, or have some college experience.

Reflecting the changes in the family composition in the U.S., we see that the fraction of married individuals declined over the years, while the fraction of those never married increased. The rest of the extract contains information about various diseases, as well as information about *activities of daily living* (ADL) and *independent activities of daily living* (IADL). This information is used in constructing a health index for individuals in the sample as is explained below.

2.3 Self-Reported Health Measure

The self-reported health (SRH) item is commonly provided in many data sets, including the NHIS. It is a self-assessment, by the individual regarding his/her health status. The variable is coded in five categories: 5 = excellent; 4 = very good; 3 = good; 2 = fair; and 1 = poor. Our analysis is devoted to looking at this particular variable from various angles. Most of the analysis is purely descriptive. Nevertheless, it provides a detailed account of the changes in observed SRH over time. Moreover, we link the changes in the underlying health index to changes in a number of variables measuring the level of economic activity. The variables we consider here are: (a) unemployment rate; (b) inflation; and (c) real per capita consumption. We first examine *How did reporting of self-reported health change over time*? In Table 2.2 we show the mean reported value over time for the whole population, and by age group. In Figure 2.1 we provide a visual representation of the time pattern in mean SRH along with two of the economic indicators, namely inflation and unemployment, for the whole population, and then for specific age groups. Over the period 1997-2018, the mean value for the U.S. population is centered around 4 – i.e., on average a person is in "very good" health. As expected, older people report that they are less healthy than younger people. In addition there is noticeable variation in the mean of the self-reported health measure over the sample period within age groups. This variation is shown in Figure 2.1a for the whole population and in Figures 2.1b–2.1d for several age groups. While for all groups there is some variation over time, the pattern is not the same for all groups, and, in general, there is more variation over time for the older groups.

The pattern of changes in mean SRH differs across the various age groups. Generally, there is a sharper decline in the mean reported health of the older cohorts. There might be many reasons for this dramatic trend. While we do not attempt to investigate the underlying reasons for these changes it is consistent with the findings in the literature (Case and Deaton 2015; Meara and Skinner 2015). As the literature suggests, significant declines in health have been observed over this time period, especially for the 45-54 year olds.

In order to evaluate whether average self-reported health are correlated with economic indicators, we run a simple OLS regression of mean self-reported health on a number of economic indicators in the U.S. These include: the unemployment rate, inflation, and real per capita consumption. The results are reported at the bottom of Table 2.2. Note that, in general, there is a very high correlation between these measures and self-reported health. The individual R^2's by age groups vary between .5 and 0.9, quite large by any standard. Specifically, the oldest age group is less sensitive to the changing economic environment compared to the youngest and middle-aged group. We return to the link between economic activity and self-reported health below, when discussing the effects of various factors on the individuals' health indices.

In the five panels of Table 2.3 we provide information about the fraction of people reporting each one of the five categories over the sample years, i.e., 1997 to 2018. Figures 2.2a – 2.2d provide the graphs that correspond to the various columns of Table 2.3. Finally, Figures 2.2e and 2.2f provide similar graphs for the female and male populations.

The most visible phenomenon revealed by the tables and figures is the systematic decline in the fraction of people reporting excellent health and a systematic increase in those reporting good health. While there are some changes in the fractions of the other three categories, they are a lot less pronounced than those for the two aforementioned categories. Overall, we can see that for the population as a whole (Figure 2.2a) there is an overall decline of about 4 percentage points in the excellent category and an increase of 2 percentage points in the good category. An interesting observation is that the decline in the fraction of people in the excellent category is sharp for the middle-aged cohorts than any other cohorts at about 6 percentage points. In addition, over the sample years the fraction of those reporting very good health remains about the same while we observe a noticeable increase in the fraction of those reporting good health at about 3 percentage points.

Comparing Figure 2.2e for the female population with Figure 2.2f for the male population also reveals significant differences. In particular, the fraction of men that report themselves in excellent health far exceeded that of the female population during the early sample years. Nevertheless, the fraction of males in that category declines significantly over time, while the decline for the female population is less pronounced. This is matched with the increase in the fraction of those reporting to be in good health. Note that the fraction of individuals reporting themselves in very good health for both men and women is similar.

In Panels A and B of Table 2.4 we provide the information about the evolution of the mean self-reported health for women and men, respectively. We also depict the changes in mean self-reported health for women and males in Figures 2.3 and 2.4, respectively. The two figures are organized in the same way as Figure 2.1; that is, in each figure we provide the time series for the unemployment rate and inflation in the U.S. along with mean self-reported health for the group.

Comparing first Figures 2.3a and 2.4a, for females and males we observe that the general patterns of change over time are quite similar for the two groups. However, there is one crucial difference that stands out. Generally, females tend to report they are in worse health than males. This is true not only for the two groups as a whole (see Figures 2.3a in comparison with Figure 2.4a), but also for each age group within the two groups (see Figures 2.3b–2.3d in comparison with the Figure 2.4b–2.4d, respectively). Nevertheless, the similar standard deviations of the mean responses over time indicate that they respond similarly to factors affecting their self-reported health (see the standard deviation reported at the bottom of Panels A and B of Table 2.4). Differences do exist nonetheless between specific age groups within each gender group and between women and men. In particular, we see in the two panels of Table 2.4 that the R^2's from the regression of the group mean on a set of economic indicators are, generally, larger for the middle-aged groups, for which the economic indicators can be associated with over 80% of the variation in the data. Also, while for the youngest and middle-aged groups is a downward trend in the way they evaluate, on average, their health (i.e., providing worse evaluation), this trend is a lot less pronounced for the oldest group.

Several differences in the pattern of changes for the two groups are worth noting. In particular, we note the differences in the pattern of changes between males and females in the youngest and middle-aged groups. In general we see that the mean of self-reported health has a more pronounced decreasing trend for males than for females, indicating that males became more similar to women in the way they assess their health, at least for those below the age of 55. For the older cohorts, men seem to be more similar to women throughout the sample period.

These findings suggest major shifts in health in the United States. True underlying health may have changed. On top of this, other potential changes exist. First, it might indicate that something major has changed in the way individuals evaluate their health. Second, there might have been a systematic change in the way people report their health. Of course, a combination of these three possible changes is also a valid possibility. As for the first possibility, the availability of more information and the faster transmission of information certainly make people learn faster about potential health problems that are associated with certain observed characteristics. For example, people are more aware of the potential hazards associated with being overweight, or the long term effects of actions that they have taken in the past, e.g., smoking. Moreover, the advancement of science in general, and of medical science in particular, allows detection of more medical problems over time as well as detection of known problems earlier in one's life.

2.4 The Model

2.4.1 The Ordered-Probit Model

What are the factors affecting individuals' self-reported health status? Do these factors have similar weights on all individuals reporting? To analyze that we estimate a simple ordered-probit model using all available data from 1997 through 2018. We allow the coefficients to vary across years and estimate the model for each year separately. We also estimate the model separately for the male and female populations, allowing the coefficients to change across these two populations in each sample year. While, for brevity, we do not report all the individual coefficients from our estimation, it is worthwhile noting that the very large data sets we use allows us to estimate the coefficients very precisely.

As noted above, the model we use here is a simple ordered-probit model. Without loss of generality we assume that there exists a latent health index given by

$$I_{it}^{*} = x_{i}'\beta_{t} + \varepsilon_{it},$$

where x_{it} is the vector of individual-specific objective health measures, that is, the determinants of the individual's overall health index. The vector β_t is a vector of unknown parameters. These parameters are to be interpreted as the weights that an individual assigns to each of the corresponding health measures. Finally, ε_{it} is an idiosyncratic error term, uncorrelated with x_{it}, that is

$$\varepsilon_{it}|x_{it} \sim N\left(0, \sigma_{\varepsilon}^{2}\right).$$

For identification purposes we assume that $\sigma_{\varepsilon}^{2} = 1$. We assume that when an individual is asked to report his/her SRH described above he/she will report the health condition according to following classification:

$$d_{it}^{SRH} = \begin{cases} 1 & \text{if } I_{it}^{*} \leqslant c_{1t} \\ 2 & \text{if } c_{1t} < I_{it}^{*} \leqslant c_{2t} \\ 3 & \text{if } c_{2t} < I_{it}^{*} \leqslant c_{3t} \\ 4 & \text{if } c_{3t} < I_{it}^{*} \leqslant c_{4t} \\ 5 & \text{if } I_{it}^{*} > c_{4t} \end{cases},$$

where $d_{it}^{SRH} = 1$ corresponds to *poor* health, $d_{it}^{SRH} = 2$ corresponds to *fair* health, etc. The parameters in $c_t = (c_{1t}, ..., c_{4t})'$ are the thresholds that determine the individual's classification of his/her health. It follows immediately that

$$P_{ij}(x_{it}) = \Pr\left(d_{it}^{SRH} = j|x_{it}\right) = \begin{cases} \Phi(c_1 - x_{it}'\beta_t) & \text{for } j = 1 \\ \Phi(c_2 - x_{it}'\beta_t) - \Phi(c_1 - x_{it}'\beta_t) & \text{for } j = 2 \\ \Phi(c_3 - x_{it}'\beta_t) - \Phi(c_2 - x_{it}'\beta_t) & \text{for } j = 3 \\ \Phi(c_4 - x_{it}'\beta_t) - \Phi(c_3 - x_{it}'\beta_t) & \text{for } j = 4 \\ 1 - \Phi(c_4 - x_{it}'\beta_t) & \text{for } j = 5 \end{cases}.$$

Note that all parameter vectors β_t as well as c_t are indexed by t, so that they are allowed to change over time. Consequently, the reporting of SRH can change depending both β_t and c_t. One of the main goals of the current paper is to explore the potential links between the actual reporting of SRH and changes in these parameters. Another goal is to examine the potential links between the changes in the model's parameters and changes in the economic environment.

2.4.2 The Control Variables

In the regression we use a number variables that control for both general individual characteristics, as well as specific objective measures of health. The demographic variables include: income, age, dummy variable for college education (some college education versus not), three regional dummy variable (North East, Mid-West and South), a dummy variable for being African American, two marital status variables (married and never married), number of children under 5 years old, number of children under 15 years old and six occupational specific dummy variables. For the objective measures of health we use: BMI and its squared term, four dummy variables for number of bed days (0 to 7, 8 to 30, 31 to 180, and more than 180 bed days), two dummy variables for having problems with ADL and IADL. We also account for having a number of specific chronic diseases, namely arthritis, asthma, bronchitis, any type of cancer, diabetes, emphysema, chronic headaches, hearing problems, ulcer, hypertension, heart disease, vision problems, sciatica, and sinusitis.

2.5 The Results

One of our main findings is that all coefficients' estimates are changing, and in very particular patterns over the sample years. That is, the way in which people evaluate their health in response to a particular factor has been changing. Moreover, for at least some of the variables, the levels have changed significantly over time. An example of this is the case of BMI as indicated above and will be further discussed below. For brevity we restrict our analysis to only some of the key factors affecting the latent health index. Further, we decompose the overall change in each of these factors to those that stem from changes in the individuals' evaluation of their own health (i.e., changes in the coefficients' estimates) and changes in the level of the corresponding variables (compositional changes over time). We first analyze the changes of some of the key variables that determine the individuals' health indices, namely: (a) age; (b) BMI; (c) hypertension; (d) diabetes; and (e) income. These variables are the top five variables that contribute to the evaluation of one's health index as seen in Figure 2.5.

Below we provide the results for these factors. For each factor we provide the results for the coefficient, the portion of the health index that is explained by the specific variable, namely $I_k^* = x_k' \beta_k$, for $k = 1, ..., K$, the change in I_k^* that is due to changes in the corresponding coefficient β_k, and the change in I_k^* due to changes in the corresponding variable x_k.

In Tables 2.5 through 2.11 we also report the results from a simple OLS regression of coefficient and threshold estimates from the model on the set of principal components of economic variables. The economic variables consist of 21 variables selected from the Federal Reserve Economic Data[7].

2.5.1 The Effect of Age

The distribution of age in the population has gradually changed, reflecting an overall aging of the population over time (see Figure 2.6).

Note first from Figure 2.7, in absolute terms, the effect of age has decreased over the sample period for both women and men. This indicates that age is perceived to be less negative to health over time. The effect for men is larger in absolute terms than that for women. When we examine the health index portion attributable to the age variable, we see that it is likewise decreasing over time (Figure 2.8). Decomposing the change in age portion of health index, we can see that it is mostly due to the change in coefficient versus the change in age composition in the population (Figure 2.9). The changes in the age composition have an opposing effect to that of the coefficient on the age variable. That is, aging of the population causes the index to be somewhat lower over time whereas the change in coefficient moves in the other direction. This is true for both women and men (Figure 2.10).

[7] 21 economic indicators are: Real GDP, CPI, Federal Funds Rate, Homeownership Rate, Real Imports, Real Exports, New One Family Houses Sold, Total Construction Spending, Advance Retail Sales, Business Sales, Business Inventory, Consumer Sentiment, New Private Housing Units Authorized by Building Permits, Manufacturing, Leading Index, Bank Prime Loan Rate, Treasury 10 Year Bond Yield, Personal Consumption, Real GDP Per Capita, Percent Changes in GDP, Inflation, and Unemployment.

Finally, note from Table 2.5 that the regression of the estimated age coefficients on the set of economic indicators yields a relatively high R^2 for men ($R^2 = .79$) but not for ($R^2 = .36$). This seems to indicate that the economic situation has a more significant effect on the way people evaluate their health for men than for women.

Specifically, age seems to have a more pronounced negative effect on how individuals rate their health, when the economy is in a downturn. In other words, individuals know how to cope better with their age in a better economic environment.

2.5.2 The Effect of BMI

The effect of BMI on self-reported health has changed dramatically over the past few decades. The effect can be decomposed into two factors. First, as we establish below, there were enormous changes in the BMI distribution in the population. These changes apply to virtually all segments of the population: For women and men, for blacks and whites, for young and old individuals. Second, there were considerable changes in the individuals' valuation of their health status at any given level of BMI. That is, people perceive a BMI level of, say, 30, differently in terms of its effect on health in 2018 as compared with their perception of that level of BMI in 1997. Below we document the contribution of the changes in the distribution of BMI to changes in overall SRH during the period from 1997 through 2018. We then discuss the effects of changes in individuals' valuation over the same period on the health index.

Starting with Figure 2.11, we note that there has been enormous change in the BMI distribution in the population, with a clear shift toward higher BMI.

The proportion of those who are obese has been increasing significantly, from 21% in 1997 to 34% by 2016. When we look further into different subgroups of the population, we see that those aged between 30 to 55

Next Figures 2.12 and 2.13 shows the estimation result of the coefficients for the BMI and BMI^2 variables. In the same figures, we also provide locally smoothed graphs of the same estimates. The figures show that there have also been dramatic changes in how people regard their BMI when assessing health. First note the changes in the coefficients for both BMI and BMI squared.

Generally women seem to place larger negative weight on their BMI in their overall health index relative to men. While earlier on in our study period larger BMI was interpreted by men as better for their health status as indicated by the positive estimate, we see that over time men also started to view it negatively. One noticeable result is the large year-to-year variation in the estimated coefficients, especially for men.

To see the overall effect of BMI on the health measure we depict in Figure 2.14 the average marginal effect of BMI on the health index; this is simply given by the average of the derivatives of the health index with respect to BMI, i.e., $\beta_{BMI} + 2\beta_{BMI^2}\left(BMI\right)_i$, for all individuals in a particular subsample.

The figures indicate that the overall effect for women is larger in absolute terms in each of the sample years. That is, women's valuation of having higher BMI is much more negative than that of men. It also seems that men's valuations are more sensitive to changes in economic factors as is seen from regression results between estimated coefficients of BMI and computed principal components from a set of economic indicators (see Table 2.6). The R^2 from the regressions of the coefficients for BMI and BMI squared on the set of economic indicators are .51 and .40, respectively, for men, but only .43 and .30, respectively, for women.

Figure 2.15 and Figure 2.35 indicates two important findings. First, the share of the health index accounted for by BMI is relatively large for both men and women. To give some perspective, we see in Figure 2.35 that approximately a negative one unit change in the health index is enough for one to downgrade one's self-reported health. Moreover, the share changes more frequently for men than for women. Overall, the increase in BMI tends to reduce individual's valuation of their health, and more pronouncedly so for women, as is clearly seen from Figure 2.15.

In order to determine which of the variation, that of coefficient change or that of composition change, is the greater source of variation in the health index, we show Figure 2.16.

Figure 2.16 indicates that changes in the impact of BMI on the health index come largely from changes in the individuals' valuation, i.e., changes in the coefficients that correspond to BMI. The impact of the composition of BMI in the sample is very smooth, representing a consistent downward shift of the health index. While there may be many reasons as to why individuals' valuation changed significantly over time, here we highlight some studies that help shed light on this issue. One is from Tiggemann (2011):

"There is a great deal of evidence that body image is experienced negatively by the majority of women and girls. Many are dissatisfied with their body, particularly with their body size and weight and wish to be thinner, so much so that weight has been aptly described as "a normative discontent" for women. There is also increasing evidence that men and boys too are beginning to experience body dissatisfaction – in the direction of wishing to be more muscular – albeit at lower rates (for the moment) than their female counterparts."

Perhaps one contributing factor is the well documented association between media consumption and its negative effects on body image, as a meta-analysis by Grabe et al. (2008) found. Holland and Tiggeman (2016) also find that social networks contribute to this trend, and perhaps this is what contributed to the increasingly negative individual valuations of BMI over time, and that so particularly for females.

It is also the case that awareness over health consequences of obesity has increased over time. At least among health-care circles, this is very evident. A meta-analysis of 230 cohort studies ranging from 1986 to 2015 indicated a clear relationship between overweight and obesity with mortality (Aune et al. 2016). On the public health side, we've seen an increase in efforts to raise public awareness to issues related to obesity, one notable example being Michelle Obama's *Let's Move* program. Perhaps telling also is the value of the US market for weight loss and diet control reaching $72 billion in 2018 (Marketdata LLC 2019).

It is puzzling, but perhaps not so much so when examining anecdotal evidence, that even though individuals become more aware of the long-term adverse impact of weight, there is continued increase fraction of the population with relatively very high BMI. This is illustrated more clearly in Figure 2.11. Across the different age groups, we see a 11 to 13% increase in obesity rates over the study period.

2.5.3 The Effect of Hypertension

The prevalence of hypertension has steadily increased over the sample years. As you can see in Figure 2.17, the prevalence rate increased from about 22.9% in 1997 to about 31.5% in 2018 for all US population above the age of 18. More noticeably, we see that hypertension among the older aged (above 55) population is prevalent, increasing from 47.6% to 56.7%.

We see that significant portions of the population are affected, amounting to one-third for all ages and more than half for those above the age of 55. Another interesting trend is that male hypertension prevalence has surpassed that of female's starting around 2010.

Nevertheless, as we will note later, there has also been significant increases in hypertension awareness. In addition there have been significant improvements in treatment through new and better medications, as well as nutrition adjustment and exercises. In fact, among all diseases that individuals suffer from, hypertension is relatively easy to control (see Figure 2.18). It is therefore interesting to examine the effect of such a problem on the general health index. The results for hypertension are presented in Figure 2.19.

Noticeably, for both men and women we see continuous reduction in the negative effect on the health measure as can be observed in the increasing coefficients for hypertension over time. In general, women seem to place less negative valuation on hypertension than men as observed by the consistent level differences. The trend in which we observed the reduction of negativity associated with this condition is more pronounced in the male population. This is interesting given that we noted earlier how prevalence of hypertension for men has surpassed that of women.

Overall, the portion of the health index that is due to hypertension is increasing over time in absolute terms (see Figure 2.20). However, it is rather small, certainly in comparison with other leading diseases. Figure 2.21 sheds light on this observation. Decomposing the changes in the health index due to hypertension, it is clear that the downward trend we observe in Figure 2.20 is mostly driven from changes in hypertension composition. This is an interesting observation. It implies that though there is a steady reduction in the negative valuation of hypertension among individuals, the negative trend in hypertension portion of the health index is explained by an increase in the prevalence of hypertension in the overall population.

Near the sample period that we are examining, major studies were conducted that advanced treatment knowledge for hypertension. Saklayen and Deshpande (2016) highlight some of these. Findings from SHEP 1991 was responsible for "open[ing] up the treatment option to 50 million people in the US alone." DASH 1997 was according to the authors "one of the best interventional diet studies in the hypertension research." ALLHAT 2002, ASCOT 2005, ACCOMPLISH 2008 were major studies that examined how best to administer different types of hypertension drugs, including newer ones. SPRINT 2015 "conclusively demonstrated the benefit of lowering systolic blood pressure goal in a non-diabetic population." These highlighted studies advanced the body of knowledge that helped treat hypertension. Consistent with this, we saw in Figure 2.18 that over the study period the prevalence of blood pressure control has steadily increased for US adults.

For a problem that can be largely controlled by medication one would expect that changes in the economic situation may not have an effect on the valuation of having this problem. Table 2.7 does seem to corroborate this. The R^2 from the regression of the coefficient on hypertension on the set of economic indicators is relatively small for both female ($R^2 = .08$) and male ($R^2 = .19$). This is not surprising since it is a condition that can be easily controlled for most people.

2.5.4 The Effect of Diabetes

The prevalence of diabetes has steadily increased over the sample years. As you can see in Figure 2.23, the prevalence rate increased from about 5% in 1997 to about 10% in 2018 for all US population above the age of 18. This comprises a significant portion of the U.S. adult population. Another noteworthy fact is that health care spending on diabetes treatment was the largest among all health spending on treatments in 2013 in the US, estimated to be at $101.4 billion (Dieleman, Baral, Birger et al. 2016). The rates of diabetes do not differ significantly across gender lines, though males do exhibit slightly higher prevalence of diabetes. When we further compare diabetes rates across different age groups, it is clear that most of the diabetes diagnosis is among the oldest age group (above 55), followed by the middle-aged group then the youngest (see Figure 2.24).

For both men and women we see continuous reduction in the negative valuations associated with diabetes over time (see Figure 2.25). In general, though the difference is not significant, women seem to place less negative value on diabetes than men.

The portion of the health index that is due to diabetes, however, continues to decrease over time (see Figure 2.26), and this is true for both men and women. The decrease in men is slightly larger over the sample period. To explain this divergence in observed trends between individual valuation of diabetes and greater negative impact of diabetes on the health index, we show in Figure 2.27. Here we see that while individual valuation of diabetes becomes less negative over time, diabetes contributes to individuals' greater negative health index due to its increasing prevalence in the population. These observations hold true across male and female populations (see Figure 2.28).

Near the sample period that we are examining, there were significant advancements in treatment knowledge for diabetes. Perhaps the one of most well-known among them all was the United Kingdom Prospective Diabetes Study (UKPDS). Its results were published in 1998 and played a significant role in informing diabetes treatment subsequently (Home 2008;

Genuth 2008). One of its major findings was that management of blood glucose levels should be a treatment target for diabetes (UKPDS 33). On top of that, metformin came to the fore as the initial choice of drug for diabetes management (Home 2008). Such examples of advancement in diabetes treatment are the likely reason for the less negative individual valuation of diabetes over time that we observed in earlier figures.

At the same time, diabetes remains a significant negative factor affecting individuals' health valuations. No cures exist, and lifestyle changes are hard to maintain. For example, among those treated with diet changes alone in the UKPDS study, only three percent of that group achieved the desired glucose level three years after initial treatment (UKPDS 13). Consistent with this, health literature on obesity find that even with health interventions, attaining and maintaining a weight loss goal is difficult (Glenny et al. 1997). Even with increased awareness and screening, diabetes remains a challenging health condition to treat (Selph et al. 2015).

Lastly, looking at Table 2.8, R^2 values from the regression of the coefficients of diabetes on the principal components of economic indicators are relatively large for both female ($R^2 = .38$) and male ($R^2 = .42$). This is not surprising given how diabetes treatment is the largest health care spending category in the US as we noted earlier.

2.5.5 The Effect of Income

The United States has gone through major changes over the past few decades that have enormously affected the wage structure, the return to schooling and experience and the likelihood of unemployment. Clearly, income is a key factor affecting each and every aspect of our life. Interestingly it seems to also directly affect the valuation of our health in a significant way, even more so than some other measures of health, as can be seen from the results presented in Figure 2.5. One thing to note is that this is true for income category "$75,000 and over." Therefore, for the income section we focus on this category.

Note first from Figure 2.5 that income has a large effect relative to other variables on the health index, and more so toward the latter years of the sample. Moreover, it seem to have greater effect on men's health index than on that of women (see Figure 2.30). This is explained by the fact that men tend to have higher valuation of income than women (see Figure 2.31) and men tend to have, on average, higher income than women as can be seen in from Figure 2.34.

When we look into what sources explain the changes in health index due to income, we find that it has less to do with individuals' valuation change and more so with changes in composition (see Figure 2.33). Specifically, Figure 2.33 indicates that changes in income, holding the valuation of income constant, contribute substantially to people's view of their health.

2.5.6 Changes in the Health Index's Thresholds

So far we have discussed variation in the contribution's factor to changes in the overall health index. Nevertheless, individuals may have the same valuation of their specific health index over time, and yet report differently when asked about their SRH. One way to interpret this is that there are perceived standards as to what level of the health index constitutes excellent health, very good health, etc. The threshold of the ordered-profit model are these standards, and they may change over time. It is therefore worthwhile analyzing how changes in the threshold affect individuals' reporting of the SRH, holding the level of their health index I_{it}^* constant. Furthermore, it would be worthwhile decomposing the overall reporting of the SRH into the part that stems from changes in the individual coefficient vectors β_t, and the part the stems from change in the thresholds c_t. We provide the analysis of the latter below. In this subsection we merely analyze changes in c_t over time, and relate these to changes in the set of economic indicators.

Figures 2.35 – 2.37 provide the results for the changes in c_t. In Figure 2.35 we provide the results for the whole population, while in Figures 2.36 and 2.37 we provide the results for the female and male populations, respectively. Note first that for men there is a very slight downward trend in all thresholds, while for women this is not as pronounced. This means that the same value of the health index I^*_{it} would lead individuals to report better SRH. Nevertheless, the thresholds' estimates for men are a lot more variable than those for women. As Table 2.10 and Table 2.11 indicate, men are more sensitive to changes in the economic environment than women. Specifically, the R^2's from the regressions of the four thresholds on the set of economic indicators for women are .39, .32, .32, and .30, respectively, while the corresponding R^2's for men are .48, .40,.37, and .35. A common feature to both men and women is that the lower thresholds are more sensitive to changes in the economic environment than the higher thresholds. That is, the economic environment seems to change individual's perception when evaluating their health, when their health is not particularly good. In contrast, the reporting of excellent health is less sensitive to changes in the economic environment, especially for women.

2.5.7 The Distribution of the Health Index

In Figures 2.38 through 2.40 we report density estimates for the deterministic part of the health index, i.e., $x'_{it}\beta_t$ for few selected years; the density of the error term is, by definition the same every period, that is $\varepsilon_{it} \sim N(0,1)$. The figures indicate that the entire distribution shifts from one year to the next over the sample period. However, the shifts in the distribution is much more pronounced for the men than for the women as can be seen from a comparison of Figure 2.40 with Figure 2.39

2.5.8 Implication for Reporting of SRH

Here we discuss several implications of the model's results. In particular we examine a few alternative counterfactuals. This is done in order to identify what has led to the vast changes in reporting of SRH in the NHIS. We first present the model's predictions, largely in order to establish that it does an incredibly good job in replicating what is observed in the data. Figures 2.41 through 2.43 present the model's predictions for the five categories of the SRH variable. Comparing Figure 2.41 with Figure 2.2a shows that for the population as a whole the model does a very good job in predicting the exact distribution across the five SRH's categories. This is also the case when we compare the subsamples of females (see Figure 2.42 in comparison with Figure 2.2e) and males (see Figure 2.43 in comparison with Figure 2.2f).

Having established the usefulness of the model, we now consider three alternative counterfactual exercises. In the first exercise we ask: What would the reporting of SRH be if the observed variables had remained fixed. To do that we fix the vector x at the average level over the years and then use the year-specific β_t and c_t to compute the model predictions of the probabilities reporting each of the five SRH categories. We refer to this experiment as the *constant x's experiment*. In the second experiment we maintain the assumption of the first experiment and in addition we fix the thresholds c_t at the average levels over the sample year. We refer to this experiment as the *constant x's fixed-standard experiment*. Finally, we fixed all the β_t at their average over the years, but let the c_t and the x be year-specific. In this experiment we examine what would happen if the valuation of the specific factors had remained constant, while the standards for determination (i.e., c_t) of SRH and the x's had not. We refer to this experiment as the *constant β's experiment*.

2.5.9 Constant x's Experiment

The results of this experiment are reported in Figures 2.44 through 2.46. Under constant x's, all that determines the probabilities of being in a particular SRH category are the parameters of the model, namely the β_t's and c_t's. As we documented above, most of the changes in the health index come from changes in the parameters. Hence, the patterns of changes in the fraction in each category are similar to those in Figures 2.41 through 2.43. Nevertheless, a few major differences are apparent. Particularly, the fraction of individuals in the poor category has been reduced significantly for both men and women, while the fraction of individuals in the fair category has remained relatively constant. At the higher end of the health distribution, a lot less individuals are in the excellent category, while more individuals are predicted to be in the very good category, and even more so for the good category.

Overall, we see that fixing the distribution of the x's gives very different results than those presented in Figures 2.41 through 2.43. Particularly, the lack of variation in the observed variables causes the distribution of reported SRH to be more condensed around "conservative" values. That is, the distribution of SRH are more heavily centered around the average and above-average values, largely the "Good" and "Very Good" category. The reason is that while the changes in the health index are dominated by changes in the valuation parameters (i.e., β_t's and c_t's) changes in the observed x's do have significant effects on the actual reporting of SRH. Consequently, if one is to incorporate the SRH variable in a model, one certainly needs to account for the changes in the observed variables over time, especially those such as income, education, and BMI, that are determined endogenously.

2.5.10 Constant x's, Fixed-Standards Experiment

The results of this experiment are reported in Figures 2.47 through 2.49. Note that if in addition to fixing the observed variables we also fixed the threshold, the predictions of the model change dramatically. First, the parameter vectors β_t and c_t are jointly determined. That is, one's valuation of the various factors affecting the latent health index are not completely independent from setting the standards for evaluation of the current health status. That is, reduced valuation of a particular factor is taken into account when the individual reports his/her value for the SRH variable. This is especially important in the reporting for men as can be seen from Figure 2.49 in comparison with Figure 2.43. Failing to account for this simultaneous process of the individuals' evaluation of their health status generates huge year to year variation in their self-reported health status, especially in the excellent health category, where small changes in the health index induce large changes in reporting excellent as their current health status. In other words, whatever the factors are that affect the valuation parameters β_t's are also affecting the standards parameters c_t. Hence, any model that would try to address the question about how individuals' change their valuation of the factors affecting the health index would also need to be able to explain the process by which the individuals update their standards for evaluation of the health index.

2.5.11 Constant β's Experiment

This experiment is in a way the opposite of the previous experiment, in that here the β_t's are fixed, while the x's and c_t's are allowed to vary across years. The results of this experiment are reported in Figures 2.50 through 2.52. While the changes induced by the experiment significantly alter the results relative to the predictions of the model in Figures 2.41 through 2.43, they are also quite different from those presented in Figures 2.47 through 2.49. This result again highlights the need to model the simultaneous changes in all parameters. Even though most of the changes over the sample period were indeed in the β_t's, failing to account for changes in the β_t's and c_t's leads to vastly erroneous conclusions regarding changes in SRH the variable.

2.5.12 Constant β's, Fixed-Standards Experiment

Lastly, we present the experiment where we hold β_t's and c_t's fixed but allow x's to vary across years. This would be hypothetical result if individual valuations and threshold did not change across the years and the only changes are from variations in the observed variables over time. The results are shown in Figures 2.53 through 2.55.

What's very clear from comparing these figures to the observed reporting of SRH in Figures 2.2a, 2.2e, and 2.2f is the lack of variation in the fraction of different categories of SRH. In the counterfactual exercise, the fractions stay more or less stable over the sample period, whereas what we observe in the real data are significant shifts in the distribution of SRH over time. The results highlight that what accounts for the significant shifts in the SRH is not due to changes in the composition of factors that affect the health index alone, but also due to changes in individual valuations of these factor as well as the thresholds for different SRH categories.

2.6 Summary and Conclusions

A question of the form: "*Would you say your health in general is excellent, very good, good, fair, or poor*?" has become a standard question in all surveys that address some aspect of health, especially in the United States. While the question is remarkably simple in content, little do we know about how respondents interpret this question. Changes over time in the distribution of the responses raise doubts that an answer to this question is a simple objective evaluation of one's true health status. Indeed, the literature has been struggling in interpreting the responses. Yet, the above self-reported health measure is used routinely in the literature, commonly assumed to be an objective assessment of one's health status. In this paper we make an attempt to answer the question: What does self-reported health really measure? In particular we estimate a model that is consistent with the type of question asked in surveys and analyze how changes in economic variable affect the parameters associated with this model. Specifically, we estimate the well-known ordered probit model in which it is assumed that there is a latent continuous health index that individuals form. The answer to the categorical question posed above is then interpreted to be a conversion of this latent index into a categorical answer. Using 22 years of data from the U.S. National Health Interview Survey (NHIS), we examine the time trend in self-reported health as it relates to changes in various factors determining the latent health index.

We devote our attention to the examining the distinct changes in the health index for the female and male populations. We document what has already been documented elsewhere in the literature, that self-reported health measure has changed dramatically over time. Moreover, it has changed quite differently for men and women. The estimated model yields parameter estimates that are very different, both across time and across gender. We find that the changes in the parameter estimates are closely related to changes in the underlying economic environment. Indeed when we run a simple OLS regression of the parameter estimates on a set of economic indicators we find, for most cases, very high R^2's. Nevertheless, we find that men are a lot more sensitive to the economic environment than women, in that the R^2's from the regressions for males are considerably higher than those for their female counterparts. Moreover, the various factors seem to affect the two groups very differently. For example the changes in the effect of BMI take very different patterns for men and women, both in terms of its affect on the overall index, and in terms of the fraction of the overall variance of the index that it explains.

The results clearly indicate that the self-reported health measure is all but a subjective measure of one's health status. In fact, the changes in the model's coefficients over time and their close link with the economic variable suggest that self-reported health measure stems from a complex process of evaluation performed by the individuals reporting it. Undoubtedly more work is needed in understanding why and how individual form their evaluation of their own health status. What we intended to provide in this paper is a framework with which one can uncovered part of this complicated structure.

Table 2.1: Raw Statistics from the National Health Interview Survey

	1997	2000	2003	2006	2009	2012	2015	2018
Self-Reported	3.830	3.824	3.787	3.764	3.740	3.739	3.743	3.745
Health	(1.060)	(1.063)	(1.068)	(1.063)	(1.073)	(1.069)	(1.066)	(1.054)
Age	44.42	44.82	45.13	45.49	45.92	46.41	46.88	47.30
	(17.41)	(17.40)	(17.44)	(17.50)	(17.67)	(17.86)	(18.03)	(18.21)
Some College	0.508	0.521	0.545	0.547	0.580	0.602	0.629	0.645
and Above	(0.500)	(0.500)	(0.498)	(0.498)	(0.494)	(0.489)	(0.483)	(0.478)
BMI	26.19	26.56	26.93	27.26	27.60	27.65	27.87	28.07
	(5.269)	(5.488)	(5.724)	(5.937)	(6.133)	(6.070)	(6.384)	(6.414)
Female	0.514	0.512	0.511	0.507	0.510	0.508	0.508	0.509
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
Married	0.588	0.582	0.579	0.567	0.544	0.529	0.530	0.522
	(0.492)	(0.493)	(0.494)	(0.496)	(0.498)	(0.499)	(0.499)	(0.500)
ADL	0.0123	0.0152	0.0180	0.0169	0.0185	0.0216	0.0232	0.0252
	(0.110)	(0.122)	(0.133)	(0.129)	(0.135)	(0.146)	(0.151)	(0.157)
IADL	0.0327	0.0321	0.0365	0.0355	0.0407	0.0407	0.0436	0.0471
	(0.178)	(0.176)	(0.188)	(0.185)	(0.198)	(0.198)	(0.204)	(0.212)
Diabetes	0.0519	0.0595	0.0658	0.0779	0.0906	0.0913	0.0976	0.103
	(0.222)	(0.236)	(0.248)	(0.268)	(0.287)	(0.288)	(0.297)	(0.304)
Emphysema	0.0164	0.0156	0.0148	0.0186	0.0210	0.0175	0.0145	0.0151
	(0.127)	(0.124)	(0.121)	(0.135)	(0.144)	(0.131)	(0.120)	(0.122)
Hypertension	0.226	0.224	0.249	0.268	0.285	0.294	0.306	0.307
	(0.418)	(0.417)	(0.432)	(0.443)	(0.451)	(0.456)	(0.461)	(0.461)
Ulcer	0.0909	0.0729	0.0674	0.0653	0.0769	0.0648	0.0603	0.0582
	(0.287)	(0.260)	(0.251)	(0.247)	(0.266)	(0.246)	(0.238)	(0.234)
Arthritis	0.324	0.305	0.373	0.354	0.388	0.363	0.392	0.396
	(0.468)	(0.460)	(0.484)	(0.478)	(0.487)	(0.481)	(0.488)	(0.489)

Means and standard deviations are reported here. Computed statistics are weighted by the sample weights

Table 2.2: Mean Self-Reported Health, by Year and Age

Year	All	Less than equal 30	Above 30 to 55	Above 55
1997	3.817	4.141	3.922	3.313
1998	3.824	4.167	3.924	3.317
1999	3.832	4.182	3.927	3.338
2000	3.811	4.168	3.908	3.309
2001	3.799	4.167	3.885	3.308
2002	3.777	4.142	3.860	3.300
2003	3.770	4.120	3.865	3.296
2004	3.757	4.139	3.841	3.288
2005	3.755	4.115	3.842	3.305
2006	3.747	4.111	3.821	3.321
2007	3.730	4.119	3.798	3.302
2008	3.735	4.119	3.800	3.323
2009	3.726	4.141	3.769	3.335
2010	3.727	4.071	3.786	3.371
2011	3.724	4.100	3.775	3.368
2012	3.724	4.099	3.770	3.390
2013	3.720	4.126	3.766	3.372
2014	3.747	4.131	3.808	3.402
2015	3.727	4.109	3.785	3.395
2016	3.725	4.113	3.769	3.420
2017	3.729	4.108	3.814	3.387
2018	3.725	4.106	3.811	3.384
Mean	3.753	4.136	3.828	3.349
St. Dev.	1.070	0.902	1.033	1.115
R^2	0.932	0.636	0.903	0.504

Computed statistics are weighted by the sample weights. The R^2 is from a regression of the mean self-reported health on a set of economic indicators: Inflation rate, unemployment rate, and real per capita consumption.

Table 2.3: Fraction of Self-Reported Health Categories, by Year and Age

	Age			
Year	All	$\leqslant 30$	>30 & $\leqslant 55$	> 55
Panel A — Poor Health				
1997	0.0275	0.00426	0.0207	0.0620
2001	0.0308	0.00518	0.0226	0.0692
2005	0.0317	0.00473	0.0253	0.0654
2009	0.0330	0.00744	0.0270	0.0623
2013	0.0331	0.00492	0.0277	0.0603
2018	0.0286	0.00553	0.0216	0.0513
Panel B — Fair Health				
1997	0.0873	0.0372	0.0683	0.170
2001	0.0878	0.0340	0.0709	0.167
2005	0.0914	0.0346	0.0751	0.167
2009	0.0968	0.0368	0.0861	0.161
2013	0.101	0.0396	0.0905	0.159
2018	0.0983	0.0450	0.0771	0.157
Panel C — Good Health				
1997	0.247	0.198	0.230	0.326
2001	0.245	0.180	0.234	0.323
2005	0.261	0.214	0.244	0.330
2009	0.266	0.200	0.263	0.324
2013	0.262	0.203	0.261	0.306
2018	0.267	0.200	0.263	0.315
Panel D — Very Good Health				
1997	0.317	0.335	0.329	0.277
2001	0.325	0.350	0.343	0.267
2005	0.321	0.335	0.343	0.273
2009	0.319	0.320	0.339	0.286
2013	0.320	0.329	0.330	0.300
2018	0.331	0.337	0.346	0.309
Panel E — Excellent Health				
1997	0.321	0.426	0.352	0.165
2001	0.312	0.431	0.329	0.173
2005	0.294	0.412	0.312	0.165
2009	0.285	0.436	0.285	0.167
2013	0.284	0.423	0.291	0.176
2018	0.275	0.413	0.292	0.167

Table 2.4: Mean Self-Reported Health, by Gender, Year and Age

		Age		
Year	All	$\leqslant 30$	>30 & $\leqslant 55$	> 55
Panel A — Females				
1997	3.763	4.072	3.890	3.282
2000	3.765	4.114	3.874	3.290
2003	3.721	4.053	3.829	3.283
2006	3.711	4.071	3.794	3.308
2009	3.703	4.139	3.739	3.336
2012	3.684	4.056	3.739	3.364
2015	3.704	4.061	3.761	3.413
2018	3.705	4.093	3.802	3.371
Mean	3.715	4.082	3.796	3.340
St.Dev.	1.076	0.912	1.046	1.110
R^2	0.872	0.230	0.841	0.669
Panel B — Males				
1997	3.874	4.211	3.956	3.351
2000	3.862	4.223	3.942	3.333
2003	3.823	4.188	3.902	3.313
2006	3.786	4.150	3.850	3.337
2009	3.752	4.143	3.800	3.334
2012	3.767	4.142	3.802	3.419
2015	3.751	4.157	3.811	3.373
2018	3.747	4.119	3.819	3.398
Mean	3.794	4.170	3.862	3.360
St.Dev.	1.063	0.888	1.018	1.122
R^2	0.930	0.678	0.897	0.050

Table 2.5: Regression of Age Coefficients on Principal Components of Economic Indicators

	(1) All	(2) Male	(3) Female
Principal Component 1	0.000302***	0.000392***	0.000231**
	(6.43)	(7.24)	(3.13)
Principal Component 2	-0.0000153	-0.00000646	0.000000341
	(-0.18)	(-0.07)	(0.00)
Principal Component 3	-0.000293*	-0.000523**	-0.0000951
	(-2.41)	(-3.74)	(-0.50)
_cons	-0.00624***	-0.00799***	-0.00495***
	(-37.23)	(-41.39)	(-18.84)
N	22	22	22
R^2	0.724	0.787	0.359

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.6: Regression of BMI Coefficients on Principal Components of Economic Indicators

	(1) All BMI	(2) Male BMI	(3) Female BMI	(4) All BMI^2	(5) Male BMI^2	(6) Female BMI^2
Principal Component 1	-0.00210***	-0.00401***	-0.00134*	0.0000217*	0.0000520**	0.0000107
	(-4.24)	(-4.30)	(-2.36)	(2.79)	(3.40)	(1.21)
Principal Component 2	0.000608	0.000302	0.000595	-0.00000635	-0.00000153	-0.00000631
	(0.68)	(0.18)	(0.59)	(-0.45)	(-0.06)	(-0.40)
Principal Component 3	-0.00269	0.000165	-0.00399*	0.0000400	-0.00000271	0.0000571*
	(-2.10)	(0.07)	(-2.73)	(1.99)	(-0.07)	(2.50)
_cons	-0.0237***	0.000561	-0.0355***	-0.0000244	-0.000434***	0.000165***
	(-13.39)	(0.17)	(-17.58)	(-0.88)	(-7.95)	(5.22)
N	22	22	22	22	22	22
R^2	0.560	0.507	0.426	0.399	0.391	0.304

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7: Regression of Hypertension Coefficients on Principal Components of Economic Indicators

	(1) All	(2) Male	(3) Female
Principal Component 1	0.00170 (1.16)	0.00272 (1.16)	0.00118 (0.80)
Principal Component 2	-0.00455 (-1.72)	-0.00637 (-1.51)	-0.00243 (-0.92)
Principal Component 3	-0.00184 (-0.48)	-0.00363 (-0.60)	0.000106 (0.03)
_cons	-0.260*** (-49.47)	-0.282*** (-33.64)	-0.243*** (-46.30)
N	22	22	22
R^2	0.201	0.181	0.0767

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.8: Regression of Diabetes Coefficients on Principal Components of Economic Indicators

	(1) All	(2) Male	(3) Female
Principal Component 1	0.00755** (3.27)	0.00854** (3.01)	0.00696* (2.11)
Principal Component 2	0.00000433 (0.00)	0.00812 (1.59)	-0.00831 (-1.40)
Principal Component 3	-0.0134* (-2.24)	-0.00862 (-1.17)	-0.0162 (-1.90)
_cons	-0.501*** (-60.72)	-0.501*** (-49.46)	-0.491*** (-41.59)
N	22	22	22
R^2	0.466	0.419	0.357

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.9: Regression of Threshold Estimates on Principal Components of Economic Indicators

	(1) Threshold 1	(2) Threshold 2	(3) Threshold 3	(4) Threshold 4
Principal Component 1	-0.0295**	-0.0216*	-0.0189*	-0.0159*
	(-3.43)	(-2.68)	(-2.49)	(-2.12)
Principal Component 2	0.0148	0.0148	0.0138	0.0162
	(0.96)	(1.02)	(1.01)	(1.20)
Principal Component 3	-0.0500*	-0.0498*	-0.0466*	-0.0455*
	(-2.25)	(-2.40)	(-2.37)	(-2.34)
_cons	-3.878***	-2.767***	-1.597***	-0.559***
	(-123.61)	(-94.17)	(-57.42)	(-20.37)
N	22	22	22	22
R^2	0.497	0.437	0.415	0.388

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Threshold 1 is the cutoff between "Poor" and "Fair," Threshold 2 is between "Fair" and "Good," Threshold 3 is between "Good" and "Very Good", and Threshold 4 is between "Very Good" and "Excellent"

Table 2.10: Regression of Threshold Estimates on Principal Components of Economic Indicators (Female)

	(1) Threshold 1	(2) Threshold 2	(3) Threshold 3	(4) Threshold 4
Principal Component 1	-0.0206 (-2.07)	-0.0128 (-1.40)	-0.0116 (-1.27)	-0.00891 (-0.98)
Principal Component 2	0.0185 (1.04)	0.0166 (1.01)	0.0153 (0.93)	0.0155 (0.95)
Principal Component 3	-0.0646* (-2.51)	-0.0546* (-2.31)	-0.0579* (-2.44)	-0.0566* (-2.41)
_cons	-4.075*** (-112.20)	-2.933*** (-87.86)	-1.755*** (-52.35)	-0.702*** (-21.21)
N	22	22	22	22
R^2	0.393	0.315	0.319	0.300

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Threshold 1 is the cutoff between “Poor” and “Fair,” Threshold 2 is between “Fair” and “Good,” Threshold 3 is between “Good” and “Very Good”, and Threshold 4 is between “Very Good” and “Excellent”

Table 2.11: Regression of Threshold Estimates on Principal Components of Economic Indicators (Male)

	(1) Threshold 1	(2) Threshold 2	(3) Threshold 3	(4) Threshold 4
Principal Component 1	-0.0562***	-0.0479**	-0.0434**	-0.0401**
	(-4.05)	(-3.40)	(-3.21)	(-3.05)
Principal Component 2	0.00535	0.00755	0.00706	0.0114
	(0.21)	(0.30)	(0.29)	(0.48)
Principal Component 3	-0.0124	-0.0229	-0.0117	-0.0102
	(-0.35)	(-0.63)	(-0.34)	(-0.30)
_cons	-3.562***	-2.474***	-1.307***	-0.277***
	(-70.37)	(-48.16)	(-26.51)	(-5.78)
N	22	22	22	22
R^2	0.480	0.401	0.369	0.349

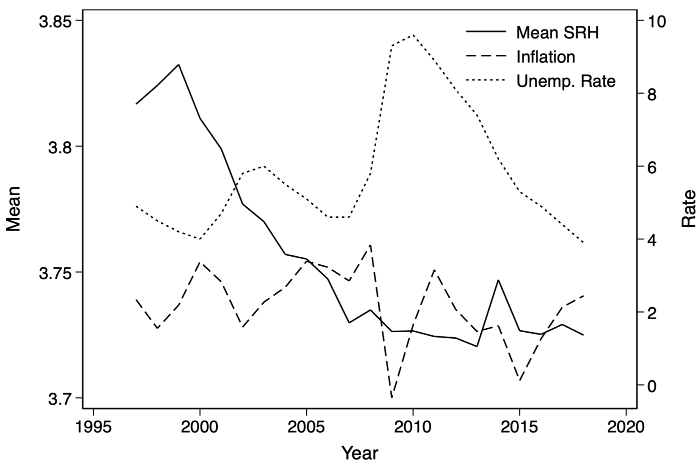
t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

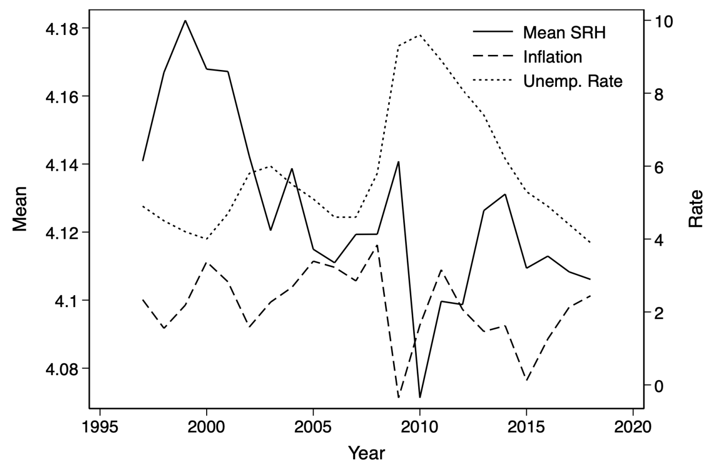
Threshold 1 is the cutoff between "Poor" and "Fair," Threshold 2 is between "Fair" and "Good," Threshold 3 is between "Good" and "Very Good", and Threshold 4 is between "Very Good" and "Excellent"

Figure 2.1: Mean of Self-Reported Health and Economic Indicators, by Age

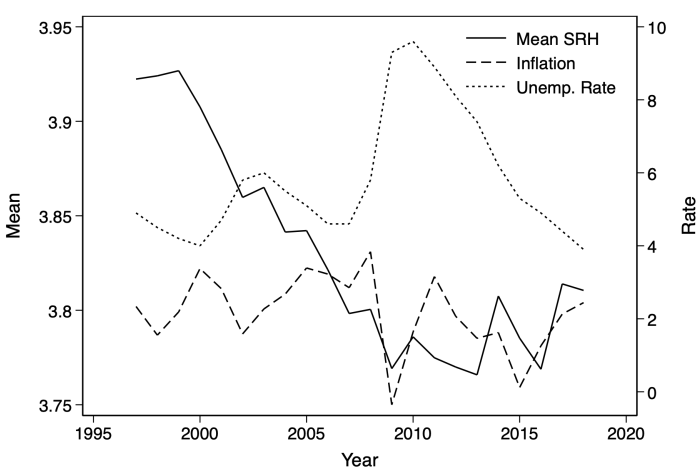
(a) Whole Population



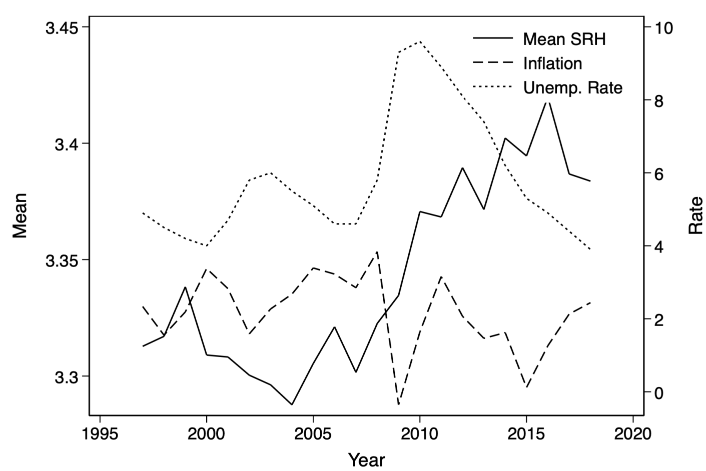
(b) Age⩽ 30



(c) 30 <Age⩽ 55



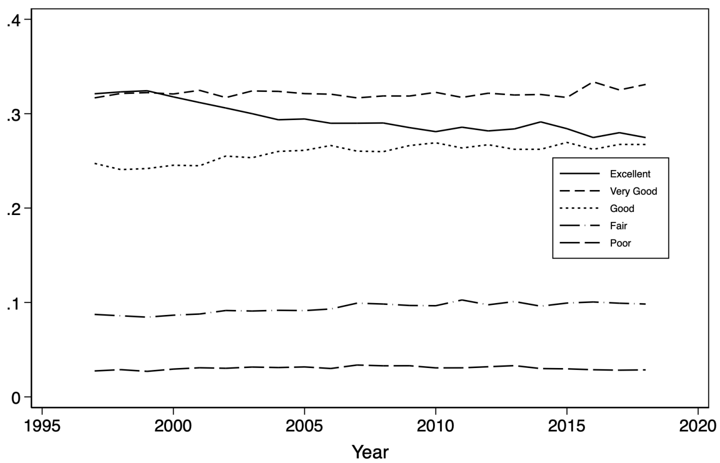
(d) Age> 55



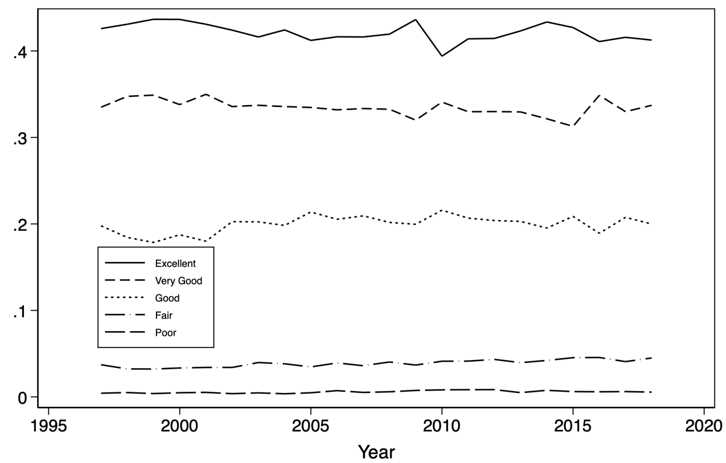
For all the figures, the left axis values correspond to the mean self-reported health and the right axis values correspond to the economic indicators

Figure 2.2: Fraction of Individuals in Health Categories, by Age and Gender

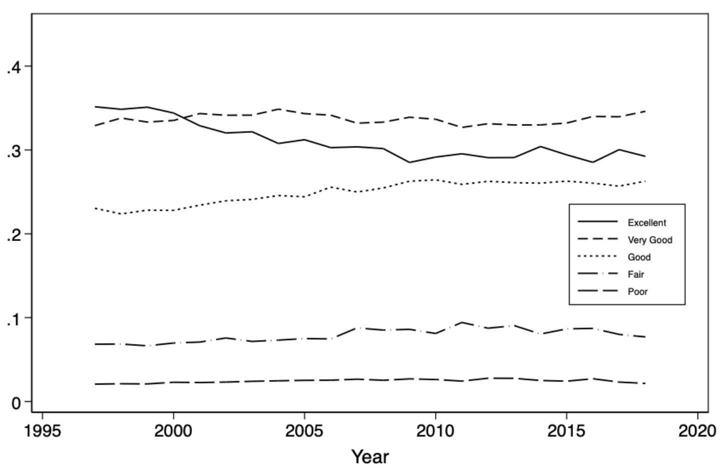
(a) Whole Population



(b) Age⩽ 30



(c) 30 <Age⩽ 55



(d) Age> 55

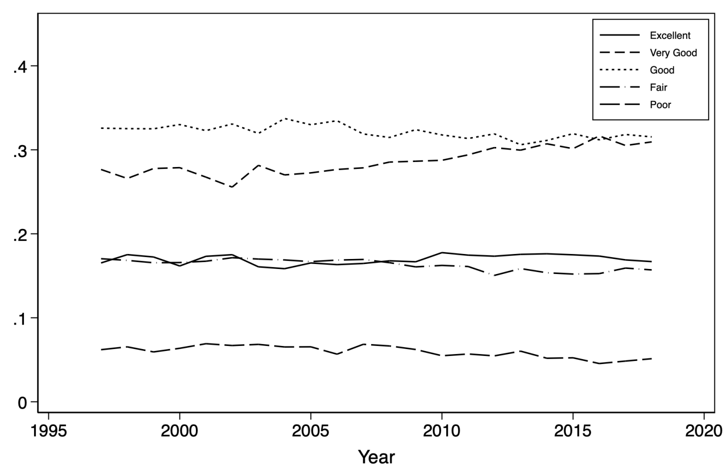
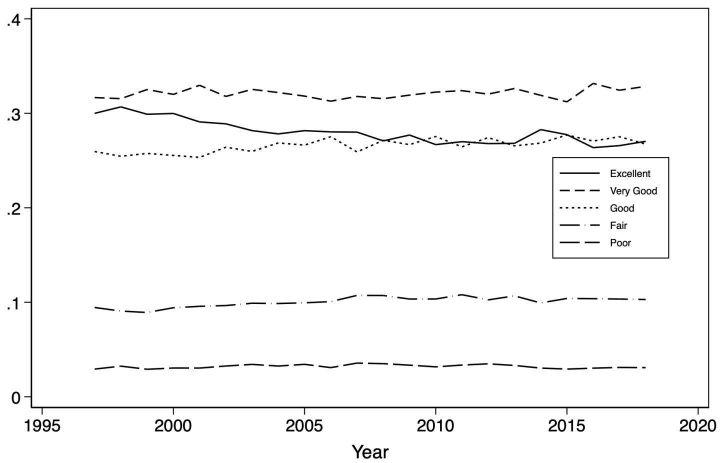


Figure 2.2: (Continued)

(e) Female



(f) Male

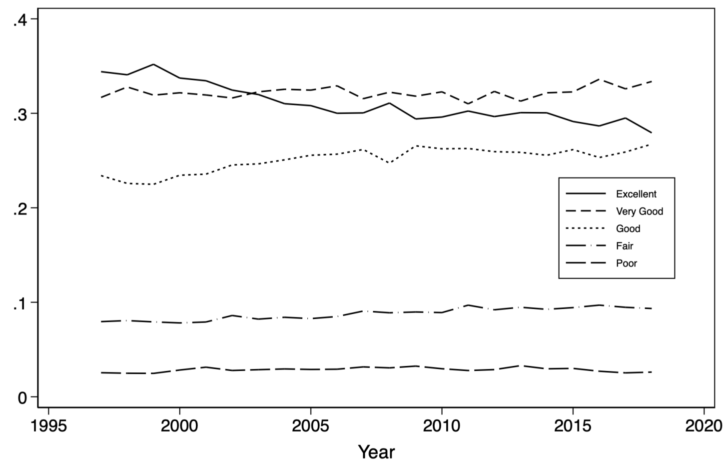
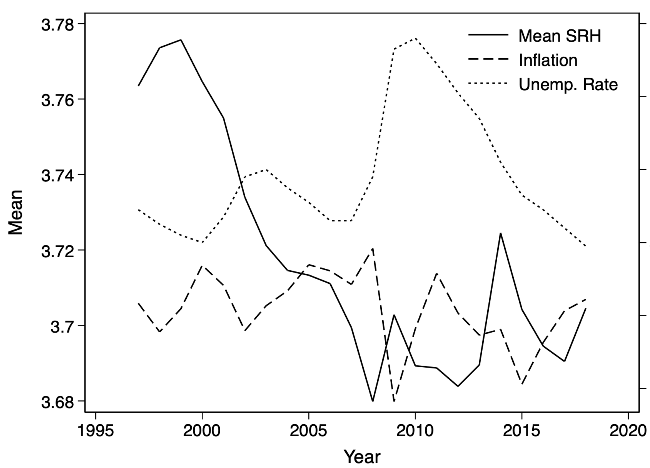
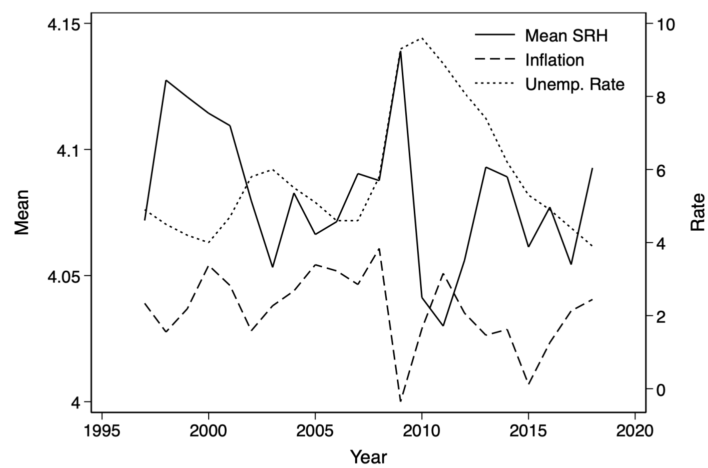


Figure 2.3: Mean of Self-Reported Health and Economic Indicators for Females

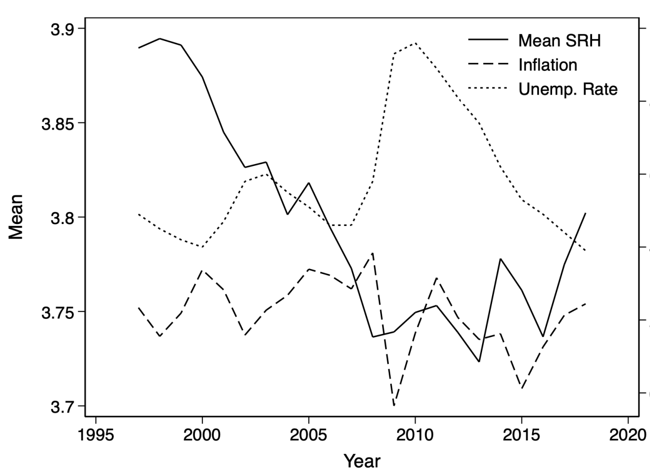
(a) Whole Population



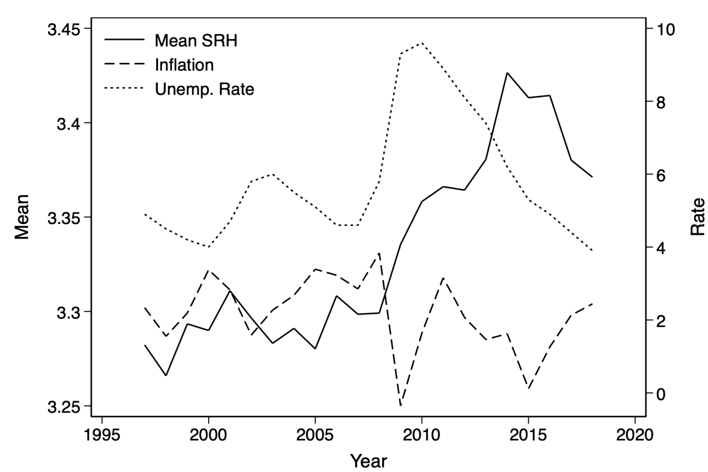
(b) Age⩽ 30



(c) 30 <Age⩽ 55



(d) Age> 55

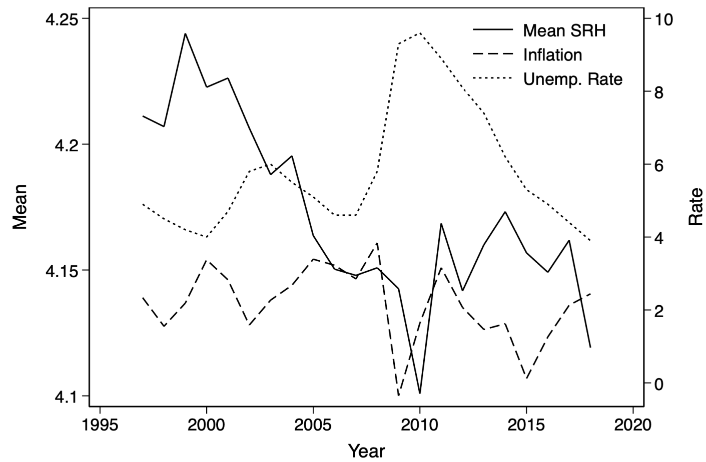


For all the figures, the left axis values correspond to the mean self-reported health and the right axis values correspond to the economic indicators

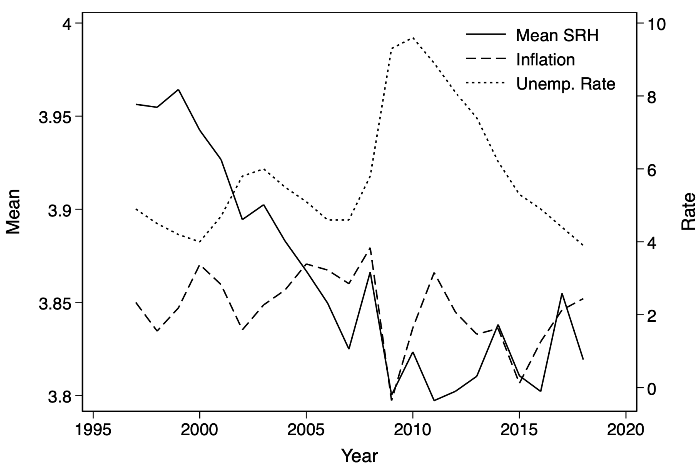
Figure 2.4: Mean of Self-Reported Health and Economic Indicators for Males

(a) Whole Population

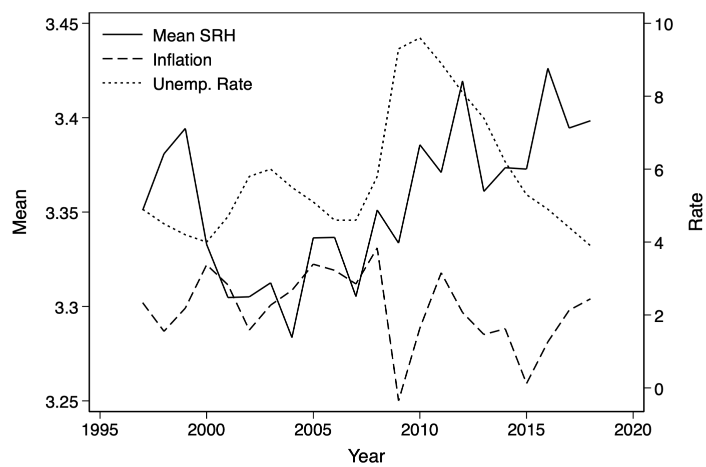
(b) Age⩽ 30



(c) 30 <Age⩽ 55



(d) Age> 55



For all the figures, the left axis values correspond to the mean self-reported health and the right axis values correspond to the economic indicators

Figure 2.5: Select Top Contributing Variables To Health Index

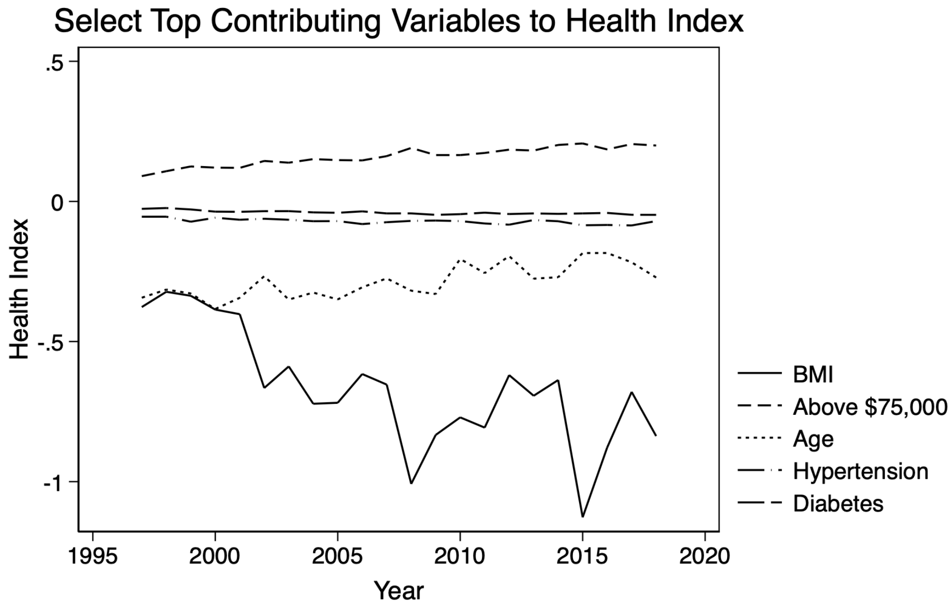


Figure 2.6: Kernel Density Plot of Age Over Time

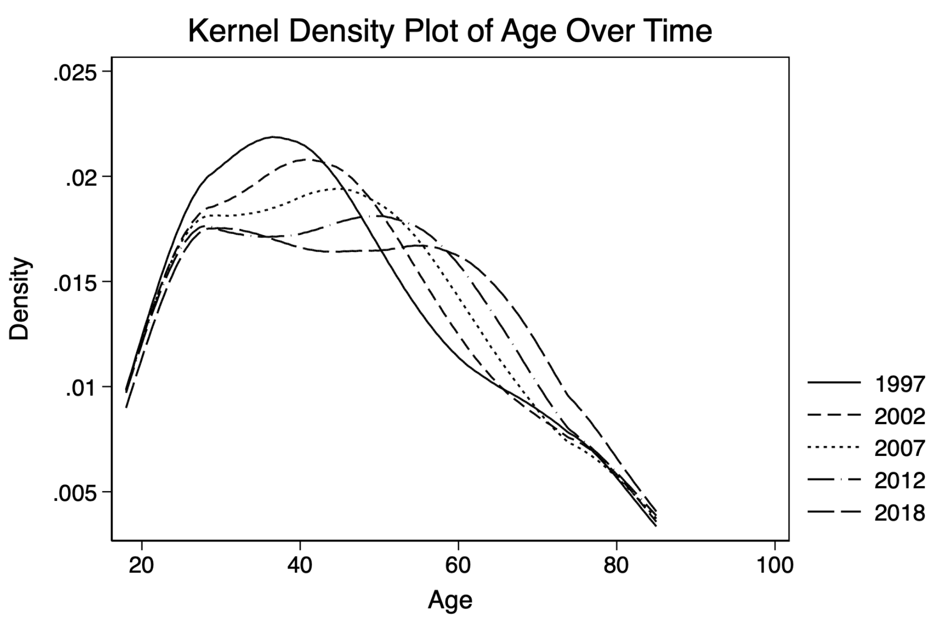


Figure 2.7: Estimated Coefficients of Age

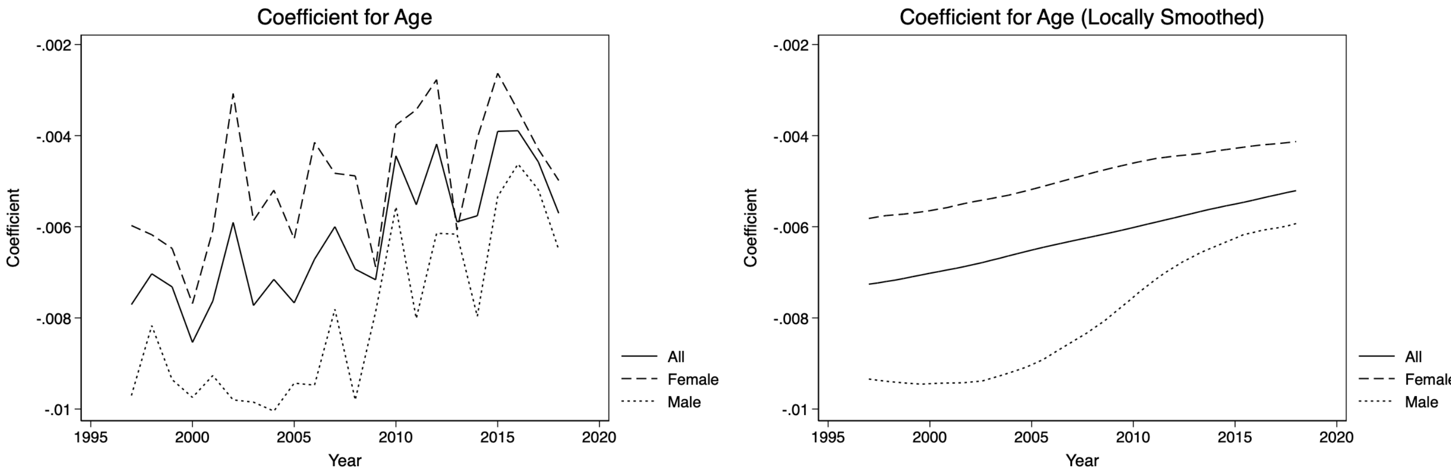


Figure 2.8: Age Portion of Health Index

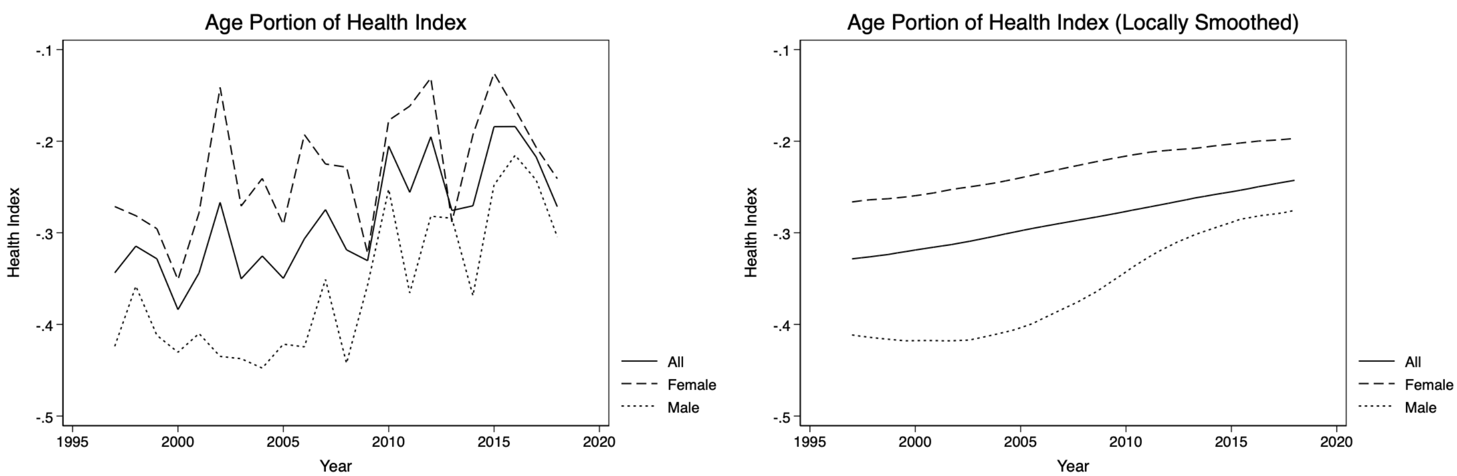


Figure 2.9: Changes in Health Index due to Age

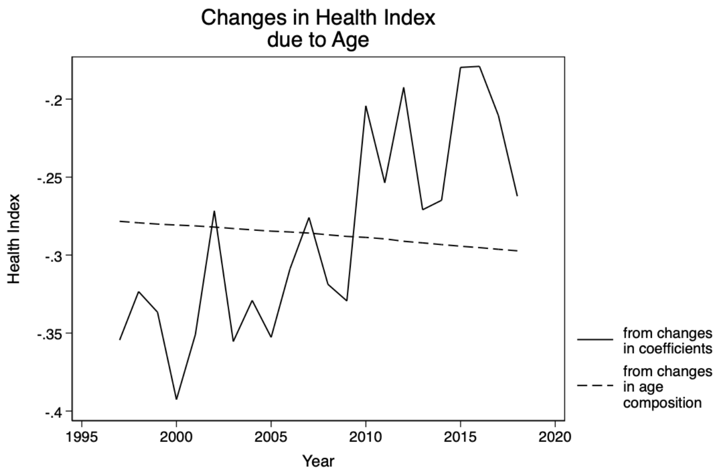


Figure 2.10: Changes in Health Index due to Age by Gender

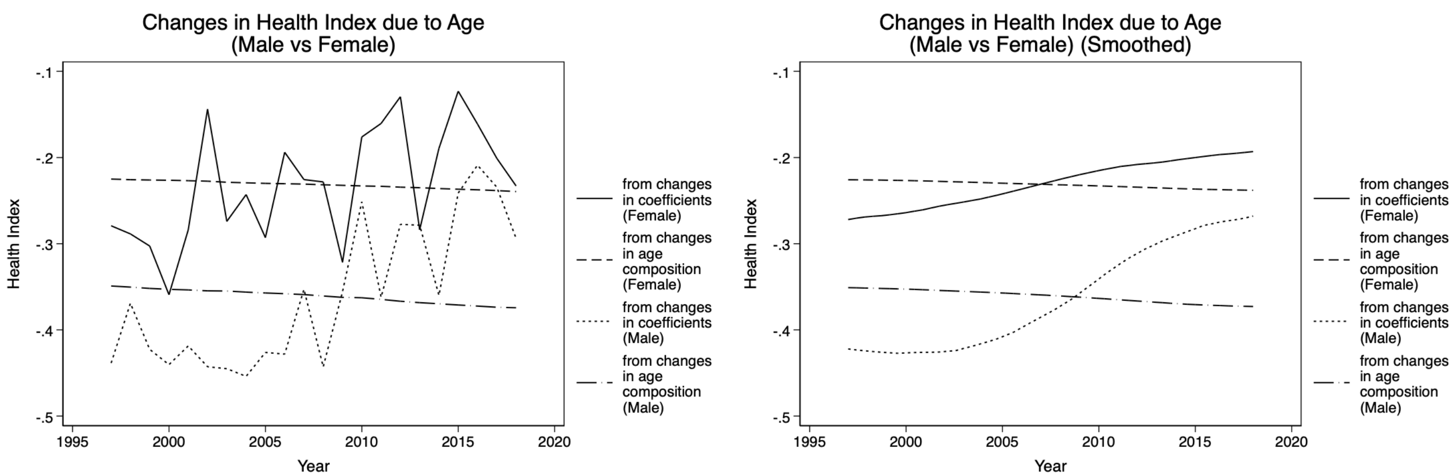
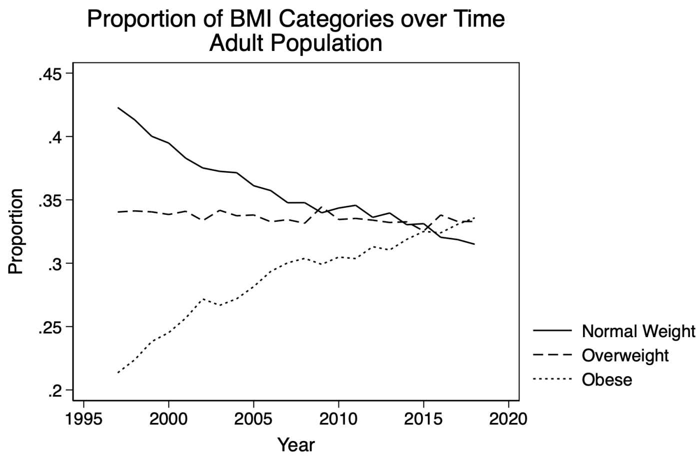
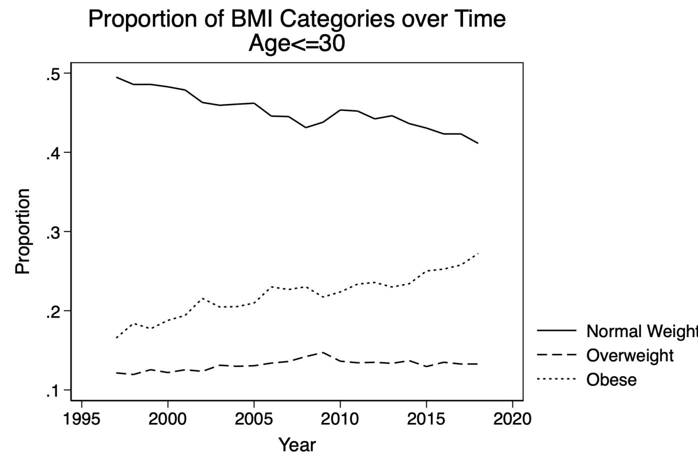


Figure 2.11: BMI Categories Over Time By Age Group

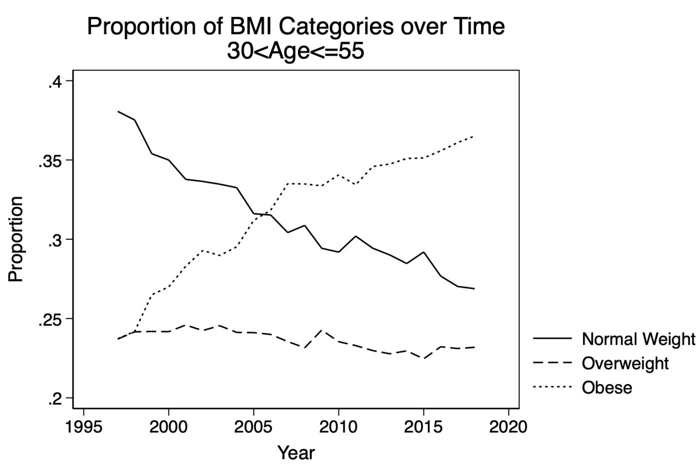
(a)



(b)



(c)



(d)

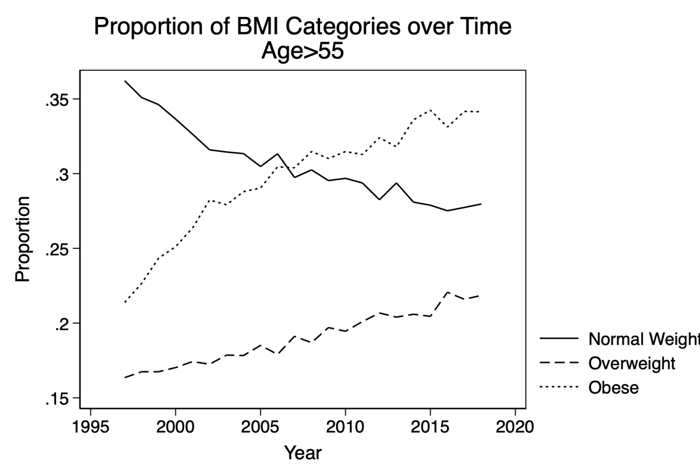


Figure 2.12: BMI Coefficient Estimates

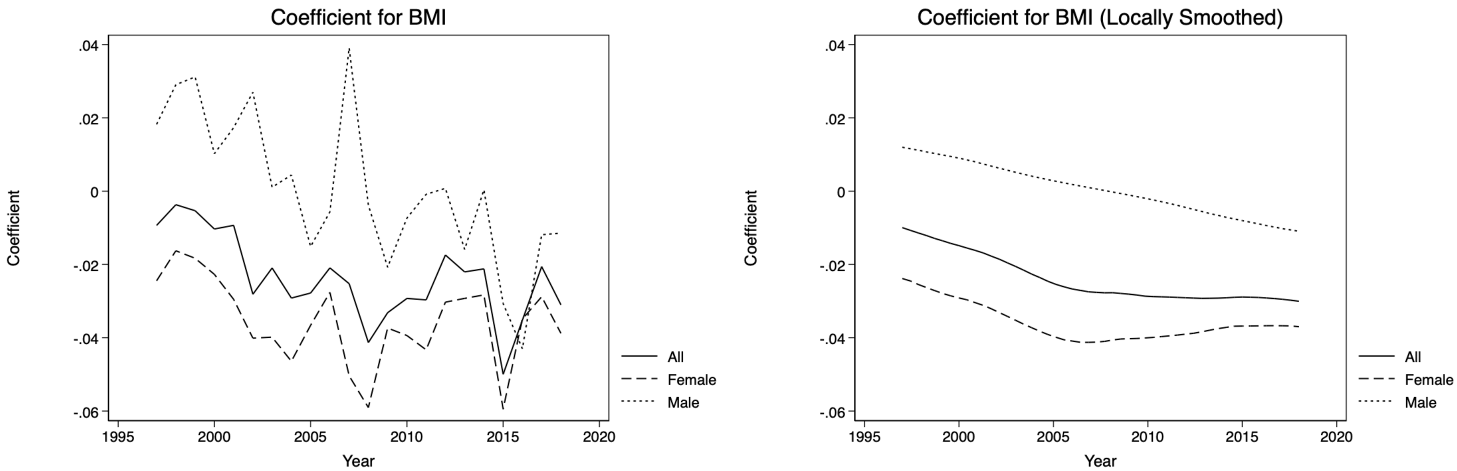


Figure 2.13: BMI^2 Coefficient Estimates

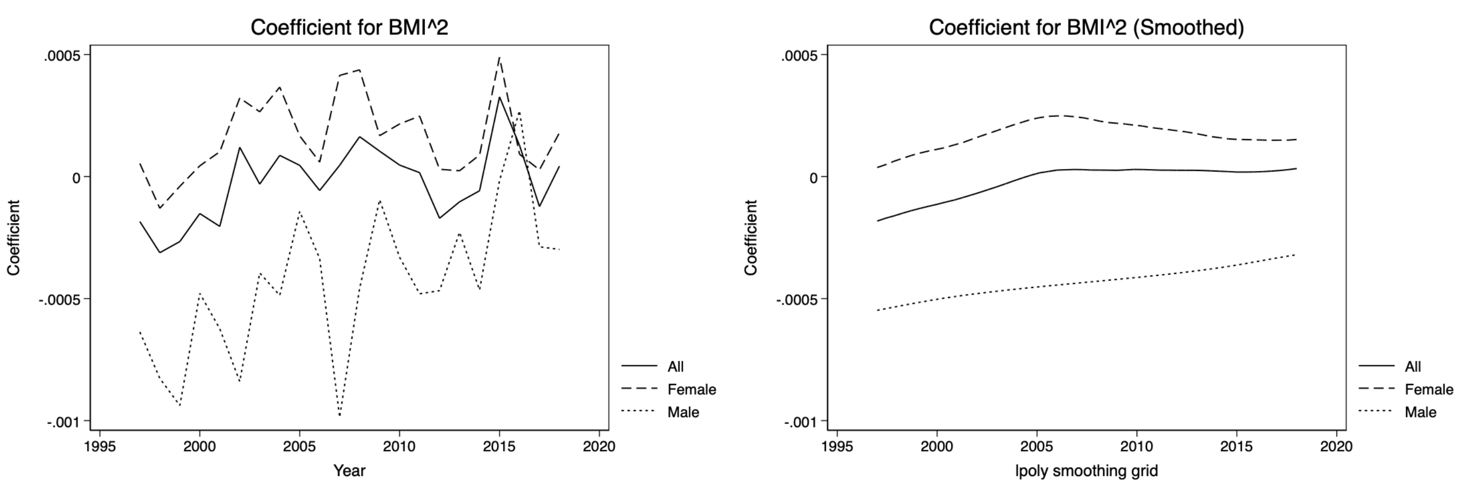


Figure 2.14: Marginal Effect Estimate of BMI

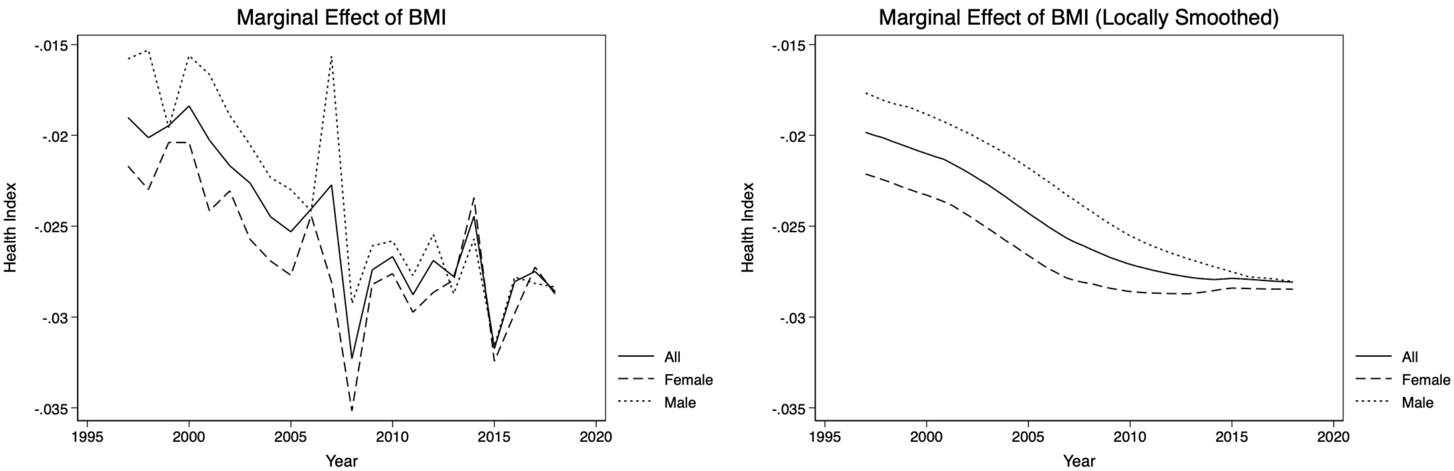


Figure 2.15: BMI Portion of Health Index

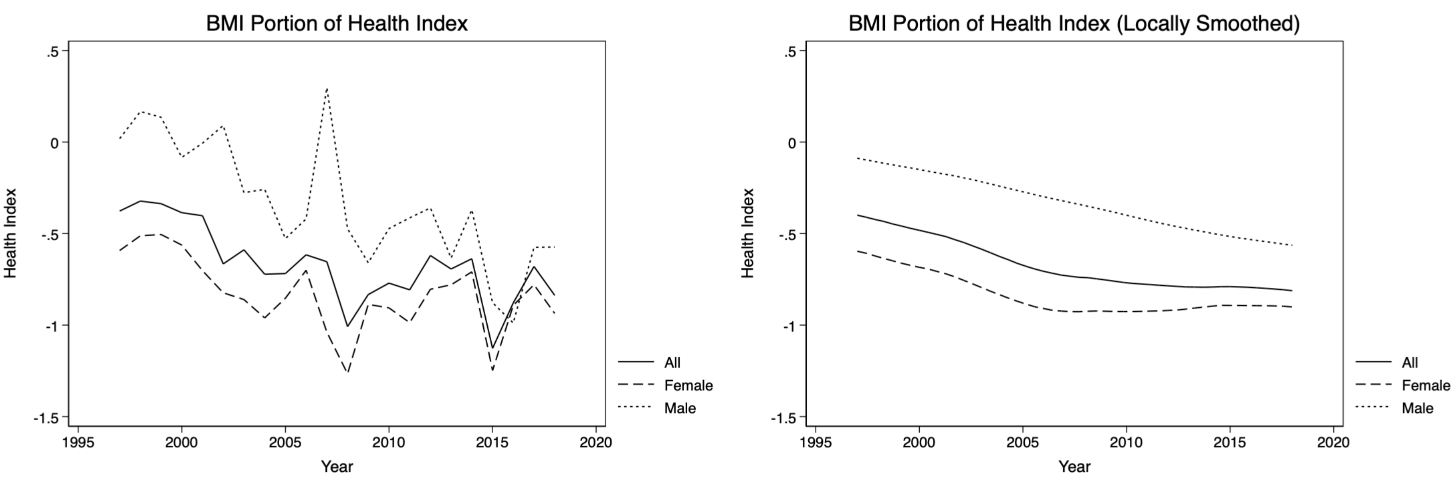


Figure 2.16: Changes in BMI Portion of HI due to Changes in BMI Coefficient and Composition

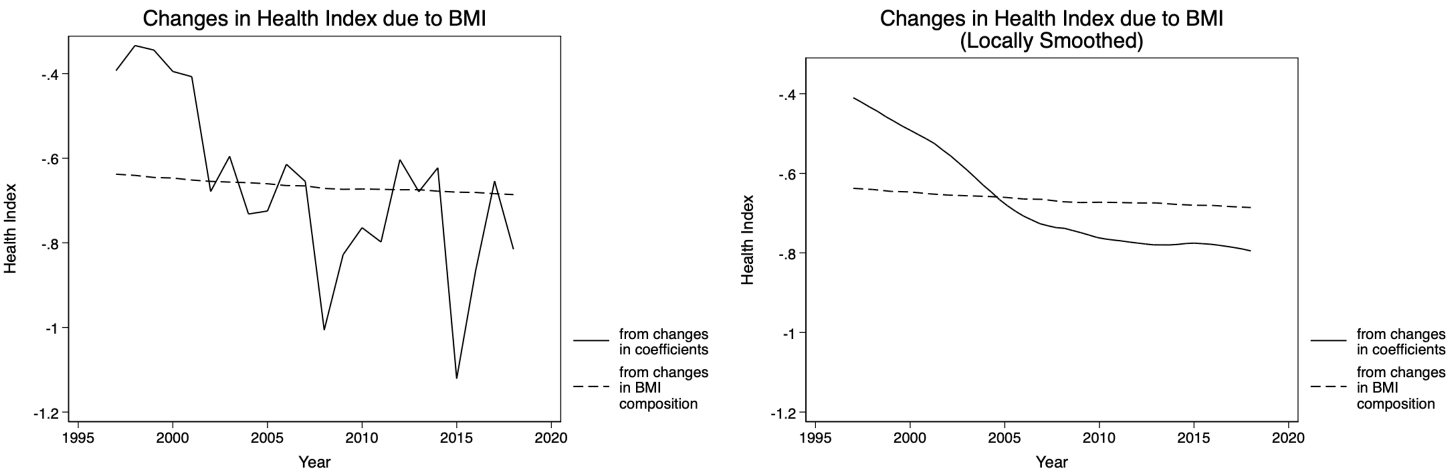


Figure 2.17: Prevalence of Hypertension

(a) (b)

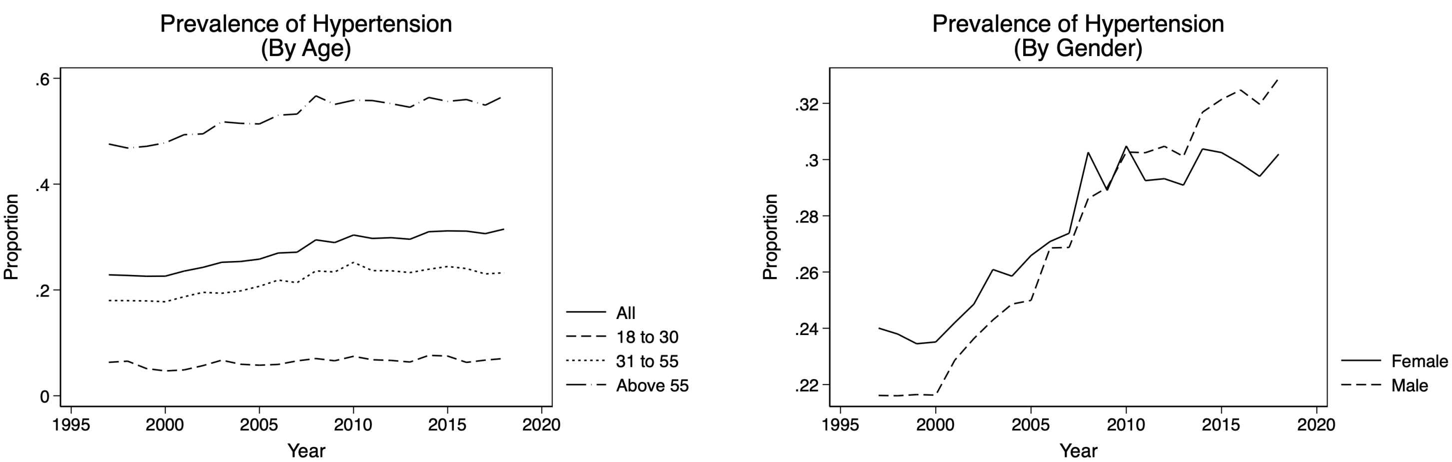


Figure 2.18: Prevalence of Blood Pressure Control

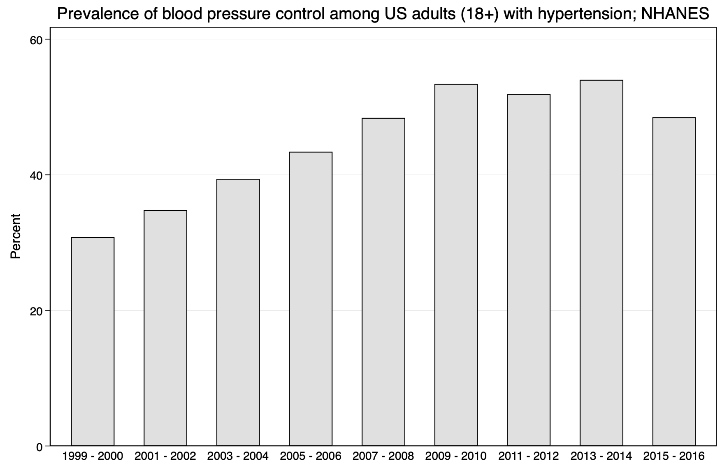


Figure 2.19: Estimated Coefficients of Hypertension

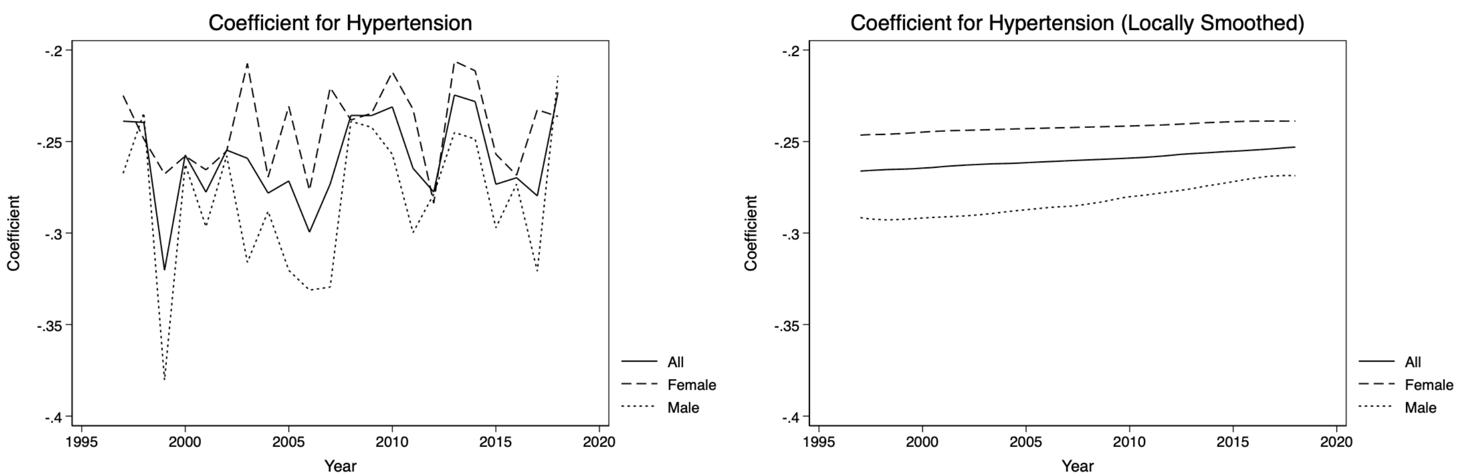


Figure 2.20: Hypertension Portion of Health Index

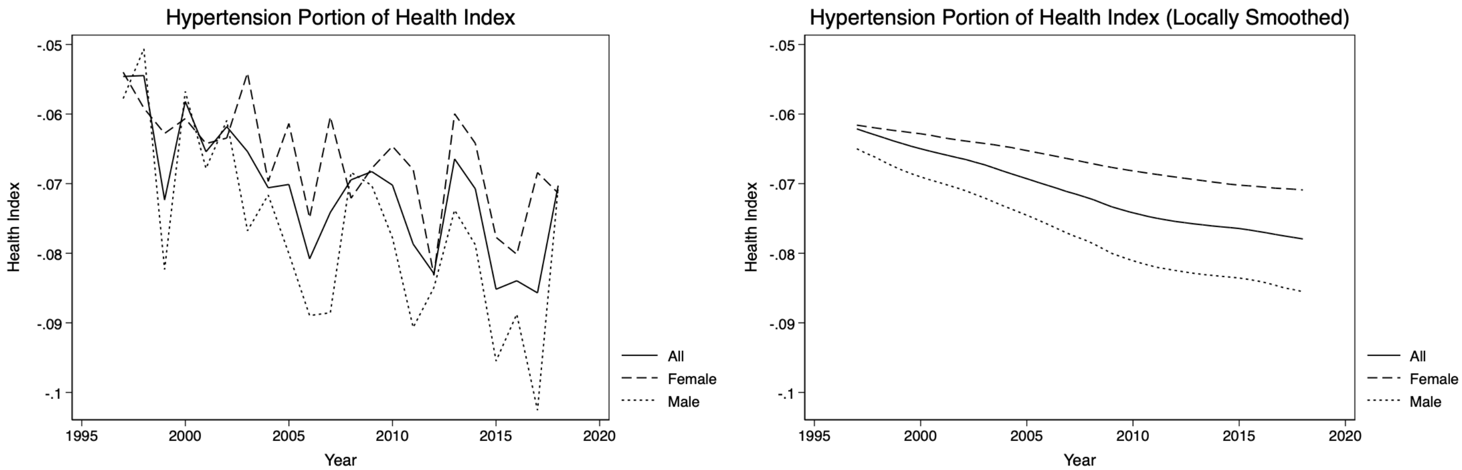


Figure 2.21: Changes in Hypertension Portion of HI due to Changes in Hypertension Coefficient and Composition

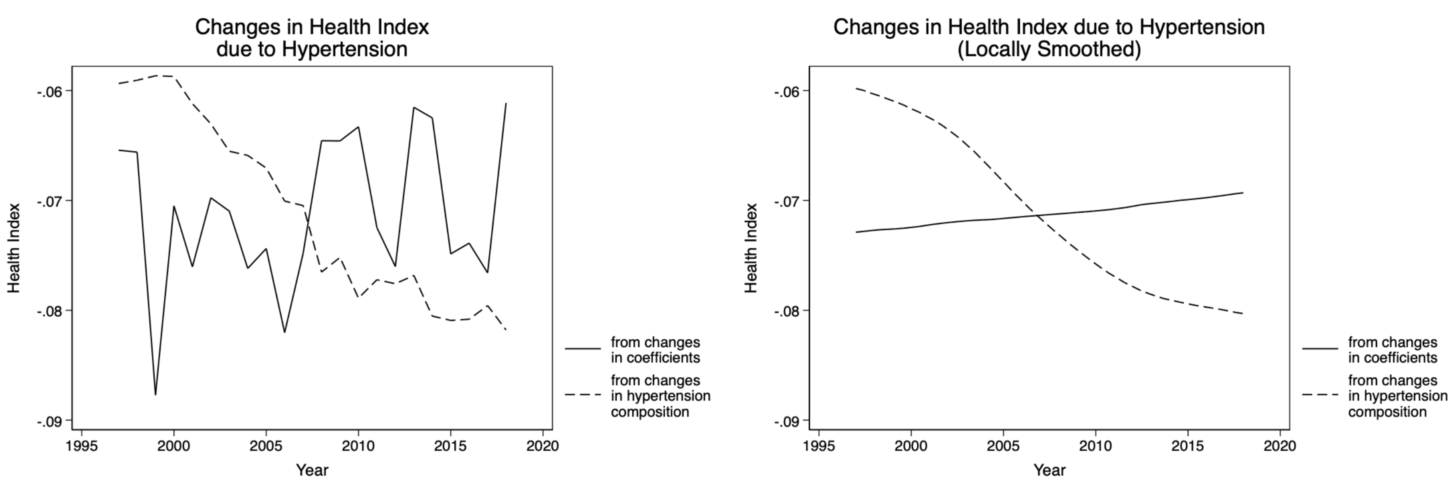


Figure 2.22: Changes in Hypertension Portion of HI due to Changes in Hypertension Coefficient and Composition (Male and Female)

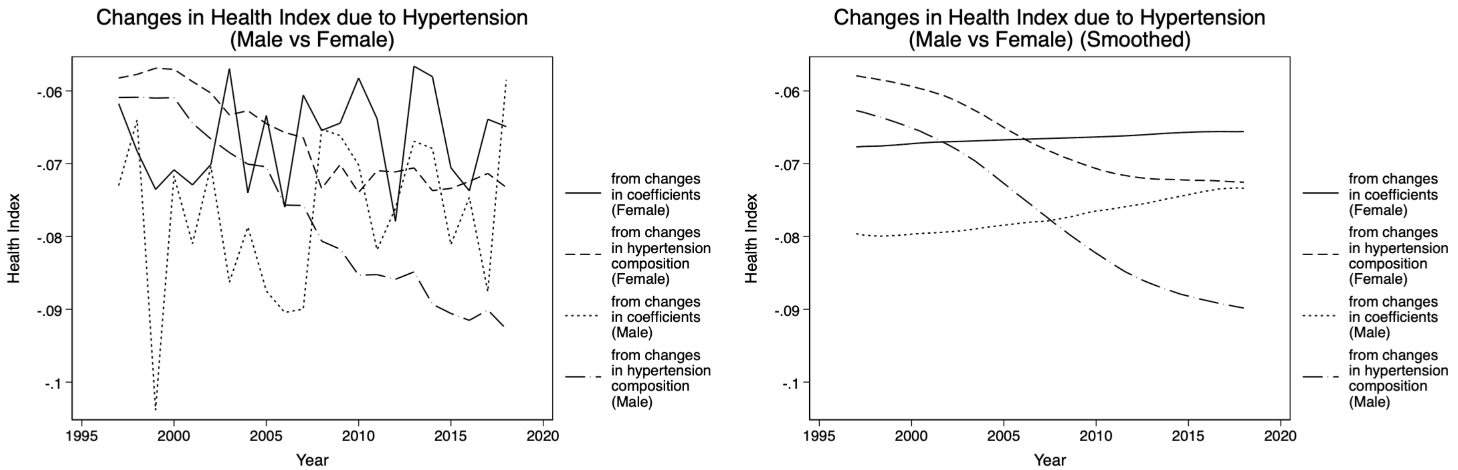


Figure 2.23: Prevalence of Diabetes

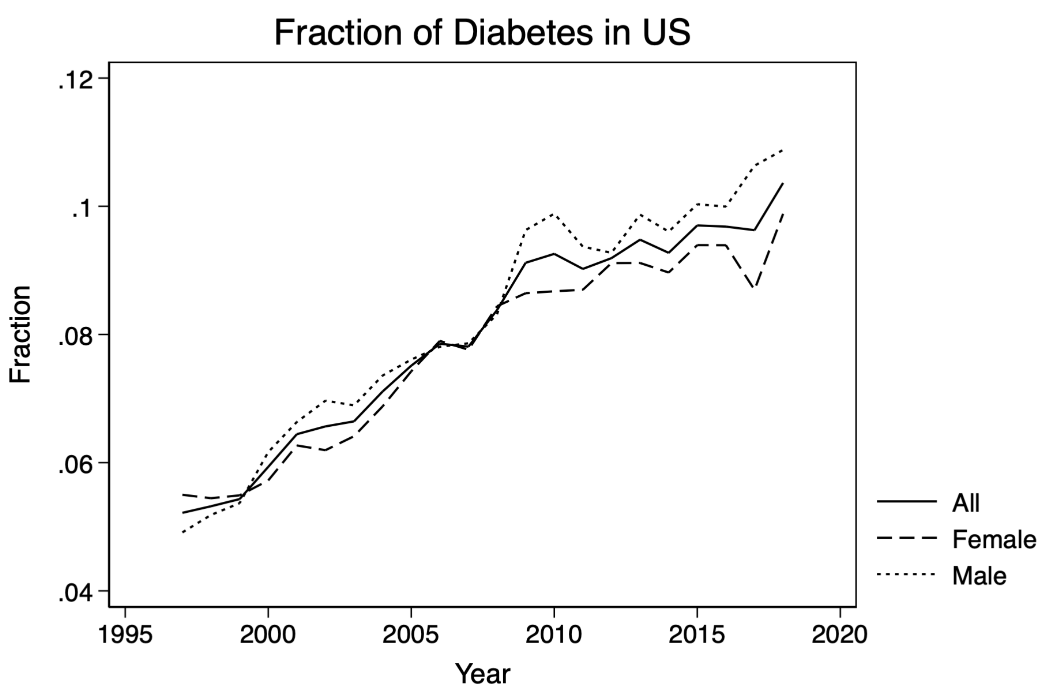


Figure 2.24: Prevalence of Diabetes by Age

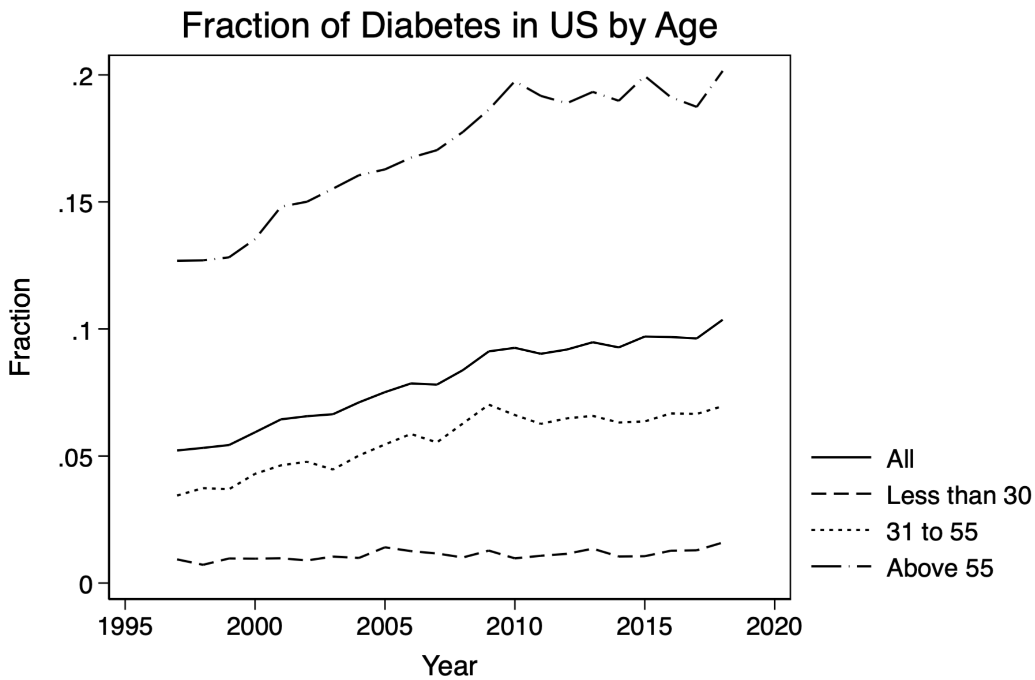


Figure 2.25: Estimated Coefficients of Diabetes

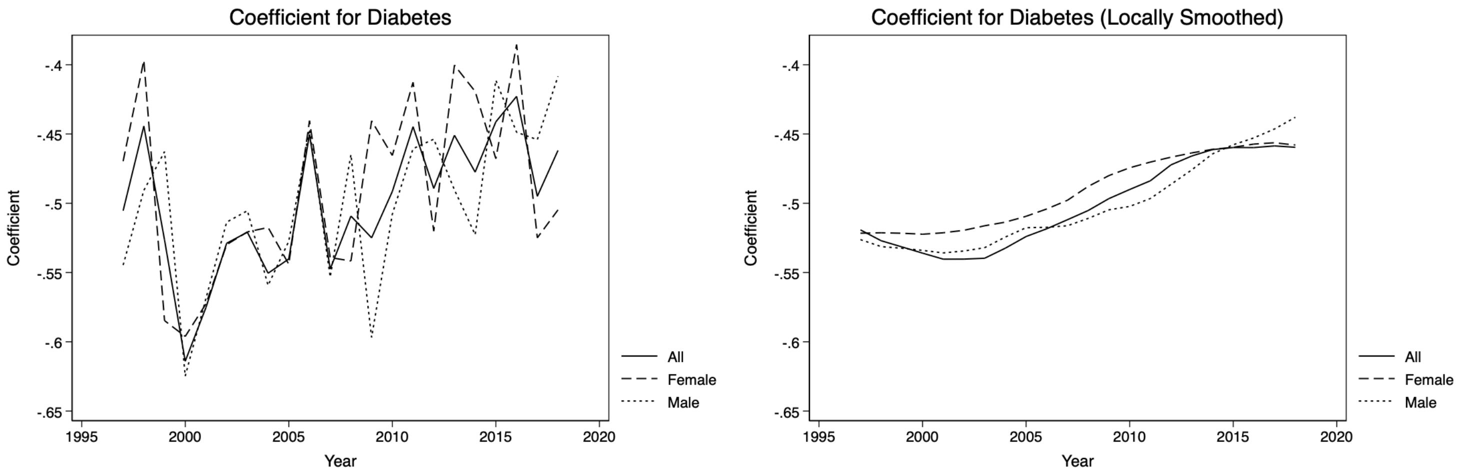


Figure 2.26: Diabetes Portion of Health Index

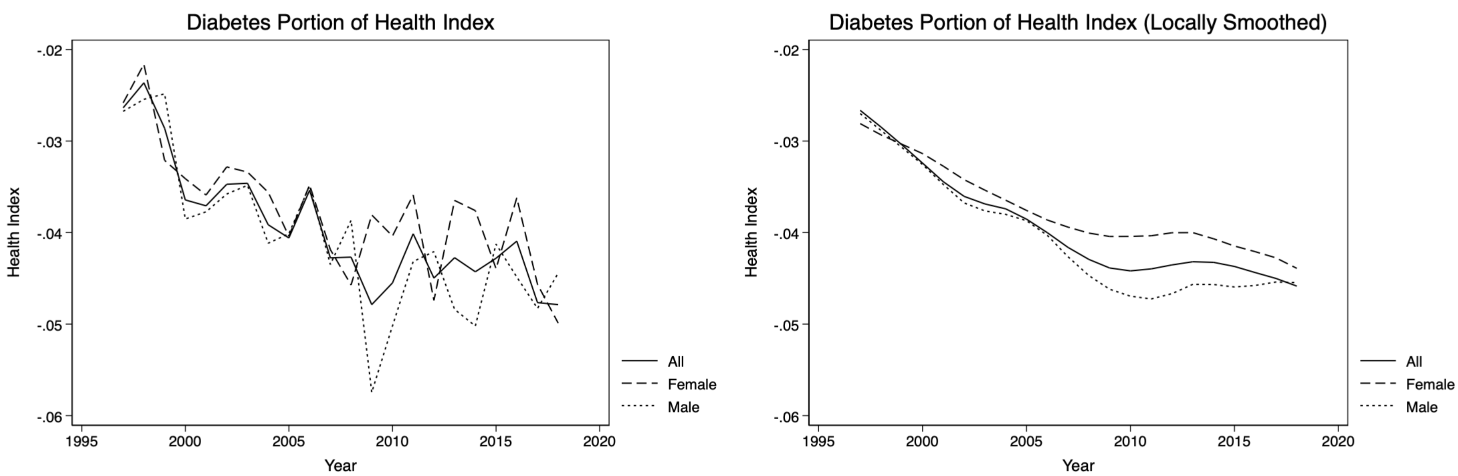


Figure 2.27: Changes in Diabetes Portion of HI due to Changes in Diabetes Coefficient and Composition

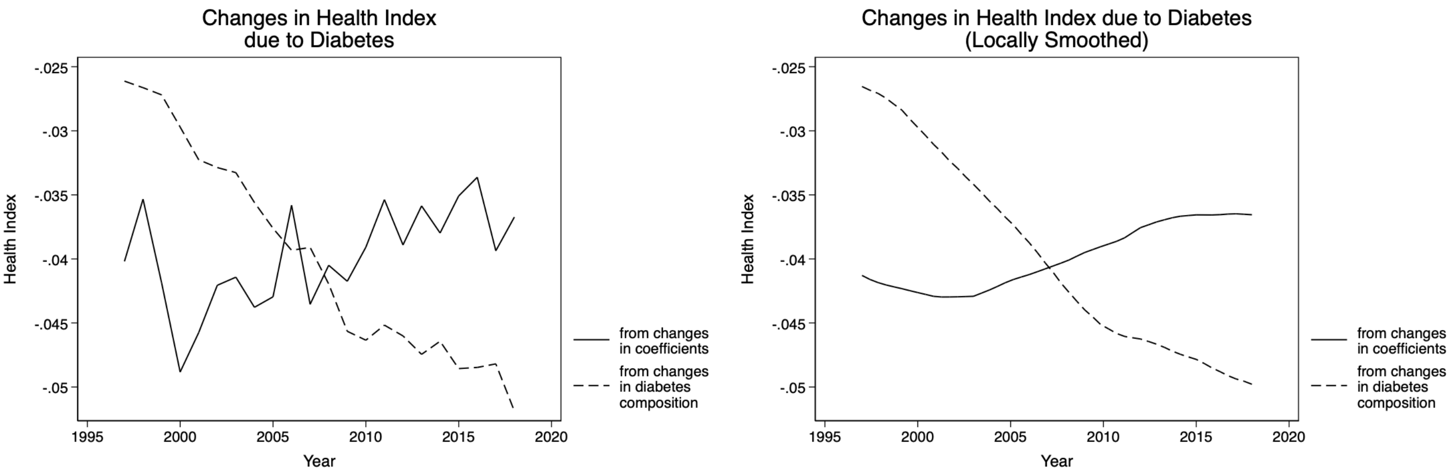


Figure 2.28: Changes in Diabetes Portion of HI due to Changes in Diabetes Coefficient and Composition (Male and Female)

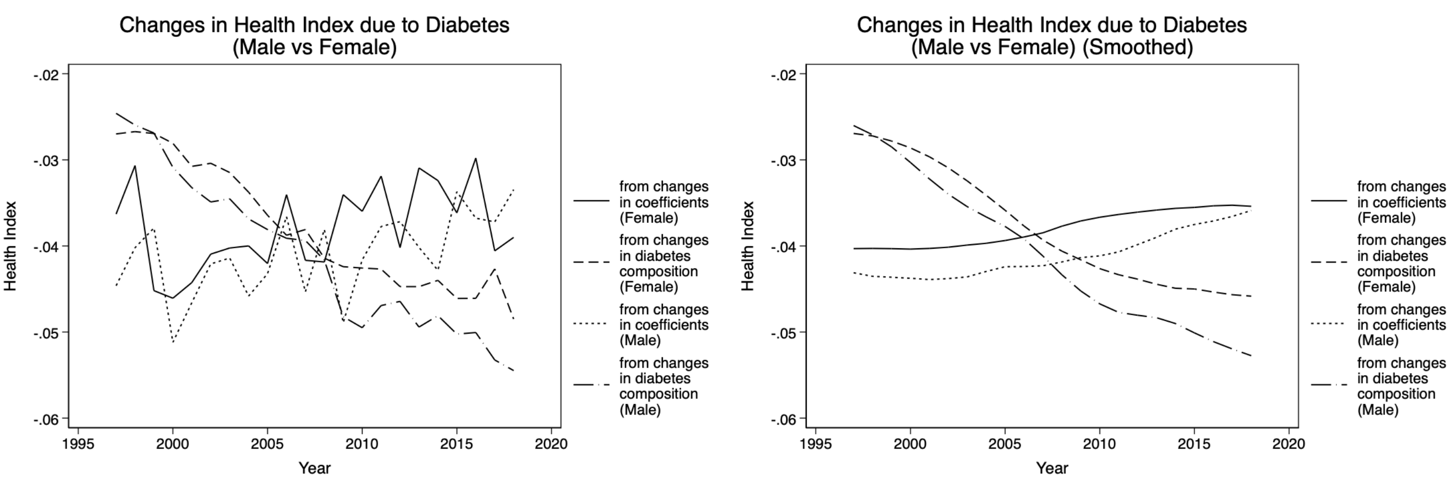


Figure 2.29: Fraction of Income Categories Over Time

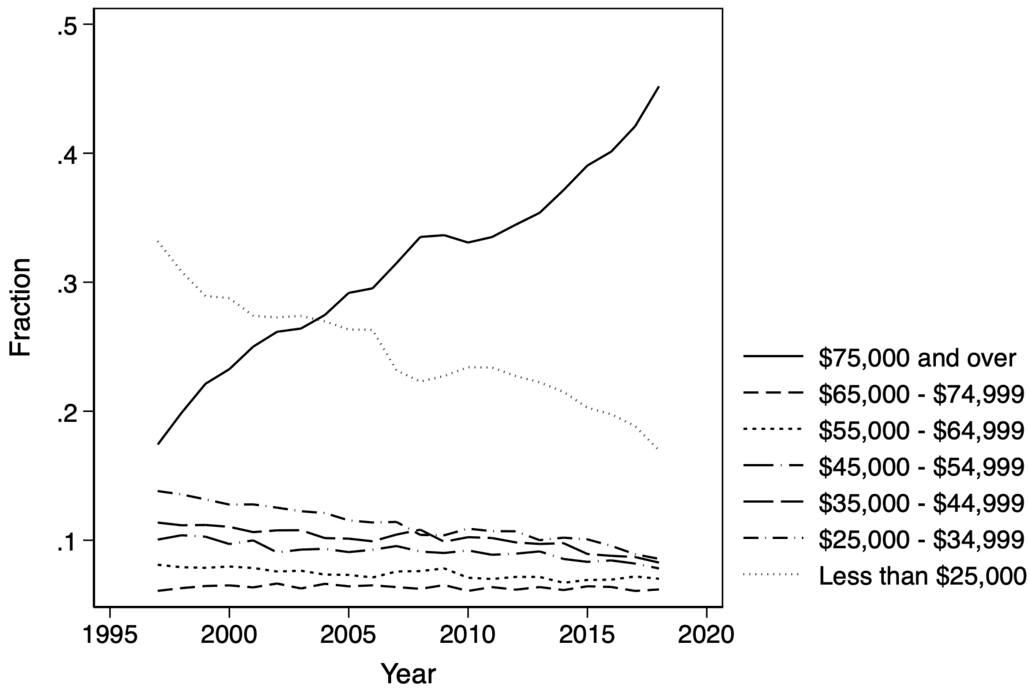


Figure 2.30: Income Portion of Health Index for Income $75,000 and Over (By Gender)

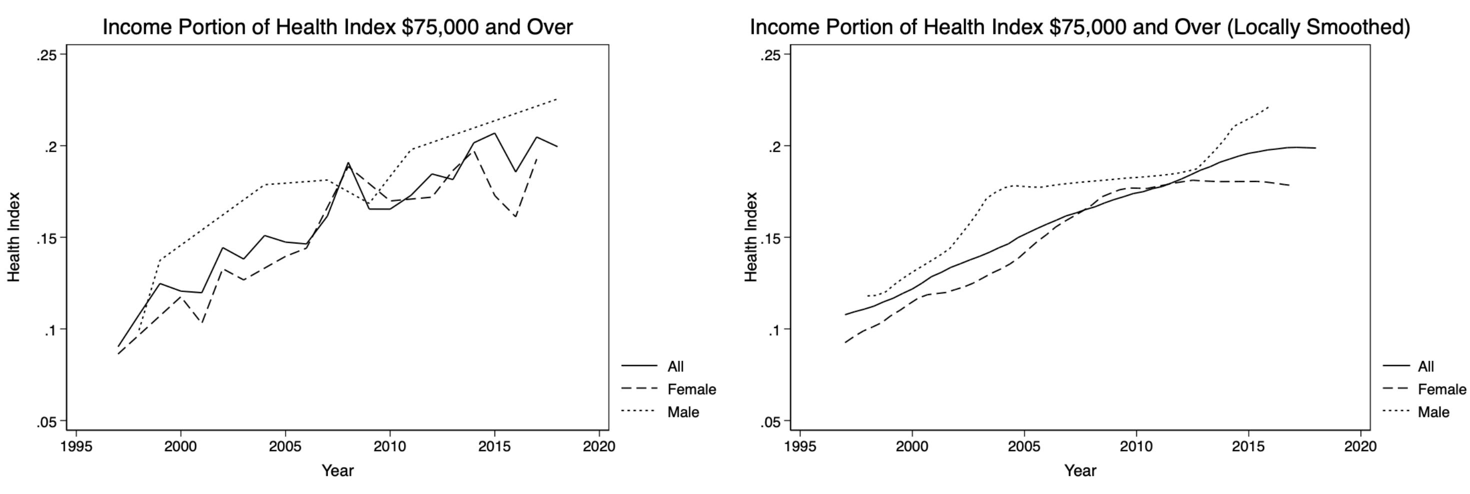


Figure 2.31: Estimated Coefficients of Income $75,000 and Over

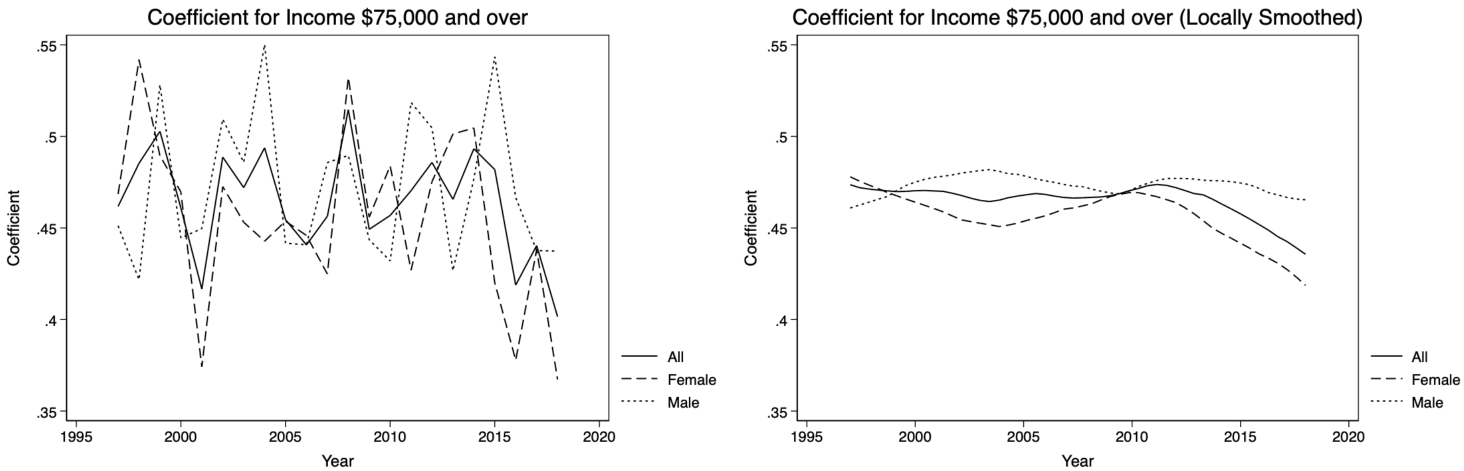


Figure 2.32: Estimated Coefficients of Income (Select Categories)

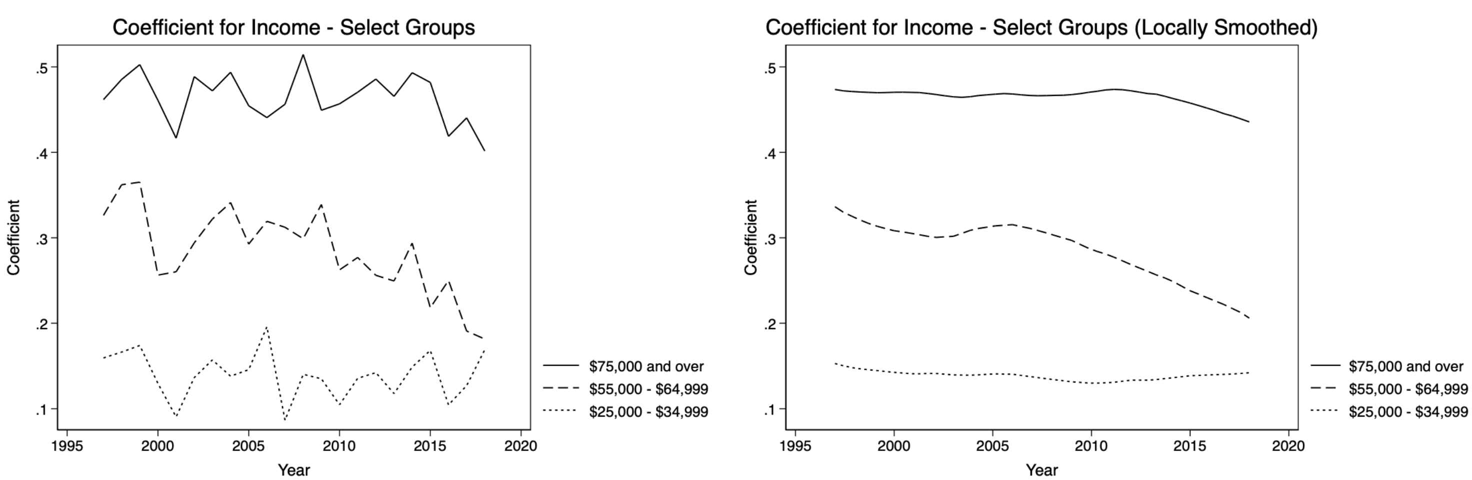


Figure 2.33: Changes in Income Portion of HI due to Changes in Income Coefficient and Composition (Income Over $75,000)

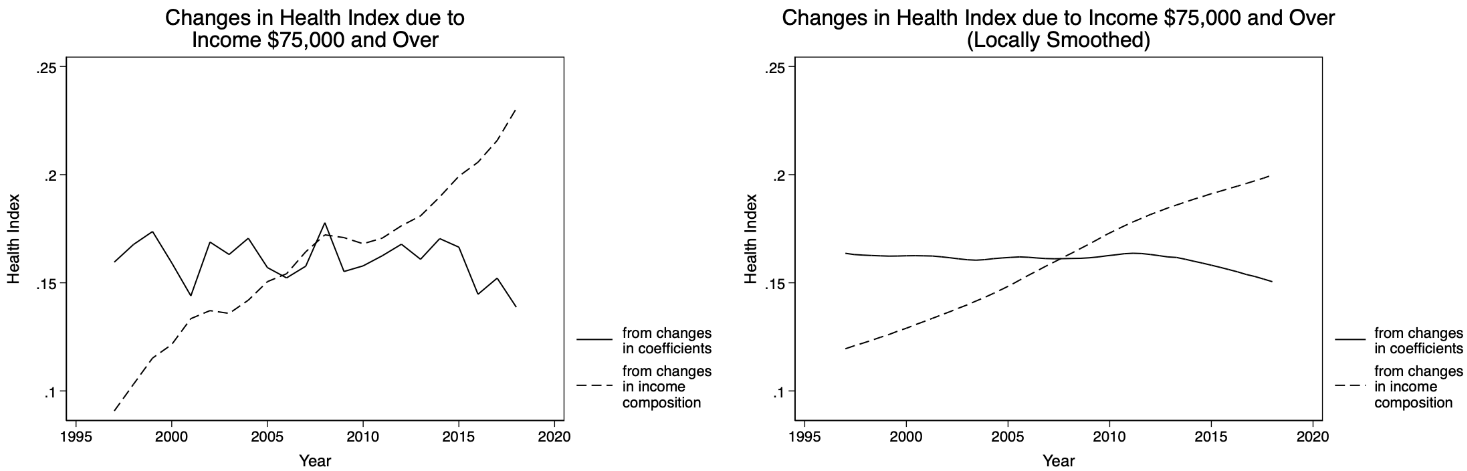


Figure 2.34: Changes in Income Portion of HI due to Changes in Income Coefficient and Composition (Income Over $75,000) - Male and Female

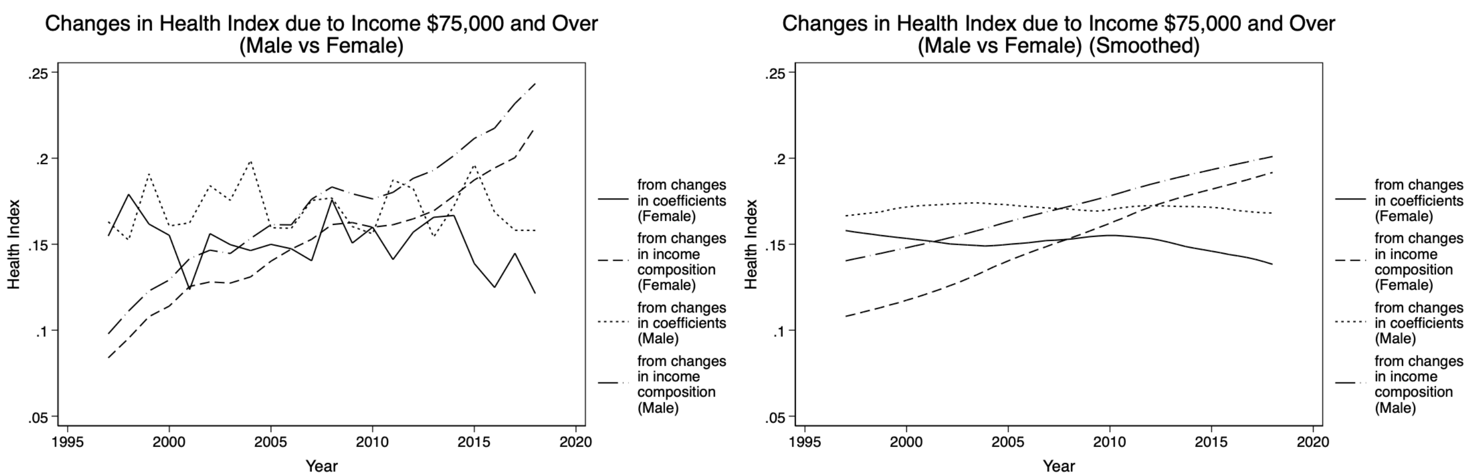


Figure 2.35: Estimated Thresholds of Health Index

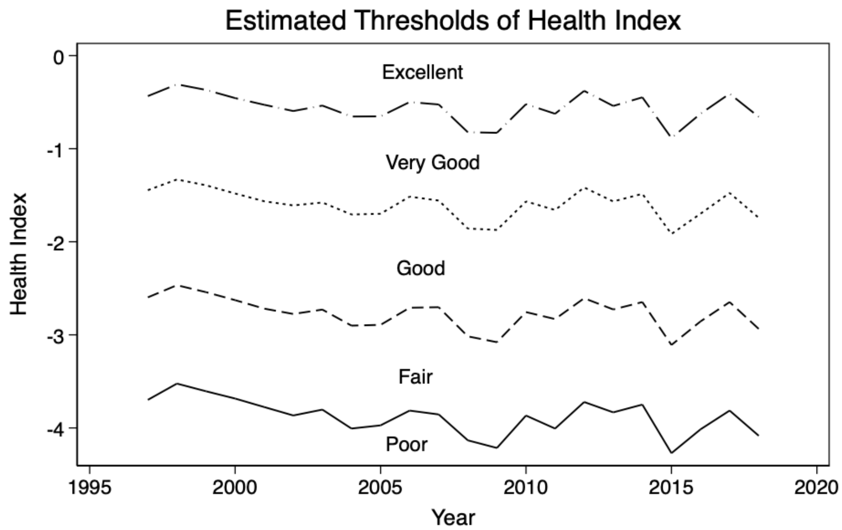


Figure 2.36: Estimated Thresholds of Health Index (Female)

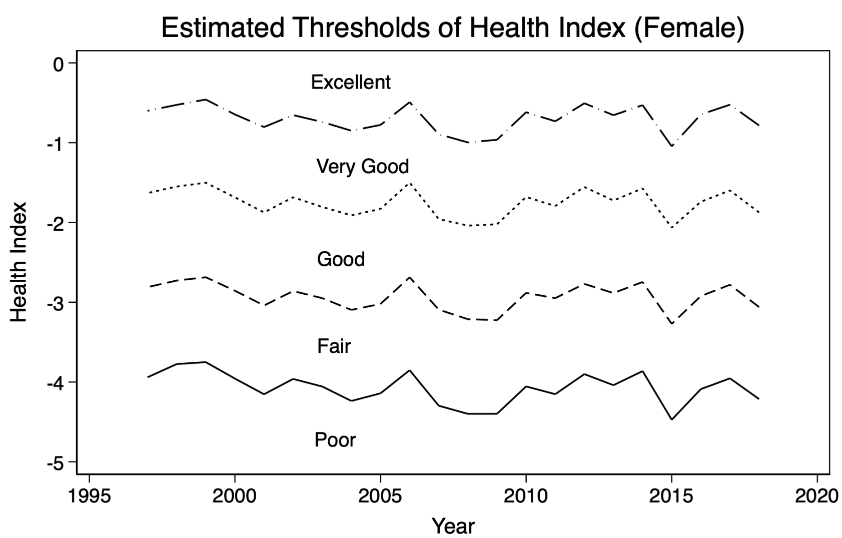


Figure 2.37: Estimated Thresholds of Health Index (Male)

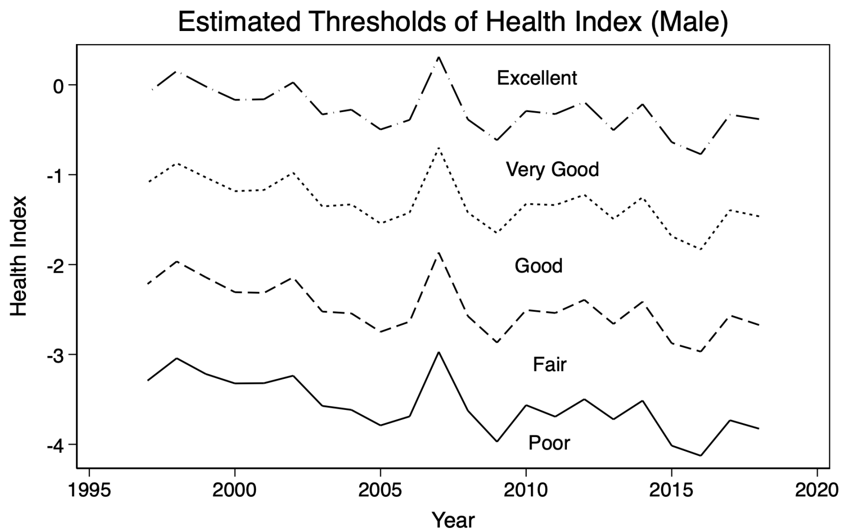


Figure 2.38: Health Index Density, for All, for Selected Years

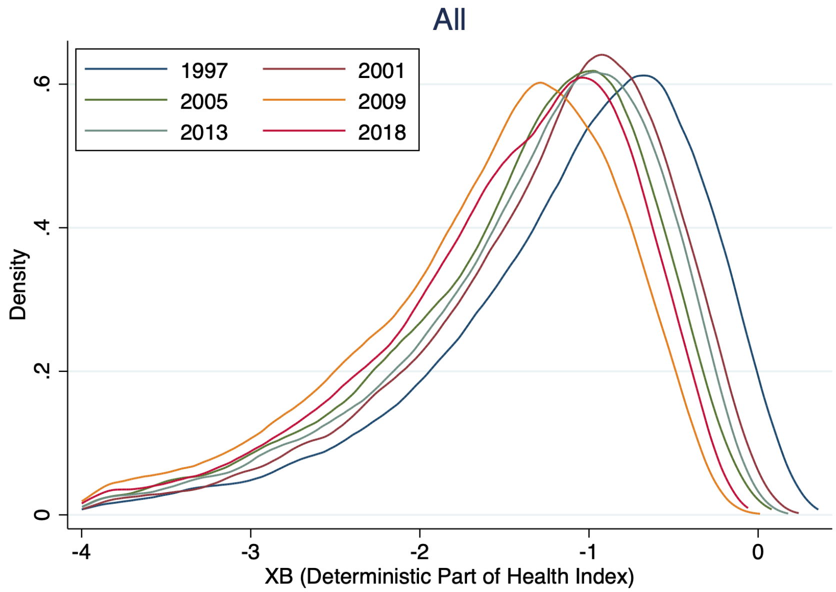


Figure 2.39: Health Index Density, for Females, for Selected Years

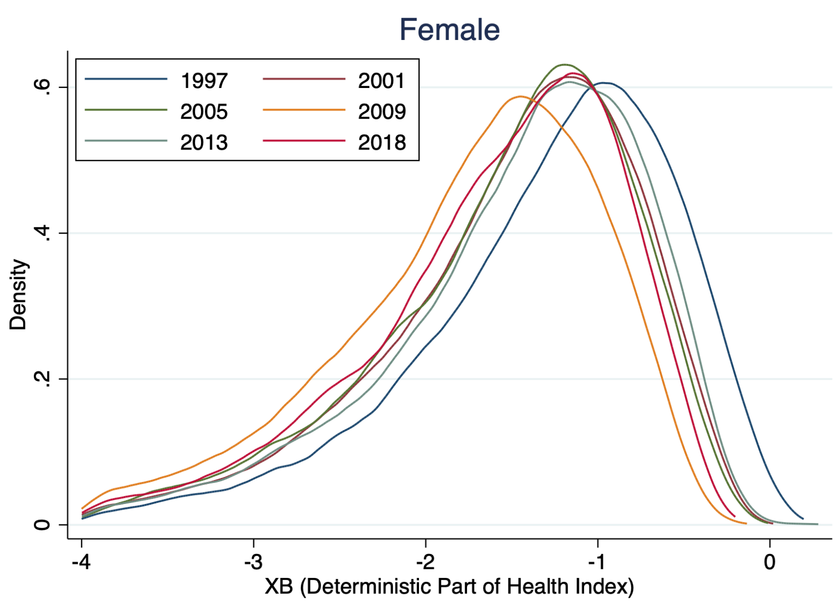


Figure 2.40: Health Index Density, for Males, for Selected Years

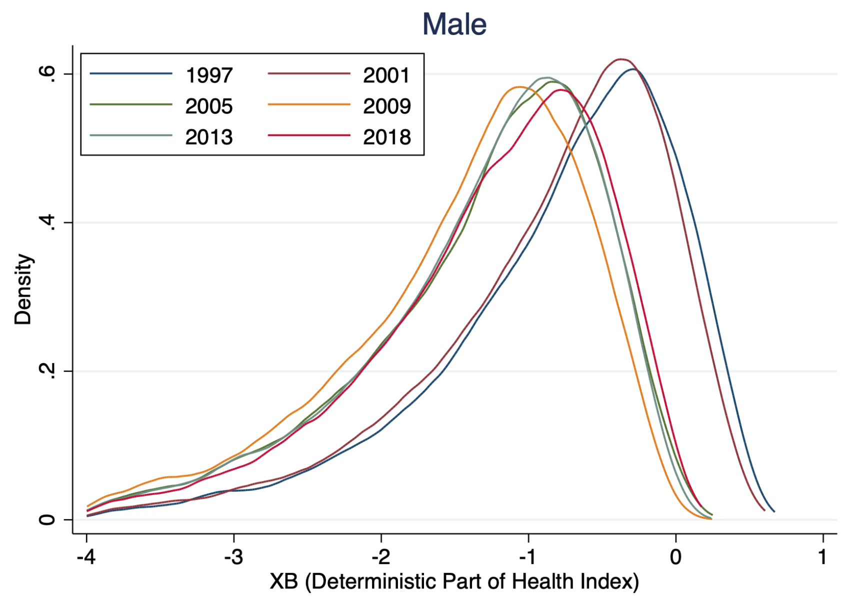


Figure 2.41: Model Predictions of Health Categories (All)

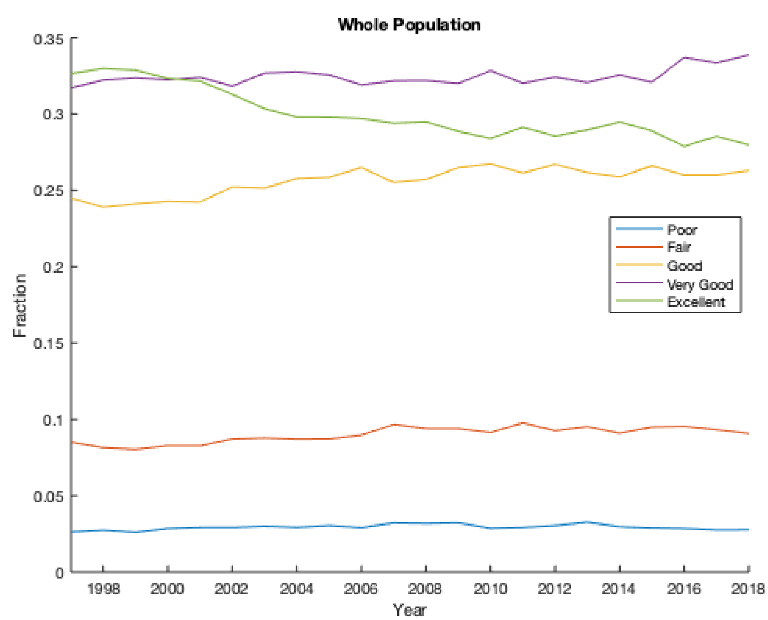


Figure 2.42: Model Predictions of Health Categories (Female)

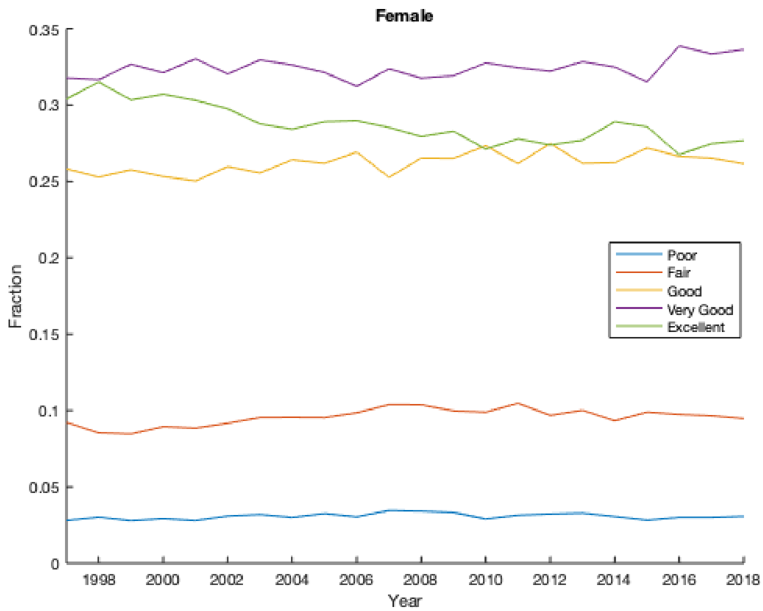


Figure 2.43: Model Predictions of Health Categories (Male)

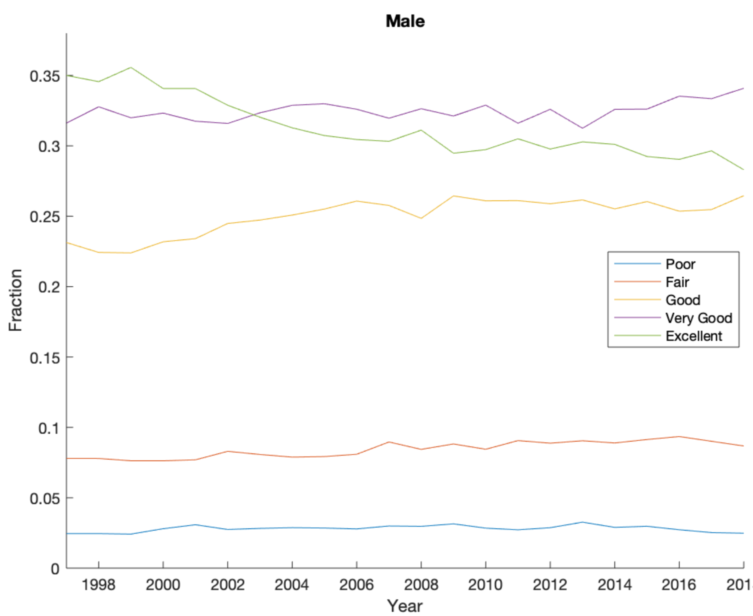


Figure 2.44: Model Counterfactuals Constant x's, Year-Specific β's and c's (All)

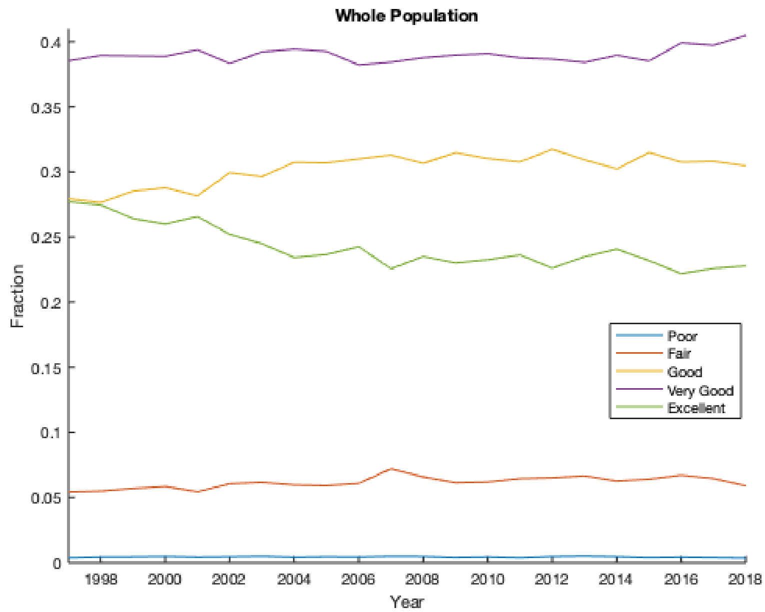


Figure 2.45: Model Counterfactuals Constant x's, Year-Specific β's and c's (Female)

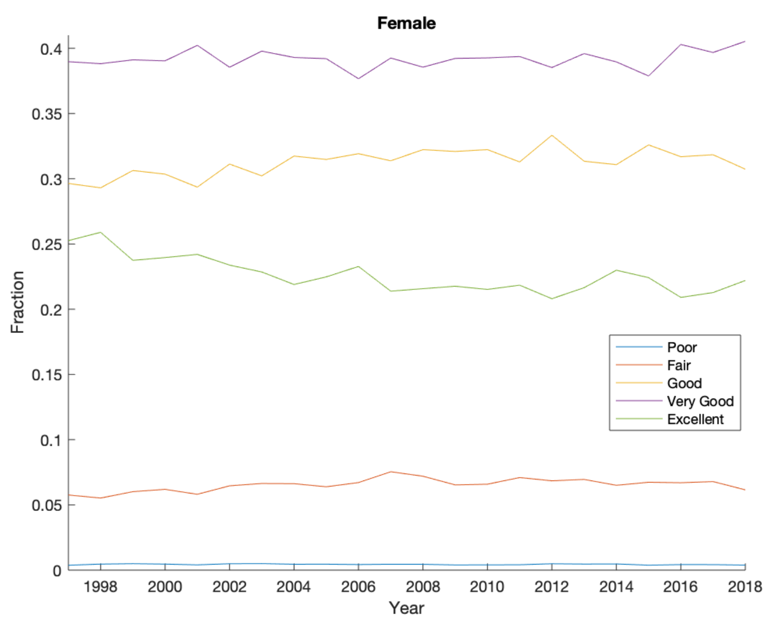


Figure 2.46: Model Counterfactuals Constant x's, Year-Specific β's and c's (Male)

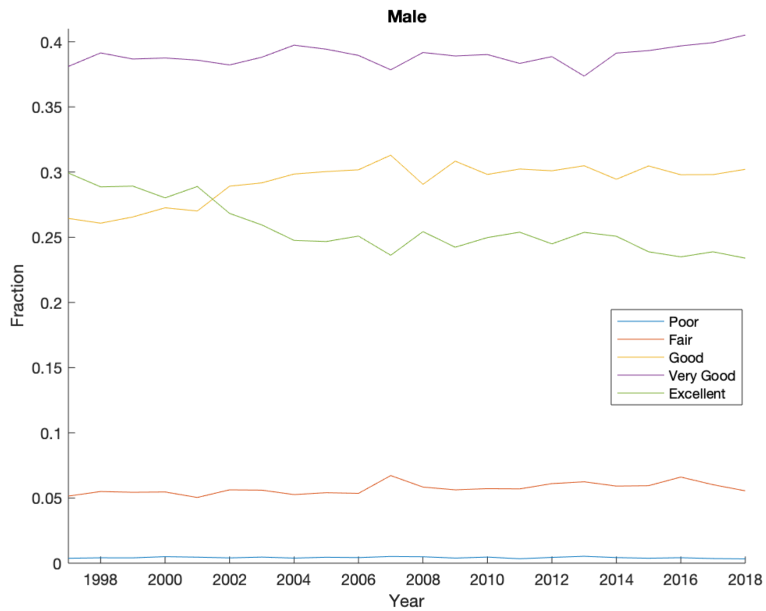


Figure 2.47: Model Counterfactuals Constant x's and c's, Year-Specific β's (All)

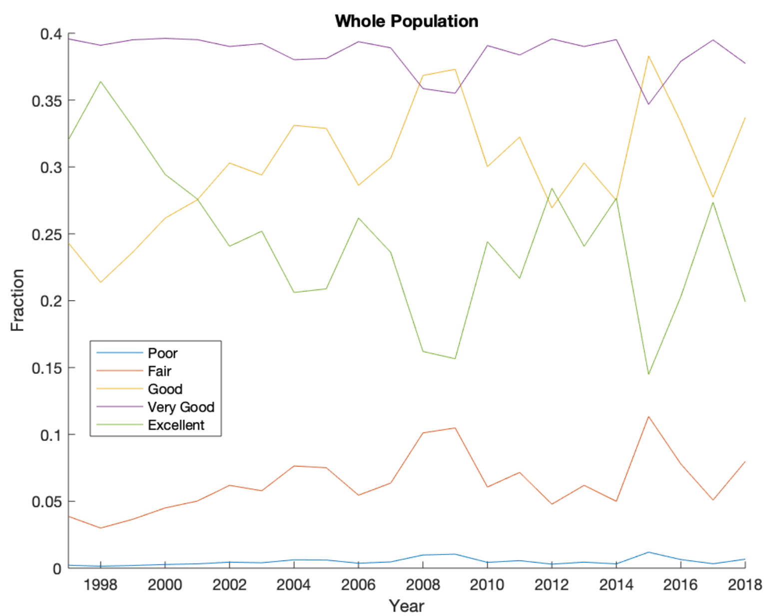


Figure 2.48: Model Counterfactuals Constant x's and c's, Year-Specific β's (Female)

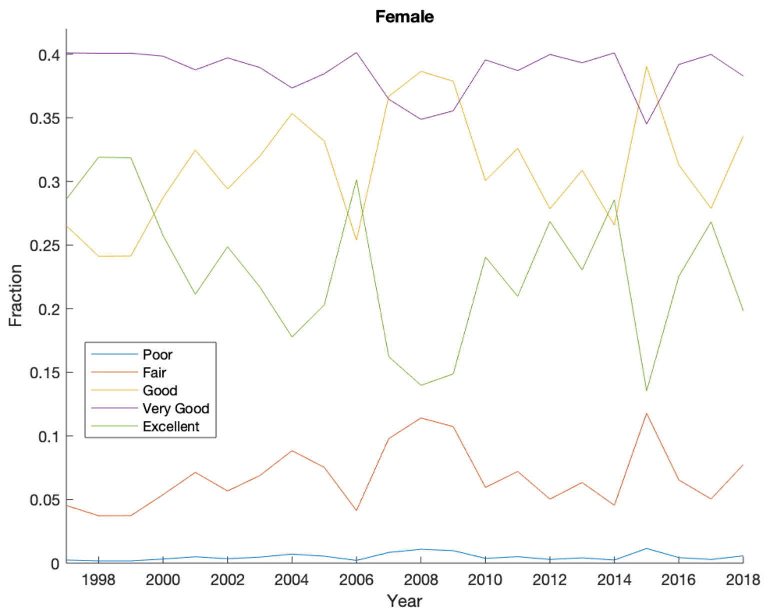


Figure 2.49: Model Counterfactuals Constant x's and c's, Year-Specific β's (Male)

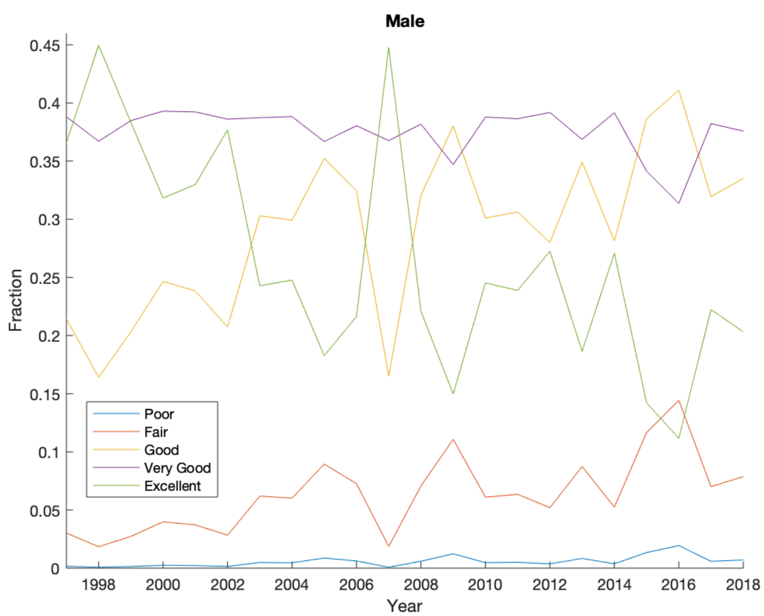


Figure 2.50: Model Counterfactuals Constant β's, Year-Specific x's and c's (All)

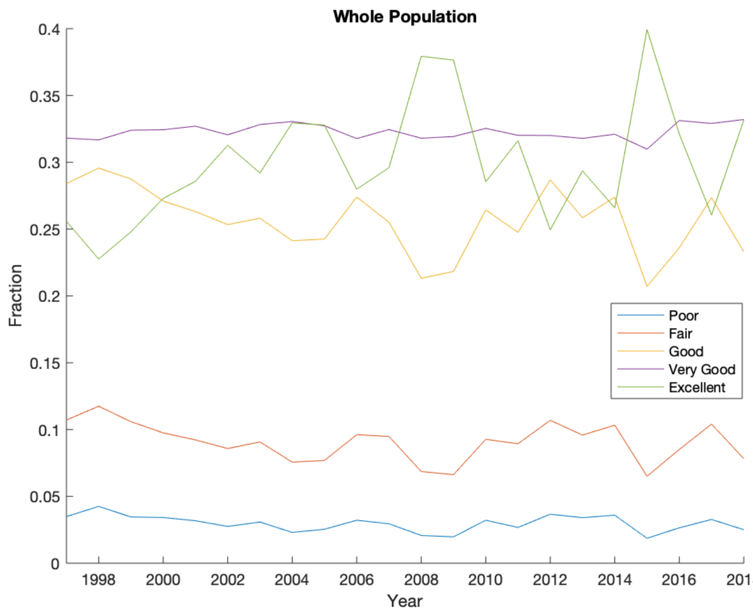


Figure 2.51: Model Counterfactuals Constant β's, Year-Specific x's and c's (Female)

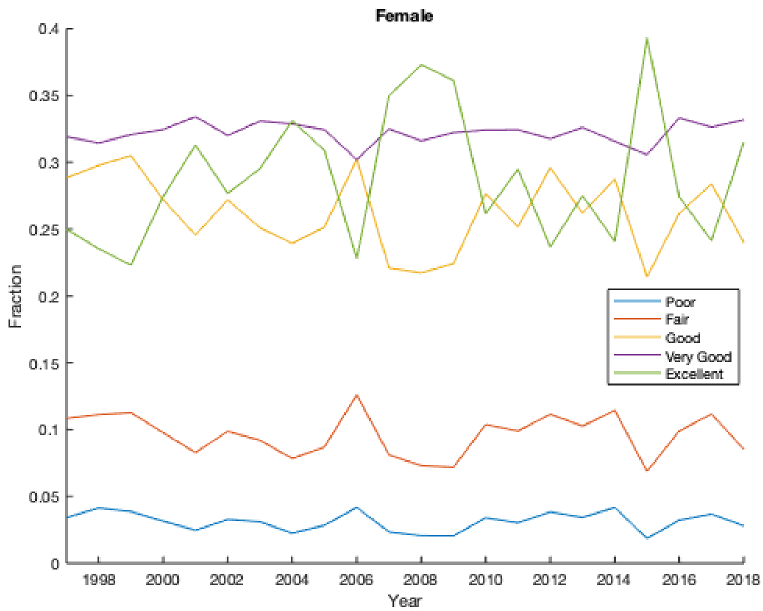


Figure 2.52: Model Counterfactuals Constant β's, Year-Specific x's and c's (Male)

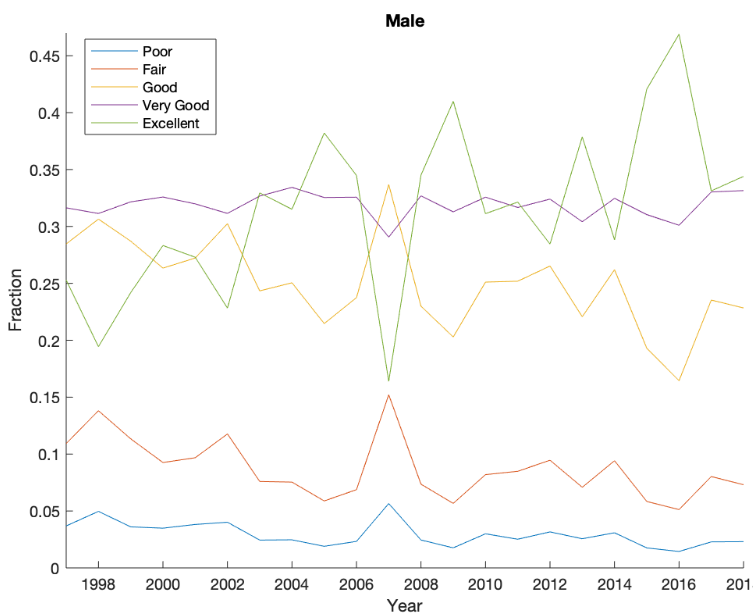


Figure 2.53: Model Counterfactuals Constant β's and c's, Year-Specific x's (All)

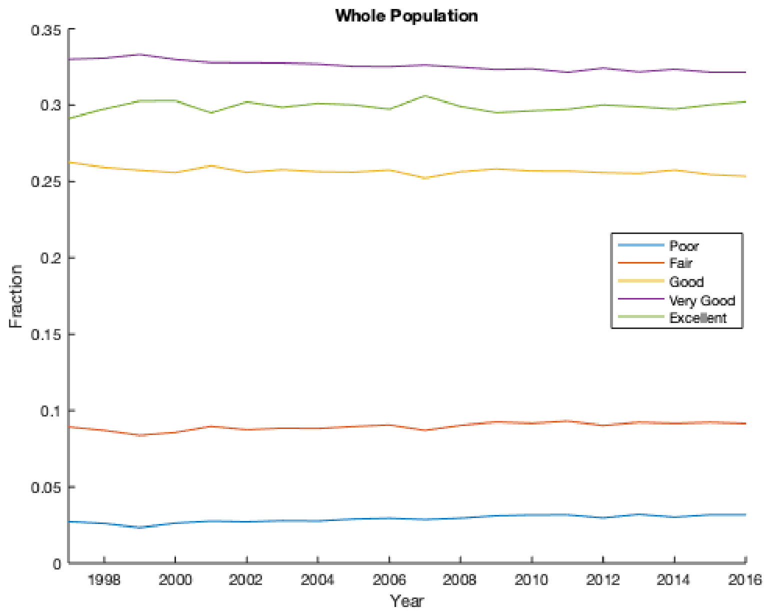


Figure 2.54: Model Counterfactuals Constant β's and c's, Year-Specific x's (Female)

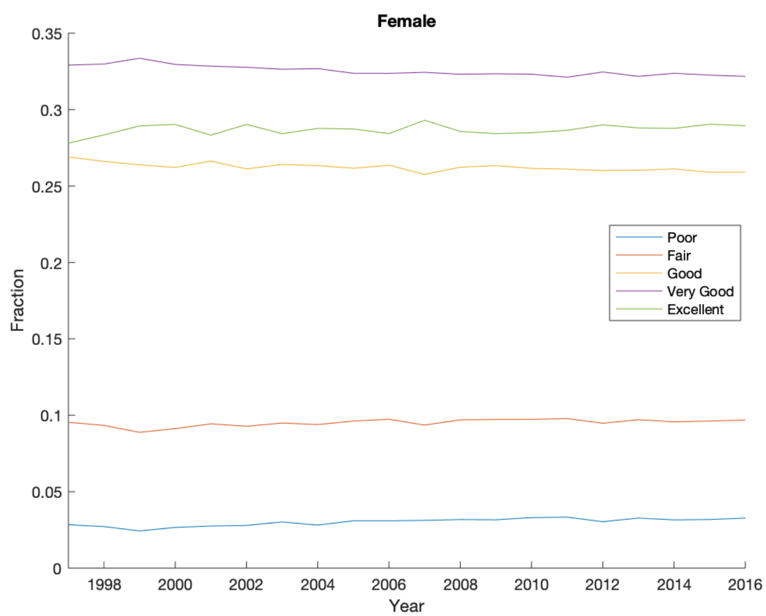
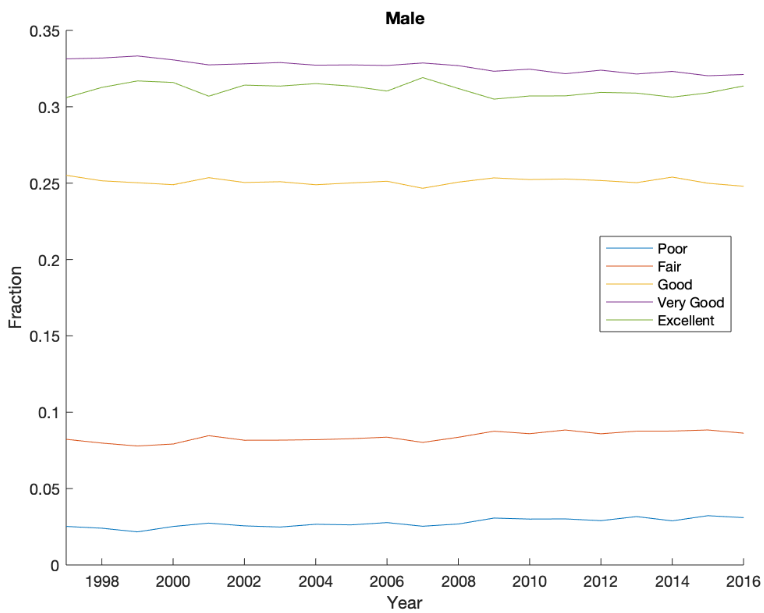


Figure 2.55: Model Counterfactuals Constant β's and c's, Year-Specific x's (Male)



2.7 References

Allison P.J., Locker D., and J.S. Feine (1997): "Quality of life: A dynamic construct," Social Science & Medicine, 45:2, pp. 221–30.

Aune, D, Sen A, Prasad M, et al. "BMI and all cause mortality: systematic review and non-linear dose-response meta-analysis of 230 cohort studies with 3.74 million deaths among 30.3 million participants." BMJ. 2016;353:i2156. Published 2016 May 4. doi:10.1136/bmj.i2156

Bjorner, J., Fayers P., and E. Idler (2005): "Self-Rated Health," in Fayers P. and R. Hays (eds): *Assessing Quality of Life in Clinical Trials*. New York: Oxford University Press, pp. 309–24.

Bound, J. (1991): "Self-reported versus objective measures of health in retirement models," *Journal of Human Resources*, 26:1, pp. 106–38.

Bound, J., Schoenbaum, M., Stinebrickner, T.R., and T. Waidmann (1999): "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers," *Labour Economics*, 6:2, pp. 179–202.

Buchinsky, M., Maestas, N., Reggio, I (2014): "The Evolution of Self-Reported Health," *Mimeo*.

Case, A. and A. Deaton (2005): "Broken Down by Work and Sex: How Our Health Declines," in Wise D.A. (ed): *Analyses in the Economics of Aging*, pp. 185–212.

Case, A. and A. Deaton (2015): "Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century." Proceedings of the National Academy of Sciences 112, no. 49 (2015): 15078-15083.

Centers, for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division for Heart Disease and Stroke Prevention. DHDSP Data Trends & Maps [online]. 2015. [accessed Jul 01, 2020]. URL: https://www.cdc.gov/dhdsp/maps/dtm/index.html.

Contoyannis, P., Jones, A.M., and N. Rice (2004): "The dynamics of health in the British Household Panel Survey," *Journal of Applied Econometrics*, 19:4, pp. 473–503.

Crimmins, E.M. (1990): "Are Americans healthier as well as longer-lived," *Journal of Insurance Medicine*, 22:2, pp. 89–92.

Currie, J. and B.C. Madrian (1999): "Health, Health Insurance and the Labor Market," in Ashenfelter, O. and D. Card (eds): *Handbook of Labor Economics*, Part 3, Chapter 50, pp. 309–416.

Diehr, P., Williamson, J., Patrick, D.L., Bild, D.E., and G.L. Burke (2001): "Patterns of Self-Rated Health in Older Adults Before and After Sentinel Health Events," *Geriatrics*, 49:1, pp. 36–44.

Dieleman JL, Baral R, Birger M, et al. US Spending on Personal Health Care and Public Health, 1996-2013. JAMA. 2016;316(24):2627–2646. doi:10.1001/jama.2016.16885

Freedman, V.A., Martin, L.G., and R.F. Schoeni (2002): "Recent Trends in Disability and Functioning Among Older Adults in the United States A Systematic Review," *Journal of the American Medical Association,* 288:24, pp. 3137-46.

Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). UK Prospective Diabetes Study (UKPDS) Group. Lancet. 1998 Sep 12;352(9131):854-65. Erratum in: Lancet 1998 Nov 7;352(9139):1558. PMID: 9742977.

FRED, federal reserve economic data. St. Louis, MO: Federal Reserve Bank of St. Louis.

Genuth, S. "The UKPDS and its global impact." Diabetic Medicine 25 (2008): 57-62.

Glenny, A. M., S. O'meara, A. Melville, T. A. Sheldon, and C. Wilson. "The treatment and prevention of obesity: a systematic review of the literature." International journal of obesity 21, no. 9 (1997): 715-737.

Golembiewski, R.T., Billingsley, K. and S. Yeager (1976): "Measuring Change and Persistence in Human Affairs: Types of Change Generated by OD Designs," *The Journal of Applied Behavioral Science*, 12:2, pp. 133–157.

Grabe, S., Ward, L. M., & Hyde, J. S. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. Psychological Bulletin, 134(3), 460-476. doi:http://dx.doi.org/10.1037/0033-2909.134.3.460

Henderson, R.M. (2005): "The Bigger the Healthier: Are the Limits of BMI Risk Changing over Time?" *Economics and Human Biology*, 3:3, pp. 339–66.

Home, P. D. "Impact of the UKPDS—an overview." Diabetic medicine 25 (2008): 2-8.

Howard, G.S. and P.R. Dailey (1979): "Response-Shift Bias: A Source of Contamination of Self-Report Measures," *Journal of Applied Psychology*, 64:2, pp. 144–50.

Idler, E.L. and Y. Benyamini (1997): "Self-rated Health and Mortality: A Review of Twenty-Seven Community Studies," *Journal of Health and Social Behavior*, 38:1, pp. 21–37.

Idler, E.L. and S.V. Kasl (1995): "Self-Ratings of Health: Do They also Predict Change in Functional Ability?" *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 50:6, pp. 344–53.

Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group

Jürges, H. (2007): "True Health vs Response Styles: Exploring Cross-Country Differences in Self-Reported Health," *Health Economics*, 16:2, pp. 163–78.

Kapteyn, A., Smith, J.P. and A. van Soest (2007): "Vignettes and Self-Reports of Work Disability in the United States and the Netherlands," *American Economic Review*, 97:1, pp. 461-73.

Kunst, A.E., Bos, V., Lahelma, E., Bartley, M., Lissau, I., Regidor, E., Mielck, A., Cardano, M., Dalstra, J.A.A., and J.J.M. Geurts (2005): "Trends in Socioeconomic Inequalities in Self-Assessed Health in 10 European Countries," *International Journal of Epidemiology*, 34:2, pp. 295–305.

Lee, Y. (2000): "The Predictive Value of Self Assessed General, Physical, and Mental Health on Functional Decline and Mortality in Older Adults," *Journal of Epidemiology & Community Health*, 54:2, pp. 123–129.

Lindeboom, M. and E. van Doorslaer (2004): "Cut-Point Shift and Index Shift in Self-

Reported Health," *Journal of Health Economics*, 23:6, pp. 1083–99.

Lokshin, M. and M. Ravallion (2005):"Searching for the Economic Gradient in Self-Assessed Health," World Bank Working Paper, WPS3698.

Marketdata LLC. 2019. "The U.S. Weight Loss & Diet Control Market."

Meara, E. and J. Skinner (2015): "Losing ground at midlife in America." Proceedings of the National Academy of Sciences 112, no. 49 (2015): 15006-15007.

Millsap, R.E. and S.B. Hartog (1988): "Alpha, Beta, and Gamma Change in Evaluation Research: A Structural Equation Approach," *Journal of Applied Psychology*, 73:3, pp. 574–84.

Organisation, for Economic Co-operation and Development. 2005. Health at a Glance: OECD Indicators 2005: Paris and Washington, D.C.: Author.

Parker, M.G. and M. Thorslund (2007): "Health Trends in the Elderly Population: Getting Better and Getting Worse," *The Gerontologist*, 47:2, pp. 150–158.

Ruhm, C.J. (2000): "Are Recessions Good for Your Health?" *Quarterly Journal of Economics*, 115:2, pp. 617-50.

Ruhm, C.J. (2003): "Good Times Make You Sick," *Journal of Health Economics*, 22:4, pp. 637–58.

Saklayen, M. G., & Deshpande, N. V. (2016). Timeline of History of Hypertension Treatment. Frontiers in cardiovascular medicine, 3, 3. https://doi.org/10.3389/fcvm.2016.00003

Selph S, Dana T, Blazina I, Bougatsos C, Patel H, Chou R. Screening for type 2 diabetes mellitus: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med. 2015 Jun 2;162(11):765-76. doi: 10.7326/M14-2221. PMID: 25867111.

Sprangers, M.A.G. and C.E. Schwartz (1999): "Integrating Response Shift into Health-Related Quality of Life Research: A Theoretical Model," *Social Science & Medicine*, 48:11, pp. 1507–15.

Tiggemann, M. (2011). Sociocultural perspectives on human appearance and body image. In T. F. Cash & L. Smolak (Eds.), Body image: A handbook of science, practice, and prevention (p.12-19). The Guilford Press.

United Kingdom Prospective Diabetes Study (UKPDS). 13: Relative efficacy of randomly allocated diet, sulphonylurea, insulin, or metformin in patients with newly diagnosed non-insulin dependent diabetes followed for three years. BMJ. 1995 Jan 14;310(6972):83-8. PMID: 7833731; PMCID: PMC2548496.

Verbrugge, L.M. (1984): "Longer Life but Worsening Health? Trends in Health and Mortality of Middle-Aged and Older Persons.," *Milbank Quarterly*, 62:3, pp. 474–519.

Wagstaff, A. and E. Van Doorslaer (1994): "Measuring Inequalities in Health in the Presence of Multiple-Category Morbidity Indicators.," *Health Economics*, 3:4, pp. 281–89.

Waidmann, T., Bound, J., and M. Schoenbaum (1995): "The Illusion of Failure: Trends in the Self-Reported Health of the US Elderly.," *Milbank Quarterly*, 73:2, pp. 253–87.

Wilson, R.W. and T.F. Drury (1984): "Interpreting Trends in Illness and Disability: Health Statistics and Health Status," *Annual Review of Public Health*, 5, pp. 83–106.

3 The Importance of Ordinal Rank Among Peers: Focusing on Height and Income

3.1 Introduction

Though highly context-specific, peer effects are important determinants in human behavior (Sacerdote 2014). A growing body of research show that peer effects are relevant in various contexts.

While the presence of peer effects are well-established, we have yet to understand them fully as peer effects can be heterogeneous (Sacerdote 2014; Mendolia, Paloyo, and Walker 2018; Tincani 2017). Of a particular interest for us is the presence of rank and status concerns that operate within peer interactions. Tincani (2017) use student data from Chile surrounding the 2010 Chilean earthquake and find that rank concerns play a significant role in determining student effort. This suggests a need for us to further investigate how peer effects are played out in the classrooms, as we learn that it is not simply the case that having high-ability peers will be beneficial. Further evidence on general rank and status concerns is found in Kuziemko et al. (2014) which provide extensive evidence of "last-place aversion" in various settings. Brown et al. (2008) also document evidence that wage rank affect job satisfaction.

The paper most closest to ours is from Murphy and Weindhardt (2018). Using British data, they look at the effect of academic rank at the final stage in elementary school on later academic performance and major choices. Their key finding is that rank is important conditional on the absolute and relative (from mean) performance on the test. This begs the question, would students respond to rank measures other than ones that are academic-related?

This is pertinent, especially as adolescents' perceived lower social hierarchy is associated with adverse outcomes. Goodman et al. (2005) find that students with perceived social disadvantage in terms of race and socioeconomic status have higher levels of stress. In an interesting twin study in Britain, Rivenbark et al. (2020) also find evidence that perceived social standing matters. They examine responses from twins of the same family and find that the one who indicated a higher perceived standing for his or her family exhibited better well-being in early and late adolescence. We investigate whether income and height, which can factor into one's calculation of social standing have any effect on student outcomes. In particular for income rank, while we are not aware of its role in an educational setting, we have ample evidence that perceived income rank matters in other contexts (Hounkaptin et al. 2015; Wetherall et al. 2015; Euteneuer 2014; Hoebel et al. 2017). As for the relevance of height, we also have ample evidence in labor market outcomes (Loh 1993; Harper 2008; Case and Paxson 2008) but we are not aware of studies that look at height ranking in education settings.

We use the South Korean education setting to investigate whether ordinal height and income rank have any effect on future academic achievement. We use the fact that between classrooms we observe variations in the distribution of height and income, which in turn generate different ordinal ranks even among students who have same height or income. Our results consistently show the importance of ordinal income rank, even after controlling for income levels, on future academic performance. We investigate potential mechanisms and find strong evidence that ordinal income rank is associated with greater parental investment in their children's education. Private tutoring is one main way we see this happening in the Korean context.

The remainder of the paper is as follows. We begin with a brief survey of the peer effects literature in Section 3.2. Then we introduce the panel data from the Korean education setting that we employ for our analysis in Section 3.3. In Section 3.4, we discuss our empirical strategy and in Section 3.5 we present our results. It is then followed by our exploration of potential mechanisms. We conclude in Section 3.6.

3.2 Review of Relevant Literature

With the help of two surveys on peer effects (Sacerdote 2014; Epple and Romano 2011), we highlight some of the growing literature here.

We first examine peer effects in classroom settings, specifically related to academic performance. Lavy and Schlosser (2011) look at gender composition in schools in Israel and find that having greater proportion of female students in a class led to improved student test scores. They observe that classrooms with greater female student composition have less disruptions, better teacher-to-student relationships, and lesser teacher fatigue. This is in line with results from Hoxby (2000) which looks at student outcomes from public schools in Texas. Other types of variations in student composition have been examined, including proportion of "low-achievers" defined as those who repeat their grade (Lavy, Passerman, and Schlosser 2008), and proportion of those who come from families that experienced domestic violence (Carrell and Hoekstra 2010). Imberman, Kugler, and Sacerdote (2012) analyze changes in classroom composition in Houston and Louisiana following evacuations from Hurricane Katrina. They conclude that while they find little or no impact of the influx of new students on student performance, on average, students do benefit from high achieving peers and perform worse with low achieving peers.

While the previous studies rely on "natural" variation in the composition of students, other studies attempt to examine peer effects in controlled assignment settings. For example, Sacerdote (2001) looks at random assignments in freshmen year roommates and dorm-mates and find evidence of peers' influence on student GPAs at Dartmouth college. Another such study is from Zimmerman (2003) who uses data from Williams College to find that having peers with higher SAT verbal scores raises one's own GPA. A potential limitation of these studies is that students will likely have peers outside of their roommate and dorm-mates. To address this, Carrell et al. (2009) use a unique setting at the United States Air Force Academy where students are grouped to be part of 30-person peer groups and have significant amount of academic and social interactions together. They find sizable peer effects, especially from having peers with higher SAT verbal scores.

Peer effects in contexts other than academic have been examined as well. The general conclusion of this literature is that oftentimes peer effects play significant roles in life decisions even outside of test outcomes (Sacerdote 2014). For example, Duncan et al. (2005) students more likely to binge drink if they have peers who demonstrate that behavior. Sacerdote (2001) finds that decisions whether to join a fraternity or a sorority are also influenced by peers. Boisjoly et al. (2006) provide evidence that white students who were assigned African American roommates demonstrated a more favorable opinion to affirmative action and diversity in the student composition. Duflo and Saez (2003) conducted an experiment and found evidence that decisions on retirement plan enrollment also are subject to peer influence. De Giorgi et al. (2010) find such evidence in major choices among college students. Even work productivity is subject to peer effects as demonstrated by Falk and Ichino (2006). These studies demonstrate that peer influences are a significant factor in one's decision-making process.

We adopt the main ideas found in the identification strategies found in Murphy and Weinhardt (2018) and Elsner and Isphording (2017). Murphy and Weinhardt use administrative data from British schools covering five cohorts and use variations in academic score distributions to find the effect of primary school academic ranking on secondary school outcomes. They use the fact that even while controlling for school-subject-cohort specific mean and relative distance from the mean, different ordinal rankings can arise from the shape of the distribution. Elsner and Isphording (2017) adopt a similar empirical strategy using data from the National Longitudinal Study of Adolescent to Adult Health to find similar results. They also rely on variations in school-cohort compositions that generate differences in ordinal ranking even if you had the same ability level. We adopt their insights into our empirical strategy in our data context.

3.3 The Data

3.3.1 Seoul Education Longitudinal Study 2010

The data for this study came from the Seoul Education Longitudinal Study 2010 (SELS 2010) available through the Seoul Education Research & Information Institute (SERII). This is a panel data that began its data collection in 2010 with the purpose of making data-driven improvements in education policy in Seoul, Korea. Data was collected from five major groups: 1) students, 2) parents, 3) teachers, 4) school principals, and 5) schools.

SELS incorporates stratification and clustering in its sample design. For the student sample 289 schools were selected, and within each school two classrooms were included in the sample. In all, sample size amounted to 16,500 students. In the first wave of this study in 2010, 5,200

4th graders in elementary school, 4,600 7th graders (in their 1st year of middle school), and 6,600 10th graders (in their 1st year of high school) were included in the sample. Afterwards these students were tracked until their graduation in high school (12th grade). That meant the sample years for students ranged from three years (for those who were in 10th grade in 2010) to nine years (for those who were in 4th grade in 2010). This is illustrated in Table 3.1.

For the student sample, wide-ranging information was collected that captured a wholistic picture of a student's life. Questions included their views toward school, teachers, and friends, attitudes in class, involvement with private tutoring, parental support, self-view, and indicators of physical health. As for the parent questionnaire, it included questions regarding family composition, socioeconomic status, their involvement in their children's education including whether information about private tutoring. The teacher questionnaire specifically asked Korean, Mathematics, and English subject teachers on their experience and how they prepare for classes in addition to methods of student evaluation. The questionnaire designed for the school principal asks questions regarding school administration and decision-making process. Finally, the school questionnaire consists of administrative information on the school.

On top of the information that was collected among these groups, separate academic assessments were conducted as part of the study. Students were tested in Korean, Mathematics, and English subjects.

3.3.2 Panel Sample

For our panel data, we combine data from waves one through six (2010 through 2015) and focus on the elementary student panel. This sample is selected so that we understand the effect of early-age income and height ranking on later outcomes. We end at wave six because of the availability of the data at the point of our analysis. For our main analysis, we end up with 2,457 students in our sample panel.

In Table 3.2 we provide descriptive statistics of the sample. It includes information on students and their parents, from personal characteristics to academic as well as information on private-tutoring usage. We also have some limited but useful information on the parents as well, like their education level and income. Perhaps not surprising but worth noting is that the students in this sample come from relatively higher socioeconomic backgrounds, as can be noted by the high proportion of parents who have some post-secondary education (0.70 for fathers and 0.60 for mothers) and above-bachelor level degrees (0.56 for fathers and 0.41 for mothers).

3.3.3 Ordinal Rank

Now for our main explanatory variable, we detail how we construct our ordinal rank variable. We use the family income and height variable to compute the students' rank variables within their classroom. For each classroom, we first construct the ordinal rank n_{icj} for student i, classroom c, and variable j. Then the ordinal rank is transformed to a percentile rank in order to have a comparable rank measure across different classroom sizes:

$$R_{icj} = \frac{N_{icj} - n_{icj}}{N_{icj}} \tag{3.1}$$

where N_{icj} is the classroom size. As noted earlier, j indicates the variable of interest, and in our case we construct different measures of academic, income, and height rank within a classroom setting. It should be noted that except for the height rank variable, because other variables like income and academic scores are not directly observable, it is a proxy measure of the perceived ranking within each classroom. We also have growing evidence that individuals are more accurate in their self-assessment of their social standing (Kraus and Keltner 2009; Anderson et al. 2001; Anderson et al. 2006). Henceforth this R_{icj} will be referred to as ordinal rank. To see the resulting distribution of the ordinal ranks of income and height, see Figures 3.1 and 3.2

3.4 Empirical Strategy

Our main specification will be as follows. Let $t = 1$ be the first wave. We run the regression five times for each period.

$y_{i,t} = \beta_1 + \beta_2 AcademicScore_{i,1} + \beta_3 AcademicScore_{i,1}^2 + \beta_4 AcademicScore_{i,1}^3 + \beta_5 ScoreRank_{i,1} + \beta_6 Height_{i,1} + \beta_7 HeightRank_{i,1} + \beta_8 HeightRank_{i,1} \times Female_{i,1} + \beta_9 Income_{i,1} + \beta_{10} IncomeRank_{i,1} + \beta_{11} IncomeRank_{i,1} \times Female_{i,1} + \beta_{12} Female_{i,1} + \sum_{c=1}^{C} \gamma_c Peer_{c,1} + \epsilon_{i,t}$ for $t = 2, 3, 4, 5, 6$

where $y_{i,t}$ is academic achievement for student i at period t, $AcademicScore_{i,1}$ is student i's overall academic test score from period 1, $ScoreRank_{i,1}$ is ordinal academic rank based on $AcademicScore_{i,1}$, $Height_{i,1}$ is student height, $HeightRank_{i,1}$ is ordinal rank of height ***by***

gender as we thought that is the proper comparison group, $HeightRank_{i,1} \times Female_{i,1}$ is the interaction term for height rank and indicator for females, $Income_{i,1}$ is income, $IncomeRank_{i,1}$ is ordinal rank of income, $IncomeRank_{i,1} \times Female_{i,1}$ is the interaction term for income rank and indicator for females, $Female_{i,1}$ is indicator for female, and $\sum_{c=1}^{C} \gamma_c Peer_{c,1}$ is the collection of classroom dummy variables and will help control for the classroom characteristics. For the last collection of peer indicators, $\sum_{c=1}^{C} \gamma_c Peer_{c,1}$, we also define peer groups in terms of classroom-gender groups as well as run a school-level peer analysis.

For the intuition behind this empirical strategy, we refer to Figure 3.3 also found in Murphy and Weinhardt (2018). It illustrates how two students who have the same income (or height) level and same distance from the mean income (or height) level can have different ordinal rank depending on the distribution of income (or height). In this figure, notice how points X and Y have the same income or height level and distance away from the class mean, yet have different ordinal rank. Same income or height levels can result in different ordinal rank measures based on the distribution of income or height in a given classroom. Figures 3.1 and 3.2 help demonstrate this as well.

In order to implement this strategy, it is important that we include the classroom indicators, which will absorb all the mean differences in sampled classrooms. As Elsner and Isphording (2017) points out, this is a within-transformation of all variables at the school-classroom level. This allows us to control for time-invariant classroom-level characteristics and factors that affect all students in the same way. We end up comparing students across all classes after removing mean differences between classes.

3.5 Results

3.5.1 Main Results

Now we come to the results. We refer to Table 3.3 for our discussion. Highlighting couple things of note, we find that ordinal income rank in Wave 1 (4th grade) has a significant impact on subsequent academic achievement. The interpretation of our results would indicate that a one percent increase in the ordinal income rank is associated with 0.08 to 0.12 increase in future overall academic achievement score (out of 100). To give context to this number, the standard deviation of ordinal income rank is around 28 and that of overall academic score is around 20. This would then correspond to having approximately 0.14 increase in standard deviation in academic scores arising from a one standard deviation increase in ordinal income rank. The results from the interaction term between income rank and female dummy variable indicates that there is no differential impact of income rank between male and female students.

When we examine the height and height ranking variables, we make note that the ordinal rank has no effect on future academic performance. Rather, the height variable plays a role in academic achievement and this is consistent with the observation from Case and Paxson (2008) that taller children demonstrate superior cognitive abilities, even during ages before schooling had any significant role in their education. As expected academic score from Wave 1 predicts future academic performance very well. Academic rank from Wave 1 is statistically insignificant in explaining future performance, and plays a minimal role.

We explore different definitions of the peer indicators. In Table 3.4 we show results from defining peer indicators by classroom-gender groups. This is to explore the possibility that the relevant reference group for students would be same-gender peers for height and income. We noted earlier that for the ordinal height rank variable, we already define it in this way, but we extend our analysis here by defining the ordinal income rank variable within classroom-gender groups and including peer dummies that reflect this change. This specification would amount to comparing variations in ordinal rank that arise in differences in the distribution of income and height between classroom-gender groupings.

Our point estimates, specifically with regards to the impact of ordinal income and height rankings, do not change that significantly compared to the previous specification. The statistical significance we observed with the height measure, however, are lessened for Waves 2 and 3, perhaps reflective of the decrease in variation once we focus our analysis between classroom-gender groupings which are smaller.

In Table 3.5, we expand our analysis by defining peers as same-school peers. This specification allows us to use added variation in income and height distributions across classrooms that even have mean-differences as shown in Figure 3.4. A snapshot of the sample illustrate this as well in Figure 3.5. This specification, however, fails to control for potential confounders across classrooms such as teacher quality, which is why our preferred analysis is the one with peers defined those within their classrooms. However, we do show the results here and we find that there are no significant deviations from our main specification.

3.5.2 Potential Mechanisms

So what can explain what we observe in our analysis? We explore them here. Specifically we focus our attention on 1) student confidence and 2) parental investment.

3.5.2.1 Student Confidence

Boosting student confidence can translate into increased effort and investment in academics as noted in Elsner and Isphording (2017) and Murphy and Weinhardt (2018). In order to examine whether these ordinal ranking measures affect student confidence, we focus on two questions that students were asked to respond to: "I believe I'm an able person" and "I believe I am a valuable person." In relation to these questions, they were asked to rate themselves from a scale of 1 to 5, where 1 corresponded to "Highly Disagree," 2 "Disagree," 3 "Average," 4 "Agree," and 5 "Highly Agree." We run ordered probit regressions and the results are shown in Table 3.6 and 3.7.

Turning our attention to these tables, we find a mixed picture. While in the later years we find that there are little to no evidence that any of the ordinal rank measures have measurable impact on students' confidence measures, we do see that income rank does seem to have somewhat of an impact on the self-ability measure as well as self-value measures during the earlier waves. These could be potential channels through which we observe improved academic performance in later years. The mixed result, however, also suggests that this channel cannot explain all of the improved academic gains for those who have higher ordinal rank in earlier years. Height seems to have no bearing on self-perception measures, nor ordinal rank of height.

3.5.2.2 Parental Investment

We next focus our attention on how income rank affects parental interest and involvement in students' education. For these we will examine students' response as well as parents' response. On the student side, we focus our attention on two questions asked: "My parents are interested in how I'm doing in school (parental interest)" and "My parents help me with my studies (parental involvement)." Students were asked to respond from a scale of 1 to 5, same as the one introduced above. Here we also run ordered probit regressions and the results are shown in Tables 3.8 and 3.9.

For the parents' response, we look at questionnaire given to parents. Similar to the student questionnaire, they were asked to respond to certain statements from a scale of 1 to 5, indicating whether they agreed with them or not. The specific statements we examine are: "I do not hold back on educational spending for my children," "I manage my children's daily schedule," "I encourage my children to study hard," and lastly "I collect information about private tutoring for my children." These results are shown in Tables 3.10 through 3.13.

First on the students' end, we turn our attention to Tables 3.8 and 3.9. We do find strong evidence, from students' response, that higher ordinal income rank is associated with greater parental interest and involvement. It is worth noting that ordinal income rank, not income, is what we find to have statistical significance in relation to these variables that indicate parental investment in their children's education. This could be a clear channel through which the ordinal income rank can affect future educational achievement as students receive more investment in their education.

When we examine responses from parents' end, we see similar findings as shown in Tables 3.10 through 3.13. In Table 3.10 we report parents' attitude towards educational spending. Not surprisingly, while higher income is associated with parents' willingness to invest more in their children's education, we see that the more significant association is with the ordinal income rank variable. In Table 3.11, we show how parents' managing of their children's schedule is affected by the ordinal income rank variable. And once again we see a strong association between the two. Likewise we see them encouraging studying more with higher income rank in Table 3.12 and them being more proactive about collecting information about private tutoring in Table 3.13. All of these point to the fact that higher income rank may affect future academic performance of students' through increased parental investment in their children's education.

Lastly, we examine private tutoring, which we believe is how parental investment in children's education is primarily manifested in the South Korean setting (Lee 2005; Kim and Bang 2016; Ko 2019). In fact, there's a term coined to express this: "education fever." Private Education Expenditure Survey in 2019 reports that total private tutoring expenditures of elementary, middle, and high school students reached 21 trillion Korean Won in 2019 (about $18.7 billion USD in 2021). The results are shown in Tables 3.14 through 3.20.

In Table 3.14 we present results on the use of private tutoring. It is pretty clear that a link exists between ordinal income rank and private tutoring usage. While Wave 1 results indicate a statistical significance only in the 0.10 level, we make note that a lot of observations are lost because of collinearity issues, where we have significant number of classrooms where all students participate in some form of private tutoring (in Table 3.2 we noted how 89.7% of students were already involved in private tutoring in Wave 1).

In order to examine whether there are private-tutoring quality differences among students, we prevent Tables 3.15 through 3.17. These results report whether parents of students with higher ordinal income rank spend more for their students, which is indicative of parental investment and quality of the private education. We find different results by subject. While for Korean and Math subjects we do not find a significant tie between expenditure on these subjects to ordinal income rank, parents of students with higher income rank seem to be willing to spend more for private education in English. This is consistent with what we know about the nature of private tutoring in Korea, as Park (2009) highlights especially the "English boom" in Korea.

When we look at results on weekly hours spent on private tutoring in Tables 3.18 through 3.20, we find similar results. We find that while for the Korean subject, ordinal income rank has no impact, we do find evidence that students spent more time in private tutoring in math and english if they had higher ordinal income rank.

3.6 Conclusion

We studied whether ordinal ranks in height and income among classmates, which could potentially help form one's self-view and motivate parental investment in their children, have any impact on future academic performance. Our results indicate that while ordinal rank in height has no measurable impact on future academic performance, ordinal rank in income has consistently been shown to have impact in this regard. We also noted that not ordinal rank in height, but height level itself had impact on future academic performance, consistent with the finding from Case and Saxon (2008).

We further investigate how ordinal rank in income could affect future academic performance, and while we find mixed results on self-esteem of students, we find strong and consistent evidence that higher ordinal income rank is associated with greater parental investment in their children's education. On both students' and parents' end, we find that parental involvement is strongly tied to ordinal income. Finally, given what we know about the South Korean's educational context with private tutoring, we find evidence that greater parental investment is manifested through increased use and expenditures in private tutoring for students.

Table 3.1: Sample Year For Each Grade

	Elementary			Middle			High School		
2010 Cohorts	4th	5th	6th	7th	8th	9th	10th	11th	12th Grade
4th Grade (Elementary)	**2010**	2011	2012	2013	2014	2015	2016	2017	2018
7th Grade (Middle School)				**2010**	2011	2012	2013	2014	2015
10th Grade (High School)							**2010**	2011	2012

In Korea, 1st to 6th graders are part of elementary school, 7th to 9th part of middle school, and 10th through 12th graders part of high school. This table is based on the

Table 3.2: Sample Descriptive Statistics - 4th Grade Panel (2010)

Variables	Mean	SD	Min	Max	Count
Height (cm)	138.7703	5.778431	122	154	2457
Female	.4867725	.4999267	0	1	2457
Private Tutoring Use	.8974042	.3034928	0	1	2427
Korean Score (Out of 100)	84.63655	14.03736	8	100	2457
Math Score (Out of 100)	72.9243	18.91783	5	100	2457
English Score (Out of 100)	85.93895	14.29054	8	100	2457
Attitude: Korean Class (1(Low) - 5 (High))	3.966517	.8390136	1	5	2449
Attitude: Math Class (1-5)	4.078208	.8971076	1	5	2455
Attitude: English Class (1-5)	4.242041	.9044759	1	5	2450
Weekly Hrs on Private Tutoring (Korean)	1.156406	1.90039	0	28	1803
Weekly Hrs on Private Tutoring (Math)	1.796619	2.337958	0	25	1952
Weekly Hrs on Private Tutoring (English)	2.07051	2.392632	0	24	1943
Class Size	21.44689	4.558582	9	30	2457
Father: Any Post-Secondary Education	.705171	.456061	0	1	2398
Mother: Any Post-Secondary Education	.5980149	.4904004	0	1	2418
Father: Bachelor and Above	.5642202	.495962	0	1	2398
Mother: Bachelor and Above	.4086022	.4916771	0	1	2418
Monthly Family Income	465.6414	280.1757	70	3000	2457
Monthly Cost: Private Tutoring (Korean)	7.615674	7.457424	0	100	823
Monthly Cost: Private Tutoring (Math)	11.8591	10.01458	0	150	1291
Monthly Cost: Private Tutoring (English)	19.1985	15.4562	0	250	1672

Monthly family income and monthly costs of private tutoring are in 10,000 Korean Won which is approximately $9 USD in 2021.

Table 3.3: Main Result

Variables (Wave 1)	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	-1.572	-1.460	-4.057	-2.291	-4.875*
	(3.531)	(3.204)	(2.907)	(2.611)	(2.812)
$Academic\ Score^2$	0.0278	0.0278	0.0581	0.0322	0.0658*
	(0.0456)	(0.0420)	(0.0389)	(0.0360)	(0.0386)
$Academic\ Score^3$	-0.000116	-0.000123	-0.000231	-0.000104	-0.000255
	(0.000192)	(0.000180)	(0.000170)	(0.000161)	(0.000173)
Academic Rank	-0.0565	-0.0443	-0.0424	-0.0874	-0.0435
	(0.0385)	(0.0369)	(0.0486)	(0.0585)	(0.0650)
Height	0.251**	0.284**	0.472***	0.493**	0.505**
	(0.108)	(0.126)	(0.176)	(0.205)	(0.200)
Height Rank (By Gender)	-0.0308	-0.0236	-0.0548	-0.0919**	-0.0562
	(0.0237)	(0.0266)	(0.0361)	(0.0423)	(0.0410)
Height Rank (By Gender) × Female	0.0106	-0.000994	-0.00118	0.0287	-0.00709
	(0.0183)	(0.0198)	(0.0272)	(0.0308)	(0.0326)
Income	-0.000480	0.000379	0.000247	-0.000190	0.00150
	(0.00143)	(0.00172)	(0.00224)	(0.00316)	(0.00294)
Income Rank	0.0846***	0.0778***	0.116***	0.118***	0.111***
	(0.0172)	(0.0199)	(0.0255)	(0.0296)	(0.0308)
Income Rank × Female	0.00910	0.00696	0.000745	0.00138	-0.00889
	(0.0203)	(0.0223)	(0.0293)	(0.0316)	(0.0343)
Female	0.814	3.480**	5.157**	5.232**	8.978***
	(1.516)	(1.513)	(2.096)	(2.315)	(2.318)
Constant	54.24	38.25	77.38	37.61	99.42
	(92.68)	(81.22)	(76.23)	(67.47)	(70.58)
Observations	2,457	2,457	2,457	2,457	2,457
R-squared	0.303	0.292	0.294	0.283	0.281

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1.

Table 3.4: Main Result (Gender-Specific Peers)

Variables (Wave 1)	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	-1.684	-0.405	-1.837	-0.354	-4.026
	(3.736)	(3.245)	(3.050)	(3.047)	(3.277)
Academic Score^2	0.0289	0.0137	0.0292	0.00625	0.0557
	(0.0485)	(0.0430)	(0.0417)	(0.0423)	(0.0453)
Academic Score^3	-0.000119	-6.07e-05	-0.000110	9.62e-06	-0.000215
	(0.000205)	(0.000186)	(0.000185)	(0.000190)	(0.000204)
Academic Rank	-0.0609	-0.0430	-0.0245	-0.0876	-0.0452
	(0.0446)	(0.0453)	(0.0600)	(0.0676)	(0.0768)
Height	0.243*	0.251	0.671***	0.754***	0.988***
	(0.143)	(0.157)	(0.217)	(0.270)	(0.272)
Height Rank (By Gender)	-0.0279	-0.0195	-0.0913**	-0.141***	-0.138**
	(0.0278)	(0.0311)	(0.0422)	(0.0523)	(0.0535)
Height Rank (By Gender) × Female	0.00744	0.000974	-0.00237	0.0350	-0.0116
	(0.0198)	(0.0214)	(0.0288)	(0.0316)	(0.0344)
Income	0.000729	0.00125	0.00183	0.000918	0.00253
	(0.00177)	(0.00207)	(0.00261)	(0.00355)	(0.00336)
Income Rank (By Gender)	0.0751***	0.0719***	0.110***	0.116***	0.104***
	(0.0189)	(0.0227)	(0.0274)	(0.0321)	(0.0333)
Income Rank (By Gender) × Female	0.00598	0.00310	-0.00627	-0.0113	-0.0108
	(0.0199)	(0.0232)	(0.0295)	(0.0321)	(0.0351)
Female	3.553*	7.465***	-2.310	-8.933***	-1.843
	(2.125)	(2.107)	(2.662)	(2.924)	(3.045)
Constant	55.58	19.29	-7.014	-46.25	7.917
	(97.69)	(83.40)	(82.04)	(81.51)	(86.61)
Observations	2,457	2,457	2,457	2,457	2,457
R-squared	0.346	0.349	0.353	0.335	0.346

Robust standard errors are in parenthesis

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Standard errors are clustered at the classroom-gender leve from Wave 1.

Table 3.5: Main Result (School-Level Controls)

Variables (Wave 1)	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	-0.711	-0.395	-2.200	-0.838	-3.246
	(4.063)	(3.743)	(3.314)	(2.462)	(2.595)
Academic Score^2	0.0164	0.0138	0.0335	0.0125	0.0441
	(0.0521)	(0.0487)	(0.0441)	(0.0339)	(0.0351)
Academic Score^3	-7.16e-05	-6.82e-05	-0.000132	-2.39e-05	-0.000163
	(0.000219)	(0.000208)	(0.000191)	(0.000151)	(0.000155)
Academic Rank	-0.0314	-0.00788	0.00419	-0.0463	-0.0263
	(0.0348)	(0.0282)	(0.0387)	(0.0485)	(0.0523)
Height	0.232**	0.310***	0.452***	0.420**	0.429**
	(0.102)	(0.111)	(0.153)	(0.174)	(0.187)
Height Rank (By Gender)	-0.0257	-0.0274	-0.0502	-0.0775**	-0.0433
	(0.0212)	(0.0240)	(0.0347)	(0.0382)	(0.0398)
Height Rank (By Gender) × Female	0.00782	-0.00308	-0.00394	0.0275	-0.00625
	(0.0182)	(0.0194)	(0.0275)	(0.0312)	(0.0343)
Income	0.000218	0.000954	0.000961	-7.75e-05	0.000603
	(0.00145)	(0.00165)	(0.00224)	(0.00291)	(0.00282)
Income Rank	0.0780***	0.0732***	0.106***	0.117***	0.119***
	(0.0173)	(0.0189)	(0.0237)	(0.0293)	(0.0299)
Income Rank × Female	0.0108	0.00756	0.00953	-0.000546	-0.0138
	(0.0189)	(0.0207)	(0.0278)	(0.0307)	(0.0330)
Female	0.802	3.457**	4.752**	5.525**	9.340***
	(1.540)	(1.494)	(2.136)	(2.384)	(2.390)
Constant	37.60	9.886	37.52	15.80	69.29
	(107.7)	(96.29)	(85.36)	(64.47)	(68.07)
Observations	2,457	2,457	2,457	2,457	2,457
R-squared	0.271	0.268	0.267	0.255	0.257

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1.

Table 3.6: Ordered Probit Regression: Dependent - Self-Ability

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	0.134	-0.0100	0.0703	-0.0806	-0.0407	-0.133
	(0.211)	(0.170)	(0.237)	(0.193)	(0.164)	(0.160)
Academic Score^2	-0.00197	-8.52e-05	-0.00108	0.000887	0.000521	0.00133
	(0.00280)	(0.00226)	(0.00316)	(0.00262)	(0.00226)	(0.00220)
Academic Score^3	1.01e-05	1.91e-06	5.30e-06	-2.28e-06	-1.88e-06	-3.05e-06
	(1.20e-05)	(9.73e-06)	(1.37e-05)	(1.14e-05)	(1.01e-05)	(9.91e-06)
Academic Rank	-0.00215	-0.000944	0.00361	-0.00243	0.00147	-0.00385
	(0.00320)	(0.00323)	(0.00338)	(0.00338)	(0.00326)	(0.00304)
Height	-0.00260	-0.00244	0.00489	0.00151	0.00569	0.00561
	(0.0105)	(0.00969)	(0.00970)	(0.00990)	(0.00989)	(0.00952)
Height Rank (By Gender)	-0.00105	0.000697	-2.68e-05	-1.10e-05	0.00105	0.000327
	(0.00207)	(0.00188)	(0.00188)	(0.00196)	(0.00204)	(0.00197)
Height Rank (By Gender) × Female	0.000746	-7.82e-05	-0.00177	0.00134	-0.000758	0.000218
	(0.00166)	(0.00153)	(0.00159)	(0.00152)	(0.00163)	(0.00154)
Income	-2.44e-05	5.27e-05	0.000107	9.17e-05	0.000245*	0.000254*
	(0.000159)	(0.000133)	(0.000148)	(0.000158)	(0.000136)	(0.000135)
Income Rank	0.00335**	0.00347**	0.00183	0.00313**	0.000699	0.000844
	(0.00166)	(0.00141)	(0.00140)	(0.00141)	(0.00128)	(0.00137)
Income Rank × Female	-0.00168	-0.00198	-0.000455	-0.00308*	0.00127	-0.000359
	(0.00176)	(0.00168)	(0.00173)	(0.00167)	(0.00166)	(0.00158)
Female	0.102	0.226*	0.0465	-0.0324	-0.118	-0.0646
	(0.133)	(0.126)	(0.131)	(0.122)	(0.125)	(0.117)
Observations	2,424	2,436	2,449	2,447	2,440	2,451

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Students were asked to what degree they agreed with the statement "I believe I'm an able person" from a scale of 1 (low) to 5 (high).

Table 3.7: Ordered Probit Regression: Dependent - Self-Value

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	-0.0442	0.0956	-0.00597	0.00765	0.0955	-0.139
	(0.220)	(0.181)	(0.224)	(0.199)	(0.171)	(0.176)
Academic Score^2	0.000323	-0.00149	-1.75e-06	-0.000189	-0.00124	0.00156
	(0.00294)	(0.00244)	(0.00300)	(0.00268)	(0.00234)	(0.00241)
Academic Score^3	4.68e-07	7.89e-06	2.00e-07	1.90e-06	5.19e-06	-4.60e-06
	(1.29e-05)	(1.06e-05)	(1.31e-05)	(1.18e-05)	(1.05e-05)	(1.08e-05)
Academic Rank	-0.00167	0.00142	0.00510	-0.000917	0.00416	-0.00273
	(0.00300)	(0.00334)	(0.00355)	(0.00345)	(0.00334)	(0.00347)
Height	-0.0202*	0.00183	0.00372	-0.00243	0.0134	0.0115
	(0.0106)	(0.0104)	(0.00980)	(0.0109)	(0.0104)	(0.0102)
Height Rank (By Gender)	0.00343*	0.000893	-0.000234	0.00119	9.89e-05	0.000135
	(0.00198)	(0.00207)	(0.00195)	(0.00207)	(0.00208)	(0.00206)
Height Rank (By Gender) × Female	0.000244	-0.00115	-0.00126	0.00241	-0.000956	0.000335
	(0.00156)	(0.00165)	(0.00162)	(0.00158)	(0.00162)	(0.00163)
Income	2.27e-05	0.000149	0.000155	2.47e-05	0.000268*	0.000225
	(0.000165)	(0.000145)	(0.000138)	(0.000167)	(0.000145)	(0.000137)
Income Rank	0.00428***	0.00305**	0.00158	0.00473***	0.000299	0.000833
	(0.00163)	(0.00150)	(0.00145)	(0.00146)	(0.00135)	(0.00145)
Income Rank × Female	-0.00402**	-0.00158	-0.00116	-0.00481***	-1.43e-05	-0.00161
	(0.00170)	(0.00159)	(0.00159)	(0.00176)	(0.00165)	(0.00159)
Female	0.266**	0.337***	0.131	0.0368	0.0492	0.0665
	(0.125)	(0.126)	(0.126)	(0.125)	(0.127)	(0.126)
Observations	2,426	2,432	2,446	2,447	2,441	2,450

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Students were asked to what degree they agreed with the statement "I am a valuable person" from a scale of 1 (low) to 5 (high).

Table 3.8: Ordered Probit Regression: Dependent - Parental Interests

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	0.0930	-0.112	0.371**	0.255	0.0818	0.150
	(0.187)	(0.187)	(0.175)	(0.182)	(0.182)	(0.191)
Academic Score^2	-0.00146	0.00137	-0.00520**	-0.00318	-0.00111	-0.00205
	(0.00258)	(0.00256)	(0.00244)	(0.00251)	(0.00250)	(0.00257)
Academic Score^3	8.28e-06	-4.53e-06	2.40e-05**	1.36e-05	5.21e-06	9.06e-06
	(1.14e-05)	(1.14e-05)	(1.11e-05)	(1.12e-05)	(1.12e-05)	(1.12e-05)
Academic Rank	-0.00434	-0.00292	0.00135	0.000113	0.000926	0.00247
	(0.00449)	(0.00363)	(0.00346)	(0.00348)	(0.00354)	(0.00360)
Height	0.00413	-0.000816	0.0132	0.00174	0.00146	0.00388
	(0.0130)	(0.0111)	(0.0108)	(0.0110)	(0.0114)	(0.0116)
Height Rank (By Gender)	4.46e-05	0.000903	0.000284	0.00120	0.000580	-0.000689
	(0.00277)	(0.00219)	(0.00211)	(0.00225)	(0.00232)	(0.00230)
Height Rank (By Gender) × Female	-0.000783	0.000325	-0.00292*	2.93e-06	-1.66e-05	0.000780
	(0.00185)	(0.00173)	(0.00170)	(0.00173)	(0.00184)	(0.00170)
Income	0.000232	-2.14e-05	-0.000149	1.16e-05	0.000240	0.000113
	(0.000155)	(0.000151)	(0.000141)	(0.000141)	(0.000198)	(0.000170)
Income Rank	0.00202	0.00389**	0.00580***	0.00542***	0.00261	0.00485***
	(0.00160)	(0.00155)	(0.00160)	(0.00138)	(0.00164)	(0.00149)
Income Rank × Female	-0.00212	-0.00135	-0.00446**	-0.00360**	-0.00232	-0.00414**
	(0.00178)	(0.00176)	(0.00189)	(0.00174)	(0.00154)	(0.00162)
Female	0.124	0.0338	0.336***	0.117	0.136	0.220*
	(0.117)	(0.130)	(0.117)	(0.110)	(0.113)	(0.117)
Observations	2,236	2,236	2,236	2,236	2,236	2,236

Robust standard errors are in parenthesis

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Standard errors are clustered at the classroom level from Wave 1. Students were asked to what degree they agreed with the statement "My parents are interested in how I'm doing in school" from a scale of 1 (low) to 5 (high).

Table 3.9: Ordered Probit Regression: Dependent - Parental Involvement

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	0.0914	0.0286	0.403**	0.0790	0.142	0.186
	(0.230)	(0.204)	(0.165)	(0.199)	(0.189)	(0.169)
Academic Score^2	-0.000880	-0.000389	-0.00558**	-0.000953	-0.00199	-0.00275
	(0.00306)	(0.00272)	(0.00230)	(0.00273)	(0.00255)	(0.00228)
Academic Score^3	2.74e-06	2.14e-06	2.49e-05**	4.35e-06	9.71e-06	1.29e-05
	(1.33e-05)	(1.19e-05)	(1.04e-05)	(1.22e-05)	(1.11e-05)	(1.00e-05)
Academic Rank	0.000124	-0.000116	0.00440	-0.000379	-0.00114	0.00261
	(0.00336)	(0.00359)	(0.00334)	(0.00315)	(0.00390)	(0.00315)
Height	-0.0106	0.00712	0.0202*	-0.00208	-0.0101	-0.00450
	(0.0108)	(0.0103)	(0.0107)	(0.0105)	(0.0108)	(0.00896)
Height Rank (By Gender)	0.000666	-0.00290	-0.00312	0.000694	0.00200	0.00131
	(0.00238)	(0.00207)	(0.00227)	(0.00203)	(0.00212)	(0.00187)
Height Rank (By Gender) × Female	-0.000933	0.000735	-0.00317*	-0.00138	-0.00101	-0.00180
	(0.00190)	(0.00163)	(0.00174)	(0.00173)	(0.00165)	(0.00173)
Income	-0.000211	0.000119	0.000124	0.000138	0.000180	0.000185
	(0.000147)	(0.000156)	(0.000144)	(0.000155)	(0.000192)	(0.000154)
Income Rank	0.00330**	0.00270*	0.00336**	0.00466***	0.00315**	0.00374***
	(0.00154)	(0.00157)	(0.00144)	(0.00156)	(0.00156)	(0.00129)
Income Rank × Female	-0.000813	-0.000644	-0.000836	-0.00157	-0.00130	-0.00165
	(0.00172)	(0.00170)	(0.00189)	(0.00181)	(0.00170)	(0.00156)
Female	0.0413	-0.0532	0.0267	0.0503	0.0624	0.116
	(0.127)	(0.120)	(0.124)	(0.130)	(0.117)	(0.113)
Observations	2,253	2,253	2,253	2,253	2,253	2,253

Robust standard errors are in parenthesis

*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Students were asked to what degree they agreed with the statement "My parents help me with my studies" from a scale of 1 (low) to 5 (high).

Table 3.10: Ordered Probit Regression: Dependent - Educational Spending

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	0.0290	-0.235	-0.118	0.115	-0.139	-0.197
	(0.158)	(0.150)	(0.154)	(0.143)	(0.191)	(0.153)
Academic Score^2	-0.000213	0.00313	0.00139	-0.00147	0.00200	0.00260
	(0.00216)	(0.00202)	(0.00208)	(0.00197)	(0.00259)	(0.00206)
Academic Score^3	1.41e-06	-1.31e-05	-4.52e-06	6.73e-06	-9.23e-06	-1.08e-05
	(9.50e-06)	(8.91e-06)	(9.20e-06)	(8.83e-06)	(1.15e-05)	(8.99e-06)
Academic Rank	-0.00692**	-0.00460	-0.00468	-0.00311	-0.000727	-0.000959
	(0.00334)	(0.00294)	(0.00303)	(0.00321)	(0.00324)	(0.00303)
Height	0.0175*	0.0104	0.00738	0.00817	0.00888	0.0139
	(0.00989)	(0.00968)	(0.00993)	(0.00937)	(0.0100)	(0.00937)
Height Rank (By Gender)	-0.00230	-0.000972	0.000647	-0.00128	-0.000104	-0.00209
	(0.00207)	(0.00208)	(0.00213)	(0.00214)	(0.00199)	(0.00215)
Height Rank (By Gender) × Female	0.000145	0.00196	-0.000923	-0.000395	-0.000410	0.00162
	(0.00152)	(0.00150)	(0.00152)	(0.00167)	(0.00150)	(0.00163)
Income	0.000332**	0.000274	0.000171	0.000420**	0.000242	0.000368**
	(0.000167)	(0.000187)	(0.000196)	(0.000194)	(0.000172)	(0.000159)
Income Rank	0.00772***	0.00877***	0.00763***	0.00606***	0.00567***	0.00574***
	(0.00149)	(0.00169)	(0.00160)	(0.00171)	(0.00172)	(0.00157)
Income Rank × Female	-0.00245	-0.00310*	0.000646	-0.000700	-0.00190	-0.00296*
	(0.00172)	(0.00166)	(0.00177)	(0.00169)	(0.00175)	(0.00168)
Female	0.0620	0.0485	0.0410	0.0582	0.141	0.0831
	(0.112)	(0.114)	(0.113)	(0.124)	(0.114)	(0.118)
Observations	2,418	2,431	2,427	2,399	2,380	2,408

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Parents were asked to what degree they agreed with the statement "I do not hold back on educational spending for my children" from a scale of 1 (low) to 5 (high).

Table 3.11: Ordered Probit Regression: Dependent - Manage Daily Schedule

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	0.377*	0.104	0.0925	0.495***	0.205	0.391**
	(0.199)	(0.143)	(0.203)	(0.187)	(0.189)	(0.152)
Academic Score^2	-0.00473*	-0.00132	-0.00135	-0.00646***	-0.00275	-0.00547***
	(0.00270)	(0.00198)	(0.00273)	(0.00249)	(0.00253)	(0.00212)
Academic Score^3	1.98e-05*	6.19e-06	7.09e-06	2.82e-05***	1.22e-05	2.53e-05***
	(1.19e-05)	(8.92e-06)	(1.19e-05)	(1.08e-05)	(1.10e-05)	(9.57e-06)
Academic Rank	0.000564	-0.00233	-0.00180	0.000145	0.00167	-0.000937
	(0.00344)	(0.00364)	(0.00336)	(0.00375)	(0.00344)	(0.00346)
Height	0.00848	0.00379	0.0124	-0.00352	0.00820	0.0115
	(0.0114)	(0.0107)	(0.0103)	(0.00932)	(0.00855)	(0.00932)
Height Rank (By Gender)	-0.00143	0.000861	-0.000778	0.00150	0.000138	-0.000982
	(0.00222)	(0.00223)	(0.00213)	(0.00188)	(0.00188)	(0.00201)
Height Rank (By Gender) × Female	0.00120	-0.000253	-0.000949	-1.31e-05	-0.000940	-0.00176
	(0.00171)	(0.00175)	(0.00163)	(0.00166)	(0.00165)	(0.00160)
Income	-0.000180	-3.36e-05	0.000111	0.000235	-0.000120	-0.000214
	(0.000116)	(0.000152)	(0.000127)	(0.000167)	(0.000138)	(0.000156)
Income Rank	0.00775***	0.00416***	0.00490***	0.00481***	0.00631***	0.00639***
	(0.00136)	(0.00152)	(0.00145)	(0.00150)	(0.00148)	(0.00153)
Income Rank × Female	-5.67e-06	0.00211	0.000813	0.000646	-0.000308	-0.000452
	(0.00157)	(0.00179)	(0.00159)	(0.00177)	(0.00160)	(0.00156)
Female	0.0156	-0.0813	-0.0452	-0.113	0.0636	0.130
	(0.120)	(0.121)	(0.115)	(0.119)	(0.111)	(0.106)
Observations	2,422	2,433	2,427	2,402	2,380	2,406

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Parents were asked to what degree they agreed with the statement "I manage my children's daily schedule" from a scale of 1 (low) to 5 (high).

Table 3.12: Ordered Probit Regression: Dependent - Encourage Studying

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	0.0497	0.0333	-0.171	0.273	0.0338	0.343**
	(0.171)	(0.202)	(0.196)	(0.170)	(0.199)	(0.143)
Academic Score^2	-0.000606	-0.000361	0.00223	-0.00326	-0.000392	-0.00468**
	(0.00238)	(0.00273)	(0.00267)	(0.00233)	(0.00267)	(0.00197)
Academic Score^3	4.05e-06	2.11e-06	-7.72e-06	1.31e-05	2.97e-06	2.21e-05**
	(1.07e-05)	(1.19e-05)	(1.18e-05)	(1.03e-05)	(1.18e-05)	(8.93e-06)
Academic Rank	-0.00753**	-0.00229	-0.00905**	0.00153	-0.00770**	-0.00476
	(0.00372)	(0.00343)	(0.00359)	(0.00374)	(0.00346)	(0.00348)
Height	0.00396	0.00182	0.0147	0.00436	0.00965	0.0108
	(0.0102)	(0.0112)	(0.0106)	(0.0103)	(0.0112)	(0.0108)
Height Rank (By Gender)	0.00197	0.00107	-0.00119	0.000776	-0.00126	-0.000949
	(0.00206)	(0.00213)	(0.00228)	(0.00198)	(0.00219)	(0.00220)
Height Rank (By Gender) × Female	-0.00182	-0.00167	-0.000330	-0.000886	0.000971	-0.000485
	(0.00175)	(0.00171)	(0.00161)	(0.00176)	(0.00163)	(0.00181)
Income	-3.12e-06	1.36e-05	0.000203	0.000438**	0.000129	3.41e-05
	(0.000135)	(0.000137)	(0.000150)	(0.000206)	(0.000163)	(0.000130)
Income Rank	0.00718***	0.00612***	0.00330**	0.00396**	0.00646***	0.00666***
	(0.00157)	(0.00165)	(0.00161)	(0.00184)	(0.00166)	(0.00143)
Income Rank × Female	-0.00308*	-0.00383**	-0.000342	-0.00355*	-0.00333*	-0.00234
	(0.00168)	(0.00166)	(0.00172)	(0.00190)	(0.00173)	(0.00175)
Female	0.327***	0.288**	0.0178	0.232*	0.234*	0.187
	(0.121)	(0.124)	(0.130)	(0.129)	(0.124)	(0.120)
Observations	2,425	2,432	2,434	2,406	2,382	2,405

Robust standard errors are in parenthesis

*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Parents were asked to what degree they agreed with the statement "I encourage my children to study hard" from a scale of 1 (low) to 5 (high).

Table 3.13: Ordered Probit Regression: Dependent - Collect Information About Private Tutoring

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	0.0835	0.140	-0.0236	-0.0838	0.124	0.0308
	(0.196)	(0.154)	(0.153)	(0.157)	(0.167)	(0.133)
Academic Score^2	-0.00106	-0.00197	6.91e-05	0.000869	-0.00199	-0.000748
	(0.00265)	(0.00214)	(0.00210)	(0.00215)	(0.00229)	(0.00186)
Academic Score^3	5.42e-06	9.47e-06	1.74e-06	-1.97e-06	1.07e-05	4.88e-06
	(1.15e-05)	(9.61e-06)	(9.41e-06)	(9.52e-06)	(1.02e-05)	(8.50e-06)
Academic Rank	-0.00706**	-0.00280	-0.00551*	-0.00491	-0.00458	-0.000446
	(0.00342)	(0.00310)	(0.00318)	(0.00316)	(0.00315)	(0.00290)
Height	0.0106	-0.000352	-0.000114	0.00683	-0.00923	0.00661
	(0.00976)	(0.0102)	(0.0112)	(0.00987)	(0.0103)	(0.0116)
Height Rank (By Gender)	-0.00202	0.000528	0.000363	-0.000967	0.00282	-0.000153
	(0.00205)	(0.00226)	(0.00225)	(0.00195)	(0.00210)	(0.00221)
Height Rank (By Gender) × Female	-9.60e-07	0.00131	0.000420	0.000876	-0.000573	-0.000275
	(0.00155)	(0.00162)	(0.00160)	(0.00150)	(0.00144)	(0.00159)
Income	-0.000271**	-0.000165	-0.000182	-0.000190	-7.80e-05	-0.000101
	(0.000117)	(0.000191)	(0.000156)	(0.000168)	(0.000180)	(0.000124)
Income Rank	0.00780***	0.00927***	0.00777***	0.00956***	0.00751***	0.00790***
	(0.00145)	(0.00160)	(0.00157)	(0.00154)	(0.00166)	(0.00152)
Income Rank × Female	0.00104	-0.000198	0.000400	-0.00229	-0.00193	-0.000472
	(0.00180)	(0.00170)	(0.00171)	(0.00169)	(0.00172)	(0.00165)
Female	-0.0484	-0.0726	-0.00975	0.0889	0.0814	0.0547
	(0.122)	(0.123)	(0.124)	(0.116)	(0.115)	(0.113)
Observations	2,418	2,435	2,428	2,399	2,377	2,408

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Parents were asked to what degree they agreed with the statement “I collect information about private tutoring for my children” from a scale of 1 (low) to 5 (high).

Table 3.14: Probit Regression: Dependent - Use of Private Tutoring

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	-0.430	0.362	-0.386	-0.218	-0.0184	-0.280
	(0.446)	(0.327)	(0.300)	(0.214)	(0.232)	(0.329)
Academic Score^2	0.00538	-0.00441	0.00564	0.00342	0.000405	0.00371
	(0.00580)	(0.00447)	(0.00408)	(0.00299)	(0.00321)	(0.00441)
Academic Score^3	-2.14e-05	1.76e-05	-2.57e-05	-1.65e-05	-7.88e-07	-1.35e-05
	(2.46e-05)	(1.98e-05)	(1.80e-05)	(1.36e-05)	(1.44e-05)	(1.92e-05)
Academic Rank	-0.00328	0.00531	-0.00210	0.00270	-0.00465	-0.00926*
	(0.00653)	(0.00574)	(0.00529)	(0.00482)	(0.00470)	(0.00485)
Height	0.0190	0.00777	0.00521	0.00936	0.00706	0.00311
	(0.0199)	(0.0167)	(0.0192)	(0.0157)	(0.0155)	(0.0128)
Height Rank (By Gender)	-0.00218	-0.00203	-0.00282	0.00282	-0.00129	-0.00152
	(0.00416)	(0.00330)	(0.00370)	(0.00308)	(0.00313)	(0.00254)
Height Rank (By Gender) × Female	-0.000481	0.000807	0.00139	-0.00412	0.00102	0.00115
	(0.00314)	(0.00296)	(0.00265)	(0.00253)	(0.00222)	(0.00240)
Income	-0.000461*	-0.000230	-0.000140	-0.000102	-8.00e-05	-9.33e-05
	(0.000253)	(0.000238)	(0.000270)	(0.000226)	(0.000284)	(0.000282)
Income Rank	0.00489*	0.00913***	0.00953***	0.00646***	0.00950***	0.00962***
	(0.00277)	(0.00263)	(0.00254)	(0.00212)	(0.00255)	(0.00247)
Income Rank × Female	0.00169	0.000340	0.00162	0.00220	0.000157	-0.00217
	(0.00316)	(0.00306)	(0.00293)	(0.00266)	(0.00245)	(0.00240)
Female	-0.163	0.213	0.355*	0.405**	0.221	0.310*
	(0.212)	(0.225)	(0.191)	(0.168)	(0.166)	(0.176)
Constant	9.491	-9.701	8.384	2.353	-1.025	5.933
	(11.46)	(7.800)	(7.452)	(5.391)	(5.829)	(8.387)
Observations	1,486	1,711	1,788	1,912	2,081	2,015

Robust standard errors are in parenthesis
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Standard errors are clustered at the classroom level from Wave 1. Students were asked to answer "yes" or "no" to whether they use private tutoring. Because usage of private tutoring is prevalent, in earlier regressions we lose a lot of observations due to collinearity with classroom indicators.

Table 3.15: Regression: Cost of Private Tutoring (Korean)

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	1.185	-4.469	1.085	-0.157	0.837	-3.971
	(2.667)	(2.766)	(3.265)	(1.830)	(1.903)	(2.893)
Academic Score^2	-0.0136	0.0662*	-0.0242	-0.00307	-0.0112	0.0565
	(0.0366)	(0.0377)	(0.0476)	(0.0247)	(0.0254)	(0.0415)
Academic Score^3	4.40e-05	-0.000317*	0.000154	4.92e-05	5.08e-05	-0.000246
	(0.000160)	(0.000167)	(0.000236)	(0.000111)	(0.000111)	(0.000188)
Academic Rank	0.0291	-0.0151	-0.0260	-0.105	-0.00966	-0.0808*
	(0.0409)	(0.0677)	(0.0680)	(0.0753)	(0.0292)	(0.0450)
Height	0.0353	-0.00476	0.0187	0.0363	-0.0130	-0.0452
	(0.133)	(0.130)	(0.165)	(0.0901)	(0.0957)	(0.136)
Height Rank (By Gender)	-0.0241	-0.0208	-0.0697	-0.00988	0.00257	0.00440
	(0.0329)	(0.0283)	(0.0655)	(0.0192)	(0.0217)	(0.0214)
Height (By Gender) × Female	0.0168	0.0551*	0.0864*	-0.00570	-0.00377	-0.00654
	(0.0231)	(0.0321)	(0.0518)	(0.0148)	(0.0153)	(0.0236)
Income	0.00298	-0.00191	0.00271	-0.00114	0.000759	-0.00494
	(0.00192)	(0.00250)	(0.00657)	(0.00112)	(0.00131)	(0.00373)
Income Rank	-0.0332	0.0461**	-0.0832	-0.00558	0.0142	0.0713*
	(0.0246)	(0.0225)	(0.124)	(0.0162)	(0.0152)	(0.0402)
Income Rank × Female	0.0318	-0.0335	0.0660	-0.0145	-0.0132	-0.0249
	(0.0242)	(0.0337)	(0.0630)	(0.0290)	(0.0157)	(0.0262)
Female	-2.616	-0.277	-9.518	0.559	-0.884	-1.193
	(2.140)	(1.845)	(7.663)	(2.270)	(1.103)	(1.078)
Constant	-22.67	112.4	-6.395	17.16	-10.93	104.7
	(70.32)	(69.57)	(83.60)	(45.73)	(50.30)	(75.22)
Observations	823	987	1,169	1,829	1,646	1,677
R-squared	0.262	0.204	0.122	0.125	0.170	0.133

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Parents were asked about how much they spend monthly in 10,000 Korean Won units (approximately $ 9 USD). While more than 2100 students respond that they are involved in private tutoring, this analysis suffers from a significant drop in response rates.

Table 3.16: Regression: Cost of Private Tutoring (Math)

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	-0.890	-3.061	4.054	-1.824	1.361	1.675
	(2.147)	(1.967)	(3.614)	(2.950)	(4.903)	(3.651)
Academic Score^2	0.0143	0.0437	-0.0629	0.0198	-0.0263	-0.0263
	(0.0288)	(0.0273)	(0.0514)	(0.0409)	(0.0693)	(0.0518)
Academic Score^3	-7.42e-05	-0.000197	0.000319	-5.93e-05	0.000151	0.000134
	(0.000125)	(0.000122)	(0.000245)	(0.000183)	(0.000317)	(0.000236)
Academic Rank	0.00260	-0.0622	-0.0294	-0.0739	-0.0279	-0.00715
	(0.0350)	(0.0546)	(0.0777)	(0.0853)	(0.0488)	(0.0564)
Height	0.0891	0.0108	0.429	-0.252	-0.244	-0.0283
	(0.115)	(0.125)	(0.264)	(0.184)	(0.200)	(0.215)
Height Rank (By Gender)	-0.0189	-0.0145	-0.125	0.0662*	0.0854*	0.0189
	(0.0244)	(0.0280)	(0.0885)	(0.0362)	(0.0461)	(0.0442)
Height Rank (By Gender) × Female	0.00323	0.0539	0.0221	-0.0315	-0.0379	-0.0197
	(0.0176)	(0.0329)	(0.0566)	(0.0327)	(0.0281)	(0.0342)
Income	0.00322	0.00314	0.00335	0.00846**	-0.000815	-0.00462
	(0.00245)	(0.00241)	(0.00320)	(0.00409)	(0.00271)	(0.00431)
Income Rank	0.0183	0.0416	-0.0429	0.0146	0.122***	0.170***
	(0.0189)	(0.0290)	(0.0713)	(0.0458)	(0.0277)	(0.0465)
Income Rank × Female	0.0116	-0.0346	0.0875	0.00649	-0.0318	-0.0658*
	(0.0191)	(0.0373)	(0.0616)	(0.0478)	(0.0266)	(0.0387)
Female	-1.507	-0.934	-8.988	-1.104	1.217	2.962
	(1.454)	(1.849)	(7.296)	(3.340)	(1.832)	(2.981)
Constant	21.34	88.76*	-105.7	132.8*	47.32	4.773
	(55.88)	(49.27)	(101.8)	(78.86)	(106.2)	(95.22)
Observations	1,291	1,648	1,808	1,952	1,872	1,927
R-squared	0.253	0.187	0.133	0.191	0.177	0.179

Robust standard errors are in parenthesis

*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Parents were asked about how much they spend monthly in 10,000 Korean Won units (approximately $ 9 USD). While more than 2100 students respond that they are involved in private tutoring, this analysis suffers from a significant drop in response rates.

Table 3.17: Regression: Cost of Private Tutoring (English)

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	-3.797	-2.456	-0.171	-3.928	-2.750	-0.736
	(2.380)	(3.301)	(4.294)	(3.391)	(3.258)	(3.648)
Academic Score^2	0.0508	0.0336	4.82e-05	0.0537	0.0351	0.00598
	(0.0325)	(0.0445)	(0.0579)	(0.0463)	(0.0453)	(0.0523)
Academic Score^3	-0.000221	-0.000145	3.22e-05	-0.000236	-0.000146	-5.94e-06
	(0.000140)	(0.000189)	(0.000253)	(0.000203)	(0.000203)	(0.000236)
Academic Rank	-0.00199	-0.0388	-0.113	-0.00914	0.00650	-0.0272
	(0.0669)	(0.0746)	(0.0909)	(0.113)	(0.0453)	(0.0612)
Height	-0.237	0.0456	0.132	-0.213	-0.331*	-0.278
	(0.157)	(0.162)	(0.186)	(0.232)	(0.174)	(0.249)
Height Rank (By Gender)	0.0618	-0.0454	-0.0286	0.0809*	0.0815**	0.00898
	(0.0406)	(0.0400)	(0.0527)	(0.0471)	(0.0375)	(0.0501)
Height Rank (By Gender) × Female	-0.0105	0.0451	-0.00113	-0.0468	-0.0211	0.0212
	(0.0285)	(0.0322)	(0.0364)	(0.0397)	(0.0254)	(0.0343)
Income	0.000460	0.00208	0.000871	0.0119*	-0.00363	-0.00366
	(0.00232)	(0.00213)	(0.00283)	(0.00643)	(0.00235)	(0.00371)
Income Rank	0.0760***	0.0808***	-0.00827	-0.00437	0.0970***	0.137***
	(0.0213)	(0.0308)	(0.0595)	(0.0609)	(0.0259)	(0.0412)
Income Rank × Female	-0.00522	-0.0246	0.100*	-0.0108	0.0120	-0.0488
	(0.0241)	(0.0466)	(0.0546)	(0.0501)	(0.0267)	(0.0342)
Female	-0.134	-0.579	-6.802	1.943	0.200	2.213
	(1.978)	(2.477)	(5.470)	(3.521)	(1.715)	(2.579)
Constant	150.5**	75.72	15.20	155.8*	142.5*	86.11
	(68.04)	(86.74)	(110.8)	(91.85)	(76.22)	(98.72)
Observations	1,672	1,852	1,923	1,951	1,833	1,895
R-squared	0.208	0.172	0.131	0.144	0.169	0.143

Robust standard errors are in parenthesis

*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Parents were asked about how much they spend monthly in 10,000 Korean Won units (approximately $ 9 USD). While more than 2100 students respond that they are involved in private tutoring, this analysis suffers from a significant drop in response rates.

Table 3.18: Ordered Probit Regression: Weekly Hours on Private Tutoring (Korean)

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	0.0980	0.223	0.107	0.0585	0.00973	0.0158
	(0.223)	(0.163)	(0.216)	(0.245)	(0.350)	(0.323)
Academic Score^2	-0.00135	-0.00398*	-0.00165	-0.000872	-0.000627	0.000192
	(0.00300)	(0.00223)	(0.00295)	(0.00329)	(0.00461)	(0.00427)
Academic Score^3	6.42e-06	2.01e-05**	7.46e-06	5.90e-06	5.25e-06	-2.04e-06
	(1.32e-05)	(9.93e-06)	(1.31e-05)	(1.44e-05)	(1.98e-05)	(1.84e-05)
Academic Rank	-0.00252	0.00697*	0.00128	-0.0107***	-0.00322	-0.00414
	(0.00404)	(0.00371)	(0.00394)	(0.00401)	(0.00465)	(0.00497)
Height	-0.0131	-0.00107	0.00808	0.0177	-0.0202	-0.0111
	(0.0132)	(0.0125)	(0.0126)	(0.0137)	(0.0141)	(0.0147)
Height Rank (By Gender)	0.00416	0.00249	-0.00162	-0.00352	0.00514*	0.00228
	(0.00269)	(0.00251)	(0.00266)	(0.00296)	(0.00305)	(0.00317)
Height Rank (By Gender) × Female	-0.00117	-0.00126	-8.16e-05	0.00189	-0.00121	-0.000454
	(0.00215)	(0.00186)	(0.00192)	(0.00211)	(0.00246)	(0.00221)
Income	0.000196	6.48e-05	-8.16e-05	-0.000519***	-0.000222	-0.000433**
	(0.000164)	(0.000171)	(0.000177)	(0.000181)	(0.000219)	(0.000214)
Income Rank	-4.36e-05	-0.00171	0.00177	0.00374*	0.000107	0.00115
	(0.00175)	(0.00171)	(0.00184)	(0.00203)	(0.00236)	(0.00233)
Income Rank × Female	-0.00209	-0.00101	-0.00148	-0.00129	0.00351	0.00382
	(0.00197)	(0.00175)	(0.00230)	(0.00221)	(0.00250)	(0.00238)
Female	0.238	0.0610	0.0657	-0.164	-0.354*	-0.566***
	(0.157)	(0.142)	(0.158)	(0.149)	(0.185)	(0.184)
Observations	1,803	1,925	1,871	1,811	1,735	1,817

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Students were asked how many weekly hours they spend on private tutoring in a certain subject. They were allowed to choose from 0 to 7, where 0 meant none, 1 meant less than 1 hour, 2 meant between 1 to 2 hours and so on. A response of 7 indicated 7 or more weekly hours spent on private tutoring on a certain subject.

Table 3.19: Ordered Probit Regression: Weekly Hours on Private Tutoring (Math)

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	0.0907	0.331	-0.0821	-0.0189	-0.0664	-0.0623
	(0.256)	(0.217)	(0.195)	(0.230)	(0.255)	(0.195)
Academic Score^2	-0.00140	-0.00479*	0.000815	8.12e-05	0.000526	0.000599
	(0.00343)	(0.00287)	(0.00265)	(0.00304)	(0.00344)	(0.00268)
Academic Score^3	6.14e-06	2.23e-05*	-1.54e-06	1.44e-06	-4.69e-07	-1.58e-06
	(1.51e-05)	(1.25e-05)	(1.17e-05)	(1.32e-05)	(1.51e-05)	(1.20e-05)
Academic Rank	0.00381	0.000589	-0.00546*	-0.00616*	-0.000789	0.00197
	(0.00352)	(0.00357)	(0.00331)	(0.00345)	(0.00326)	(0.00329)
Height	-0.00530	-0.000425	0.00875	0.0163	-0.00311	0.0144
	(0.0110)	(0.0109)	(0.0108)	(0.0116)	(0.0119)	(0.0115)
Height Rank (By Gender)	0.00234	0.000947	-0.000443	-0.00198	0.00110	-0.00249
	(0.00210)	(0.00203)	(0.00223)	(0.00230)	(0.00243)	(0.00234)
Height Rank (By Gender) × Female	0.000352	-0.000203	0.000798	0.000397	0.00101	0.000857
	(0.00175)	(0.00169)	(0.00168)	(0.00192)	(0.00181)	(0.00191)
Income	-0.000105	0.000206	4.90e-05	0.000136	0.000162	0.000158
	(0.000204)	(0.000151)	(0.000192)	(0.000155)	(0.000181)	(0.000127)
Income Rank	0.00454**	0.00245	0.00526***	0.00485***	0.00160	0.00173
	(0.00196)	(0.00179)	(0.00169)	(0.00163)	(0.00174)	(0.00166)
Income Rank × Female	-0.00155	-0.000896	-0.00123	-0.000326	0.00174	0.00373**
	(0.00186)	(0.00190)	(0.00179)	(0.00191)	(0.00208)	(0.00187)
Female	0.0694	-0.0303	-0.00923	0.0422	-0.261*	-0.290**
	(0.140)	(0.131)	(0.134)	(0.147)	(0.158)	(0.140)
Observations	1,952	1,989	1,983	1,886	1,798	1,861

Robust standard errors are in parenthesis
*** p<0.01, ** p<0.05, * p<0.1

Standard errors are clustered at the classroom level from Wave 1. Students were asked how many weekly hours they spend on private tutoring in a certain subject. They were allowed to choose from 0 to 7, where 0 meant none, 1 meant less than 1 hour, 2 meant between 1 to 2 hours and so on. A response of 7 indicated 7 or more weekly hours spent on private tutoring on a certain subject.

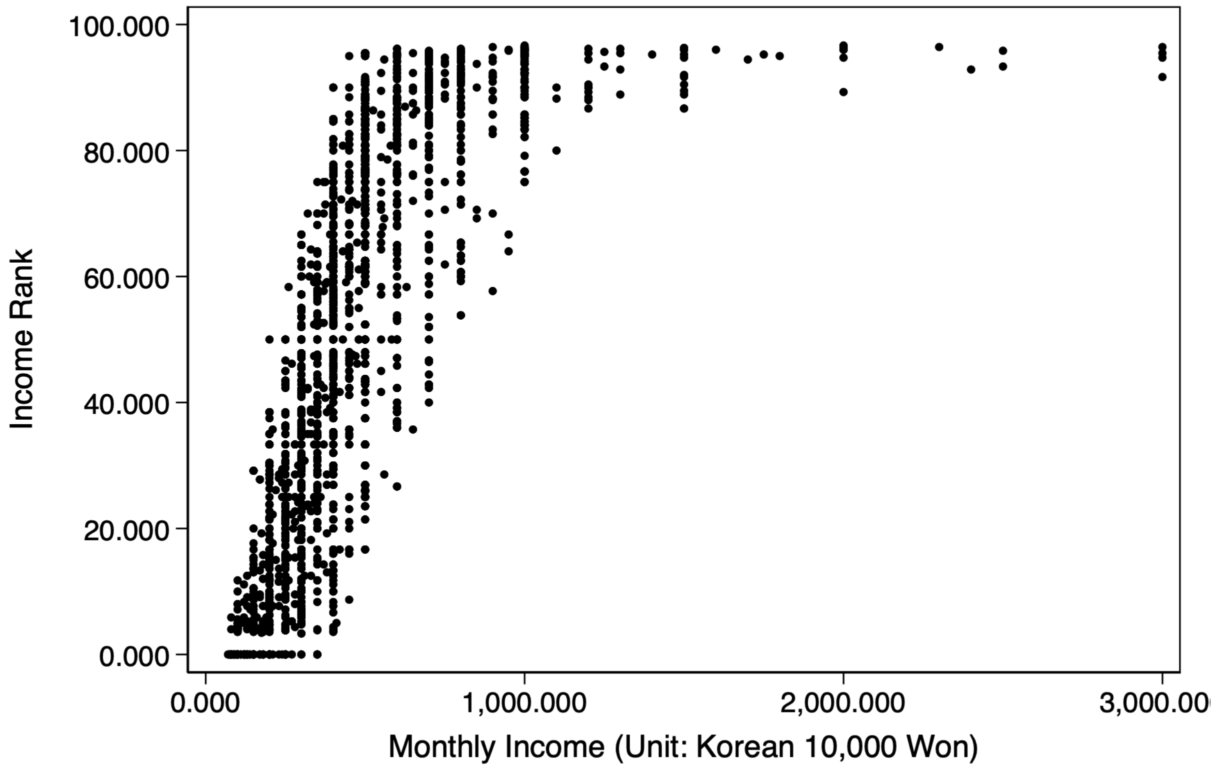
Table 3.20: Ordered Probit Regression: Weekly Hours on Private Tutoring (English)

Variables (Wave 1)	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6
Academic Score	-0.0304	0.304	-0.298	-0.0546	-0.00999	-0.183
	(0.263)	(0.201)	(0.217)	(0.196)	(0.199)	(0.288)
Academic Score^2	0.000546	-0.00443*	0.00373	0.000774	-0.000200	0.00212
	(0.00351)	(0.00265)	(0.00292)	(0.00265)	(0.00262)	(0.00382)
Academic Score^3	-1.76e-06	2.09e-05*	-1.46e-05	-2.37e-06	2.28e-06	-7.21e-06
	(1.52e-05)	(1.14e-05)	(1.28e-05)	(1.18e-05)	(1.13e-05)	(1.64e-05)
Academic Rank	-0.00487	0.000989	-0.00204	-0.00413	0.00158	-0.00194
	(0.00393)	(0.00335)	(0.00363)	(0.00335)	(0.00359)	(0.00362)
Height	0.00353	0.00840	0.01000	0.0247**	0.00409	0.00747
	(0.0108)	(0.00998)	(0.0105)	(0.0103)	(0.0119)	(0.0116)
Height Rank (By Gender)	0.000737	0.000136	0.00115	-0.00372*	-0.00105	-0.00175
	(0.00217)	(0.00206)	(0.00227)	(0.00225)	(0.00250)	(0.00232)
Height Rank (By Gender) × Female	-0.000520	-0.00116	-0.00131	0.000346	-0.000194	-0.000299
	(0.00180)	(0.00166)	(0.00175)	(0.00175)	(0.00193)	(0.00178)
Income	0.000147	-7.30e-05	4.88e-05	-4.39e-05	-2.67e-05	0.000263*
	(0.000190)	(0.000168)	(0.000154)	(0.000169)	(0.000168)	(0.000135)
Income Rank	0.00556***	0.00522***	0.00593***	0.00410**	0.00197	-0.000323
	(0.00181)	(0.00165)	(0.00162)	(0.00162)	(0.00179)	(0.00159)
Income Rank × Female	-0.00142	3.10e-05	-0.000221	0.000790	0.00270	0.00542***
	(0.00162)	(0.00180)	(0.00180)	(0.00181)	(0.00211)	(0.00188)
Female	0.210*	0.0637	0.0975	-0.0411	-0.189	-0.244*
	(0.123)	(0.135)	(0.131)	(0.140)	(0.154)	(0.138)
Observations	1,943	1,993	1,984	1,874	1,782	1,848

Robust standard errors are in parenthesis
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

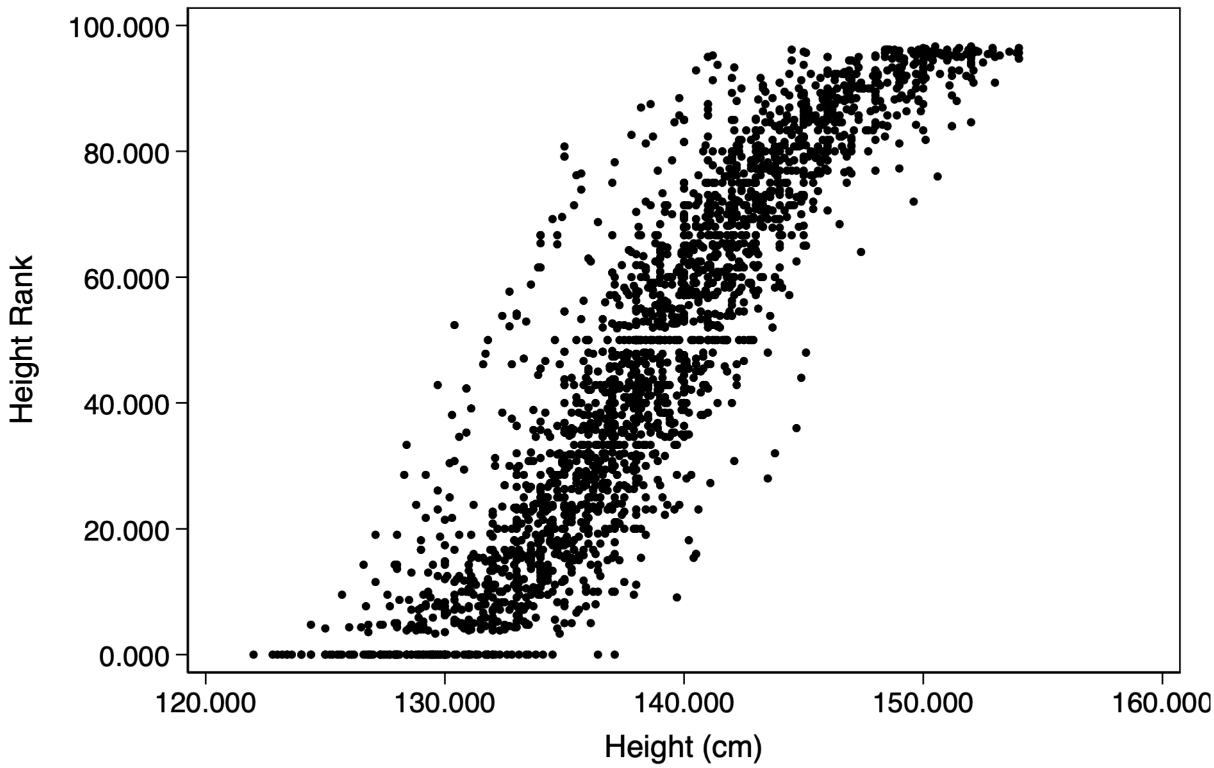
Standard errors are clustered at the classroom level from Wave 1. Students were asked how many weekly hours they spend on private tutoring in a certain subject. They were allowed to choose from 0 to 7, where 0 meant none, 1 meant less than 1 hour, 2 meant between 1 to 2 hours and so on. A response of 7 indicated 7 or more weekly hours spent on private tutoring on a certain subject.

Figure 3.1: Income Rank vs Income Levels



Ordinal income rank plotted against income levels. This figure illustrates how students with the same monthly family income can have different ordinal income ranks based on who their classmates are.

Figure 3.2: Height Rank vs Height Levels



Ordinal height rank plotted against height levels. This figure illustrates how students with the height can have different ordinal height ranks based on who their classmates are.

Figure 3.3: Rank and Distribution of Income (or Height)

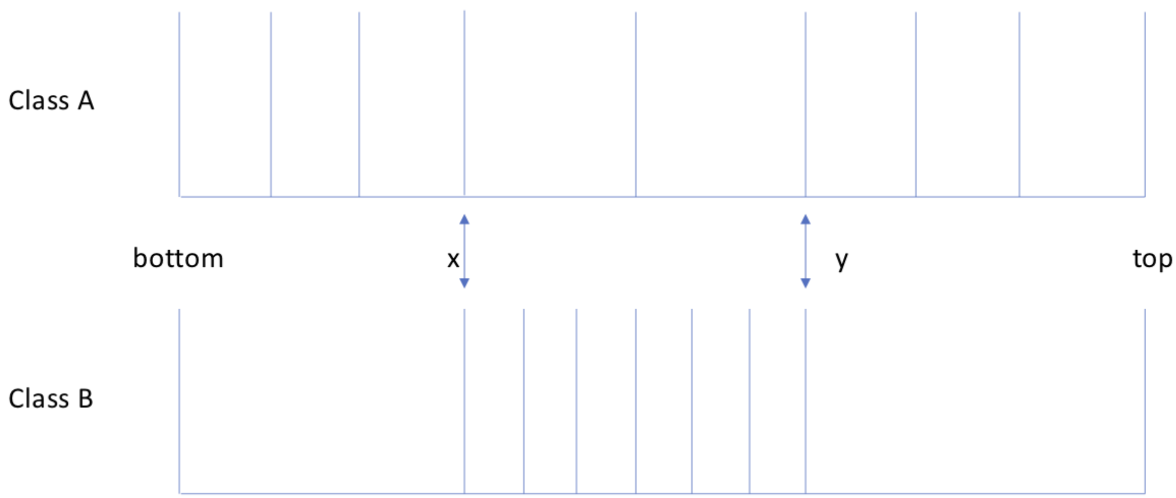


Figure adopted from Murphy and Weinhardt (2018). This figure illustrates how ordinal rank is dependent on the income and height distribution of a classroom. Same levels of income (or height) could have different ordinal ranks as illustrated here (see X and Y, for example).

Figure 3.4: Rank and Distribution of Income (or Height)

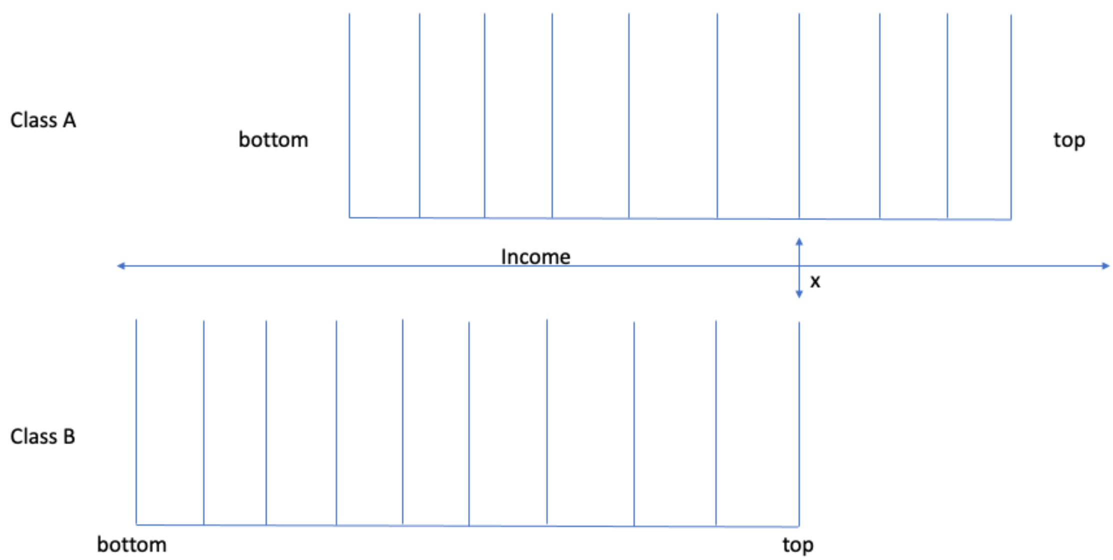
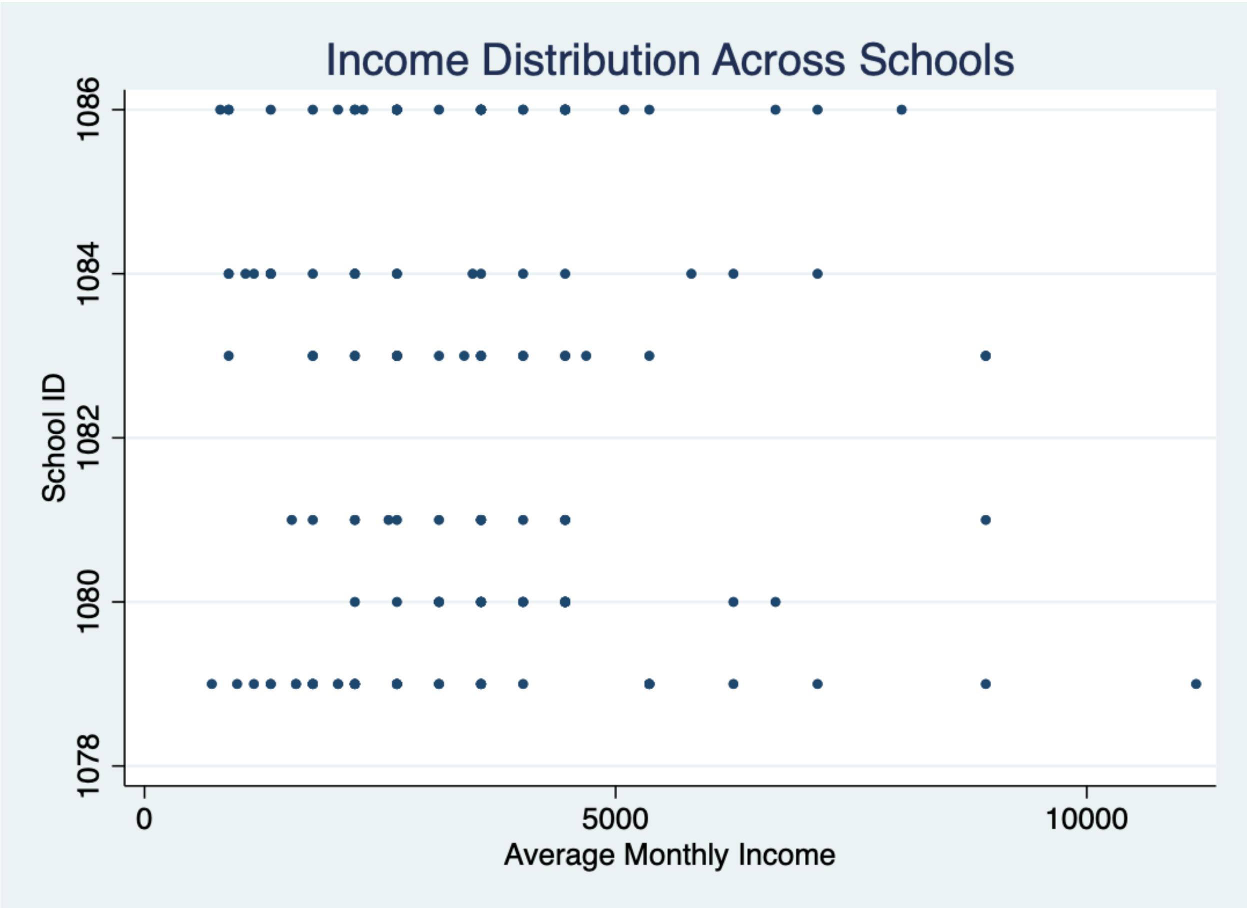


Figure adopted from Elsner Isphording (2017). This figure illustrates how ordinal rank is dependent on the income and height distribution of a classroom. Here we allow for differences in means across classrooms.

Figure 3.5: Income Distribution Across Schools



On

3.7 References

Anderson, C., O. P. John, D. Keltner, and A. M. Kring. Who attains social status? effects of personality and physical attractiveness in social groups. *81*(1), 116–132.

Anderson, C., S. Srivastava, J. S. Beer, S. E. Spataro, and J. A. Chatman. Knowing your place: Self-perceptions of status in face-to-face groups. *91*(6), 1094–1110.

Bramoullé, Y., H. Djebbari, and B. Fortin. Peer effects in networks: A survey. *12*(1), 603–629.

Brown, G. D. A., J. Gardner, A. J. Oswald, and J. Qian. Does wage rank affect employees' well-being? *47*(3), 355–389.

Card, D., A. Mas, E. Moretti, and E. Saez. Inequality at work: The effect of peer salaries on job satisfaction. *102*(6), 2981–3003.

Carrell, S. E. and M. L. Hoekstra. Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *2*(1), 211–228.

Case, A. and C. Paxson. Stature and status: Height, ability, and labor market outcomes. *116*(3), 499–532.

De Giorgi, G., M. Pellizzari, and S. Redaelli. Identification of social interactions through partially overlapping peer groups. *2*(2), 241–275.

Duflo, E. and E. Saez. The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment.

Duncan, G. J., J. Boisjoly, M. Kremer, D. M. Levy, and J. Eccles. Peer effects in drug use and sex among college students. *33*(3), 375–385.

Elsner, B. and I. E. Isphording. A big fish in a small pond: Ability rank and human capital investment. *35*(3), 787–828.

Epple, D. and R. E. Romano. Chapter 20 - peer effects in education: A survey of the theory and evidence. In J. Benhabib, A. Bisin, and M. O. Jackson (Eds.), *Handbook of Social Economics*, Volume 1, pp. 1053–1163.

Euteneuer, F. Subjective social status and health. *27*(5), 337–343.

Falk, A. and A. Ichino. Clean evidence on peer effects. *24*(1), 39–57.

Gibbons, S. and S. Telhaj. Peer effects: Evidence from secondary school transition in england. *78*(4), 548–575.

Goodman, E., B. S. McEwen, L. M. Dolan, T. Schafer-Kalkhoff, and N. E. Adler. Social disadvantage and adolescent stress. *37*(6), 484–492.

Guo, G., Y. Li, H. Wang, T. Cai, and G. J. Duncan. Peer influence, genetic propensity, and binge drinking: A natural experiment and a replication. *121*(3), 914–954.

Harper, B. Beauty, stature and the labour market: A british cohort study. *62*, 771–800.

Hoebel, J., U. E. Maske, H. Zeeb, and T. Lampert. Social inequalities and depressive symptoms in adults: The role of objective and subjective socioeconomic status. *12*(1).

Hoxby, C. Peer effects in the classroom: Learning from gender and race variation.

Imberman, S. A., A. D. Kugler, and B. I. Sacerdote. Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *102*(5), 2048–2082.

Kim, J.-S. and H. Bang. Education fever: Korean parents' aspirations for their children's schooling and future career. *Volume 25*(2).

Ko, C. Essays in applied microeconomics.

Kraus, M. W. and D. Keltner. Signs of socioeconomic status: A thin-slicing approach. *20*(1), 99–106.

Kremer, M. and D. Levy. Peer effects and alcohol use among college students. *22*(3), 189–206.

Kuziemko, I., R. W. Buell, T. Reich, and M. I. Norton. Last-place aversion: Evidence and redistributive implications. *129*(1), 105–149.

Lavy, V., M. D. Paserman, and A. Schlosser. Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *122*(559), 208–237.

Lavy, V. and A. Schlosser. Mechanisms and impacts of gender peer effects at school. *3*(2), 1–33.

Lee, C. J. Korean education fever and private tutoring - ProQuest.

Loh, E. S. The economic effects of physical appearance. *74*(2), 420–438.

Mendolia, S., A. R. Paloyo, and I. Walker. Heterogeneous effects of high school peers on educational outcomes. *70*(3), 613–634.

Murphy, R. and F. Weinhardt. Top of the class: The importance of ordinal rank.

Osafo Hounkpatin, H., A. M. Wood, G. D. A. Brown, and G. Dunn. Why does income relate to depressive symptoms? testing the income rank hypothesis longitudinally. *124*(2), 637–655.

Park, J.-K. 'english fever' in south korea: its history and symptoms.

Rivenbark, J., L. Arseneault, A. Caspi, A. Danese, H. L. Fisher, T. E. Moffitt, L. J. H. Rasmussen, M. A. Russell, and C. L. Odgers. Adolescents' perceptions of family social status correlate with health and life chances: A twin difference longitudinal cohort study. *117*(38), 23323–23328.

Sacerdote, B. Chapter 4 - peer effects in education: How might they work, how big are they and how much do we know thus far? In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education*, Volume 3, pp. 249–277. Elsevier.

Sacerdote, B. Experimental and quasi-experimental analysis of peer effects: Two steps forward? *6*(1), 253–272.

Sacerdote, B. Peer effects with random assignment: Results for dartmouth roommates. *116*(2), 681–704.

SERII. Seoul education longitudinal study 2010.

StatisticsKorea. Private education expenditures survey of elementary, middle and high school students in 2019.

Tincani, M. Heterogeneous peer effects and rank concerns: Theory and evidence.

Wetherall, K., M. Daly, K. Robb, A. Wood, and R. O'Connor. Explaining the income and suicidality relationship: Income rank is more strongly associated with suicidal thoughts and attempts than income. *50*.

Wetherall, K., K. A. Robb, and R. C. O'Connor. Social rank theory of depression: A systematic review of self-perceptions of social rank and their relationship with depressive symptoms and suicide risk. *246*, 300–319.

Zimmerman, D. J. Peer effects in academic outcomes: Evidence from a natural experiment. *85*(1), 9–23.