

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Understanding Risk Aversion in Older Americans: New Approaches Using Genetic Data

Permalink

<https://escholarship.org/uc/item/1fp2k1kp>

Author

Harrati, Amal Cherifa

Publication Date

2014

Peer reviewed|Thesis/dissertation

**Understanding Risk Aversion in Older Americans:
New Approaches Using Genetic Data**

by

Amal Cherifa Harrati

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Demography

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Kenneth Wachter, Chair
Professor Ronald Lee
Professor William Dow

Fall 2014

**Understanding Risk Aversion in Older Americans:
New Approaches Using Genetic Data**

Copyright 2014
by
Amal Cherifa Harrati

Abstract

Understanding Risk Aversion in Older Americans:
New Approaches Using Genetic Data

by

Amal Cherifa Harrati

Doctor of Philosophy in Demography

University of California, Berkeley

Professor Kenneth Wachter, Chair

In this dissertation, I explore the nature and role of risk aversion among older Americans from a variety of perspectives. Risk preferences are important to demographers for several reasons. First, risk preferences are fundamental to most individual-level demographic events, including to financial decision-making, health behaviors, labor market decisions, migration, and marriage and family-formation. Second, there is substantial evidence that risk aversion increases with age. With age also comes increased responsibility in terms of making specific financial and health decisions. In the age of decreasing pensions, older persons must make significant decisions about their financial portfolios and finances in light of pending retirement decisions. In fact, the decision to retire is itself one in which risk plays a role. Third, health behaviors, which are a function of one's riskiness, often display their effects at older ages.

I explore the genetic nature of risk aversion through a number of approaches. Taking advantage of a newly-released database with over two million pieces of genetic variants, I examine the specific genetic nature of risk aversion through two genomic techniques: a Genome-Wide Association study (GWAS) and a Genome-Wide Complex Trait Analysis (GCTA). I provide evidence that risk aversion is a highly complex trait that is a function of a large number of possibly interactive genetic variants. Through the GWAS, I show that the number of genetic variants influencing individual-level differences in risk aversion is numerous and that these variants are likely to be scattered across the genome. The GCTA, while using a separate methodological approach, confirms this finding. I argue that the intricate nature of the genetic underpinnings to risk aversion should be better understood in order to more precisely model economic decisions involving risk preferences.

I also characterize risk aversion from a non-genetic perspective. Using panel data of risk aversion collected over nearly two decades, I use longitudinal methods to explore the extent to which the relationship between hypothetical risk and measurable risky behaviors remain consistent across both time and among individuals. As a follow-up, I examine the specific time period following the 2008 recession to examine any change in the relationship in portfolio allocations relative to stated risk tolerance for individuals after the global financial

crisis. I conclude that the relationship between measured risk and risky behaviors remains relatively constant across the 15 years prior to the global financial crisis. The analysis also shows that the relationship between risk and financial assets does in fact change slightly after the global financial crisis, though the statistical evidence is not very strong.

This dissertation provides a contribution to the understanding of the complex nature of risk aversion and is one of the first to characterise its genetic influences. This research helps to answer questions on the economic, social and biological drivers and consequences of risk aversion among older Americans.

To my daughters, Mayssa Nour Harrati and Nayla Yasmine Harrati.

Make it your duty and pleasure to seek the answers to all your questions.

I love you more than all the stars in the sky.

Contents

Contents	ii
List of Figures	iv
List of Tables	v
1 Introduction	1
2 Data and Methods	4
2.1 The Health and Retirement Study	4
2.2 Phenotype Data	8
2.3 Genotyped Data	11
3 Genome-Wide Association Study	21
3.1 Introduction and Motivation	21
3.2 Background	22
3.3 Data	27
3.4 Empirical Strategy	28
3.5 Results	30
3.6 Discussion	39
4 GCTA	41
4.1 Introduction	41
4.2 Background	42
4.3 Data and Methods	45
4.4 Results	46
4.5 Discussion	48
5 Do Stated Risk Preferences Predict Behavior Over Time?	50
5.1 Introduction	50
5.2 Background	50
5.3 Data and Methods	54
5.4 Results	59

5.5 Discussion	75
6 Conclusion	77
6.1 Summary	77
6.2 Implications	78
6.3 Limitations	79
6.4 Further Study	80
7 References	82

List of Figures

2.1	The X and Y intensities are calculated for each sample as the mean of the sum of the normalized intensities of the two alleles for each probe on those chromosomes. Sample sizes are given in the axis labels. X heterozygosity is the fraction of heterozygous calls out of all non-missing genotype calls on the X chromosome for each sample. Source: University of Washington, 2012	15
2.2	Principal component analysis of 12,507 study subjects with 1230 HapMap controls. Color-coding is according to self-identified race, while symbol denotes ethnicity (Hispanic or not). Axis labels indicate the percentage of variance explained by each eigenvector. Source: University of Washington, 2012	17
2.3	Principal component analysis of 12,419 unrelated study subjects without HapMap controls. Color-coding is according to self-identified race, while symbol denotes ethnicity (Hispanic or not). Axis labels indicate the percentage of variance explained by each eigenvector. Source: University of Washington, 2012	18
3.1	Manhattan "Skyline" Plot. Data source: HRS GWAS data	36
3.2	Power Calculation Against the Alternative Hypothesis	39
5.1	Distribution of Risk Aversion using Cardinal Measure.	58

List of Tables

2.1	Counts of responses by race and ethnicity by survey year. Data source: Author's tabulation of the HRS	9
2.2	Summary Statistics of Risk Preference Question, 1992-2006. Data source: Author's tabulations of the HRS.	10
2.3	Distribution of Risk Aversion in Study Sample. Data source: Author's tabulation from the HRS RAND data set.	20
3.1	Summary Statistics for GWAS Sample	31
3.2	Risk Aversion by Control Sample	33
3.3	Top SNPs for Test Sample	34
3.4	Top SNPs for Full Sample	37
4.1	Results from GCTA Analysis for Risk Aversion, Height, and Cognition	47
5.1	Descriptive Statistics for Study Sample. Data Source: Author's tabulations from the HRS.	55
5.2	Summary Statistics for Cardinal Measure of Risk Aversion. Data Source: Author's tabulation of HRS data.	57
5.3	Tabulations for Risk Aversion Across Different Risky Behaviors, 1992 to 2006. Data source: Author's tabulation of the HRS.	60
5.4	Tabulations of Risk Preferences For Wealth Across HRS Survey Waves: 1992 to 2006. Data source: Author's tabulations of the HRS.	64
5.5	Consistency of Responses Across Survey Waves. Source: Author's tabulation of the HRS.	66
5.6	Associations of Individual Risk Preferences Survey Responses Across Survey Waves	67
5.7	Relationship between Risk and Smoking, Drinking: 1992 to 2010. Data Source: HRS.	70
5.8	Regression Results for Decision to Purchase Life Insurance and Self-Employment: 1992 to 2010. Data source: HRS.	71
5.9	Regression Results for Financial Portfolio Allocation: 1992 to 2010. Data source: HRS.	73

5.10 Changes in Associations to Risk Aversion Before and After the Recession. Data source: HRS. 75

Acknowledgments

I have many great people to thank.

I am very fortunate to have Ken Wachter as the chair of my dissertation. I cannot imagine a better advisor to have along this journey. His guidance and support were indispensable and I can truly say this dissertation would never have come together without him. Ken first introduced me to this subject matter, and over the course of a number of years, taught me about the subject with patience, knowledge and humor. He was a perfect guide through the dissertation process, providing guidance through my progress, a gentle push when I needed it, and giving me time to step away from my work when life demanded it. I am so very lucky to have benefited from and been inspired by his tutelage. If I can be half the mentor and scholar he is, I will consider my career a great success.

I also wish to thank Ron Lee, who has been a very important mentor during my time as a graduate student. Ron was a great support when I was deciding on a dissertation topic, providing great input and encouraging me to explore even my most nascent ideas. He was also played an important role in my job search; he was always so generous with his time and ideas and I thank him deeply for this. I am also extremely grateful to Ron for providing much of my graduate school funding.

Thanks also to Will Dow for being a great outside member and advisor. Will provided important ideas and suggestions that were smart, challenging and pragmatic. His insight and time were invaluable in making this dissertation what it is.

Jenna Johnson-Hanks has been a wonderful mentor in providing an important perspective to my work and career both as a scholar and a mother of young children. Her enthusiasm towards the field of Demography is infectious and I always left my meetings with her feeling brighter, happier and more motivated.

I extend a big thank you to Carl Mason, who was an integral part of setting up the computing system for the genetic data and who answered countless (stupid) computer questions with patience and humor.

I am so grateful to have had as dissertation group partners and close friends Sarah Cowan and Margaret Frye. What began as a dissertation workshop partnership quickly formed into some of my most important friendships. My work and life is so much richer because of them. I am filled with gratitude towards the Demography Community at UC Berkeley. I am also grateful for the support and shared experiences of the rest of my cohort, and particularly Reid Hamel, Carol Thomas, Ellen Smith, and Fiona Willis-Nunez. Thanks also to Alison Gemmill who was my regular "study buddy" and friend through the final lap of the dissertation writing.

Many friends enriched my life during these long years of graduate school. To Jama Adams, whose pride surpassed even my own in my humble accomplishments. She is truly my biggest cheerleader and I am lucky to have her in my life. To Hang Phan, who kept me sane and balanced through the ups and downs of graduate school and in the 15 years prior and 100 years to follow. To Jenny Chu, who patiently and kindly endured long periods of my absence when I was hard at work.

I am forever indebted to my mother, Samia Bendimerad, who provided endless hours of childcare on a weekly basis for many long years. It is said that it takes a village to raise a child, and she nearly single-handedly helped raise not only my two babies, but was a crucial part in helping me grow my dissertation as well. I don't know that I'll ever be able to thank her enough all of her help and support and I love her dearly for everything she has given me in life.

Thanks also so much to the rest of my family. My father, Fouad Bendimerad, provided support and encouragement both as a father and as a fellow scholar. My sister, Soumeya Bendimerad, is not only one of my greatest cheerleaders, but also acted as my masterful editor, often acting on last-minute requests. Finally my "baby" brother, Karim Bendimerad, was anything but in his constant encouragement, checking in, and true curiosity about my work and future.

Finally, a very special thank you to my husband, Jallel Harrati, who has been a steadfast, patient and loving partner through this process. We were married during my first week of graduate school, and through the birth of two kids, countless conference trips, late-nights work sessions, dinners not made and vacations not taken, this experience has been very much a shared journey.

Chapter 1

Introduction

This dissertation explores the nature of risk aversion among older Americans. Risk preferences are important to demographers for several reasons. First, risk preferences are fundamental to most individual-level demographic events, including to financial decision-making, health behaviors, labor market decisions, migration, and marriage and family-formation. Second, there is substantial evidence that risk aversion increases with age. With age also comes increased responsibility in terms of making specific financial and health decisions. In the age of decreasing pensions, older persons must make significant decisions about their financial portfolios and finances in light of pending retirement decisions. In fact, the decision to retire is itself one in which risk plays a role. Third, health behaviors, which are a function of one's riskiness, often display their effects at older ages. For example, the deleterious effects of smoking, drinking, or poor nutritional choices translate into a higher prevalence at older ages of cancer, diabetes, and organ failure. To boot, because the horizon within which one can "correct" for any errors that are made is shorter, the optimality of decisions is that much more important.

While it is evident that studying risk preference in older Americans is important, how to study such a complex phenomenon is less clear. When I began researching risk aversions, I followed the very good advice of my advisor, Ken Wachter. I answered the same questions that measure risk aversion from the Health and Retirement Study (HRS) that I had planned to use for my work. After all, the dissertation, while far from life and death, is a long-term investment that includes a series of decisions and questions that are themselves risky; for example: Will there be any positive findings to my inquiry? Will this particular line of work take longer than expected? Taking my advisor's advice, I decided to start my foray into the topic in the shoes of the respondent rather than the researcher.

The questions from the HRS are themselves fairly straightforward; however, from the perspective of someone who is researching the topic, they are fraught with complexity. When I answered the questions, I couldn't help but think about the different context in which I might have answered the questions. Would I have given the same answer if I were not at the same income level? To what extent were my answers a result of a careful thought process versus a gut reaction? Would my 20-year-old self, or my future 60-year-old self, have

answered the questions in the same manner? This short exercise highlights the myriad ways in which studying individual risk aversions is a tricky business.

Risk aversions are an integral part of most human decision-making, yet they remain poorly understood. Empirical research has examined the important heterogeneity in risk preferences across populations, including differences across gender (Eckel & Grossman, 2002; Powell & Ansic, 1997; Schubert, Brown, Gysler, & Brachinger, 1999); across family background (Hartog et al., 2002); across work characteristics (Praag & Cramer, 2001); across educational attainment levels (Brunello, 2002); and across different contexts of risk (Soane & Chmiel, 2005; Weber, 2002). Given the particular importance of risk aversion in decisions made at older ages, much of the empirical work has focused on the implications for retirement decisions and savings (Bodie, Merton, & Samuelson, 1992; Hurd, 1990; Karatzas, Lehoczky, & Shreve, 1987; Sunden & Surette, 1998). More importantly, the common finding across the literature examining risk aversion, using any one of these measures, is that heterogeneity in risk aversion is real and substantial and furthermore cannot be entirely explained by observed characteristics typically used in empirical models.

This heterogeneity is evident in the ways older Americans vary in their preparation for the financial burdens of retirement and old age. For example, gender, age, and education levels are all predictors of risk aversion but often can explain only a small portion of the variation Americans display in risk aversion (Barksy, 1997). However, while such differences in preparation cannot be attributed entirely to income, education, cognitive ability, family background, or other factors alone, evidence does suggest that differences in risk aversion are an important driver of the variation in wealth accumulation of individuals.

Also central to the study of risk aversion is the debate around the extent to which risk preferences are immutable and fixed. Economists argue that risk preferences are fixed and that any variation in stated risk across time are due to measurement error. Others point to evidence that risk preferences can vary, sometimes drastically, across domains and time. In other words, even though risk preferences have been well-studied empirically, large amounts of heterogeneity that remain unexplained.

A plethora of measurements for risk have been created to accommodate the heterogeneity and the complexity of the research. These measures can be categorized into one of three forms: (1) survey-based assessments, such as respondents' answers to hypothetical lottery gambles; (2) experimental evidence; or (3) inference from observed decision-making in financial, health, or insurance markets. This dissertation relies heavily on a set of hypothetical questions about risk first introduced by Barsky et al. (1997). These questions are elaborated upon in Chapter 2 of this dissertation. A number of recent studies have attempted to validate the Barsky measure of risk preference and other similar instruments and concluded that hypothetical questions track closely, albeit imperfectly, with actual risk-taking behaviors (Dohmen et al., 2005; Falk & Heckman, 2009; Guiso & Paiella, 2005). While the debate remains, hypothetical measures of risk are used in good standing and are useful when behavior is difficult to observe. Falk and Heckman (2009) go even further to argue that objections against experimental or hypothetical measures are "misguided" and that the issue of generalizability is no more a concern than is field data.

Using both hypothetical and behavioral measures of risk, I investigate the gaps in risk preference research among older Americans from a few different perspectives. First, I characterize the genetic nature of risk aversion and using a large dataset with a number of genetic markers referred to as Single-Nucleotide Polymorphisms (SNPs). Second, I use another technique called the Genome-Wide Complex Trait Analysis (GCTA) to explore the heritability of risk aversion from another perspective. Third, I step away from the genetic nature of risk aversion and turn instead to the longitudinal study of measured risk, and in particular, to what happens to investments when traditionally risk-free assets become more risky. A more detailed summary of the dissertation is as follows.

In Chapter 2 I describe the data. In Chapter 3 I report a Genome-Wide Association Study on risk aversion that examines the genetic influences on risk aversion. I use a novel data set from the Health and Retirement Study (HRS) with over 2.5 million pieces of genetic material for each of approximately 10,000 respondents. Specifically, in light of the evidence of heritability from twin studies and candidate gene studies, I explore the polygenicity of risk aversion. I find that risk aversion, like many other socio-behavioral traits, is not driven by a few strong causal genetic variants, but appears instead to be a function of a large number of variants across the genome, each with relatively small effects.

In Chapter Four, I complement the GWAS findings with a GCTA analysis. GCTA estimates the proportion of phenotypic variability within the sample explained by genome-wide SNPs. The analysis in this chapter, thus, is not focused on associations to specific individual SNPs but rather on attempts to explain the total share of heritability in the aggregate effects of all the SNPs. I find additive contributions from heritability estimates near zero, suggesting that the heritability in risk aversion is hidden in SNPs not yet made available based on the chip technology that exists.

In Chapter Five, I move away from the genetics to examine the longitudinal nature of stated risk aversion. Taking advantage of the multiple wave survey design of the HRS, I explore two main ideas. First, I look to see if the cross-sectional relationship between measured risk and risky behaviors in various domains remains consistent over time. Second, I use the change in riskiness to various assets before and after the Great Recession of 2008 to examine to what extent stated risk preference remains consistent. Chapter Six summarizes and critically discusses the main results of the dissertation and critically discusses them. In this final chapter, I also present my future research plan to expand the work presented in this dissertation.

Chapter 2

Data and Methods

In this chapter, I provide information on the data set that I used in the data analysis contained in Chapters 3, 4, and 5 of this dissertation. The dissertation utilizes the Health and Retirement Study (HRS), a well-known longitudinal survey dataset. In addition to the well-utilized survey data, much of the analysis also relies on newly-released genotype data. Both are described below.

2.1 The Health and Retirement Study

The University of Michigan Health and Retirement Study (HRS) is a longitudinal panel study that surveys in two-year study waves a representative sample of more than 26,000 Americans over the age of 50. Supported by the National Institute on Aging (NIA U01AG009740) and the Social Security Administration, the HRS explores the changes in labor force participation and the health transitions that individuals undergo toward the end of their work lives and in the years that follow.

Since its launch in 1992, the study has collected information about income, work, assets, pension plans, health insurance, disability, cognitive functioning, health care expenditures, and physical health functioning. The HRS, which is collected through an in-person survey, has become an important source of multidisciplinary data in addressing important questions about the challenges and opportunities of aging. Since its inception, over 2,500 academic and policy papers have been published using HRS data.

The current sample includes over 26,000 persons in 17,000 households. The HRS is well-known for its high quality measurement of many key socio-economic and labor market outcomes, including wealth, income, and retirement decisions. With each biennial wave of data collection, the HRS incorporates rich experimental modules with detailed assessments for specific topics. This study uses a similar repeated experimental module for risk aversion particularly well-suited for this study.

Design History

The HRS and the Asset and Health Dynamics Among the Oldest Old (AHEAD) studies were created as separate but related surveys. The original HRS study was supported through a cooperative agreement between the National Institutes on Aging (NIA) and the University of Michigan, with additional funding from the Social Security Administration, the Assistant Secretary for Planning and Evaluation (ASPE) in the U.S. Department of Health and Human Services (DHHS), and the Pension and Welfare Benefit Office (see Juster & Suzman, 1995). The HRS was joined in 1993 by the companion study, AHEAD, which studies persons born before 1924 who were aged 70 and over in 1993. It was funded as a supplement to the HRS (see Soldo et al., 1997). In its original conceptualization, the HRS study was designed to follow age-eligible individuals and their spouses as they made the transition from active worker into retirement; the AHEAD study was designed to examine the dynamic interactions between health, family, and economic variables in the post-retirement period at the end of life. The HRS study spanned three waves of data collection: 1992, 1994, and 1996. The AHEAD study included two waves: 1993 and 1995.

The HRS and AHEAD sample designs provided for ‘exit interviews’ with a surviving spouse, child, or other informant concerning medical expenditures and family interactions with the deceased during the final stages of life. Exit interviews were also designed to provide information about the disposition of assets following death.

Both studies obtained detailed information in a number of domains, including demographics, health status, housing, family structure, disability, retirement plans, net worth, income, employment of respondent, work history, current employment, and health and life insurance. In addition, there were several important linkages between HRS and AHEAD survey data and between information from employers and administrative data. HRS supplementary data included administrative data from Social Security earnings and benefits records, National Death Index data, Medicare claims record data, and employer pension data.

In 1998 the HRS and AHEAD studies were merged with respondents from each forming a cohort in a combined interview. At the same time, two new cohorts were added: the Children of the Depression Era (CODA), born between 1924 and 1930, and War Babies (WB), born between 1942 and 1947.

Sample

The target population for the original HRS cohort included all adults in the contiguous United States born between the years 1931 and 1941 who reside in households and included a 2:1 oversample of African-American and Hispanic populations. Following conventional practice for population surveys, institutionalized persons (i.e., those in prisons, jails, nursing homes, and long-term or dependent care facilities) were initially excluded from the survey population. However, individuals were still followed if they moved from the household population into any one of these institutional settings during the survey period. The original

sample was refreshed with new birth cohorts (51 to 56 years of age) every six years and has been expanded over the years to include a broader range of birth cohorts. Again, panel members that moved to an institution or a nursing home during the study were kept in the sample.

The HRS observational unit is an eligible household financial unit. Throughout this document, the term “household” is used interchangeably for convenience with the more precise definition “household financial unit.” The HRS household must include at least one age-eligible member from the 1931 to 1941 birth year cohorts. Age eligible members can be: (1) a single, unmarried age-eligible person; (2) a married couple in which both persons are age-eligible; or (3) a married couple in which only one spouse is age-eligible. If a sample housing unit (HU) contains more than one unrelated age-eligible person, one of these persons is randomly selected as the financial unit to be observed. If an age-eligible person has a spouse, the spouse is automatically selected for HRS even if he or she is not age-eligible.

Since 1998, the objective of the HRS has been to provide information about the U.S. population over age 50 through biennial surveys with samples of that population. Prior to 1998, the target populations were more limited: the original HRS target population was limited to those born between 1931 and 1941, and that of the AHEAD study was limited to those born in 1923 or before. For practical reasons, the decision was made to add new cohorts every six years rather than at each wave of data collection. Therefore, in 1998, the target population was defined as those born in 1947 or before and thus approximately those age 51 and older. Since new cohorts were not added in 2000 or 2002, the target populations were approximately 53 and older in 2000 and 55 and older in 2002. In 2004, a supplementary sample was added to make the total sample representative of those born in 1953 or before, and thus, once again, approximately age 51 and older. In the 2010 wave, the mid-baby boom cohort (born 1954 to 1959) was added, and in 2016 the late baby boom cohort (born 1960 to 1965) is scheduled to be added.

Two of the five samples interviewed by the HRS to date, and a majority of a third sample, came from a screening conducted in 1992 of 69,337 housing units. That sample of housing units was generated using a multi-stage, clustered area probability frame. Of those housing units, 14% (9,419) were determined to be non-sample (i.e., unoccupied or non-households). In all but 214 of the 59,918 identified households, the eligibility of the household members for inclusion in the HRS, AHEAD, or WB samples was determined, for a screening response rate of 99.6%.

At the baseline data collection for the HRS sample in 1992, a total of 15,497 individuals were eligible for interviews. This total included persons identified in the household screening, plus their spouses or partners, regardless of year of birth. Of those identified in this way, interviews were obtained with 12,652 respondents (7,704 households) for an overall response rate of 81.6%.

The second sample was generated for what began as the AHEAD study. This sample consisted of individuals born in 1923 or before. Those born between 1914 and 1923, and about half of those born in 1913 or before, were identified through the 1992 household screening operation. The other half of those born in 1913 or before were identified using the

Medicare enrollment files maintained by the Health Care Financing Administration (HCFA, since renamed the Centers for Medicare, Medicaid Services, or CMS). For the AHEAD sample, interviews were obtained with 8,222 respondents (6,046 different households) with a response rate of 80.4%.

In 1998, the HRS and AHEAD studies were merged through a single interview schedule. At the same time, the third and fourth samples were added. The War Baby (WB) sample consists of those born between 1942 and 1947, inclusive, and was obtained from the same 1992 household screening. The Children of the Depression Age (CODA) sample consists of those born between 1924 and 1930 (the ‘missing’ birth cohorts between the HRS and AHEAD samples). These individuals were identified from the Medicare enrollment file. Since many members of these birth cohorts were already part of the study they were current or former spouses and partners of those in the HRS and AHEAD cohort the new samples excluded those individuals with spouses or partners who were born in 1923 or before, or between 1931 and 1947. The baseline response rates for the CODA and WB samples in 1998 were 72.5% and 70%, respectively.

In 2004, a new sample cohort of individuals born between 1948 and 1953 (age 51 to 56 in 2004) was introduced, which carries forward the steady-state aspect of the HRS. The Early Baby Boomer (EBB) sample was obtained through the screening of 38,385 households. Eligibility was determined in 91.3% of the screened households, and a total of 4,420 individuals in 2,755 households were found to be eligible. Interviews were completed with 3,330 individuals in 2,154 households for individual and household interview response rates of 75.3% and 78.2%, respectively. The EBB sample factored in screening response rate yields overall baseline response rates of 68.7% for individuals and 71.4% for households.

RAND Data

The HRS is a longitudinal household survey dataset deployed for the study of retirement and health among the elderly in the United States. It is extraordinarily rich and complex. With the goal of making the data more accessible to researchers, the RAND Center for the Study of Aging, with funding and support from the National Institute on Aging (NIA) and the Social Security Administration (SSA), created the RAND HRS data files.

The RAND HRS is a user-friendly version of a subset of the HRS. It contains cleaned and processed variables with consistent and intuitive naming conventions, model-based imputations and imputation flags, and spousal counterparts of most individual-level variables.

The data include any individual interviewed at least once. This set includes individuals who were age-eligible (i.e., born in eligible years) at the time of their first interview, spouses that were not age-eligible at baseline, and spouses who married an age-eligible respondent between survey waves. The HRS public release files provide imputations for many asset and income types, but the imputation method is not consistent across all waves. The RAND HRS data contain imputations of all asset and income types using a consistent method for all waves. Beginning with HRS 2006, RAND has provided the income and asset imputations for the HRS. The RAND HRS data file contains summary measures of income and assets.

All analyses in this dissertation that are based on the HRS survey data uses the RAND data files.

2.2 Phenotype Data

The HRS also included an experimental module, first proposed by Barsky et al. (1997), in the first wave of the survey in 1994 and was included in a total of six subsequent data waves for the assessment of risk aversion. The measure is derive from answers to the following questions:

Suppose that you are the only income earner in the family, and have a good job guaranteed to give you your current (family) income every year for life. You are given the opportunity to take a new and equally good job, with a 50-50 chance it will double your (family) income and a 50-50 chance that it will cut your (family) income by a third. Would you take the new job?

If the first job is chosen in the second question again, then: Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by twenty percent. Would you take the first job or the second job?

If first job is chosen in the second question again, then: Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by 10 percent. Would you take the first job or the second job?

If second job is chosen in the first question, then: Suppose the chances were 50-50 that the second job would double your lifetime income, and 50-50 that it would cut it in half. Would you take the first job or the second job?

If second job is chosen in the second question again, then: Suppose the chances were 50-50 that the second job would double your lifetime income and 50-50 that it would cut it by seventy-five percent. Would you take the first job or the second job?

Category 1: Respondent would take a job with even chances of doubling income or cutting it in half.

Category 2: Respondent would take a job with even chances of doubling income or cutting it by a third.

Category 3: Respondent would take a job with even chances of doubling income or cutting it 20%.

Category 4: Respondent would take or stay in the job that guaranteed current income given any of the above alternatives.

Birth Cohort	Years of Data Collection								
	1992	1993/94	1995/96	1998	2000	2002	2004	2006	2008
1890-1923									
Hispanic	29	433	357	309	277	215	177	136	113
Black	40	1,041	850	707	567	440	346	271	205
White	137	6,007	5,134	4,281	3,554	2,868	2,313	1,809	1,381
Other/Unknown	7	92	74	60	56	38	27	19	16
1924-30									
Hispanic	79	119	114	246	229	209	203	178	154
Black	122	183	166	377	334	304	273	246	222
White	806	1,285	1,213	3,064	2,818	2,602	2,367	2,171	1,936
Other/Unknown	13	24	21	65	54	50	43	33	35
1931-41									
Hispanic	912	773	758	735	691	694	682	621	602
Black	1,688	1,504	1,401	1,334	1,248	1,197	1,153	1,047	1,013
White	7,048	6,486	6,245	6,045	5,716	5,531	5,291	5,098	4,839
Other/Unknown	169	154	133	126	122	108	103	93	91
1942-47									
Hispanic	105	96	98	245	236	227	227	221	215
Black	164	150	146	462	426	429	420	397	393
White	899	874	854	2,336	2,232	2,197	2,117	2,047	1,994
Other/Unknown	27	24	25	58	54	57	52	50	52
1948-53									
Hispanic	39	41	41	82	79	79	162	406	405
Black	49	45	43	93	88	85	124	512	495
White	196	188	185	481	492	503	633	2,099	2,045
Other/Unknown	11	11	11	19	18	15	47	91	78
1954+									
Hispanic	11	10	14	35	38	46	1,927	147	164
Black	17	12	13	32	43	49	2,879	127	126
White	78	84	90	185	192	207	14,952	611	605
Other/Unknown	5	5	5	7	13	15	371	39	39
TOTAL									
Hispanic	1,175	1,472	1,382	1,652	1,550	1,470	1,927	1,709	1,653
Black	2,081	2,936	2,619	3,005	2,706	2,504	2,879	2,600	2,453
White	9,164	14,924	13,721	16,392	15,005	13,909	14,952	13,835	12,800
Other/Unknown	232	310	269	335	317	283	371	325	311

Table 2.1: Counts of responses by race and ethnicity by survey year. Data source: Author's tabulation of the HRS

Table 2.2: Summary Statistics for Risk Aversion question					
Respondents					
Survey Year	N	Mean	Standard Deviation	Min	Max
1992	11707	3.28	1.09	1	4
1998	5117	4.64	1.59	1	6
2000	1359	4.73	1.54	1	6
2002	6093	4.66	1.54	1	6
2004	2956	4.64	1.51	1	6
2006	6414	4.7	1.48	1	6
Spouses					
Survey Year	N	Mean	Standard Deviation	Min	Max
1992	9086	3.29	1.08	1	4
1998	3582	4.64	1.55	1	6
2000	859	4.71	1.51	1	6
2002	4546	4.63	1.54	1	6
2004	2058	4.66	1.49	1	6
2006	4712	4.72	1.45	1	6

Table 2.2: Summary Statistics of Risk Preference Question, 1992-2006. Data source: Author's tabulations of the HRS.

The questions posed in the experimental model separated the respondents into four or six distinct risk preference categories (depending on the survey wave), from least risk-averse to most risk-averse, and allows one to estimate specific relative-risk coefficients for sample individuals. In the original paper evaluating this measure, Barsky et al. (1997) found that this measure of risk is highly correlated with a number of risk behaviors, including smoking, failing to have insurance, and holding stocks rather than treasury bonds.

Changes to variable over survey waves

The phenotype data on risk aversion were collected in the following waves (or years) of the HRS survey: Wave 1 (1992), Wave 4 (1998), Wave 5 (2000), Wave 6 (2002), Wave 7 (2004), and Wave 8 (2006). Proxy respondents were not asked this question in all survey waves, and the sample of the population asked these questions varies across waves. Specifically, in Wave 1, all self-reporting respondents were asked these questions. In Wave 4, AHEAD cohort respondents were not asked, but all self-reporting CODA and War Babies respondents were, along with all new HRS cohort spouses. One of 10 HRS cohort respondents was also randomly selected for these questions. In Wave 5, the questionnaire indicates that respondents were selected based on whether they were asked the question in 1998 and their experimental module assignment in 1996, in addition to random selection among those under 65. But the criteria involving 1998 and 1996 do not appear to be accurate, nor is the selection based on

age. It appears instead that one of 12 respondents was randomly selected for these questions, regardless of age. All entry cohort subsamples were eligible for selection. In Waves 6 and 8 if the person was 65 or older the questions were skipped. Otherwise, all other self-reporting respondents were asked these questions. In Wave 7 only the new EBB cohort was asked the question.

There were two important changes to the survey question over the course of its inclusion in the HRS. First, the classification of individual responses into risk categories expanded from four categories to six.

From Waves 4 forward, additional questions were asked that allow two more categories from the original four risk categories:

Category 1a: Less risk-averse than 1 above: Respondent would take a job with even chances of doubling income or cutting it by 75%.

Category 3a : Between categories 3 and 4 above: Respondent would take a job with even chances of doubling income or cutting it by 10%.

The second major change, a result of concerns over status quo bias, was a difference in the wording of the question in which individuals were more likely to choose their current circumstances when presented with a choice. In Wave 1, the pair of jobs presented were a hypothetical current job and a new one. To eliminate any concerns of status quo bias, from Wave 4 forward the pair of jobs presented are both new jobs, given that the respondent will need to move and find a new job.

To recapitulate, in Wave 1 the question wording was: Suppose that you are the only income earner in the family, and you have a good job guaranteed to give you your current (family) income every year for life. You are given the opportunity to take a *new and equally good job*, with a 50-50 chance it will double your (family) income and a 50-50 chance that it will cut your (family) income by a third. Would you take the new job?

(In Waves 2 and 3, these questions were not asked.) From Wave 4 forward the question wording is: Suppose that you are the only income earner in the family. Your doctor recommends that you move because of allergies, and *you have to choose between two possible jobs*. The first would guarantee your current total family income for life. The second is possibly better paying, but the income is also less certain. There is a 50-50 chance the second job would double your total lifetime income and a 50-50 chance that it would cut it by a third. Which job would you take the first job or the second job?

2.3 Genotyped Data

In 2012, the HRS has recently released a set of genetic markers suitable for a Genome-Wide Association Study (GWAS) whereby it genotyped 2.5 million single nucleotide polymorphisms (SNPs) on 12,507 respondents.

History of Genetic Data in the HRS

Although HRS data collection began in 1992, it was only in 2004 that practical discussions about including genetic data began. The 2005 HRS renewal application (requesting funding for the 2006 to 2011 period) proposed the collection of biomarkers, including DNA collection extracted from saliva samples as part of the in-home interview, but no funds were requested for genotyping or analysis at that time. The biomarkers collection began on the first half of a sample in 2006 and followed on the other half in 2008. Meanwhile, there was ongoing discussion with NIA staff, the NIA HRS Data Monitoring Committee, and co-investigators about what studies to perform on the collected DNA. In September of 2010, the National Academy of Sciences hosted a workshop titled “Using Genome-Wide Association Studies (GWAS) to Explore Fundamental Questions About Aging in the Health and Retirement Study (HRS) Sample,” which discussed the key themes and possible challenges of integrating genotype data with the HRS. The National Institute on Aging (NIA) commissioned the National Research Council Committee on Population to convene a two-day expert meeting to consider what data to collect on which traits and endophenotypes to optimize the HRS GWAS information as well as to explore ways in which the HRS can be harmonized with other types of large-scale studies to help uncover complex phenotypes attributable to genetics. Toward this end, more than 30 leaders in the fields of gerontology, economics, sociology, demography, genetics, population genetics, epidemiology, and psychology from throughout the United States and Europe convened in Washington, D.C. on September 23-24, 2010. The rationale for the decision to include genetic data into the HRS and the goals thereby can be summarized in a quote from the expert meeting:

“ Linking rich genotyping with the deep phenotyping available in a ongoing multi-disciplinary longitudinal study creates uniquely valuable opportunities for research on the genetics of disease, cognitive and physical function, longevity, and social and economic behavior and decision-making. Longitudinal measurement permits multiple observations on stable traits and the modeling of trajectories of change in age-related traits or age at onset in discrete disease states. The breadth of measurement will enable investigation of correlated genetic patterns in multiple domains, and sophisticated modeling of gene-environment interactions. A genotype database from a large nationally-representative sample will be an important reference point on allele frequencies and ancestry admixtures in the US population. Finally, the results of genetic analysis can inform future waves of HRS to sharpen measurement of relevant traits. Equally important, the HRS is built for comparability with other studies, creating opportunities for replication and pooling that are crucial for future advance in genetic discovery. This resource creates new horizons for research in behavioral and health sciences.” (National Institute of Aging, 2010)]

Beginning in 2006, the study added direct measures of physical function, biomarkers of cardiovascular risk, social networks, and expanded measurement of psychological traits (e.g., big 5 personality measures, affect, and sense of control). As part of the new measures added

in 2006, the study also began asking respondents to donate DNA samples to be held in repository for future research.

Collection and Genotyping Procedure

A total of 12,507 study subjects were genotyped. The study was genotyped in two phases. In 2006, samples were collected using a mouthwash method. In 2008, the study switched to collection using Oragene DNA self-collection kits, which provide samples with higher DNA concentration and yield. Based on prior rates of consent, the HRS expects an additional 3,000 Oragene samples to be added in 2010, including a substantial expansion of the minority sample.

In 2006, a random one-half of the HRS survey sample was pre-selected to complete an enhanced face-to-face (EFTF) interview, which in addition to the core HRS interview included a set of physical performance tests, anthropometric measurements, blood and saliva samples, and a psychosocial self-administered questionnaire. The sample was selected at the household-level. In 2008, an EFTF interview was conducted on the remaining half of the sample. Respondents who consented to the saliva collection in either 2006 or 2008 were included in the 2012 GWAS data release.

The genetic material was collected using Illumina's Human Omni2.5-Quad (Omni2.5) BeadChip methodology. Saliva was collected on half of the HRS sample from each wave starting in 2006. In 2006, saliva was collected using a mouthwash collection method. In 2008, the data collection method switched to the Oragene kit. Saliva completion rates were 83% in 2006 and 84% in 2008.

The genotyping was performed by the NIH Center for Inherited Disease Research, using the Illumina Human Omni-2.5 Quad beadchip, with coverage of approximately 2.5 million single nucleotide polymorphisms (SNPs).

Quality Control

Genotypic data that passed initial quality control at CIDR were released to the Quality Assurance/Quality Control (QA/QC) analysis team at the University of Washington, the study investigators' team and dbGaP. These data were analyzed by all four groups and the results were compiled into a quality control document by the University of Washington (2012).

The document provides details on a number of quality control measures that were applied to the genotype data, including the following list: gender identity, chromosomal anomalies, relatedness, population structure, missing call rates, batch effects, sample quality, duplicate sample discordance, Mendelian errors, Hardy-Weinberg equilibrium, minor allele frequency, duplicate SNP probes, sample exclusion and filtering summary, and SNP filtering summary.

While much of the technical detail of the quality control is best left to the quality control documentation, there are a few key portions of the quality control procedures of the genotype

that are worth elaborating on as they form important parts of the method in the proceeding analytical chapters of this dissertation.

Overall Sampling and Data Issues

In the following, the term “sample” refers to a DNA sample and, for brevity, “scan” refers to a genotyping instance (including genotyping chemistry, array scanning, genotype calls, etc.). A total of 13,129 samples (including duplicates) from study subjects were put into genotyping production, of which 12,857 were successfully genotyped and passed CIDR’s QC process. The subsequent QA process identified 12 subjects with unresolved identity issues. Of these 12, seven were unexpected duplicates identified by CIDR prior to release, two were determined to have questionable identity by the SI based on their HRS IDs, one was a respondent dropped from the HRS sample, and one was found to be an unexpected relative of another subject. The set of scans to be posted include 12,845 study participants and 411 HapMap controls.

The 12,845 study scans derive from 12,507 subjects and include 336 subjects with duplicate scans (334 subjects with two scans each and two subjects with three scans each) (Table 2.3). The 411 HapMap control scans derive from 88 subjects of which 87 are replicated two or more times. The study subjects occur as 12,335 singletons and 84 families of two or three members each. The study families were discovered during the analysis of relatedness. The HapMap controls include 25 trios as well as 13 singletons.

The reported median call rate is 99.7%. The first phase consists of DNA from buccal swabs collected in 2006 and extracted using the Qiagen Autopure method. The second phase consists of saliva samples collected in 2008 and extracted with Oragene. Although the two phases were genotyped separately, the raw data were clustered and called together. The samples were genotyped in batches corresponding to 96-well plates. Each plate contained from one to four HapMap controls, as well as an average of two study sample duplicates. The HapMap is a catalog of common genetic variants that occur in human beings. It describes what these variants are, where they occur in our DNA, and how they are distributed among people within populations and among populations in different parts of the world. Because the HapMap samples have gone through an extensive clinical and phenotypic investigation, they are often used, as they are here, as a standard for quality control measurements of other genotype data.

Gender Identity

One of the quality control checks performed included the verification of sex and gender for individuals. This check was employed to verify that an individual’s self-identified gender was the same as their genetic sex. Since risk aversion —primary phenotype studied in this dissertation —varies significantly by gender, I will briefly highlight the quality controls deployed for this data.

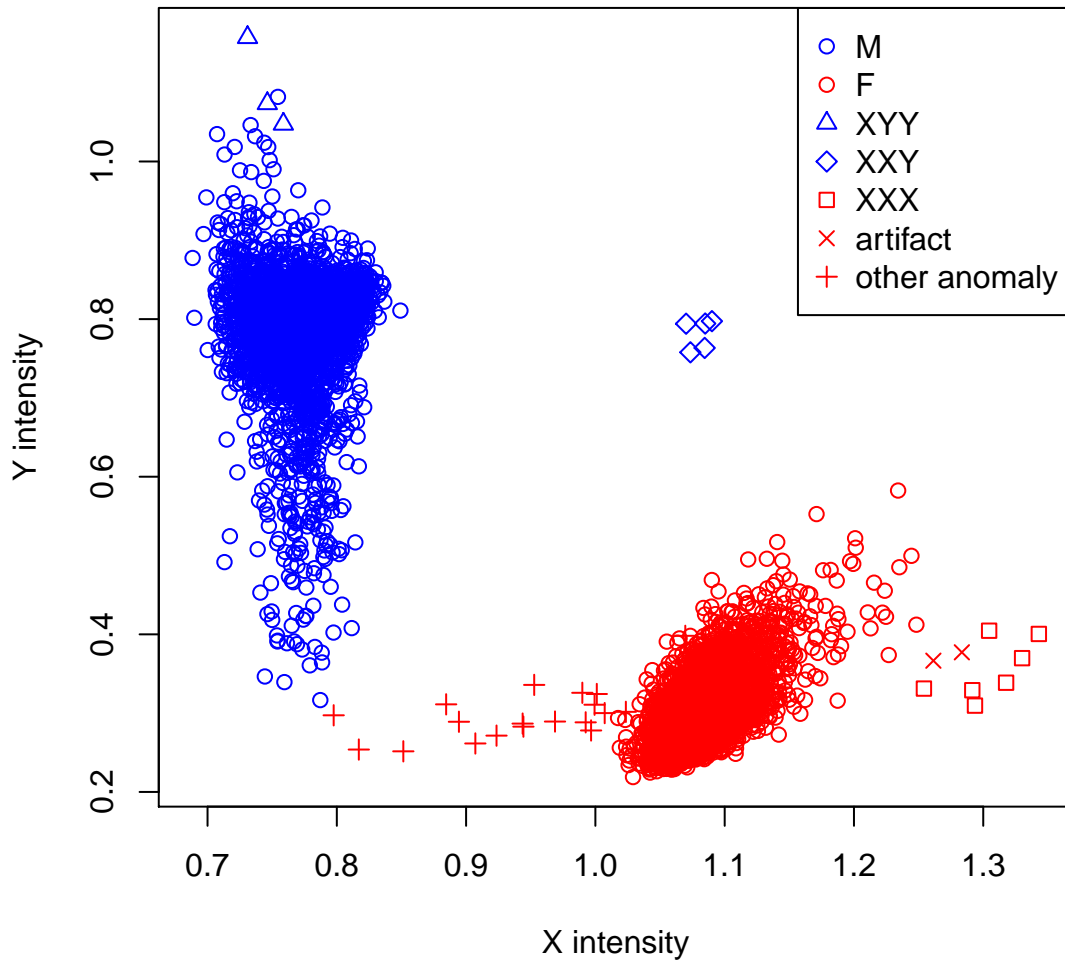


Figure 2.1: The X and Y intensities are calculated for each sample as the mean of the sum of the normalized intensities of the two alleles for each probe on those chromosomes. Sample sizes are given in the axis labels. X heterozygosity is the fraction of heterozygous calls out of all non-missing genotype calls on the X chromosome for each sample. Source: University of Washington, 2012

To check gender identity, the quality control groups examined both X chromosome heterozygosity and the means of the intensities of SNP probes on the X and Y chromosomes. The expectation is that male and female samples will fall into distinct clusters that differ markedly in X and Y intensities. Figure 2.1 indeed shows that there are two main clusters, as expected, and no gender mis-annotations are apparent. There are tails of low Y intensity for males and low X intensity for females, which research has found is not unusual for elderly subjects (University of Washington, 2012).

Population Stratification

Population stratification is the presence of a systematic difference in allele frequencies between subpopulations in a population, possibly due to different ancestry, especially in the context of association studies. Population stratification becomes important in studies using genetic data, and particularly in a GWAS study such as that presented in Chapter 3, because of the possibility of confounding (Knowler, 1988). In other words, spurious allelic associations can be made between a genetic marker and a phenotype that are actually a result of the differing frequencies of alleles within subpopulations of the sample.

In the 1990s, concern over population stratification led some researchers to question the validity of population-based genetic association studies. Currently, the impact of population stratification is less contentious. The general consensus is that population stratification will likely be a small source of bias in well-designed studies of Caucasian populations. Still, nationally-representative random sample surveys such as the HRS do not fit this criterion; the HRS in particular does not fit this criterion because of its deliberate over-sampling of African-Americans.

While a number of methods to control for population stratification are used, including restricting a sample to be Caucasian-only, one of the most common methods to control for population stratification is the use of principal components (Price et al., 2006). The HRS quality control applied the oft-used technique developed by Patterson et al. (2006) to apply a principal components analysis at the individual level. The result of this analysis is 20 components, or eigenvectors, which are linearly uncorrelated variables that account for the genetic variance among individuals. The first component has the largest achievable proportion of variance, the second component has the second greatest share of variance, and so it continues for all 20 components. The HRS makes a file of individual-level principal components available to users of the GWAS data. As I elaborate in Chapters 3 and 4, these principal components are used as control variables as part of the analysis for both of these chapters.

Figure 2.2 exhibits the principal components method used in the HRS data. The x-axis presents the values for the first principal component and the y-axis as the values for the second principal component. Each individual is plotted on these coordinates based on their values. These figures also include the HapMap control samples, which have been identified in large part because of their relatively pure genetic ancestry; they therefore provide a good comparison. Readers will note that the first component, which has the highest share of genetic variance across individuals, contributes less than 5% of total variance. There are clear delineations along self-identified racial categories, which is an indication that ancestry-based genetic variance does indeed exist and that the principal components method may in fact be an appropriate control measure. Figure 2.3 provides the same analysis without the inclusion of the HapMap control sample.

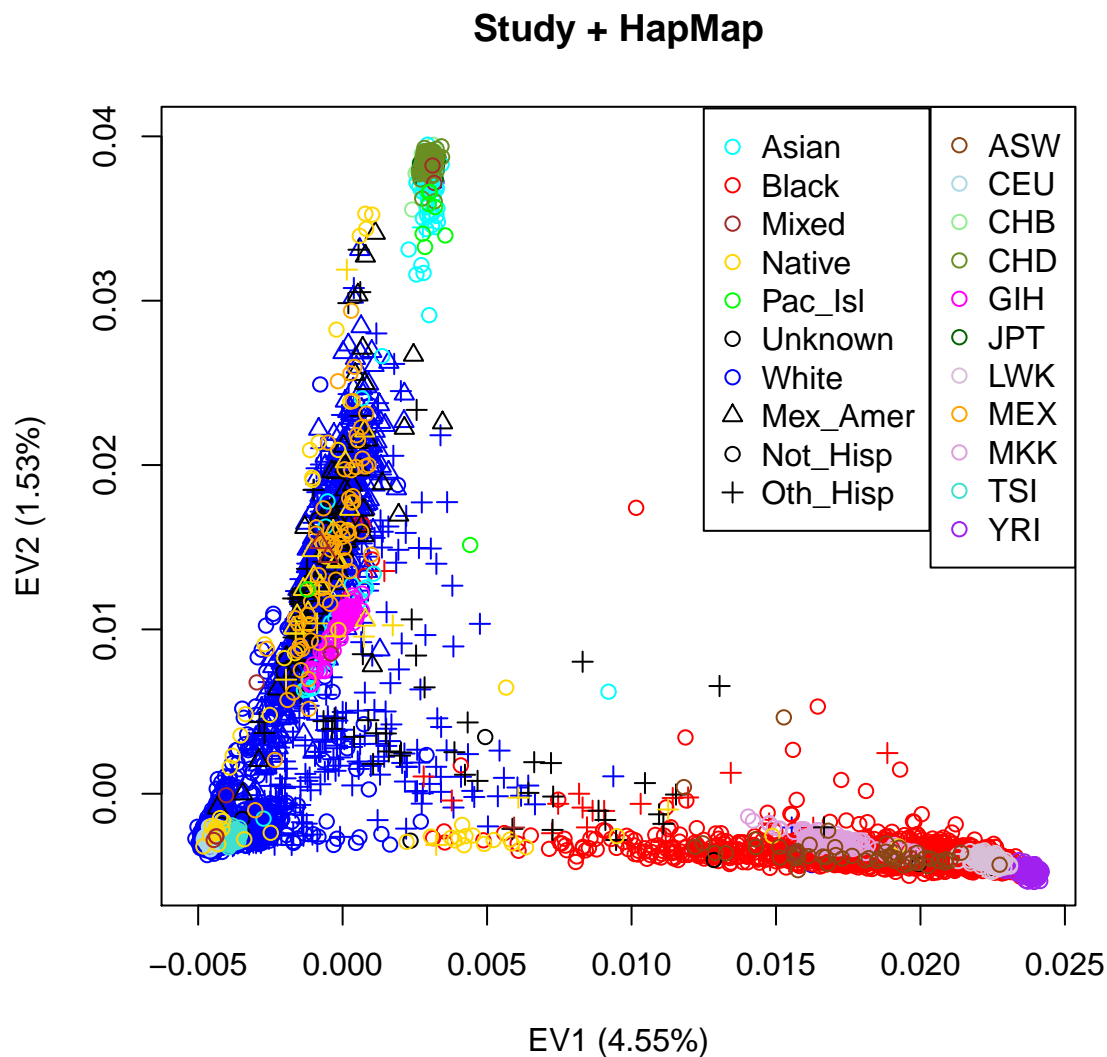


Figure 2.2: Principal component analysis of 12,507 study subjects with 1230 HapMap controls. Color-coding is according to self-identified race, while symbol denotes ethnicity (Hispanic or not). Axis labels indicate the percentage of variance explained by each eigenvector. Source: University of Washington, 2012

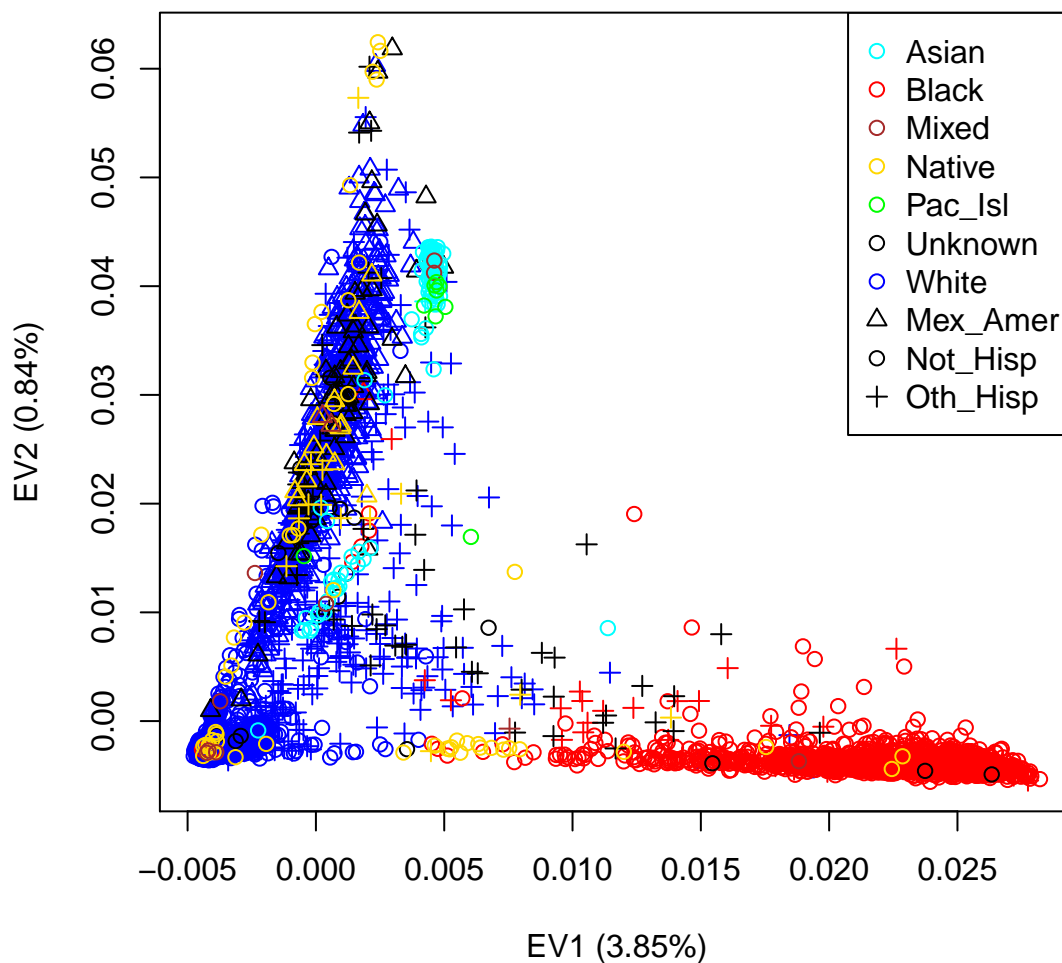


Figure 2.3: Principal component analysis of 12,419 unrelated study subjects without HapMap controls. Color-coding is according to self-identified race, while symbol denotes ethnicity (Hispanic or not). Axis labels indicate the percentage of variance explained by each eigen-vector. Source: University of Washington, 2012

Confidentiality Protections

The genotype data and a limited set of phenotype measures were deposited in the NIH GWAS repository (dbGaP), which provides a convenient method of distribution to researchers who meet NIH requirements for access. My access to the data was approved as an internal collaborator under a subproject of the Center for the Economics and Demography of Aging (NIA Grant 2P30AG012839-18), with Ronald Lee, Principal Investigator, and Kenneth Wachter, Project Leader. The protocol approved by the DbGAP Data Access Committee is number 3630, under request 14779. The first approval and the most recent approval came on July 10, 2014. came on June 1, 2012. Approval for linking genotype and phenotype data was granted by the HRS under their Genetic Data Access Use Agreement. The research plan was approved by the University of California Committee for the Protection of Human Subjects, protocol 2011-10-3707, dated Nov. 18, 2011 and renewed Oct. 4, 2012, Nov. 29, 2012, November 13, 2013, and Sept. 2, 2014. The process to request access to any dbGaP study is done via the dbGaP authorized access system. Once all application procedures were completed, the data were downloaded onto a dedicated secure server with an encryption system. In order to merge together the confidential genetic data with the publicly-available HRS phenotype data, I made use of two data “cross-walks,” one made available with the genetic data and one supplied separately by the HRS.

Value	1992	1998	2000	2002	2004	2006
65+, not asked				9468		10380
Skip, not selected		8631	15927			
Cohort skipped		4984			15198	
Don't know	109	335	160	406	113	325
Other missing	194	123	3	23	15	28
Refuse		151	68	139	45	62
Proxy (skip)	642	2043	2062	2036	1802	1260
1. Least Risk Averse	1503	323	71	363	157	283
2	1279	375	95	378	199	426
3	1358	483	122	590	271	628
4	7567	773	196	954	519	980
5		863	239	1160	601	1349
6. Most Risk Averse		2300	636	2648	1209	2748
Total	12652	21384	19579	18165	20129	18469

Table 2.3: Distribution of Risk Aversion in Study Sample. Data source: Author's tabulation from the HRS RAND data set.

Chapter 3

Genome-Wide Association Study

3.1 Introduction and Motivation

Older Americans vary in their preparation for the financial burdens of retirement and old age. These differences cannot be attributed entirely to income, education, cognitive ability, family background, or a number of other factors. Evidence suggests that differences in risk aversion are an important driver of the variation in wealth accumulation of individuals. Risk preferences are not entirely understood; though they have been well-studied empirically, there are still large amounts of heterogeneity that remain unexplained. Gender, age, and education levels are all predictors of risk aversion but often can explain only a small portion of the variation Americans display in risk aversion (Barksy, 1997). Risk aversion is known to be a heritable trait (Benjamin et al., 2012; Cesarini et al., 2009, 2010) implying that risk aversion is partly driven through variations in individual genetic propensities (see Section 2.2 for more detail).

Until recently, this idea was difficult to test empirically. However, there has been growing enthusiasm for the use of molecular genetic data in social science research (Benjamin et al., 2012) and the recent introduction of genetic data in social science surveys has provided researchers with a new chance to link survey responses with individual's biological measures. As such, a growing body of research has emerged that correlates genetic heritability to observed social and economic behavior. These studies borrow from long-standing techniques from the biological and medical science that have shown that genetic variation helps explain individual risk's for Alzheimer's disease, various cancers, and a number of other diseases.

As an important and often difficult to measure parameter, risk preferences have been a primary focus of economists in this area of study. Still, little is known presently about the nature of the relationship between an individual's biology and their risk preferences. Genetic variants that lead to differences in observed traits typically take two forms. The first is that one or a few genetic mutations cause direct observable differences in traits. A variant on the APOE gene that has been linked to Alzheimer's disease is a good example, as is the BRCA gene, which is known as the "breast cancer gene." While these often get our attention in

both academic and popular media because of the genes' direct corollary to observable traits, they remain scientifically very rare. A more common expression of genetic variability is the second form, that lower-level activity at different areas across the human genome collectively translates to differences in observed behavior.

Using the newly-released genetic data from the U.S. Health and Retirement Study, this study is one of the first of its kind to run a genome-wide association of risk aversion using a large, nationally-representative sample. Specifically, this study seeks to answer the question: Is the heritable portion of risk aversion in individuals driven by a few genetic variants with large causal effects or by a high number of variants that all contribute small effects separately leading to large individual variation in risk aversion collectively? In light of the evidence of substantial heritability of risk preference, this study seeks to use the individual-level genetic data to explore the nature of this heritability.

To preview the findings, this study concludes that risk aversion is not driven by a few specific genetic variants with strong causal effects and is likely to be highly polygenetic in nature. The analysis finds no individual genetic variant to be statistically significant at a specific threshold of detectability with a sample size of over 7,000 individuals. The whole sample was initially divided into an exploratory subsample of 90% of the respondents and a confirmatory subsample of 10% respondents. This implies that risk aversion is likely to be the result of a series of small and compounding biological pathways and does not operate like cognitive decline or like certain diseases driven strongly by specific genetic mutations. The observed heterogeneity in risk aversion among older Americans, then, is partially the result of a large number of smaller genetic factors that collectively lead to large variations in observed behavior.

This paper is organized as follows. Section 2 reviews the literature that provides compelling evidence of the heritability of risk preferences. Section 3 describes data used and Section 4 details the empirical strategy. Section 5 discusses the results. The paper will end with a discussion in Section 6 and plans for further research in Section 7.

3.2 Background

There is evidence that variations in risk preferences are partially driven by genetic variation, or in other words, have an aspect of heritability. A number of twin studies provided early evidence that risk aversion was heritable. By observing and measuring monozygotic (identical) and dizygotic (fraternal) twins, these studies were able to create estimates of heritability while controlling for shared environment factors. These studies, described in Section 3.2, lend substantial support to the idea that risk preferences are heritable, with estimates ranging from 20-60 percent, but cannot point to any specific genetic activity explaining this heritability. A subsequent body of work has shown correlations between experimentally-elicited risk preferences to a few specific genes, most notably the dopamine and serotonin transporters that are associated with reward systems in the brain. These studies, also described in Section 3.2, provide evidence that risk preferences do have a genetic component,

but are unable to provide information on the biological pathways that might drive this heritability or the possible influence of any other genetic effects.

Since the conclusions of these studies, the introduction of genome-wide data has provided researchers the opportunity to scan large portions of the genome rather than just a few specific pieces of genetic material. Genome-wide data collects hundreds of thousands or even millions of individual single-nucleotide polymorphisms (SNPs), which are areas on the genome where variation can exist. The human DNA is composed of base pairs of alleles that lie along the twenty-three chromosomes. These base pairs of alleles make up the “ladder rungs,” so to speak, of the double-helix form of the DNA, and are passed down from parent to offspring. Single-nucleotide polymorphisms (SNPs) are specific base pairs where there is some variation in the combination of alleles, which can lead to phenotypic differences among individuals. (A comprehensive explanation of molecular biology written for economists can be found in Beauchamp et al., 2011.)

From a methodological perspective, the introduction of these data offers the potential for more sophisticated and high-powered empirical work in this area; prior heritability work was limited to a small sample of twins or a subset of genes commonly used in the biomedical literature. Genome-wide studies allow researchers to test every individual measured SNP in the genome against the outcome of interest and as such, move down to a lower level of aggregation: from humans in twin studies, to genes, to individual SNPs. If one were to consider this within the context of standard regression, this is equivalent to obtaining information on hundreds of thousands of additional covariates for a model. The promise of these data is impressive. Still, while the samples on the genetic material have grown substantially in size, some reaching 2 million pieces of genetic data, samples of respondents, who are more costly and difficult to gather, has not grown in step and remain underpowered. As a result, many of the early findings correlating genetic variation to economic preferences have failed to replicate and have produced false positives.

Measures of risk aversion vary largely

Given its centrality in everyday life, empirical research has examined the important heterogeneity in risk preferences across populations including differences across gender (Eckel & Grossman, 2002; Powell & Ansic, 1997; Schubert, Brown, Gysler, & Brachinger, 1999); family background (Hartog et al., 2002); work characteristics (Praag & Cramer, 2001); and educational attainment levels (Brunello, 2002), as well as across different contexts of risk (Soane & Chmiel, 2005; Weber, 2002). Given the particular importance of risk preferences in decisions made at older ages, much of this empirical work has focused on the implications for retirement decisions and savings (Bodie, Merton, & Samuelson, 1992; Hurd, 1990; Karatzas, Lehoczky, & Shreve, 1987; Sunden & Surette, 1998). With all of these studies has come a plethora of measures of risk preferences. These measures typically take one of three forms: survey-based assessments such as respondent’s answers to hypothetical lottery gambles, experimental evidence, or inference from observed decision-making in financial, health or insurance markets.

This study uses a measure of risk aversion first introduced by Barsky et al. (1997) that categorizes risk aversion through responses to a series of hypothetical gambles on lifetime income through a specific job. Barsky and his coauthors concluded that this measure is positively related to a number of risky behaviors and provide evidence about the validity and usefulness of measures of this preference parameter. Since introduction, this measure of risk preference has been used in a number of studies attempting to measure risk preferences (Evan & Smith, 2010; Schulhofer-Wohl, 2007; Smith et al., 2004), has inspired similar measures in other surveys (Dohmen et al., 2011; Falk et al., 2005) and the original paper confirming the validity of the measure has been cited over a thousand times (Google Scholar).

Survey-based evidence on attitudes towards risk has been criticized, mainly on the grounds that hypothetical measures may not be well-connected to behavior exhibited within a set of real constraints. A number of recent studies have attempted to validate Barsky et al.'s measure of risk preference and other similar instruments and concluded that hypothetical questions track closely, albeit imperfectly, with actual risk-taking behaviors (Dohmen et al., 2005; Falk & Heckman, 2009; Guiso & Paiella, 2005). While the debate remains, hypothetical measures of risk are used in good standing and are useful when behavior is difficult to observe. Falk and Heckman (2009) go even further to argue that objections against experimental or hypothetical measures are “misguided” and that the issue of generalizability is no more a concern than in field data.

More importantly, the common finding across the literature examining risk preferences using any one of these three measures is that heterogeneity in risk preferences is real and substantial and cannot be entirely explained by observed characteristics typically used in empirical models, further underscoring the importance of this study exploring genetic correlates to risk preferences.

Twin studies provide estimates of heritability of risk preferences

Twin studies are the mainstay of behavioral genetics and play a crucial role in establishing heritability. Evidence suggesting that risk aversion might be partially hard-wired arose from twin studies using the comparison of monozygotic (MZ, or identical) and dizygotic (DZ, or fraternal) twins to estimate heritability. These studies exploit the fact that MZ twins share one hundred percent of their genetic material and DZ twins only half in order to estimate a percentage of the variation in observed risk aversion that can be attributed to genetic variation.

Given the difficulty in obtaining large samples of twins, risk preferences in this body of work are most often elicited through experimental techniques. Cesarini et al. (2009) use a twin study design to show that genetic differences explain about 20 percent of individuals' variation in experimentally-elicited preferences for risk. The authors also used measures of survey-based hypothetical questions similar to those used in this study as well as additional measures of risk based on observed experimental behaviors with financial incentives attached. Using a sample of 314 identical twin pairs and 141 non-identical twin pairs, the authors conclude that there is strong evidence that preferences for risk are broadly heritable (Cesarini

et al., 2009). Around the same time, Zhong et al. (2009) published a twin study from a sample of 232 twin pairs in China (167 were monozygotic and 65 were dizygotic) and provide evidence of the heritability of economic risk attitude to be as high as 57 percent. The authors measure risk attitude through a hypothetical measure that asks subjects to pick from a series of gambles with different probability distributions and payoffs. Zyphur et al. (2009) study a subset of 200 male twin pairs from the Minnesota Twin registry and attribute approximately half of the variation in risk preferences to be of a genetic nature. One of the largest twin studies exploring risk preferences was done on a total of 1,875 twin pairs in Australia, comprising 867 pairs of identical twins and 1,008 pairs of non-identical twins (Le et al., 2010). They found that approximately 20 percent of the variation in attitudes towards risk are linked to genetic differences by using responses when individuals were asked to rank their risk aversion on a scale from 1 to 10.

Some twin studies do exist using financial decision-making to measure risk preferences and find similar results. A follow-up study by Cesarini et al. (2010) using the Swedish twin registry showed that 25 percent of variations in the difference in riskiness of portfolio allocation could be attributed to genetic variation. Barnea and Cronqvist (2010) also use the Swedish twin registry on identical and fraternal twins' financial portfolios to examine heritability of risk as measured through the relative amount of portfolio invested in equities as well as overall portfolio volatility. They find that a genetic factor explains about one-third of the variance in stock market participation and asset allocation and conclude that there are "innate differences in factors affecting stock market participation costs" that can be attributed to genetic variation in risk preferences.

Research based on twins has been used previously by economists to good effect in the study of both earnings and educational attainment (Le et al., 2012). In the study of earnings, the framework has been used to address the issue of genetic influences on earnings as well as the bias in the conventional estimate of the return to schooling. Due in large part to Ashenfelter and Krueger (1994), this approach has stimulated considerable interest and has now been applied to data from the US (Ashenfelter & Rouse, 1996), Australia (Miller, Mulvey, & Martin, 1995, 2006), the UK (Bonjour, Cherkas, Haskel, Hawkes, & Spector, 2003), and Sweden (Isacson, 2003). This replication across countries has generated additional confidence in the findings.

Candidate genes point to specific areas on the genome but lack precision of strength of effect

While twin studies provide estimates of overall heritability, candidate gene studies provide complementary evidence through the association of measured risk preferences to specific genes. The candidate studies point to four or five genes for which genetic variants have repeatedly been associated with multiple measures of risk preferences.

Recent findings in neuroscience suggest that the neurotransmitters dopamine (Schultz, 2007) and serotonin (Daw, Sham, & Dayan, 2002) have important roles in decision making.

Genes that regulate these neurotransmitters impact the processing of information about rewarding stimuli (Yacubian et al., 2007) and harmful stimuli (Frank et al., 2007; Kim et al., 2006); are related to personality traits such as extraversion (Ruetter & Henning, 2005), novelty seeking (Ebstein, 1996), and anxiety; and are associated with the development of addiction (Kreek, 2005). Moreover, activity within parts of the brain regions activated by serotonergic and dopaminergic neural pathways has been shown to relate to individuals' financial risk taking behavior (He et al., 2010).

Among dopamine receptors, DRD4 has been cited a number of times as a gene associated with risk taking. The DRD4 gene regulates dopamine uptake in the brain and individuals with a specific polymorphism of the DRD4 gene have been found through numerous studies to be more risk seeking. Those with the specific polymorphism in question have been measured to take more risk when choosing between a risky and riskless asset in a number of experimental settings including playing investment games (Kuhnen & Chiao, 2009) and financial gamble (Dreber et al., 2010), hypothetical lottery questions (Zhong et al., 2009), or self-reported general risk taking or behavior in risk-related activities (Dreber et al., 2010).

In addition to being so under general conditions, DRD4 is associated with differences in individuals under specific risk conditions, such as decisions under ambiguity. Carriers of specific variants of the DRD4 gene are more likely to increase the amount of risk they incur when the outcomes become ambiguous or when potential losses are allowed compared to risk baseline (Carpenter et al., 2011), although in some samples this finding was significant only among female (Chew, Ebstein, & Zhong, 2012) or male (Roe et al., 2009) subjects. Moreover, those individuals seem to not only be less risk averse in general, but also appear to be better decision-makers under risk; a study of male bridge players showed that those with specific variants of the gene take more good risks and fewer bad risks, while the opposite is found for those without the variant (Dreber et al., 2011).

Recent studies using behavioral and neuroimaging techniques have also examined the effect of a serotonin receptor—specifically—the 5-HTTLPR gene on economic decision making (Crisan et al., 2009; Homberg et al., 2008; Kuhnen & Chiao, 2009; Roiser et al., 2009; Stoltenberg & Vandever, 2010; van den Bos et al., 2009; Zhong et al., 2009), revealing that individuals with the short allele (versus the long) are more likely to exhibit risk aversion. These studies revealed that the short allele was associated with less investment in risky assets (Kuhnen & Chiao, 2009), less engagement in actively making investments decisions, and having fewer credit lines (Kuhnen, Samanez-Larkin, & Knutson, 2011). Similar to dopamine, the short allele is associated with taking less risk under ambiguity (Crisan et al., 2009), as well as other non-standard decision-making under risk such as familiarity bias (Chew, Ebstein, & Zhong, 2012) and higher loss aversion (He et al., 2010).

Other studies have linked risk preferences to a polymorphism in monoamine oxidase A gene (MAOA). Studies find that carriers of the MAOA-L polymorphism subjects with the high activity (4-repeat) allele versus subjects with the low activity (3-repeat) allele were more likely to take financial risks (Frydman et al., 2011), purchase less insurance and prefer a longshot lottery (Zhong et al., 2009). Associations to risk aversion have also been found in nicotine receptors (Roe et al., 2009), oxytocin (Apicella, 2010), and testosterone (Zethraeus

et al., 2009).

This work on specific genes provides an important starting point for the exploration of a full genome-wide analysis. The candidate gene literature provides evidence of specific areas on the genome that are correlated to variations in risk aversion but stops short of making any inference on the biological pathways at work since the genetic material used in these studies is limited. Taken as a whole, this body of work provides evidence that a variety of risk attitudes have a strong genetic component. Still, the genes that have been identified are not SNPs but another less common kind of genetic variation called “tandem repeats.” Tandem repeats represent only a small percent of the heritability, estimated at approximately 10% (Monolio et al., 2009), and are not available in the HRS data. The analysis in this paper uses the large data on SNPs to run a genome-wide analysis study (GWAS), which tests each individual SNP against the measure of risk aversion. This dataset, with more than a million SNPs, represents an important contribution towards the understanding of the heritability of risk preferences.

3.3 Data

This study uses both phenotype and genotype data from The Health and Retirement Study (HRS), which has been detailed in Chapter 2 of this dissertation. As a review, the University of Michigan Health and Retirement Study (HRS) is a longitudinal survey of a representative sample of Americans over the age of 50. The target population for the original HRS cohort includes all adults in the contiguous United States born during the years 1931-1941 who reside in households, with a 2:1 oversample of African-American and Hispanic populations. The HRS includes rich experimental modules with each wave that have detailed assessments of specific topics. This study will use such a repeated experimental module featuring risk aversion that is particularly well-suited for this study.

The HRS has recently released a set of genetic markers suitable for a Genetic Wide Association studies whereby it genotyped 2.5 million single nucleotide polymorphisms (SNPs) on respondents, which is the genotype data for this chapter. The total sample size for respondents for the GWAS is 12,595 respondents. A careful procedure of quality control was applied to the genetic data, which is described in more detail in the Results section of this paper.

For the phenotype data of this chapter, I used the risk aversion assessment question detailed in Chapter 2 of this dissertation. As a brief reminder, the HRS has included an experimental module for the assessment of risk aversion proposed by Barsky et al. (1997) that was introduced in the first wave of the survey in 1994 and included in a total of six data waves with a lottery question. The questions separate the respondents into four distinct risk preference categories, from least risk-averse to most risk-averse, and allow one to estimate specific relative-risk coefficients for sample individuals.

3.4 Empirical Strategy

In this section, I provide more details on the methods use for the GWAS of risk aversion. In a GWAS, tens or hundreds of thousands of genetic markers are individually tested for association with a trait of interest. In this study, I analyzed data on approximately 7,314 individuals from the Health and Retirement Study who have been genotyped at over two million SNPs and searched for SNPs that correlate with the specific measure of risk aversion detailed earlier.

Quality Controls for the Genetic Data

As mentioned above, the original sample of 12,595 was reduced to 7,314 after subsetting the sample to include only those with a non-missing value for risk aversion as well as reserving part of the sample for validation. Further filters were then applied to control for the quality of the genotype data. Following usual practices (Sullivan & Purcell, 2008), I applied four quality control measures to the sample.

First individuals were dropped because they had a “missingness” larger than 0.05. An individual’s missingness is the fraction of the SNPs in the array with missing data for the individual. A high missingness can be suggestive that some problem occurred in the genotyping procedure for this individual, and therefore that the nonmissing genotypic data might not be accurate enough. A requirement of less than 5% missingness is customary in the molecular genetics literature (Sullivan & Purcell, 2008). Only one individual was dropped due to this criterion.

Next, SNPs with a missing data frequency greater than 5% were deleted. A high missingness can be suggestive that some problem occurred in the genotyping procedure for the SNP. Third, I eliminated SNPs for which the least common allele had an incidence smaller than 1% (called the “minor allele frequency”). Coefficients on these SNPS will generally be imprecisely estimated and can thus be misleading.

Finally, I excluded SNPs which failed a test of Hardy-Weinberg equilibrium at the 10⁻⁴ level. The null hypothesis of this test is that the observed genotype frequencies are equal to their theoretical expectations under random mating. A large departure from Hardy-Weinberg equilibrium may be an indication of genotyping errors or the consequence of population stratification in the sample.

These four quality control measures are widely used by convention in the molecular genetics literature (Pearson & Manolio et al., 2008; Sullivan & Purcell, 2008). From the original 2.5 million SNPs on our array, 19,542, did not satisfy the missingness criteria, 567,254 did not satisfy the minor allele frequency criteria, and 627,763 did not pass the Hardy-Weinberg test. Applying all filters leaves a total of 1,222,014 SNPs and 7,313 individuals for analysis.

Population Stratification

Population stratification refers to differences in allele frequencies across subpopulations. Such differences can occur in the absence of random mating between subpopulations as a consequence of founder effects, genetic drift, and differences in natural selection pressures. When both the frequencies of alleles and environmental factors affecting a trait of interest vary across subpopulations, spurious associations between those alleles and the trait might result (Beauchamp et al., 2011). Without population stratification controls, markers which differ significantly in frequency between racial subpopulations could be found to be associated with a specific outcome measure, but those associations will usually be partly due to cultural, social or environmental differences, not to genetic differences. Population stratification has been shown to be a concern even in samples of European Americans (Campbell et al., 2005), so will be of particular concern for the HRS, which has an oversampling of African Americans.

As part of the quality control measures, the HRS applied principal component analysis to the genotypic data. In following standard procedure, I include the component scores on the leading ten principal components as control variables in the main regression specification. These values contain information about population structure, so including them in an association test partly controls for population stratification. Because principal component analysis assumes independent observations and because the HRS is not a family-based study, I use all observations included in the GWAS data.

Association Analysis

While this study takes the behavioral genetics perspective, the empirical models will be familiar to demographers. For each individual SNP that passed the filters, I ran the following regressions using an Ordinary Least Squares method:

$$\text{Risk Aversion} = \alpha_i + \beta_1 \text{SNPS} + \beta_2 \text{PC} + \beta_3 \text{X} + \epsilon, (1)$$

where Risk Aversion is a measure taking values from 1-4, with 4 being the most risk-averse, SNPS is the number of copies of the minor allele (0, 1, or 2) an individual has at SNPs, PC is a vector of the 10 top principal components of the sample (to control for population stratification as mentioned above), and the vector X includes age measured in the same year as last measure of risk aversion and gender. I ran a total of 1,222,014 regressions, one for each SNP that passed my quality control screening, using the GenABEL library in R.

The computation of p-values in the GWAS analysis uses the “Gaussian” rather than the “Binary” specification of traits. In my case, the trait is measured not as a continuous but on a four-point discrete scale. However, with over 7,000 respondents, standard normal theory arguments would lead us to expect that the coefficients for effect sizes would be very close

to asymptotic normality and the calculated p-values would be appropriate. I checked this assumption with simulations based on a two-way table of risk aversion by allele count for one SNP in the tail of the distribution of observed p-values and found, as expected, that the simulated estimates of effect sizes conformed closely to asymptotic normality well out into the tail of the distribution. Thus, I feel confident that the calculation of p-values has not been distorted by the discrete character of my trait measure.

Inference Under Multiple Hypothesis Testing

As is apparent from the above specification, the difficulty in interpreting the results of a GWAS comes from the repeated number of regressions that are run on the same sample of respondents. Because the sample of respondents is much smaller than the number of SNPs and because a very large number of hypotheses are being tested, many SNPs will inevitably turn out to be statistically significant at conventional levels just because of sampling variation. In other words, if my genetic sample includes 500,000 SNPs, one would expect 12,500 SNPs to be significant at the 5% level through random chance alone. As such, standard levels of significance for p-values are not appropriate for this technique. Instead, several methods have been proposed to address this issue. The most utilized threshold in the literature for large GWAS's based on 500,000-SNP array data was set by the Wellcome Trust Case Control Consortium, to set a p-value of 5×10^{-7} (Burson et al., 2007). In following with standard practice, this was the threshold used for the results discussed below in Section 5.

The most stringent solution is to use the Bonferroni correction, in which the conventional significance threshold is divided by the number of tests performed to obtain a Bonferroni-corrected significance threshold or, equivalently, all p-values are multiplied by the number of tests performed to obtain Bonferroni-corrected p-values. In the first stage study with the HRS data, 1,222,014 tests were performed (one for each SNP that passed the quality-control filters), thus yielding a Bonferroni-corrected significance threshold of $p=0.05/1,222,014=4.09 \times 10^{-8}$. However the Bonferroni approach is generally agreed to be overly conservative, because SNPs that are close to one another are generally correlated and thus not statistically independent.

As has been discussed in previous studies of this nature, prior experience with false positives in the field of medical genetics has led researchers to be cautious in interpreting any result that has not been replicated in an independent sample. Hence, the above significance thresholds must be seen as suggestive (Beauchamp et al., 2011).

3.5 Results

Descriptive Analysis

Table 3.1 details some of the descriptive statistics of the sample for the GWAS. The total sample size is 10,455 respondents who have been genotyped and have at least one response

Gender	N	%
Male	4,311	41%
Female	6,144	59%
Age		
0-49	503	5%
50-54	1,941	19%
55-59	3,258	31%
60-64	3,483	33%
65-69	707	7%
70-79	536	5%
80+	27	0%
Risk Aversion		
Least Risk Averse	1,189	11%
2nd Most Risk Averse	1,037	10%
3rd Most Risk Averse	1,482	14%
Most Risk Averse	6,747	65%
N	10,455	

Table 3.1: Summary Statistics for GWAS Sample

for the hypothetical risk measure. The test sample is a random subset of 70% of the full sample, for a total of 7,314 respondents. As is expected, the distribution of demographic characteristics for both the subset and the full sample are nearly identical. The sample is nearly 60 percent female and, as is consistent with the sampling procedure, the majority of the sample is between the age of 50 and 65. Although the HRS sample is restricted to individuals aged 50 and over, spouses are included, and as such the youngest sample member is 25. The oldest person in the sample is 92. As noted earlier, the set of risk aversion questions allows a categorization of four progressive categories of risk averse, with the value of four representing the most risk averse. In this sample, approximately 65% of the sample is categorized as most risk averse. The other respondents are fairly even split between the remaining three categories, with just over 10% being the most risk-seeking.

The risk aversion question was asked in Waves 1 and 4-8, with the number of respondents in each wave ranging from 748 in Wave 5 to 5,451 in Wave 1. For the analysis, I pooled all respondents who had ever responded to the hypothetical risk aversion question. For those who have been sampled more than once, I used the most recent response. Not surprisingly,

risk aversion tends to increase with age, although not by a marked amount here since the sample consists primarily of persons over the age of 50 years. Also in line with the literature on risk aversion, there are slightly higher shares of females categorized as most risk averse. Finally, because I pooled responses across different waves, I ran some analyses to ensure that there were no systematic differences in responses across waves. Because the sample includes some spouse pairs, it is possible that these results are influenced by some assortative mating on risk preferences. Further work to explore this issue is planned.

Risk Aversion Measure				
	Least Risk Averse	2nd Most Risk Averse	3rd Most Risk Averse	Most Risk Averse
Sex				
Male	436 15%	345 12%	410 14%	1,790 60%
Female	384 9%	399 9%	633 15%	2,948 68%
Age Category				
0-49	37 11%	40 12%	55 16%	213 62%
50-54	167 12%	157 12%	226 17%	800 59%
55-59	242 10%	219 9%	299 13%	1,547 67%
60-64	280 11%	251 10%	355 14%	1,567 64%
65-69	52 11%	45 9%	70 14%	325 66%
70-79	40 11%	31 8%	36 9%	272 72%
80+	2 11%	1 5%	2 11%	14 74%
Most Recent Wave With Data				
1994	229 12%	208 11%	210 11%	1,259 66%
2000	65 9%	66 10%	84 12%	471 69%
2002	35 12%	19 7%	36 13%	193 68%
2004	148 12%	118 10%	184 15%	749 62%
2006	11 18%	7 11%	8 13%	36 58%
2008	332 10%	326 10%	521 16%	2,030 63%

Table 3.2: Risk Aversion by Control Sample

Results of GWAS Analysis

In Table 3.3, I report results for the 20 SNPs which attained the highest statistical significance for the specification in Equation 1 for the test subsample. The first column gives the “name,” or rs or kpg number, of each SNP with the chromosome on which it is located. In the second column, I report the effect size of the SNP. In the third column I report the p-value for each SNP. None of the approximately 1.2 million of the SNPs reached the conventional significance threshold of 5×10^{-7} established by the Wellcome Trust Case Control Consortium. Likewise, since the Bonferroni is a stricter threshold of significance, none of the SNPs meet the Bonferroni significance of 4.09×10^{-8} . In fact, none of the SNP reached a significance of 10^{-7} ; seven of the top twenty SNPs reach a significance level of 10^{-6} .

SNP	Effect size	P-value	allele	chromosome	position
kpg4560299	0.14694087	3.15E-06	GA	5	435256
kpg1703572	0.14580118	3.87E-06	AG	5	435273
rs818185	0.07931359	4.02E-06	GA	2	10648857
kpg5696875	-0.079052	4.73E-06	AG	19	9102742
rs10099909	-0.13249798	5.01E-06	AG	8	11287371
rs186493	0.08415274	6.47E-06	AC	16	3289331
kpg7673403	-0.18416093	8.39E-06	AG	12	58918789
kpg4043337	0.09221066	1.06E-05	AG	5	478655
rs7657627	-0.08238064	1.30E-05	AG	4	16412051
rs7762279	-0.13126482	1.48E-05	AG	6	32755290
rs957792	0.13700457	1.54E-05	GA	5	429989
kpg9153906	0.07763928	1.54E-05	CA	19	7924957
kpg910500	0.07500762	1.70E-05	GA	12	61497958
kpg12522368	0.07538613	1.86E-05	AG	12	61537215
kpg6975417	0.11141721	1.94E-05	GA	5	496730
kpg5076136	-0.34611196	2.02E-05	AG	12	22951899
kpg9495611	-0.28826115	2.09E-05	AG	14	67446965
rs9471770	-0.10335409	2.11E-05	AC	6	42081642
rs2079134	0.08690109	2.23E-05	AG	4	106006036
kpg9614205	0.13442179	2.34E-05	AG	5	429031

Table 3.3: Top SNPs for Test Sample

The SNPs are found on a number of chromosomes, suggesting that risk aversion is not driven by a small number of SNPs with large effects but rather, that there may be low-level genetic activity across different parts of the genome. This can be seen in the fourth column which reports the minor allele for each SNP. Columns five and six represent the

chromosome and the position on the chromosome where the SNP is found. The collection of SNPs from the list of the top twenty is spread over nine chromosomes. Six of the SNPs are concentrated near each other on chromosome 5 and there is another concentration of SNPs on chromosome 12. Otherwise, the rest of the SNPs are somewhat scattered across a number of chromosomes. These results would suggest that risk aversion is not driven by a small number of SNPs with large effects but rather, that there may be low-level genetic activity across different parts of the genome. Figure 3.1 shows the plot of the top-ranked SNPs analyzed against the $-\log_{10}$ of the p-value. This version of a “Manhattan plot” as it is often called, only displays the top of the “skyline”—the subset of SNPs with $\log_{10}(\text{p-value})$ greater than 4. In other words, this plot includes only the top 113 SNPs; most of the SNPs lie below this threshold and therefore would be at the bottom of this plot if the y-axis were extended to zero. I’ve truncated The axis is truncated to put the top-ranked SNPs into context relative to significance. As is evident, none of the SNPs makes their way up to the significance threshold of 10^{-7} , but there are a large number of SNPs that “rise above” with p-values in the range of 10^{-6} . With this sample size, it is impossible to distinguish between what are SNPs with actual effects and what is statistical noise. As sample size grows and our statistical tools become more powerful, the genetic components of risk aversion will likely reveal themselves.

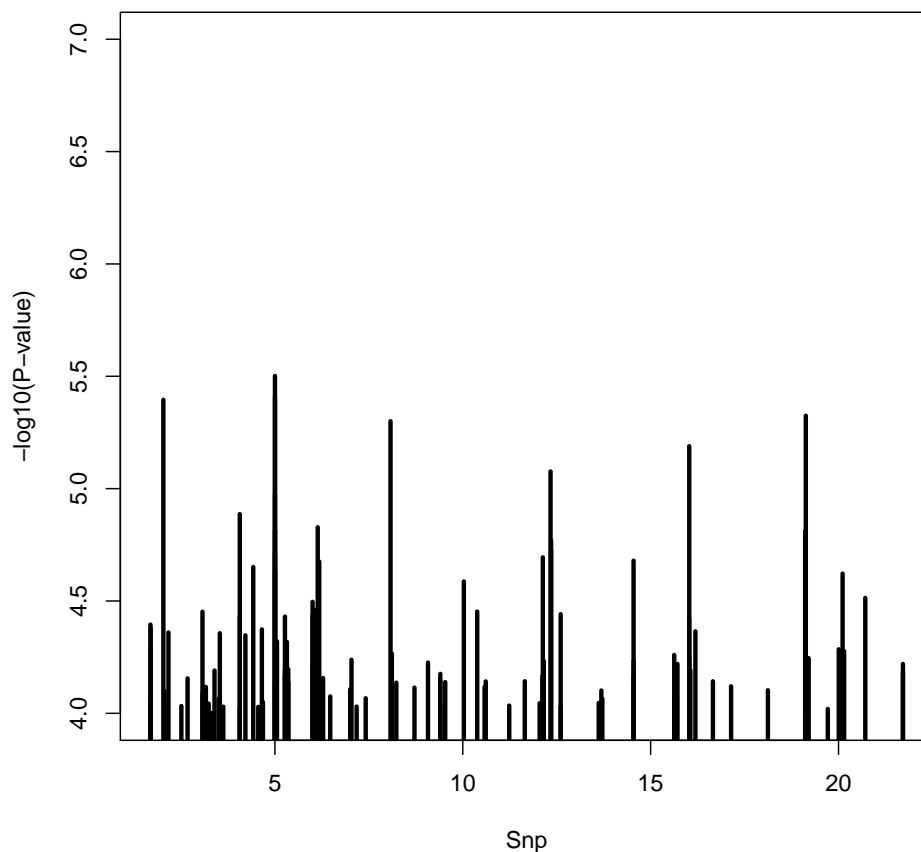


Figure 3.1: Manhattan "Skyline" Plot. Data source: HRS GWAS data

The results for Table 3.3 are based on a subset of the whole sample since part of the sample was reserved as a validation sample for positive results. Since there were no positive results to validate, the full sample can be re-run for the analysis. Table 3.4 shows the results of the GWAS run on the full sample of 10,455 individuals that have been genotyped and have at least one response to the risk aversion measure. Again, no SNPs meet the traditional significance threshold of 5×10^{-7} . So, the increase in sample size does not add enough power to detect any SNPs that meet the significance criteria. It does, however, change the ranking of top SNPs. Interestingly, only two SNPs (kcp5696875 and rs10099909) are on the list of top hits for both runs. Moreover, the chromosome with the highest concentration of top SNPs move from chromosome 5 to chromosome 7. This might suggest that these may be the two SNPs that have true effects on risk aversion.

The results from Table 3.3 are based on the specification that controls for age and gender as well as population stratification through the inclusion of the top ten principal components. I re-ran the analysis without the inclusion of the principal components. While the exact order of SNPs moves around slightly, the results remain largely unchanged, implying that population stratification is not too much of a concern in this sample.

SNP name	P-value	allele	chromosome	position
kgp6835378	3.85E-06	CA	2	19298409
kgp4319112	5.71E-06	AG	13	110916213
kgp12230322	6.40E-06	CA	2	19287912
rs2064389	7.54E-06	GA	6	69425811
kgp5070982	9.77E-06	GA	13	110917782
kgp11108464	9.83E-06	GA	7	81374817
rs9363963	9.86E-06	GA	6	69439107
kgp9523505	1.02E-05	CA	13	110917768
rs10208636	1.04E-05	GA	2	19279238
rs10099909	1.05E-05	AG	8	11287371
kgp4233366	1.07E-05	AC	14	20951094
kgp5696875	1.35E-05	AG	19	9102742
rs686642	1.38E-05	CA	13	110917768
kgp8478356	1.40E-05	AG	20	49690694
kgp11028672	1.57E-05	CG	12	133696785
kgp8459226	1.59E-05	GA	2	19299787
kgp7522087	1.69E-05	GA	10	5361694
kgp3375800	1.74E-05	AC	8	13484030
rs11895316	1.88E-05	AG	2	19284343
kgp12799769	1.98E-05	AC	11	5692543

Table 3.4: Top SNPs for Full Sample

Detection Bounds

The results from the GWAS are consistent with a picture in which large numbers of small genetic effects combine to account for the known heritability of risk aversion. Hiding within the random noise produced by sampling error with my sample of 7,313 respondents, there might be a few moderately large causal effects, but not too many and not too large. The analysis allows some quantification of these terms “too many” and “too large.” One good strategy is to calculate the statistical power function for a GWAS test statistic against a family of alternative hypotheses.

Of course there are a number of different ways of constructing families of alternative hypotheses. Here I adopt a simple approach. I develop a standard of what is meant by “large” by considering the p-statistics for SNPs that show up in the tail of the distribution, specifically the batch B of SNPs in the data with estimated p-values smaller than 10^4 . These

p-values are the ones based on the chi-square test on one degree of freedom appropriate to an additive genetic model as already shown in Figure 3.1. The strength of effect for SNP_i is measured on a log scale by

$$G_i = -\log_{10}(p_i)$$

Under the null hypothesis the true causal effects for all SNPs are zero, so that all the variation being seen is due to sampling error, I expect one in ten-thousand of all SNPs to appear in the batch B. Indeed they do. The expected number of SNPs to appear in this batch is 122; I observe 107 in our data.

Under the null hypothesis, G_i are independent exponential random variables. If I had used natural logarithms they would have been exponential variables with unit means. Since I am using logarithms to the base 10, they are exponential random variables with mean $1/\log_{10}$. The observed mean for SNPs in batch B is 4.329. The expected mean would be $4 + 1/\log_{10} = 4.434$, very close to the observed mean of 4.329. For our standard for “large”, I take effects whose true strengths measured on the G-value scale are as large as the observed mean for SNPs in B, namely 4.329. For every choice of $k = 5, 6 \dots 15$, I consider an alternative hypothesis that k SNPs have true effects this large and all the others have true effects of zero. All SNPs are subject to the sampling error found in my sample, making observed estimates vary around the corresponding true values. This construction gives us a family of alternative hypotheses H_k indexed by k .

The customary GWAS test statistic for a test of the null hypothesis of no causal effects (all noise) is $\max(G_i)$. As I have discussed, the customary rejection region for the test is $R1 = \max(G > 7 - \log_{10}(5))$, based on the traditional p-value criteria set by the Wellcome Trust Case Control Consortium. With my data, I would reject the null hypothesis with any rejection region inside the regions $R2 = \max(G > 5.501)$, which is the maximum value in my data. Figure 3.2 shows the statistical power functions for test R1 and test R2 against our family of alternative hypothesis H_k as a function of the posited possible number of causal alleles k . We see that for k bigger than 10 the customary test R1 has power greater than 92.74%. For test R2 the power is greater than 98.9% for any k bigger than 10, and power of greater than 90.26% for k bigger than 3.

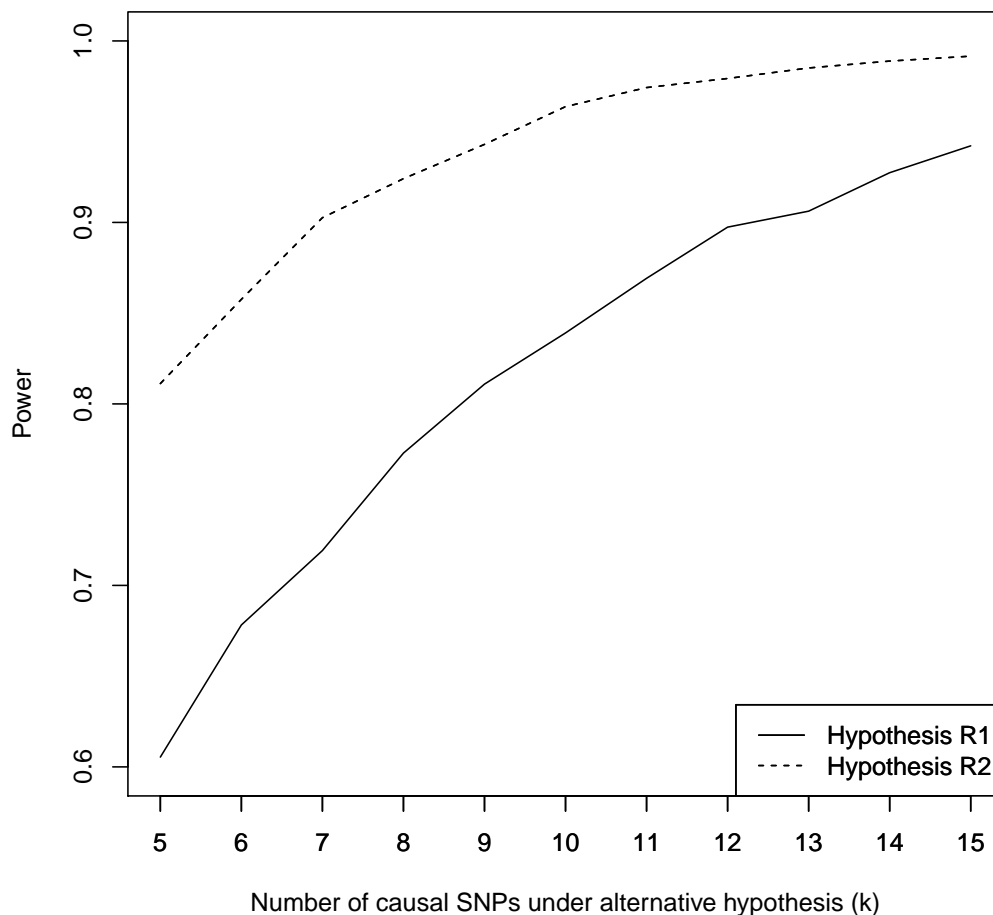


Figure 3.2: Power Calculation Against the Alternative Hypothesis

We see that it is unlikely that we would be rejecting the null hypothesis if there were more than a dozen or so SNPs with true causal effects on risk aversion as strong as the apparent effects in the batch we have described.

Higher sample sizes would allow us to strengthen these bounds. But our power function calculations already point strongly to a highly polygenic character for the heritable component of risk aversion.

3.6 Discussion

As the results have shown, my analysis reveals that risk aversion is likely to be highly polygenic in nature and it not driven by a few genetic variants with large causal effects. The results of the GWAS using a large, nationally-representative survey were unable to find any associations of single SNPs significant at the conventional threshold of genome-wise

significance required for the sample size of 7,313. The interpretation of these results is that the heritability of risk aversion must be driven by large numbers of genetic variants with causal effect sizes small enough that they cannot be detected with the current sample sizes.

I confirm these results through some analysis of detectability bounds to eliminate a family of alternative hypotheses in which a few large causal SNPs drive genetic variation in risk aversion. My findings suggest that much of the “missing heritability”, the gulf between the cumulative explanatory power of specific common variants identified to date and the overall heritability estimated through twin studies, reflects the fact that risk preferences have a complicated genetic nature that require still-larger sample sizes to identify.

These findings add to a body of accumulating evidence from studies exploring a number of economic, political and social preferences (Beauchamp et al., 2011; Benjamin et al., 2012; Fowler and Dawes, 2013) that suggest that the effects of common genetics variants explored through candidate gene studies on complex outcomes are small. The introduction of genome-wide data allows the inclusion of genetic variants that are less common and previously unexplored, and as a result, allows for a much more precise understanding of the nature of the preferences that underlie economic parameters. A study exploring a similar measure of risk aversion using a sample size of 2,900 (Benjamin et al., 2012) showed no significant SNPs associated with risk aversion; my study has expanded this scope by exploring whether some of these genetic variants would come to light using a sample size more than two and a half times the size. Continued investment in these data types and subsequent studies will ultimately reveal more about the complex genetic nature of risk preferences and other important economic parameters.

Genetic variation is an important component of risk preferences and until very recently, has been largely neglected. As large cohorts of older workers are moving into retirement, and as retirement savings are driven in growing shares by private savings, there is a growing imperative to understand the unexplained heterogeneity in individual risk aversion. The recent introduction of genetic material into social science surveys presents scholars with a unique opportunity to capture sources of variation that until recently have been nearly impossible to measure. This finding is consistent with the hypothesis that two individuals who are identical in terms of income, education, wealth, and age may still make very different portfolio investment choices. That is, faced with the same budget constraints and optimization problem, individuals with different genetic endowments may still make very different investment choices. My results suggest that genetic markers may well ultimately help us shed light on the fundamental question of why individuals differ in their willingness to take risks.

Chapter 4

GCTA

4.1 Introduction

The results of Chapter 3 suggest that risk aversion is, like many complex traits, a highly polygenic trait in which causal SNPs are not detectable at the current HRS sample size. Based on this finding, in this chapter, I extend the analysis to another genomic technique to estimate the genetic contribution to observed variation in risk aversion. Using a technique called the Genome-Wide Complex Trait Analysis (GCTA) (Yang et al., 2010), this chapter seeks to answer the straight-forward question of whether there is evidence of heritability for risk aversion.

To answer this question, the GCTA technique gathers statistical strength by estimating all SNPs simultaneously, rather than by examining the effects of each SNP individually, and thus the total genetic variance that can be attributed to phenotypic variation. The GCTA technique depends on measures of genetic variants that are near each other through a genetic phenomenon called linkage disequilibrium. Actual causal variants remain unknown, but the “common” variants that are linked to causal variants and that are captured on the commercial chips used to genotype individuals and be shared by individuals to the similar degree to the unknown causal variants. As such, there is an important distinction in the reading of this and other studies between the terms “causal” SNP and “common” SNP, where the latter refers to the actual SNPs used in the analysis. The results give an estimate of narrow-sense heritability, meaning the total amount of variation in risk aversion that can be attributed to the aggregate additive effects of all the common SNPs in the data. (I will refer to this technique exclusively as GCTA but it is also referred to as the Genomic-Relatedness-Matrix Maximum Likelihood (GREML)).

The technique starts by building a genetic relatedness matrix (GRM) between all possible pair-wise combinations of individuals using the available SNP data. Then, using maximum-likelihood approach, it calculates the total share of phenotypic variation among individuals that can be attributed to the genetic variation GRM. The result is an estimate of heritability that measures the percentage of total variation in risk aversion that can be attributed to

genetic variation among sample individuals. The advantage of this technique is that it does not rely on the statistical significance of one or many SNPs but instead takes all the SNP data as a whole. Given the lack of significant SNPs found in the GWAS in Chapter 3, this technique provides an excellent alternative.

A preview of the findings is as follows: I find no detectable levels of heritability for risk aversion using the HRS data. I posit that the most likely explanation is that the polygenetic nature of risk aversion is one that is characterized by moderate numbers of causal SNPs whose correlation with the full set of common SNPs is not sufficiently high to be captured in the “common” SNPs found in the data. This, and other possible explanations for this finding, are discussed in detail in later sections of this chapter.

4.2 Background

Since its introduction in 2010, the GCTA technique has been widely adapted among behavioral genetics. It has been applied to height (Yang, 2010), intelligence (Chabris et al. 2012; Davies et al., 2011), personality (Vinkhuyzen et al., 2012), several common diseases (Lee, Wray, Goddard, & Visscher, 2011), schizophrenia (Lee et al., 2012), and a number of political and economic phenotypes (Benjamin et al., 2012).

Two main findings emerge from this body of work. First, heritability estimates tend to be approximately half of the estimates produced in twin studies. The gap between these estimates is often referred to as “missing heritability” and has been hotly debated in the field of behavioral genetics for a number of years. The most common explanations for the missing heritability are either that the gap in estimates between twin studies and GCTA studies is due to upward bias in the twin studies or to inaccurate tagging of SNPs in the genetic data. A more detailed discussion of this issue is found towards the end of this chapter.

Second, the amount of genetic variance explained by any one chromosome is proportional to chromosomal length. The longer the chromosome is in length, the more genetic variability is associated with it. This has been found in intelligence, economic and political phenotypes, and a number of other complex traits that have been tested.

The GCTA technique was first tested on human height because of its strong known heritability (Fisher, 1928; Hewitt, 1999; Wood et al., 2014; Yang et al., 2010). Height in humans is a classic quantitative trait, is easy to measure, and studies for well over a century have used it as a model for investigating the genetic basis of complex traits (Fisher, 1918; Galton, 1886). Rare mutations that cause extreme height have been found, but these do not explain much of the variation in the general population. The heritability of height has been estimated at approximately 80% (Fisher, 1918; Macgregor et al., 2006; Silventoinen et al., 2003); and the GCTA estimate found in Yang et al. (2010) is 45%.

Findings traits are heritable have been fairly ubiquitous and a number of studies have used GCTA-like techniques to calculate these heritability estimates. Perhaps most relevant to this study is the work of Benjamin et al. (2012) who used a Swedish sample to study 10 traits measuring various economic and political preferences, including risk aversion. Using

GCTA, they did not find a significant finding on a heritability estimate for risk aversion. That is, the narrow-sense heritability estimate for risk aversion was measured at 13%, but the p-values showed the estimate could not be distinguished as being statistically different from zero. However, significant heritability estimates were found for educational attainment (16%), trust (24%), and political views on economic policies (34%).

A wide range of other complex socio-behavioral traits have also been studied. Estimates of the narrow-sense heritability of extraversion and neuroticism taken from SNP data from approximately 12,000 unrelated individuals are 12% and 6%, respectively (Vinkhuyzen et al., 2012). Using GCTA analysis, Boardman, Domingue, and Daw (2014) provided evidence that education has a heritability of 33%, BMI of 43%, depression of 19%, and self-rated health of 18%. Subjective well-being has common narrow-sense heritability estimates between 12% and 18% when using GCTA techniques (Reitvald et al., 2013). Estimates of the heritability of postpartum depression range from 6% to 22% (Byrne et al., 2014); and the ability of individuals to recognize various facial expressions ranged from 30% to 40% (Dickie et al., 2014).

Studies on intelligence suggest narrow-sense heritability estimates of 40% for crystallized-type intelligence and 51% for fluid-type intelligence (Davies et al., 2011). Authors of this study stated that their “results unequivocally confirm that a substantial proportion of individual differences in human intelligence is due to genetic variation, and are consisted with many genes of small effects underlying the additive genetic influences on intelligence” (Davies et al., 2011, p. 256). Interestingly, estimates of the heritability of intelligence increases from about 20% in infancy to as high as 80% in later adulthood, depending on the construct of intelligence (Plomin & Deary, 2014).

A similar number of reports can be found that study traits proximal to disease. Perhaps the most notable of these studies comes from the author of the seminal GCTA paper on height (Yang, 2010), who used a sample of Koreans to analyze 49 human quantitative traits that are relevant to human diseases. These traits included measures of obesity, liver and kidney functions, diabetes, blood counts, and others (Yang et al., 2014). They found that 43 of the 49 traits had non-zero estimates of narrow-sense heritability, ranging from 7.8% to 76.8%.

Taken as a whole, this literature shows heritability estimates to vary widely. This is to be expected since the underlying “true” heritability, which is unknown, differs across phenotype. For that reason alone, heritability estimates using any technique will and should differ. However, methodological constraints may also impact estimates to a second-order degree in two ways. First, the accuracy of the measure of the phenotype differs; some phenotypes are more or less precisely measured. For example, specific diseases may be measured more precisely than some economic or social preferences, which may have multiple ways in which to measure or which may be subject to greater measurement error. Second, difference in the size of the SNP sample can contribute to less precise estimates. Because the method depends on capturing genetic variance from SNPs that are in linkage disequilibrium to the unknown causal variants, both the size and the nature of the SNP sample can have marginal effects on the heritability estimates.

Since the initial introduction of the GCTA technique, studies have emerged that either extend or critique this method. The partitioning of the genome-wide variation into individual chromosomes, for example, led to the common finding discussed above that the amount of genetic variation explained by any given chromosome is proportional to the length of that chromosome. In addition, studies have used GCTA methods on subsets of SNPs to measure the share of genetic variance across different sets of SNPs, which helps to create boundaries on the total number of causal SNPs that may exist for highly polygenic traits. For example, in estimating the heritability of height using a sample of over 250,000 individuals, Wood et al. (2014) tested a different number of variants, and showed that the most strongly associated SNPs 2,000, 3,700, and 9,500 —explained 21%, 24%, and 29% of phenotypic variance, respectively. The results of this analysis led them to conclude that height is not likely to be driven by more than 1,000 causal SNPs.

A number of newer studies have also begun to look more closely at the proper conditions necessary for the validity of SNP-based heritability estimation (Lee & Chow, 2014) as well as for testing the assumptions implicit in the GCTA technique (Conley et al., 2014; Plomin, 2014). For example, a key assumption in the approach that uses a genetic relatedness estimation through the maximum likelihood model, as GCTA does, is that genetic relatedness between individuals is completely independent to environmental similarity. A recent study using the HRS (and two other surveys) showed that two unrelated individuals are in fact more likely to have been reared in a similar environment if they are genetically similar and that this effect was not eliminated by controls for populations structure (Conley et al., 2014). However, when the authors included this environmental confound in the GCTA models, heritability estimates did not change substantially, and thus they concluded that the potential bias in GCTA estimates is probably minimal.

Da and Wang (2013) have introduced a technique for a joint prediction of heritability for genetic variation attributed to both additive and dominance effects, along with an associated software package called GVCBLUP. This technique has been applied to animal breeding models but also has the potential for adaptation to human genetics. Zuk et al. (2012, 2014) have released a series of papers looking to expand the estimates of narrow-sense heritability. First, they tackle the issue of epistasis by testing the assumption implicit in GCTA methods that there is no genetic interaction among variants (Zuk et al., 2012). They use Crohn's disease to show that up to 80% of the missing heritability could be attributed to genetic interactions among three genetic pathways. In a more recent paper, they described an analytical framework for the design of rare variants association studies (Zuk et al., 2014), rather than the common variants used in the current GCTA techniques.

4.3 Data and Methods

Data

As detailed in Chapter 2, both the genotype and the phenotype data used in this chapter come from the HRS.

The genetic data are the same genome-wide association data used in Chapter 3, and the same number of quality control measures that were applied in the GWAS analysis in Chapter 3 were also applied to this data. In addition, and in keeping with standard practice, individuals with a genetic relatedness of greater than 0.025 were excluded from the analysis.

The phenotype data for risk aversion use the continuous measure of risk first developed by Miles et al. (2008) and discussed in detail in Chapter 2 of this dissertation. As a reminder, this measure is a composite measure based on individual-level responses from the ordinal risk aversion question asked in the HRS. It is a measure of risk tolerance where 0 is most risk averse and 1 is most risk seeking. The mean value for this sample is 0.23 (s.d. 0.08) (details on this measure can be found in Chapter 2). Because the construction of the risk measure requires responses in at least two survey waves, the sample size is restricted to 5,411 individuals.

As a point of comparison, the GCTA analysis is also run for two other phenotypes, height and cognition. Height is measured in inches and is measured in either 2006 or 2008, corresponding to the same year the individual was genotyped. Cognition is a composite measure from the HRS that is intended to capture a total cognition score based on a series of individual cognition questions asked in the HRS. It sums the responses of a series of recall questions, where respondents were asked to recite back a list of random words both immediately and five minutes after they were given, as well as some self-reported mental-status questions. Total cognitive scores range from 0 to 35, with higher scores indicating stronger recall and self-reported mental status. The values are again taken from responses from either the 2006 or 2008 survey to correspond to the same year the individual was genotyped. The mean cognitive score in the sample is 22.02 and the standard deviation is 5.16.

Controls in the analysis include respondent sex, age in the year of genotyping, and the first ten principal components used to control for population stratification which are provided by the HRS with the GWAS data.

Methods

The GCTA method uses SNP data to estimate narrow-sense heritability through two steps. First, it builds a genetic relatedness matrix (GRM) between all possible pair-wise combinations of individuals using the available SNP data. The result is a $N \times N$ matrix, where N is the number of individuals in the sample and every entry in the lower or upper diagonal represents the genetic relatedness between each pair of individuals. Then, using a restricted maximum-likelihood technique, it calculates the total share of phenotypic variation

among individuals that can be attributed to the genetic variation found in the GRM, resulting in an estimate of narrow-sense heritability.

The power of the method comes from comparing not just two groups, like monozygotic and dizygotic twins, but from the millions of pair-by-pair comparisons in samples of thousands of individuals. GCTA provides a lower-bound estimate of narrow heritability that does not rest on the same set of assumptions relied on in twin studies. A key assumption behind GCTA is that among individuals who are not in the same extended families, environmental factors are uncorrelated with differences in the degree of genetic similarity, or “relatedness.” In this analysis, genetic relatedness is directly estimated from the SNP data, unlike in behavior genetic studies, where expected relatedness (inferred from the family pedigree) is used.

GCTA detects only those genetic effects tagged by the common SNPs that are incorporated in commercially available DNA arrays used in GWA studies. Because these SNPs are only imperfectly correlated with the causal variants, relatedness with respect to the causal variants is measured with error. Consequently, the estimated relationship between phenotype and genetic relatedness is attenuated, and hence the estimator is a lower bound for narrow-sense heritability (Yang, 2010). Additionally, as mentioned above, GCTA is limited to detecting the additive effects of SNPs; it cannot detect gene-gene or gene-environment interaction.

The team of authors of Yang et al. (2010) have developed an accompanying open-source software package by the same name which I used for this analysis. I ran the analysis in the following steps. First, I create the GRM matrix for all individuals with a relatedness greater than 0.025. I ran the GRM for each one of the 22 autosomal chromosomes separately and then merged them together because of the large file. Then, I performed the restricted maximum likelihood analysis (REML) for each one of the three phenotypes described above (risk aversion, height, cognition) using the merged GRM. For each one of the three REML runs, I included the following controls: sex, age at year of genotype, and the first 10 principal components. The results are described below.

4.4 Results

Table 4.1 describes the output from the GCTA analysis for the main phenotype of interest, risk aversion, as well as two comparison phenotypes, height and cognition. The first row, $V(G)$, refers to the additive genetic variance for the trait. The second row, $V(E)$ refers to environmental variance for the phenotype. The assumption implicit in GCTA is that the total phenotypic variance is due either to genetic variance or to environmental. As such, V_p , is the sum of both components of variance, since the variance is assumed to be additive. The estimate of interest, $V(G)/V_p$, measures the share of total variance attributed to genetic variance, or in other words, narrow-sense heritability. The remaining rows describe the standard error and associated p-values for the heritability estimate (ie. $V(G)/V_p$) and the total number of individuals included in the analysis.

	Risk Aversion	Height	Cognition
V(G)	0	0.001774	7.60639
V(E)	0.005108	0.002773	12.470267
Vp	0.005108	0.004548	20.076657
V(G)/Vp	0.000001	0.390133	0.378867
S.E.	0.03898	0.037349	0.037574
P-value	0.5	6.92E-13	3.89E-16
N	5411	7171	7253

Table 4.1: Results from GCTA Analysis for Risk Aversion, Height, and Cognition

As Table 4.1 shows, the heritability estimate for risk aversion in this sample is nearly zero and not statistically significant. This corroborates an earlier finding from Benjamin et al. (2012) which also found no statistically-significant heritability estimates for risk, albeit using a very different sample and a different measure of risk. Because the restricted maximum likelihood does not, by default, allow non-negative values, the estimated variance is constrained to be at or above zero. (Additional runs without the constraint did not change the findings.)

In contrast, both height and cognitive score have positive heritability estimates that are strongly statistically significant. The heritability estimate on height is 39%, which is well within a reasonable range of recent GCTA estimates on other samples. This is only slightly lower than the typical findings in the literature that the SNP-based approach to heritability is approximately half of the estimates based on twin studies. Twin study estimates are around 80% and Yang et al. (2010) found a GCTA estimate of 45%.

Likewise, cognition, which has heritability estimate of 37.9%, is statistically significant. Interpretation of these heritability estimates within the wider context of other heritability estimates in the literature is tricky for a few reasons. Firstly, cognition itself is a broad concept and can incorporate a number of definitions including fluid intelligence, memory, spatial or verbal reasoning, and different definitions of intelligence. Moreover, cognition and cognitive capacities vary largely over the life course, so that the age of the sample is particularly important. The measure here is a composite intended to capture immediate and delayed recall (memory) as well as self-reported mental status. As such, there are two studies that may be most relevant. A twin study by (Finkel et al., 1994) used a sample of adult twins aged 65 to 85 and found the heritability for a general cognitive score to be 54%. Likewise, a twin study McClearn et al. (1997) use a Swedish sample of individuals over 80 years old without any major cognitive or motor impairment and, using a similar

composite measure of cognition, estimate heritability to be 62%. The SNP-based estimate of 37.9% that I report here supports the common finding of being approximately half of the twin-based estimates.

Despite the twin studies literature demonstrating moderate levels of heritability, risk aversion has no significant heritability estimate. There are a number of possible explanations for this finding, all of which are discussed in the following section.

4.5 Discussion

Given the evidence I summarize in Chapter 3, there is strong reason to believe that risk aversion is indeed heritable. A number of twin studies and candidate gene studies have all shown evidence of heritability in risk aversion. As such, the estimate in Table 4.1 of zero heritability is a bit puzzling. There are a number of possible reasons for a zero estimate of heritability of risk aversion from this sample. I will discuss the most likely primary reason and will also discuss other explanations that may either work also contribute to this null estimate.

My approach for Chapters 3 and 4 is to examine the genetic nature of risk aversion from two extreme perspectives. First, in Chapter 3, I take a granular approach by characterizing the genetic nature of risk aversion from the individual SNP perspective, examining one SNP at a time. In this chapter, I take the opposite approach, taking the effects of the entire set of SNPs among which the causal SNPs are assumed to be hidden. Since the heritability does not appear through either extreme approach, it is feasible that the heritability is somewhere in the middle; the genetic nature of risk aversion is one that is too polygenic to be found with the single SNP GWAS approach but too sparse or uneven for the GCTA approach. A likely conclusion is that the genetic architecture of risk aversion is of a moderate polygenicity in which causal SNPs are numerous, but not too numerous, and perhaps spread unevenly across the genome. In such circumstances, it may be very difficult to uncover heritability of risk aversion through current genomic techniques.

Previous papers have found that the SNP-based heritability estimates are between one-quarter and one-half the size of twin study estimates. One interpretation of the gap is that genotyped SNPs tag less than half the additive genetic variation in those traits. If the causal alleles are unevenly scattered across the genome, it is possible that there is a greater likelihood that the common SNPs used in the GCTA analysis miss these causal alleles, and as such, the heritability estimates of risk aversion using this technique is not captured.

Of course, the GCTA approach used here and the twin studies do rely on different assumptions. The merits and critiques of twin studies are hotly debated in behavioral genetics and this paper is not the place for a treatise of the matter. Still, a few things are worth noting. The twin studies rely on a shared environment assumption that states that monozygotic and dizygotic twins share the same similarity in environment and as such, any difference between and among twins, regardless of zygosity, can be entirely attributed to genetic variation. The issue of the shared environment assumption is one that is generally debated. It

is feasible that this shared environment may hold up well with certain phenotypes and less well with others. Given the complex nature of risk aversion, perhaps the shared environment assumption fails and as such, the heritability estimates from twin studies are overstated.

Moreover, risk aversion may be a broad phenotype in which more precise measurement is needed in order to capture the true phenotypic variation. The various forms of measurement for risk aversion have been debated at length (Falk and Heckman, 2009). Economists and psychologists have developed a variety of methodologies to elicit risk preferences that encompass experimental methods, field experiments, and observing behaviors deemed risky. The measures of risk tolerance take a number of forms and much debate exists around what these measures capture and whether the concepts are identical, overlap, or are domain-specific. Importantly, heritability estimates based on twin studies are also based on a number of different measures of risk. It may be that without a more precise definition of the phenotype, heritability estimates are too noisy.

Chapter 5

Do Stated Risk Preferences Predict Behavior Over Time?

5.1 Introduction

Older Americans vary in their preparation for the financial burdens of retirement and old age. These differences cannot be attributed entirely to income, education, cognitive ability, family background, or a number of other factors. Evidence suggests that differences in risk aversion are an important driver of the variation in wealth accumulations of individuals. Risk preferences are not entirely understood; though they have been well-studied empirically, there are still large amounts of heterogeneity that remain unexplained. Gender, age, and education levels are all predictors of risk aversion but often can explain only a small portion of the variation Americans display in risk aversion (Barksy, 1997).

What is less well understood is to what extent these relationships hold over time, both across and within individuals. This chapter seeks to answer two questions. First, does the relationship between hypothetical risk and measurable risky behaviors remain consistent across both time (cross-sectionally) and among individuals (longitudinally)? Secondly, does the change in the riskiness of bonds following the 2008 recession change the relationship in portfolio allocations relative to stated risk tolerance for individuals? This paper is organized as follows. Section 2 reviews the literature related to measures of risk, as well as a brief description of the nature of risky assets during the Global Financial Crisis. Section 3 describes data used and Section 4 discusses the results. The paper will conclude in Section 5.

5.2 Background

Measures of risk aversion vary largely

Given its centrality in everyday life, empirical research has examined the important heterogeneity in risk preferences across populations including differences across gender (Schu-

bert, Brown, Gysler and Brachinger, 1999; Powell and Ansic, 1997; Eckel and Grossman, 2002); family background (Hartog et al. 2002); work characteristics (Praag and Cramer, 2001); and educational attainment levels (Brunello, 2002), as well as across different contexts of risk (Weber, 2002; Soane and Chmiel, 2005). Given the particular importance of risk preferences in decisions made at older ages, much of this empirical work has focused on the implications for retirement decisions and savings (Karatzas, Lehoczky and Shreve, 1987; Hurd, 1990; Bodie, Merton, and Samuelson, 1992; Sunden and Surette, 1998). With all of these studies has come a plethora of measures of risk preferences. These measures typically take one of three forms: survey-based assessments such as respondent's answers to hypothetical lottery gambles, experimental evidence, or inference from observed decision-making in financial, health or insurance markets.

This study makes use of a measure of risk aversion first introduced into the HRS by Barsky et al. (1997) that categorizes risk aversion through responses to a series of hypothetical gambles on lifetime income through a specific job and that is detailed in Chapter 2. The work of this paper is built upon the work put forth by Barsky et al. (1997) that established the measure of risk aversion used in this analysis. Since its inception, a number of other studies have validated its use. Additionally, other measures of risk preferences using answers relating to hypothetical questions have been developed and used to see how predictive they are of actual risk behavior.

Studies risk show consistency but longitudinal studies are sorely lacking

The Barsky et al. (1997) measure has been used to predict a variety of behaviors with mixed results

First introduced in the HRS and analyzed in Barsky et al. (1997), these questions are now included on several other large household surveys, and are widely used in empirical studies of risk-taking behavior. Further, these questions have been externally validated by evidence of statistically significant associations with a wide range of behaviors including financial investment, insurance demand, smoking, drinking, education, marriage and fertility. Variations of these questions have also appeared in the Panel Study of Income Dynamics (Luoh and Stafford, 2007), a 1997 survey of French households (Arrondel, 2002) and a 1998 survey of Dutch households (Kapteyn and Teppa, 2002).

Lusardi (1998) found that the categorical measure of risk aversion was significantly associated with wealth accumulation and Rosen and Wu (2004) reported that a risk-taking indicator was associated with risky asset ownership. Dave and Saffer (2007) found a negative and significantly relationship of risk aversion and alcohol consumption.

Anderson and Mellor (2008) use the Barsky measure to predict a variety of health behaviors including purchase of health insurance, use of preventative medical care, and engaging in behaviors that increase mortality risk including cigarette smoking and seat belt use. They find that after controlling for a number of demographic and economic traits, the Barsky mea-

sure of risk is negatively and significantly associated with cigarette smoking, heavy episodic drinking, being overweight or obese, and seat belt non-use.

This measure of risk exists in other nationally-representative datasets

A nearly identical measure to the HRS was introduced into the Panel Study of Income Dynamics (PSID) in 1996. Using this measure, Charles and Hurst (2003) find a strong correlation between parental risk tolerance and child risk tolerance and show that adult children with higher levels of risk tolerance were more likely to own a business and more likely to own stocks than those with the lowest levels of risk tolerance. Several studies show similar relationships to non-financial decisions. Kan (2003) found that risk aversion was negatively and significantly associated with job changes and residential moves. Brown and Taylor (2007) show that risk aversion is negatively associated with educational attainment and Schmidt (2008) reported higher levels of risk tolerance associated with delayed marriage, earlier births at young ages and delayed fertility at older ages.

Likewise, studies from the National Longitudinal Study of Youth (NLSY), that also has a nearly identical measure of risk, show similar relationships. Spivey (2007) use data from NLSY79 to demonstrate that risk aversion is associated with the timing of first marriage. In the NLSY, evidence regarding the predictive power of this measure is more mixed as it relates to health behaviors and healthcare. Lahiri and Song (2000) reported that risk aversion had a negative effect in a model of smoking initiation but had no effect on smoking continuation. Sloan and Norton (1997) reported no relationship for risk aversion in a model of long-term care insurance demand, and Picone et al. (2004) found that risk tolerance had either insignificant effects or effects of the wrong signs on demand for preventative medical tests.

Other measures of hypothetical risk have been developed

Guiso and Paiella (2005) use hypothetical willingness to pay for a risky asset to examine decision making in the 1995 Bank of Italy Survey of Household Income and Wealth. Risk aversion indicators have a significant effect on the likelihood of having a chronic disease but unexpected negative effects in models of health insurance ownership.

Perhaps the most well-known additional measure of risk preference was established by Dohmen et al. (2005). They designed a hypothetical question that asked how much of a 100,000 euros that had just been won in a lottery would be invested in an investment project that either doubles or halves the amount invested. Respondents were also offered choices between a “safe value” and a specified lottery; successively increasing the safe values. They find a significant correlation between a set of behaviors defined as risky and their hypothetical measure of risk and conclude that “in this sense, we qualify the conclusions derived by Barsky et al. (1997) and Guiso Paiella (2005).” Their findings also show that individual risk perceptions vary significantly across domain and that careful prediction of behaviors related to risk should consider domain-specific risk questions.

Dohmen et al. (2005) also used a question on self-reported attitudes as another means of measuring individual risk preferences. More than 20,000 subjects from German households were surveyed about their general willingness to take risks on a scale of 0 to 10. A binary measure of risk tolerance constructed from this scaled was used to predict various behaviors and were found to have positive and significant effects on smoking, stock investment and self-employment.

Kruse and Thompson (2003) studied a hypothetical willingness to pay for a risk mitigation device. They compared subject responses about willingness to pay for a lock that reduced the chance of burglary to subject play in an experiment with identical losses and probabilities, but with real money at stake. The experimental and survey-based measures were consistent on average, but on an individual level they were consistent for less than one-quarter of the subjects. Lusk and Coble (2005) compared subject responses to survey questions on the willingness to consume genetically modified foods to subject decisions in a lottery choice experiment. Their main finding is that the experimental measure was significantly associated with risk-taking behavior.

The HRS risk measure has been used few times in longitudinal work

Though multiple waves of the HRS data are available, surprisingly few studies have traced both the stability of the responses themselves as well as their predictive power over time. The only study I know of to date is that by Sahm (2007). Sahm finds that, while risk tolerance changes with age and macroeconomic conditions, individuals have relatively constant relative risk aversion. No follow-up work to this has been done to my knowledge.

Guiso et al. (2013) draw on a repeated survey of a large sample of Italian Bank customers to measure individual risk attitudes using a Holt and Laury (2002) strategy. Respondents were asked to choose between a fixed lottery and different safe amounts. They investigated its stability based on the 2008 financial crisis. These authors found that changes in risk aversion after the crisis were correlated with portfolio choices, but not with wealth, consumption habits or background risk.

Mandal and Roe (2014) pair the HRS data with the same measure from the NLSY between the years of 1993 and 2007 to examine the age effects of risk tolerance. However, they do not examine the validity of the predictive power of this measure over time. Jung and Triebiech (2014) use data from the Osaka Panel Survey to examine whether self-reported measures of risk aversion vary over time and find that there exists a component of risk aversion that is in fact time-variant and susceptible to macroeconomic and personal shocks. However, while their survey does include the Barsky measure, they use another self-reported measure of risk tolerance as their measure.

Buccioli and Miniaci (2011) use the US Survey of Consumer Finance from 1998 to 2007 to study households' portfolio risk bearing and find that risk bearing fell in the 2000s. However, they use risk based on financial portfolios, not self-reported measures.

The relative paucity of studies examining the relationship between elicited and behavioral measures of risk provides an ample opportunity to leverage the multiple waves of the HRS

data to extend the analysis. Just as importantly, the newest wave of the HRS provides information after the global financial crisis, allowing the opportunity to explore to what extent, if at all, these relationships changed.

The Great Recession changed the nature of risk/non-risky assets

The financial crisis that began in 2008 has a major impact on retirement finance. Because of the economic situation, the role of traditionally non-risky assets, such as treasury bills, suddenly turned. Traditionally, treasury securities are considered the most conservative, or “safe” investment, along with other forms of bonds and money market funds. These instruments have stood for decades as a bastion of safety in the turbulence of investment markets. Importantly, the guarantees that stand behind these securities are indeed regarded as the cornerstone of both domestic and international economy, making them attractive to both individual and institutional investors. The greatest advantage of treasury securities is that they are unconditionally backed by the full faith and credit of the US government; investors are guaranteed the return of both their interest and principal that they are due, as long as they hold them to maturity. Treasury Bills specifically have a yield considered to be the definitive risk-free rate of return by financial analysts and market technicians. As such, individuals are counseled to move a greater share of whatever financial assets they have into these investment vehicles. Empirical evidence largely supports this, as greater shares of older Americans with financial holdings have higher shares in these traditionally conservative assets. Traditional lore is that investors with short-term horizons, like older Americans at or nearing retirement, should minimize their risk, therefore investing a majority of their financial portfolio in T-bills and other low-risk assets. The greatest risk for a T-bill is the default risk, that the United States government would default on its debt obligations. Historically, default risk was not very likely and the risk-free rate is rarely called into question. That is, until the economic environment falls into disarray, as in the economic climate of 2008 and the financial collapse that followed.

5.3 Data and Methods

This study uses two separate measures of risk. The first has been detailed in Chapter 2 and is the same measure of risk used in the GWAS analysis used in Chapter 3. As a brief reminder, this measure is a hypothetical question asking respondents to choose between two jobs with varying probabilities on income. The questions separate the respondents into four distinct risk preference categories, from least risk-averse to most risk-averse, and allows one to estimate specific relative-risk coefficients for sample individuals.

Barsky et al. concluded that this measure is positively related to a number of risky behaviors and provide evidence about the validity and usefulness of measures of this preference parameter (Barsky et al., 1997). Since introduction, this measure of risk preference

has been used in a number of studies attempting to measure risk preferences (Smith et al., 2004; Schulhofer-Wohl, 2007; Evan and Smith, 2010).

Gender	N	%
Male	8,973	44%
Female	11,309	56%
Age		
0-49	241	1%
50-54	3314	16%
55-59	6340	31%
60-64	6700	33%
65-69	1368	7%
70-79	1543	8%
80+	120	1%
Risk Aversion		
Least Risk Averse	2,432	12%
2nd Most Risk Averse	1,991	10%
3rd Most Risk Averse	2,626	13%
Most Risk Averse	13,233	65%
N	20,282	

Table 5.1: Descriptive Statistics for Study Sample. Data Source: Author’s tabulations from the HRS.

The second measure used is a cardinal measure that was developed by the same authors who introduced the original ordinal measure mentioned about (Kimball et al. 2008). Following the release of this set of questions and the original paper that evaluated its performance (Barsky 1997), a proxy measure of relative risk aversion was developed from the original ordinal responses (Kimball et al., 2008). This relative risk measure had two distinct features. First, it created a cardinal measure from the original four categories so as to create a range of values rather than a set of distinct values. Secondly, it uses multiple responses of individuals across survey waves to correct for measurement error in the original values. This chapter uses the measure of risk aversion, where a higher value represents higher risk aversion. Values run from 1.2 to 16.6. The statistical model and associated details can be found in Kimball et al., 2008. Initial evaluation of this cardinal measure shows that results

are consistent with Barsky's original findings. In the multivariate analysis, this paper uses this cardinal measure as a measure of risk tolerance.

Table 5.1 details some of the descriptive statistics of the sample. The total sample size is 20,282 respondents and includes any respondent or spouse with at least one wave of response for the risk survey question. The sample is nearly 56 percent female and, as is consistent with the sampling procedure, the majority of the sample is between the age of 50 and 65. Although the HRS sample is restricted to individuals aged 50 and over, spouses are included, and as such the youngest sample member is 25. The oldest person in the sample is 92. As noted earlier, the set of risk aversion questions allows a categorization of four progressive categories of risk averse, with the value of four representing the most risk averse. In this sample, approximately 65% of the sample is categorized as most risk averse. The other respondents are fairly evenly split between the remaining three categories, with just over 10% being the most risk-seeking.

Given the evidence that risk aversion increases with age, it is not surprising that there is a high share of individuals categorized as strongly risk averse. Still, these distributions are not wildly different from younger populations or the population as a whole. Work from the NLSY that uses an identical question to elicit risk aversion for a younger population shows that the share of individuals categorized as most risk averse ranges between 46 to 54 percent, depending on the year of analysis (Spivey, 2010). Likewise, the share of individuals between the ages of 20 and 69 considered most risk averse using the identical measure from the PSID is 49.1 percent (Kimball et al. 2009).

Gender		N	%
Male		5,218	44%
Female		6,398	56%
Age (2010)			
0-49		1,250	11%
50-54		3,988	34%
55-59		4,010	35%
60-64		1,939	17%
65-69		339	3%
70-79		84	1%
80+		84	1%
N		11,616	

Risk Aversion						
Obs	Mean	Std. Dev	Median	Min	Max	
11616	0.206	0.068	0.168	0.087	0.732	

Table 5.2: Summary Statistics for Cardinal Measure of Risk Aversion. Data Source: Author’s tabulation of HRS data.

Table 5.2 shows the summary statistics for the sample that have the cardinal measure. Because of the manner in which the cardinal measure was constructed, which required that every respondent have a response in at least two of the survey waves, the sample size is smaller at 11,616. The distribution of the demographics is similar to that for the ordinal measure; the higher shares of younger individuals most likely reflects the increased likelihood that spouses are included because of the need to include only those with multiple wave responses.

Figure 5.1 shows the distribution of risk tolerance for the sample using the proxy measure. As a reminder, this is a measure that ranges from 0 to 1, with the value of 1 representing the least risk averse. The minimum value of risk tolerance in this sample is 0.087 and is shared by two respondents while the maximum is at 0.732. The mode can be seen clearly at 0.168, with nearly 40% of the sample sharing that value of risk tolerance. This value is nearly the median as well. Still, there are a number of relatively higher values shared among some of the respondents, which can be seen as high points in the histogram. Nearly 1,500 respondents have a risk tolerance between 0.232 and 0.261 and another 905 respondents have a risk tolerance of 0.326. The distribution is clearly skewed to the left; at the risk tolerance value of 0.5, less than .5% of the sample remain.

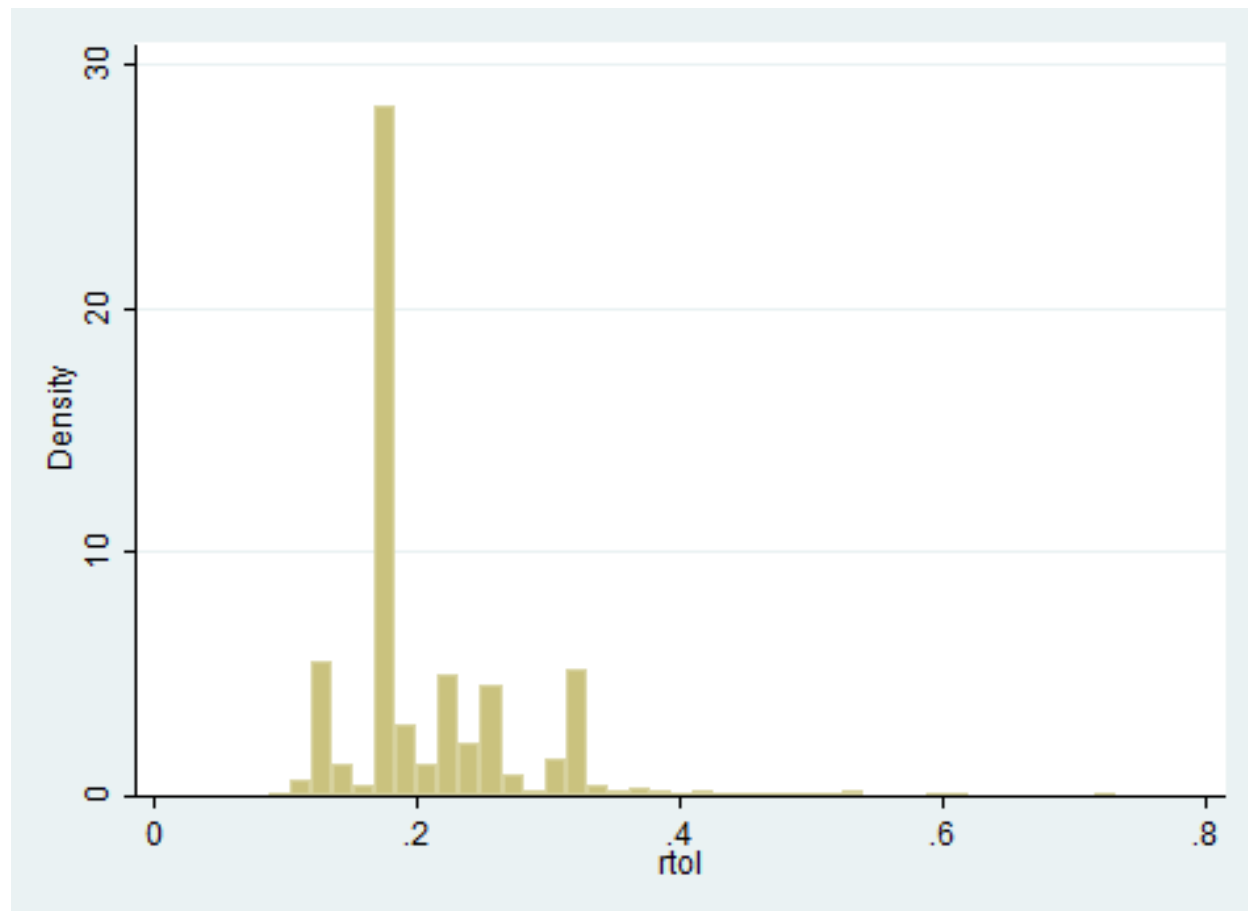


Figure 5.1: Distribution of Risk Aversion using Cardinal Measure.

The empirical strategy for this paper follows three parts. First, I will replicate and expand upon the original analysis of this measure detailed in the Barsky paper (1997). With six waves of data, I am able to see to what extent these stated preferences remain constant. Secondly, the analysis will examine the stability of these preferences over time for those that have multiple responses. Given that there were changes to the number of individuals asked the risk preference question in each wave, this analysis is limited to a subset of total respondents. There is a lively discussion in the literature as to what extent risk preferences are fixed and immutable over time (Reynaud et al., 2012; Sahm, 2012; James, 2007) and this section attempts to contribute to this knowledge. Finally, the analysis will focus on the measures of financial assets in later waves to see to what extent there were changes in the relationship between stated risk preference in anticipation of the changing financial climate.

5.4 Results

Cross tabulations replicate Barsky's original finding and, cross-sectionally, remain very consistent over time (1992 to 2006).

In this section I examine the extent to which measured risk tolerance predicts behavior over time. A body of empirical work has identified a number of behaviors that are associated with risk preferences. I examine to what extent these correlations hold true over time. These behaviors include smoking behavior, drinking, educational attainment, self-employment, whether one is a Westerner, and whether one has purchased health insurance and life insurance. The justification for the choice of these particularly risk behaviors is described in greater detail in each of the subsections below. Table 5.3 shows correlations between the measures of risk and a variety of risky behaviors for all of the waves in which the risk measure was asked. This table shows that the relationship between risk preferences and risky behavior remains fairly stable over time. The distribution of individuals categorized in risk aversion does not change over the six survey waves in which the question was included. One might have expected an age effect might appear, in which higher shares of individuals are categorized as more risk averse as the panel ages. However, because the HRS study design is such that the panel is refreshed with a younger cohort every 6 years, the sample is not necessarily older in later waves than it is in earlier waves. Moreover, because the question was not asked to the entire sample every time, the average age of the sample is not necessarily two years older in every subsequent survey wave. The effects of age will be examined in closer detail in later sections of this chapter.

Tables 5.3 and 5.4 also demonstrate that the direction of the relationship is as one would expect a priori; higher shares of individuals who identify as less risk averse (more risk tolerant) engage in more risky behavior. There are a few exceptions in which the relationship appears to be more U-shaped than linear, with higher shares of risk tolerant individuals engaging in moderate but not high levels of risky behavior. Such an example includes drinking. These relationships will be examined in more detail in the subsequent sections.

It is important to note that the results from Table 5.3 are cross sectional, and as such, it is possible that there is actually a lot of individual-level movement across categories that is not captured in the aggregate. While there is evidence that this might be the case, which I present in a later section of this paper, it is also true the nature of the panel data of the HRS is such that most of the respondents remain the same across any given time period.

% Choosing Response	Wave 1: 1992				Wave 4: 1998				Wave 5: 2000				Wave 6: 2002				Wave 7: 2004				Wave 8: 2006			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Total Sample	12.8	10.9	11.6	64.6	13.7	9.38	15	61.9	12.5	8.85	14.3	64.4	12.29	9.57	15.6	62.54	12.14	9.12	17.46	61.28	11.08	9.74	15.21	63.97
Never smoked	11.5	11	11.3	66.3	12.1	8.15	16.1	63.6	12.4	10.4	12	65.13	10.65	10.07	16.97	62.31	10.98	9.36	18.57	61.09	10.03	9.01	15.91	65.06
Quit smoking	12.9	11.3	11.9	63.9	15.1	10	15.3	59.6	12.4	8.62	15.1	63.85	13.83	8.52	15.38	62.28	13.41	9.78	17.03	59.78	11.57	11.11	15.66	61.66
Smokes Now	14.6	10.4	11.7	63.3	14	10.5	12.6	62.9	12.8	6.39	16.5	64.29	12.44	10.69	13.4	63.48	12.75	7.79	15.86	63.6	12.41	8.85	12.93	65.81
Zero Drinks per Day	12.2	10.3	9.49	68.1	12.5	8.92	14.2	64.5	11.4	8.75	12.6	67.23	10.95	8.78	14.83	65.45	10.75	8.37	16.31	64.57	9.57	9.06	13.97	67.4
Between 0-1	12.2	11.6	13	63.3	13.1	9.86	17.2	59.9	10.8	7.95	22.7	58.52	12.65	11.58	19.5	56.26	13.32	9.93	18.51	58.24	11.98	11.55	18.29	58.18
1-2 Drinks	15.4	11.5	13.5	59.6	19.1	12.1	17.3	51.5	18.5	9.93	14.6	56.95	15.98	11.6	16.97	55.45	16.74	7.37	20.09	55.8	13.94	11.85	18.82	55.4
2-5 Drinks	17.2	9.07	11.8	61.9	16.5	8.55	16.2	58.7	16	9.33	13.3	61.33	16.93	9.13	13.36	60.58	11.55	13.98	19.15	55.32	14.51	8.32	15.09	62.09
More than 5 Drinks	20.2	10.1	12.4	57.3	15.3	8.16	11.2	65.3	14.3	14.3	19.1	52.38	11.82	14.55	10.91	62.73	10.28	9.35	14.02	66.36	17.86	10	7.86	64.29
Less than 12 years ed	14.4	10.9	8.98	65.8	11.2	8.09	13.5	67.3	10.5	11.1	9.7	68.7	10.27	9.1	13.86	66.78	10.13	6.75	15.58	67.53	11.32	9.4	10.72	68.55
12 Years	10.3	10.5	11.5	67.7	10.1	8.91	14.1	66.9	9.65	6.78	13.6	70.02	10.78	9.38	14.23	65.61	8.77	6.54	14.57	70.12	8.68	8.28	12.6	70.43
13-16 years	13.4	11.3	13.5	61.8	16.2	9.97	16.2	57.7	14.1	8.53	18.4	58.93	13.02	10.4	16.62	59.96	13.29	10.51	18.39	57.81	12.31	10.05	17.08	60.56
Over 16 years	16.1	11.5	14.6	57.9	22	11.5	17.2	49.3	21.7	11.2	17.1	50	17.52	8.89	19	54.58	16.63	12.18	22.01	49.18	12.63	12.51	20.71	54.15
Not working	14.7	11.5	11.3	62.5	10.9	9.72	14	65.4	11.3	9.19	12.7	66.8	11.62	10.21	14.39	63.77	10.02	7.7	14.95	67.33	11.15	9.37	12.97	66.5
Employee	11.3	10.5	12.1	66.1	14	8.72	16.6	60.7	12.6	8.3	17	62.13	11.65	8.95	16.64	62.75	11.04	9.3	18.46	61.2	9.31	9.28	16.39	65.03
Self Employed	14.5	11.1	10.4	64	25	10.7	12.6	51.8	19.1	8.82	13.2	58.82	17.41	10.32	14.79	57.47	21.53	10.63	16.62	51.23	18.92	12.83	16.53	51.72
Western	15	11.9	13.2	59.9	15.9	9.53	15.8	58.8	14.8	10	16	59.2	14.4	9.47	14.49	61.64	14.62	9.02	16.8	59.56	11.07	10.53	16.72	61.69
Non-Western	85	88.1	86.9	40.1	84.1	90.5	84.2	41.2	85.2	90	84	40.8	85.6	90.53	85.51	38.36	85.38	90.98	83.2	40.44	88.93	89.47	83.28	38.31
Not Employed: Yes Insu	81.1	88.3	89.6	41	85.5	92.2	89.1	33.2	92.9	89.4	85.9	31.76	88.03	91.2	88.03	32.75	88.12	92.5	83.75	35.62	88.37	91.41	88.09	32.13
Not Employed: No Insu	18.9	11.7	10.4	59	14.5	7.77	10.9	66.8	7.06	10.6	14.1	68.24	11.97	8.8	11.97	67.25	11.88	7.5	16.25	64.38	11.63	8.59	11.91	67.87
Employed: Yes Insuran	83.2	87.5	87.3	42	90.8	92.8	85.5	30.9	88.9	90.7	88.9	31.48	88.55	90.31	85.9	35.24	85.04	89.37	87.4	38.19	90.68	90.35	87.78	31.19
Employed: No Insuran	16.8	12.5	12.7	58	9.18	7.25	14.5	69.1	11.1	9.26	11.1	68.52	11.45	9.69	14.1	64.76	14.96	10.63	12.6	61.81	9.32	9.65	12.22	68.81
Self-Employed: Yes Insu	84.7	89.7	90.4	35.2	76.5	91.6	82.4	49.6	85.3	88.2	79.4	47.06	82.29	95.83	82.29	39.58	78.35	87.63	88.66	45.36	83.46	88.72	90.23	37.59
Self-Employed: No Insu	15.3	10.3	9.62	64.8	23.5	8.4	17.7	50.4	14.7	11.8	20.6	52.94	17.71	4.17	17.71	60.42	21.65	12.37	11.34	54.64	16.54	11.28	9.77	62.41
Life Insurance: Yes	15.7	11.7	11.5	61	11.5	10.6	13	64.9	12.2	9.29	10.8	67.73	12.62	10.59	13.04	63.76	14.54	8.99	14.21	62.26	11.8	9.34	13.06	65.79
Life Insurance: No	11.6	10.6	11.7	66.2	14.4	8.95	15.8	60.9	12.4	8.69	15.8	63.14	12.14	9.26	16.51	62.09	10.98	9.27	18.88	60.88	10.74	9.96	16.14	63.15

Table 5.3: Tabulations for Risk Aversion Across Different Risky Behaviors, 1992 to 2006. Data source: Author's tabulation of the HRS.

Drinking and Smoking

The first rows of Table 5.3 show the distribution of risk tolerance conditional on smoking and drinking. The deleterious effects of both smoking and drinking are well-known. Individuals who have ever smoked are more risk tolerance than those who never smoked and those who smoke now are more risk tolerant than those who do not smoke now. Of particular interest are those who say they once smoked, but do not smoke now. The sample is largely composed of middle-aged to older individuals. Hence, those who quit smoking would have done so during a period of increasing public awareness of the risks associated with cigarette smoking. Those who quit smoking are somewhat more risk tolerant than those who never smoked, but less risk tolerant than current smokers (Barsky et al., 1997).

Drinking alcohol is also related to measured risk tolerance and often used as a measure of risky health behavior. Specifically, risky drinking is categorized as relatively heavy drinking; moderate drinking is not generally believed to be a health risk. Risk tolerance is higher for those who drink than for those who do not drink. The difference in risk tolerance between drinkers and nondrinkers is about the same as between smokers and nonsmokers. Table 5.3 shows risk tolerance by drinks per day. Those who take less than one drink per day have a willingness to accept the moderate gambles relatively often. As drinks per day increase, there is a monotonic increase in mean risk tolerance. Moreover, in the case of both smoking and drinking, risk tolerance remains relatively constant over time.

Educational Attainment, Employment, Immigration

The second portion of Table 5.3 shows a U-shaped relationship between years of schooling completed and the measure of risk tolerance that stays constant over time. Individuals with exactly twelve years of schooling are the least risk tolerant. Those with some post-college education (years greater than sixteen) have substantially greater than average risk tolerance.

Among the biggest risks voluntarily taken by a large segment of the population is self-employment. The self-employed generally face a riskier overall income stream than their wage-earning or salaried counterparts. Thus, one would expect risk tolerance to be positively associated with the decision to undertake self-employment. Table 5.2 shows that the self-employed are more risk than employees. There is no obvious prediction about the risk tolerance of those not working with this older sample, as they are mainly retired individuals and spouses not in the labor force.

One of the boldest risks is that of immigration; to move to a new country in search of a better life. The idea that immigrants, who move to another country or another region in search of a better life, are more daring than non-immigrants is well-documented (Barsky et al., 1997). In following with Barsky et al. (1997), I also examine the time-trend of risk tolerance and being from the Western part of the United States. The western United States has in the past been an internal frontier to which one might argue the more daring have migrated (Barsky et al., 1997). Some of the attitudes from the frontier past may have persisted to the present. Table 5.3 shows that Westerners are continually categorized as

having higher risk tolerance across time.

Health and Life Insurance

Understanding the role of risk tolerance in purchase of insurance is complex. Anyone with positive risk aversion should be fully insured against purely financial risks when insurance is actuarially fair. In the presence of a load factor, however, those who are most risk averse should be willing to buy insurance against financial risks (Barsky et al. 1997). A complication arises because the kinds of insurance purchases on which we have information are health and life insurance, where the risks are not purely financial. However, since most of these individuals are married or have children, I argue that financial responsibility for the support of others would explain why that the financially more risk averse are more likely to purchase both medical and life insurance.

The bottom portion of Table 5.3 examines respondents' stated risk preference to insurance purchase. Given that many receive insurance coverage through employment, I follow Barsky's convention and report insurance coverage for employed, self-employed and the not employed. Tabulations exclude individuals already covered by Medicare. Perhaps surprisingly, there are higher share of individuals with health insurance that state that they are relatively less risk averse. There are fewer individuals that have health insurance in the fourth category, categorized as the most risk averse. The U-shaped relationship between the four risk categories and insurance coverage also holds steady across time as well; the proportion of individuals with and without insurance across all risk categories remains very similar across all HRS waves. This is a departure from the original Barsky analysis, that shows that more risk tolerant individuals are less likely to hold insurance. Unlike Barsky, the effect of risk tolerance on the propensity to be insured is smaller for the non-employed and the employed than it is for the self-employed. Between groups, the self-employed have a higher risk tolerance and have a much lower average propensity to be insured than employees.

Similar patterns hold true for life insurance. Like many of the other behaviors detailed, there are higher shares of individuals that identify as highly risk averse than in any other category. However, there is very little difference in the categorization of risk aversion across those with or without life insurance. This holds true cross-sectionally across all of the waves of HRS.

Wealth and Income

Table 5.4 shows risk tolerance by quintiles of risk. Risk tolerance decreases with income and wealth until the middle of the distributions, and then increases. This pattern is similar to those for a number of other behavior discussed above, as well as the pattern across age. Like the original Barsky findings, risk tolerance rises at the high end of wealth and age distributions. Home equity is a major component of wealth for most individuals. The 20 percent of individuals who do not live in homes they own are substantially more risk tolerant than those who own their homes. The most risk-tolerant individuals are much less likely to

own homes than the least risk tolerant individuals. Although home prices are volatile and houses are often highly leveraged (Belsky et al., 2002), owning a house insulates individuals from local changes in the cost of shelter, and thus provides some consumption insurance (Barsky et al., 1997).

% Choosing Response					
Risk Category		<u>Wave 1: 1992</u>			
		1	2	3	4
Wealth Quintiles					
	1	17.46	11.48	8.9	62.16
	2	12.14	10.95	11.41	65.5
	3	11.48	10.63	11.52	66.37
	4	10.15	11.19	13.07	65.58
	5	13.13	10.38	13	63.49
<u>Wave 4: 1998</u>					
Wealth Quintiles					
	1	14.91	7.22	13.88	63.99
	2	11.52	10.65	14.49	63.34
	3	12.01	8.74	15.29	63.97
	4	13.24	9.52	14.1	63.14
	5	16.94	10.41	17.03	55.62
<u>Wave 5: 2000</u>					
Wealth Quintiles					
	1	11.66	10.6	12.72	65.02
	2	8.61	10.11	12.73	68.54
	3	10.65	6.46	12.55	70.34
	4	13.8	8.08	14.48	63.64
	5	17.36	9.06	18.87	54.72
<u>Wave 6: 2002</u>					
Wealth Quintiles					
	1	12.13	9.91	13.15	64.81
	2	10.2	8.68	14.99	66.13
	3	10.66	9.65	16.22	63.47
	4	12.65	9.69	15.45	62.21
	5	15.88	10.22	17.62	56.28
<u>Wave 7: 2004</u>					
Wealth Quintiles					
	1	12.28	8.19	13.64	65.89
	2	10.05	9.25	15.63	65.07
	3	11.44	8.46	18.74	61.36
	4	12.98	9.69	20.11	57.22
	5	14.66	10.56	21.12	53.66
<u>Wave 8: 2006</u>					
Wealth Quintiles					
	1	11.28	9.6	13.2	65.92
	2	10.11	8.22	12.8	68.87
	3	10.24	9.91	15.03	64.83
	4	9.75	10.13	16.83	63.29
	5	14.05	10.99	18.29	56.67

Table 5.4: Tabulations of Risk Preferences For Wealth Across HRS Survey Waves: 1992 to 2006. Data source: Author's tabulations of the HRS.

Addressing Survey Response Error

One of the challenges of working with this data is that there is a sizeable amount of movement across categories of risk within individuals and on the aggregate. Survey waves are every two years and responses to hypothetical gambles likely provide a noisy signal of true risk tolerance. This issue has been documented extensively in literature related to risk preferences and remains true for these data. (See Camerer and Hogarth (1999) for a review of several experiments with varying financial incentives.) Table 5.5 demonstrate the shares of individuals that remain consistent or change responses across two consecutive waves. These tables are cross-tabulations of aggregate values; as such, the number of individuals that give consistent responses are in the diagonal of the table and inconsistent values are in the off-diagonals of the table. The total number of individuals in each table is a subset of the total sample, since it includes only respondents that are in both survey waves. (There are shared no observations between survey waves from 2002 and 2004).

Consistency of Responses across survey waves between 1992 and 1998

	Wave 4 (1998)				
Wave 1 (1992)	1	2	3	4	Total
1 (Least risk averse)	19	9	24	52	104
2 (3rd most risk averse)	14	16	14	49	93
3 (2nd most risk averse)	14	14	24	58	110
4 (Most risk averse)	75	47	73	300	495
Total	122	86	135	459	802

Consistency of Responses across survey waves between 1998 and 2000

	Wave 5 (2000)				
Wave 4 (1998)	1	2	3	4	Total
1 (Least risk averse)	35	14	18	61	128
2 (3rd most risk averse)	16	6	15	47	84
3 (2nd most risk averse)	19	14	37	87	157
4 (Most risk averse)	36	44	68	389	537
Total	106	78	138	584	906

Consistency of Responses across survey waves between 2000 and 2002

	Wave 6 (2002)				
Wave 5 (2000)	1	2	3	4	Total
1.least risk averse	21	10	13	32	76
2.3rd most risk avers	6	4	8	28	46
3.2nd most risk avers	9	7	21	58	95
4.most risk averse	20	28	43	216	307
Total	56	49	85	334	524

Consistency of Responses across survey waves between 2004 and 2006

	Wave 8 (2006)				
Wave 7 (2004)	1	2	3	4	Total
1.least risk averse	107	52	39	109	307
2.3rd most risk avers	36	50	44	102	232
3.2nd most risk avers	37	67	126	231	461
4.most risk averse	107	121	211	1,084	1,523
Total	287	290	420	1,526	2,523

Table 5.5: Consistency of Responses Across Survey Waves. Source: Author's tabulation of the HRS.

Across all of these survey waves approximately half of the respondents have the same response as in the previous wave. The greatest share of consistent responses occurs between 2004 and 2006, with 54 percent of individuals keeping the same response. The lowest share, at 45 percent, occurred between 1992 and 1998, which is not surprising given that the elapsed time period is longer. Of the remaining individuals, approximately 25 percent of individuals have a stated preference of more risk averse; the range is 23 percent (from 2004 to 2006) to 28 percent (from 2002 to 2004). Given that risk aversion does increase with age, this is perhaps not surprising. Still, evidence from the risk aversion literature does imply that risk aversion might not change that significantly in such little time. Moreover, approximately twenty percent of respondents move from more to less risk averse.

Still, while there is movement in the risk categories from year to year, there is also a consistency of responses across time. Regression results show that past responses for risk do in fact predict future responses with very high significance. Table 5.6 illustrates results that predict the outcome for risk in the final survey wave in which it was collected (2006) as a function of all other responses from previous survey waves, as well as a specification in which sample size is larger. All are statistically significant. A number of additional specifications yield similar results; in all cases, previous risk preferences are strongly associated with future risk preferences.

Associations to Responses for Risk Preference in Wave 8 (2006)				
	(1)	(2)	(3)	(4)
Risk 1992	0.111 (0.0704)			
Risk 1998	0.282 (0.079)	0.195 (0.059)		
Risk 2000	0.165 (0.773)	0.137 (0.060)	0.137 (0.058)	
Risk 2002	0.202 (0.091)	0.203 (0.066)	0.288 (0.065)	0.207 (0.017)
r^2	0.289	0.1627	0.1261	0.0465

Note: Responses from 2004 were omitted because there are no shared observations

Table 5.6: Associations of Individual Risk Preferences Survey Responses Across Survey Waves

Regressions for early waves of risk behaviors show consistency over time

Statistical Model

For the analysis of risky behaviors in this chapter, the data are pooled across waves. There are multiple observations for each individual respondent. Behaviors are measured separately at each wave; each outcome is a report for a given respondent, indexed by i , at a given wave (or year of time) indexed by t , e.g. current smoking for individual i in year t . The predictor of central interest, however, is the composite cardinal measure of risk aversion described in previous sections, which takes a single value for each respondent and remains fixed across waves. Some other covariates, gender, race, and religion, also remain fixed across waves, but most covariates, like age and the logarithm of income, vary across waves. When data have this kind of structure, an ordinary regression model is inappropriate, because such a model treats errors of prediction around each data point as independent, identically distributed random variables, whereas many unobserved determinants of an individual's behavior are likely to be strongly correlated from wave to wave. For this reason, "multi-level" or "hierarchical" linear models of the kind described, for instance, by Gelman (2006) are employed.

In each of the linear multi-level models, the Level One equation takes the form

$$Y_{it} = \alpha_i + \sum_k \beta_k X_{itk} + \epsilon_{it}$$

where Y_{it} is the outcome, α_i is a random intercept term specific to the individual and constant across waves, and β_k and X_{itk} are coefficients and covariate values for respondent-wave pairings for covariates (indexed by k) that vary across waves. Given the individual-level values of α_i , the errors ϵ_{it} are considered to be independent and identically distributed with some unknown variance $\sigma^2(\epsilon)$. The Level Two equation expresses α_i as a linear function of the composite cardinal risk aversion measure and other covariates (indexed by m) that remain constant across waves:

$$\alpha_i = \mu + \sum_m \gamma_m W_{im} + \delta_i$$

Here μ is an intercept, γ_m are the coefficients, and W_{im} the covariates fixed across waves. The parameters to be estimated are the β_k , γ_m , μ , $\sigma^2(\epsilon)$, and $\sigma^2(\delta)$.

These formulas apply to linear versions of the models. Similar formulas apply to logistic regression variants with the same multi-level structure, used for some of the analyses.

The multi-level models are fitted using the STATA routines called `xtmixed` and `xtlogit`.

In these models, the measure of risk aversion is the continuous, cardinal measure detailed above and those taken from Kimball et al. (2001), which is detailed in Chapter 2. The cardinal measure ranges from 0 to 1. Individuals closer to 0 are least risk averse and those closer to 1 are most risk averse. In other words, a positive coefficient implies a positive

relationship between greater risk aversion and frequency or likelihood of risk behavior. Also, because the increments of change in the risk tolerance measure are very small, the coefficients on the regression results are often very large. Importantly, this risk measure captures the time-invariant measure free of measurement error. Control variables include respondent race, religion and gender, as well as age in the year of the behavior measured.

Drinking and Smoking

Table 5.7 shows results from multivariate analyses for drinking and smoking behavior. The two panels on the left show the association between drinking behavior and risk aversion. The dependent variable in the panel furthest on the left refers to the average number of days in a week that the respondent drinks any amount of alcohol, ranging from 0 to 7. The model is a linear regression and the coefficients can be interpreted as slopes. The measure of risk aversion is statistically associated with the frequency of drinking days, in the direction one might expect—the greater one’s risk aversion, the fewer days, on average, include alcohol consumption. The panel immediately to the right measures drinking behavior as a continuous variable equal to the total number of drinks consumed when drinking. The range is from 0 to 23, plus a response for 99, where 99 refers to the response, ‘I drink all day,’ which was excluded. These specifications include all respondents, including those who do not drink at all. Again, the coefficient on risk aversion is statistically significant and in the expected direction. However, as with the previous panel, the coefficient is relatively small, more than one order of magnitude smaller than race and gender.

The second half of Table 5.7 displays the results of a logistic model examining the association between risk aversion and smoking behavior. The coefficients are odds-ratios. Here, there are two outcomes measured. The first measures whether an individual used to be a smoker and quit at some point before the survey wave in which they are interviewed. The second outcome is whether a respondent is currently a smoker. Both of these measures might be behaviors involving elements of risk. Within the time period spanned by this panel, the deleterious effects of smoking became increasingly known and publicized, and with it, a number of public health campaigns aimed to encourage current smokers to quit smoking. As such, the decision to quit or to continue smoking is one that involves trading off known information about the potentially harmful effects of smoking. Still, and perhaps surprisingly, the measure of risk aversion here is not significantly associated with either quitting or continuing to smoke.

	DRINKING BEHAVIOR		SMOKING	
	Num. Days Drinking	Number of Drinks When Drinking	Quit Smoking	Currently Smoking
Gender (ref. male)	-0.667*** (0.0348)	-0.619*** (0.0275)	-0.956*** (0.167)	-0.683*** (0.125)
Race (ref. White)				
Black/African-American	-0.505*** (0.0462)	-0.111** (0.0365)	-0.695** (0.236)	0.568*** (0.166)
Other	-0.775*** (0.106)	-0.334*** (0.0838)	-0.673 (0.621)	-0.580 (0.375)
Religion (ref. Protestant)				
Catholic	0.258*** (0.0387)	0.213*** (0.0305)	0.209 (0.179)	0.162 (0.137)
Jewish	0.180 (0.127)	-0.0114 (0.100)	2.313*** (0.516)	-1.158* (0.477)
None	0.624*** (0.0800)	0.360*** (0.0633)	-0.704 (0.371)	1.321*** (0.305)
Other	0.277 (0.193)	-0.0511 (0.153)	1.307 (0.846)	-0.547 (0.681)
Age	-0.0137*** (0.00306)	-0.0236*** (0.00242)	0.195*** (0.0156)	-0.112*** (0.0114)
Year	0.0284*** (0.00646)	0.0246*** (0.00512)	0.133*** (0.0337)	-0.305*** (0.0264)
Risk Aversion	-0.0314*** (0.00703)	-0.0148** (0.00556)	-0.0347 (0.0343)	-0.0152 (0.0255)
Constant	3.045*** (0.191)	3.074*** (0.151)	-7.382*** (0.968)	1.417* (0.713)
Observations	66,890	66,844	33,766	54,485
Number of groups	10,560	10,558	7,388	11,593

Standard errors in parentheses
 *** p<0.001, ** p<0.01, * p<0.05

Table 5.7: Relationship between Risk and Smoking, Drinking: 1992 to 2010. Data Source: HRS.

Life Insurance and Employment

Table 5.8 displays the results for two logistic multi-level models for the decision to be self-employed. The decision to be self-employed, particularly in this sample of individuals over the age of 50, is one that is relatively risky as it does not offer the more certain aspects of health insurance, and a regular salary. In the panel, 21.68% of those working are self-employed. The odds-ratio on self employment is -0.109, statistically significant and in the expected direction—as one individual’s risk aversion score increases, the odds of self-employment decreases. Table 5.8 also displays results for the decision to purchase life insurance. Approximately 34% of the panel does not have life insurance. The results here also show a positive and

statistically significant relationship between risk aversion and the decision to purchase life insurance. An increase in risk aversion increases the odds of purchasing life insurance.

VARIABLES	(1) selfemp	(3) lifein
Gender (ref. male)	-1.174*** (0.127)	-1.121*** (0.0698)
Race (ref. White)		
Black/African-American	-1.460*** (0.178)	0.0588 (0.0915)
Other	-0.129 (0.401)	-1.385*** (0.206)
Religion (ref. Protestant)		
Catholic	-0.714*** (0.143)	-0.535*** (0.0771)
Jewish	2.071*** (0.536)	-1.153*** (0.254)
None	0.974** (0.307)	-1.090*** (0.158)
Other	0.433 (0.759)	-1.688*** (0.377)
Age	0.0817*** (0.0123)	-0.0269*** (0.00613)
Year	-0.00629 (0.0254)	-0.0993*** (0.0128)
Risk Aversion	-0.109*** (0.0260)	0.0639*** (0.0141)
Constant	-7.433*** (0.731)	5.476*** (0.385)
Observations	40,758	88,375
Number of hhidpn	8,714	11,577

Standard errors in parentheses
 *** p<0.001, ** p<0.01, * p<0.05

Table 5.8: Regression Results for Decision to Purchase Life Insurance and Self-Employment: 1992 to 2010. Data source: HRS.

Financial Decisions

The final portion of this analysis explores the relationship between risk and financial assets and, specifically, to what extent the relationship holds over time. Special attention

is paid to analysis from the most recent wave of survey data, which covers the time period after the global financial crisis of 2008, to examine whether this relationship changes.

Table 5.9 explores the association between the cardinal measure of risk aversion and the share of various financial instruments in the sample member's portfolio allocation. As with the previous models, the analysis is a multi-level linear regression with random effects. The outcome measures represent the share of total non-housing wealth that are held in each of the various financial holdings, which include stocks, bonds, savings accounts or money markets, a traditional retirement IRA account, a ROTH IRA account, and treasury bills. These outcomes include all possible non-housing and non-pension wealth that are asked in the survey. As such, the sum of shares for any given individual in any given year across these outcomes sums to 100%. The sample is restricted to those with at least \$1,000 in reported wealth. Covariates include self-reported race and religion, gender, age, as well as income and wealth in the corresponding survey year, both measured on a logarithmic scale.

The results for the analysis show that risk aversion plays a significant role in some but not all financial portfolio allocations. The most significant and strongest effect can be seen in the relationship between risk aversion and higher shares of wealth savings held in traditional savings accounts. Savings accounts are considered an essentially risk-free asset, as there is no investment per se, and as such, no risk of default or large risk of loss. Conversely, it is also the investment vehicle with the least amount of possibility of wealth growth, as savings accounts traditionally have low interest rates. Another relatively low-risk investment is the treasury-bill, as discussed in the background section of this chapter. As such, there is a significant and positive effect of risk aversion on the share of non-housing wealth that is invested in treasury bills across the panel. Likewise, significant effects are seen for the share of wealth held in ROTH IRAs, which hold greater tax-sheltering benefits than traditional IRAs. Finally, there is a negative association between shares of non-housing wealth invested in stocks and risk aversion; individuals with greater risk aversion are less likely to invest in stocks. All of these relationships operate in the expected direction and mirror what one might expect to hear from a financial planner. Still, the effect sizes are small. They are approximately equal to the effect of an additional year of age, but are at least two orders of magnitude smaller than the effects of financial decisions arising from income and existing wealth.

	Share Stock	Share Bond	Share Savings	Share IRA	Share Roth	Share T-Bill
Gender (ref. male)	0.00885* (0.00346)	0.00317*** (0.000891)	-0.00756 (0.00517)	-0.0182*** (0.00478)	0.00305 (0.00232)	0.0115*** (0.00260)
Race (ref. White)						
Black/African-American	-0.0428*** (0.00498)	-0.00338** (0.00130)	0.140*** (0.00742)	-0.0687*** (0.00684)	-0.0143*** (0.00336)	-0.0123** (0.00375)
Other	-0.0193 (0.0113)	-0.00283 (0.00294)	0.0771*** (0.0168)	-0.0474** (0.0155)	0.00802 (0.00763)	-0.0137 (0.00851)
Religion (ref. Protestant)						
Catholic	-0.00400 (0.00377)	-0.00168 (0.000962)	0.0151** (0.00565)	0.00956 (0.00522)	-0.00839*** (0.00251)	-0.0113*** (0.00283)
Jewish	0.0477*** (0.0122)	0.0170*** (0.00309)	-0.0885*** (0.0183)	0.0219 (0.0169)	0.0270*** (0.00808)	-0.0279** (0.00915)
None	-0.000434 (0.00773)	0.00209 (0.00197)	0.0216 (0.0116)	-0.0255* (0.0107)	0.0161** (0.00514)	-0.0152** (0.00581)
Other	0.0465* (0.0193)	-0.000431 (0.00498)	-0.0279 (0.0289)	-0.00997 (0.0266)	0.00771 (0.0130)	-0.0203 (0.0145)
Age	-6.43e-05 (0.000306)	0.000306*** (7.88e-05)	0.000935* (0.000457)	-0.00176*** (0.000422)	-0.00155*** (0.000205)	0.00213*** (0.000230)
Wealth (log)	0.0407*** (0.000879)	0.00555*** (0.000261)	-0.113*** (0.00124)	0.0505*** (0.00112)	0.00838*** (0.000652)	0.00846*** (0.000676)
Income (log)	0.0141*** (0.000863)	0.00256*** (0.000276)	-0.00647*** (0.00119)	-0.0145*** (0.00107)	0.00336*** (0.000675)	0.00234*** (0.000670)
Year	-0.00551*** (0.000667)	-0.00112*** (0.000181)	0.00956*** (0.000982)	0.00117 (0.000904)	-0.00108* (0.000463)	-0.00302*** (0.000505)
Risk Aversion	-0.00176* (0.000692)	-0.000323 (0.000176)	0.00348*** (0.00104)	-0.00119 (0.000961)	-0.00168*** (0.000460)	0.00158** (0.000520)
Constant	-0.452*** (0.0230)	-0.0933*** (0.00625)	1.739*** (0.0337)	-0.0648* (0.0310)	0.0456** (0.0160)	-0.192*** (0.0174)
Observations	72,465	72,465	72,465	72,465	72,465	72,465
Number of groups	10,730	10,730	10,730	10,730	10,730	10,730

Standard errors in parentheses
 *** p<0.001, ** p<0.01, * p<0.05

Table 5.9: Regression Results for Financial Portfolio Allocation: 1992 to 2010. Data source: HRS.

Financial Decisions After the Recession

Table 5.10 displays the results of a set of models that explore the relationship between risk aversion and financial portfolio allocation after the great recession. The specifications of the models are identical to those in Table 5.9 and discussed in the above text, with one important distinction: the inclusion of an interaction term of risk aversion with the post-recession years. Because the Great Recession officially began in December 2007 (though,

arguably, effects were felt for individuals sooner) and ended in July of 2009 according to the Board of Governors of the Federal Reserve System, and given the precise timing of the HRS fielding of the survey, the recession in this analysis was coded as the survey years for 2008 and 2010. The interaction term is defined as the individual-level risk aversion measure multiplied by an indicator variable on the recession years.

The results displayed in Table 5.10 show that the relationship between financial instruments and risk aversion remains similar across time. Importantly, there does not appear to be a change in relationship between risk aversion and investment in treasury bills, even if the risky nature of treasury bills changed dramatically during the recession. The coefficient on the interaction term remains positive and significant, implying that any multiplicative effects of risk aversion after the recession remain positive. One possible interpretation of these results is that individuals did not internalize or comprehend the changing nature of these financial instruments and thus, did not react appropriately. Other possible interpretations are that individuals were either too scared or too optimistic to make changes to their financial portfolios, or that they made diversification choices in non-financial wealth, such as housing decisions.

	Share Stock	Share Bond	Share Savings	Share IRA	Share Roth	Share T-Bill
Gender (ref. male)	0.00900** (0.00346)	0.00317*** (0.000891)	-0.00759 (0.00517)	-0.0181*** (0.00478)	0.00298 (0.00232)	0.0114*** (0.00260)
Race (ref. White)						
Black/African-American	-0.0432*** (0.00498)	-0.00338** (0.00130)	0.140*** (0.00742)	-0.0690*** (0.00685)	-0.0141*** (0.00336)	-0.0120** (0.00375)
Other	-0.0193 (0.0113)	-0.00283 (0.00294)	0.0771*** (0.0168)	-0.0474** (0.0155)	0.00800 (0.00763)	-0.0137 (0.00851)
Religion (ref. Protestant)						
Catholic	-0.00403 (0.00377)	-0.00168 (0.000962)	0.0151** (0.00565)	0.00954 (0.00522)	-0.00838*** (0.00251)	-0.0113*** (0.00283)
Jewish	0.0479*** (0.0122)	0.0170*** (0.00309)	-0.0886*** (0.0183)	0.0220 (0.0169)	0.0269*** (0.00809)	-0.0280** (0.00915)
None	-0.000403 (0.00773)	0.00209 (0.00197)	0.0216 (0.0116)	-0.0255* (0.0107)	0.0161** (0.00514)	-0.0152** (0.00581)
Other	0.0470* (0.0193)	-0.000434 (0.00498)	-0.0280 (0.0289)	-0.00966 (0.0267)	0.00747 (0.0130)	-0.0206 (0.0145)
Age	5.79e-06 (0.000306)	0.000306*** (7.89e-05)	0.000919* (0.000457)	-0.00171*** (0.000422)	-0.00158*** (0.000205)	0.00208*** (0.000230)
Wealth (log)	0.0403*** (0.000880)	0.00555*** (0.000261)	-0.113*** (0.00124)	0.0502*** (0.00113)	0.00856*** (0.000652)	0.00874*** (0.000676)
Income (log)	0.0142*** (0.000862)	0.00256*** (0.000276)	-0.00649*** (0.00119)	-0.0144*** (0.00107)	0.00330*** (0.000675)	0.00225*** (0.000670)
Year	-0.00373*** (0.000701)	-0.00113*** (0.000196)	0.00924*** (0.00102)	0.00213* (0.000939)	-0.00214*** (0.000496)	-0.00441*** (0.000533)
Risk Aversion	-0.00138* (0.000694)	-0.000325 (0.000177)	0.00341** (0.00104)	-0.000990 (0.000962)	-0.00190*** (0.000462)	0.00129* (0.000521)
Post-Recession Risk Aversion	-0.00264*** (0.000321)	1.73e-05 (0.000109)	0.000479 (0.000434)	-0.00146*** (0.000389)	0.00154*** (0.000261)	0.00206*** (0.000251)
Constant	-0.461*** (0.0230)	-0.0933*** (0.00626)	1.740*** (0.0338)	-0.0704* (0.0310)	0.0512** (0.0160)	-0.185*** (0.0175)
Observations	72,465	72,465	72,465	72,465	72,465	72,465
Number of groups	10,730	10,730	10,730	10,730	10,730	10,730

Standard errors in parentheses
 *** p<0.001, ** p<0.01, * p<0.05

Table 5.10: Changes in Associations to Risk Aversion Before and After the Recession. Data source: HRS.

5.5 Discussion

The extent to which measured risk predicts risky behavior over time remains poorly understand. This paper has addressed two questions. First, does the relationship between hypothetical risk and measurable risky behaviors remain consistent across both time and

among individuals? Secondly, does the change in the riskiness of bonds following the 2008 regression change the relationship in portfolio allocations relative to stated risk tolerance for individuals after the global financial crisis? This analysis shows that the relationship between measured risk and risky behaviors remains relatively constant across the 15 years prior to the global financial crisis. Moreover, the relationship between risk aversion and financial decisions does not appear to change after the Global Financial Recession, though it is possible that other tradeoffs in wealth protection were made by individuals during this time period that are not captured in these data.

Chapter 6

Conclusion

With this chapter, I conclude my dissertation on risk aversion in older Americans. First, I summarize the main results of the study. Second, I discuss the relevance of the findings and the limitations of my approach. Last, I propose some avenues for future research.

6.1 Summary

This study provided a quantitative assessment of risk aversion in older Americans by using both genotype and phenotype data from the Health and Retirement Study (HRS). I used this data to examine heritability at an overall genome-wide level and at an individual genetic marker level. I then provided a longitudinal analysis of the role of risk tolerance in behaviors over time.

In Chapter 2, I explained the data in sufficient detail to facilitate understanding of the analysis used in subsequent chapters. I described the measures of risk tolerance used throughout the dissertation and also spent some time discussing the genetic data, including a discussion of the format, the confidentiality issues, and the quality control measures that made it relevant for analysis in this dissertation.

Chapter 3 detailed a genetic-wide association study that looked for genetic markers called SNPs that might be associated with risk aversion. I found no single SNP that passed significance at the threshold required for this large sample of genetic data. I argued that a likely interpretation of this result is that risk tolerance is a highly polygenic trait. This means that the heritability of risk is the result of the aggregation of activity from a large number of genetic markers along the genome, each with effect sizes too small to detect with current sample sizes. I then ran a crude analysis of detectability in order to quantify how many genetic markers this might imply, given an assumed effect size. I discounted the possibility that there was likely to be fewer than 20 SNPs with moderate effect sizes that were undetected in my sample.

In Chapter 4, I pulled the view-finder back, so to speak, to examine the overall heritability of risk aversion using genetic variance across the entire genome by way of a method

called Genome-Wide Complex Trait Analysis (GCTA). While Chapter 3's analysis focused on specific spots along an individual's genome that might be associated with risk tolerance, the GCTA technique used a maximum likelihood method to calculate total heritability by examining the total phenotypic and genetic variance across the data sample. The results of the GCTA analysis indicated that, despite the evidence from twin studies that risk aversion is heritable, there does not appear to be any evidence of heritability in risk aversion in this HRS sample. The GCTA results showed estimates of near zero, with high statistical significance. I proposed a number of possible explanations that might explain the lack of heritability resulting from this study.

Chapter 5 moved away from the genetic view of risk aversion and examined the longitudinal nature of risk tolerance. In this chapter, I aimed to answer two questions. First, I examined whether risk aversion was a significant indicator of various health-related, labor-market, and financial decisions for older Americans. I found that for most behaviors examined, risk tolerance was indeed a significant contributor to these decisions, though its effect was often smaller than other factors, including factors such as gender or religiosity. Second, I examined the extent to which this relationship holds true over time. I used the longitudinal nature of the HRS to examine if the association between risk and behavior varies across time, where social norms and economic climates might differ. I used a cardinal measure of risk aversion that took into account the possible problems with measurement error. I found that the relationship between risk tolerance and behavior was more or less significant in different years, often with intuitive possible explanations, though not always.

6.2 Implications

The results of the analysis from this dissertation have a number of broader implications.

First, the results of this analysis cast a bit of a shadow on the idea that risk tolerance is in fact fixed and immutable, as is often assumed in the economic literature and elsewhere. The results from Chapters 3 and 4 showed that risk tolerance may not be as heritable as earlier studies suggested, and that these preferences may instead be driven by other forces. Likewise, the high levels of change in responses between survey waves for individuals may indeed be attributed to measurement error, or it may indicate that risk tolerance is more flexible than previously assumed. At the very least, a better understanding of what might be the more "fixed" portion of these preferences versus the time-varying component of risk within individuals should be better understood.

There are also implications for older Americans in regards to retirement and labor-market decisions. Results from Chapter 5 show that individuals do not necessarily adjust their financial portfolios with regards to their stated risk tolerance as economics conditions change. Given the increasing onus on individuals to make proper financial decisions as retirement plans move away from pensions and towards self-directed IRAs and other financial instruments, it is more important than ever that risk tolerance and its role in financial planning be well understood. Scholarly research needs to carefully identify the role of risk tolerance

and decision-making in order to properly inform financial advisors and individuals on these important decisions.

Likewise, the same holds true for health policies. Possible interpretations of the results of Chapter 5 show that there is in fact some elasticity in risk tolerance with regards to the social norms and public health campaigns of risky activities like smoking and heavy drinking. Highly risk-tolerant individuals do not seem to change their health behaviors even in times when the deleterious effects of these behaviors become well-known. Perhaps targeted interventions that account for risk tolerance may be more effective in curbing dangerous health behaviors.

Finally, this study has a more general implication towards survey measurement of risk tolerance. There is still much to be learned about what we are actually measuring when analyzing the survey measures used. One possible interpretation of the null findings from the GCTA analysis on heritability that I discussed in Chapter 4 is that the measure of risk used from the phenotype does not approximate whatever “inherent” or “hard-wired” risk tolerance may actually exist among individuals. There is a lively debate in a number of fields about the issue of risk tolerance and its measurement and I hope that the results from these analyses have made a contribution.

6.3 Limitations

In this section, I discuss the major drawbacks of my analysis. First, I address general issues regarding the data. Second, I move to specific limitations of the analysis covered in the analytical chapters of the dissertation.

Perhaps the largest drawback from using this particular data in the analysis is the low variation in the responses to the risk aversion question. This is a problem for two reasons. First, because the HRS is a survey of older populations, the large majority of respondents are categorized as the most risk averse. Second, because the question was not asked to all persons in all years, there is sometimes little overlap between survey waves. This makes doing some longitudinal work difficult. However, I believe I successfully circumvented this issue by using the cardinal measure I introduced in Chapter 5, but this does mean it is nearly impossible to perform any longitudinal analysis on the original categorical data. The GWAS study was also limited by the issue of a lack of overlapping samples between waves. A sizeable share of the original sample was lost when I had to restrict the sample to individuals who had both a response for the risk aversion question and also had genetic data available.

Part of the analysis from Chapter 5 focused on any changes in potentially risky behavior that might occur in different economic or financial climates, with particular attention paid to shifts made in financial portfolios. However, because the HRS starts at age 50, it is not possible to gather much information about these changes earlier in life. It is possible that individuals may be making more shifts in portfolio allocation earlier in life when they anticipate that the horizon for their working life is becoming smaller. If that is the case, my analysis will not capture these actions.

6.4 Further Study

This dissertation has been an exciting learning process, and this process will not end with the completion of this chapter. As with all research, answers to one question lead to many more questions, and this endeavor has proven to be no exception. Over the course of my study, research questions have multiplied and with them my enthusiasm to address each one. Below I propose some avenues of research that I will pursue in the near future. One possible area of refinement concerns the GWAS results. As I mentioned in Chapter 3, because it is likely that risk aversion is the function of numerous SNPs and therefore highly polygenic, it would require a very large number sample of individuals to identify the SNPs involved. This is a direct consequence of the difference in the extremely large number of SNPs available (nearly 2 million) relative to the small number of individuals available (approximately 10,000). There are two possible improvements that can be made to reduce this statistical issue: one must either reduce the number of SNPs or increase the number of individuals in the study. I propose one future research project for each of these two possibilities.

First, I would like to run another GWAS with the same sample of individuals from the HRS but on a smaller, more refined number of SNPs. There are no known SNPs for risk aversion to date, but there have been a number of positive findings in GWAS studies on phenotypes that might be associated with risk. These include educational attainment, extraversion, and neuroticism. These phenotypes have all been shown to be related to the risk aversion. It may be, then, that they share a set of genetic markers as well. A careful meta-analysis of the GWAS literature may narrow down a list of SNPs associated with related phenotypes on which to run the GWAS for risk. With a sample of individuals around 10,000, this may be sufficient to detect any SNPs that have small effects. Running a GWAS on the full panel of available SNPs assumes no prior knowledge of the genetic correlates to risk, which seems appropriate given there had been no prior genetic study of risk.

A second project in this vein involves increasing the sample size of individuals. As I mentioned in previous chapters, there are attempts to increase sample sizes for complex traits by consolidating various social science surveys through consortium efforts, similar to what has been done in the psychology and medical literature for a number of years, most notably through the establishment of the Social Science Genetic Association Consortium established in 2011. As the number of social surveys with genetic material increases, so does the ability to test for associations with larger sample sizes. Still, the number of phenotypes has been limited to a short list, which includes self employment and educational attainment, that can be universally measured. For the time being, risk aversion is still measured in a myriad of ways. However, this might not always be the case. Both Ad Health and the PSID have the identical measure of risk aversion used in the HRS, and Ad Health already has genetic material available and the PSID has plans to include it. It is possible that the ability to combine samples of individuals who have an identical measure of risk aversion to run a GWAS may not be in the distant future. If and when that time arrives, I hope to be the first to expand on my original GWAS analysis with a larger sample.

Additionally, the results from the GWAS and the GCTA analysis are somewhat at odds with previous research using twin studies and candidate genes that suggest a heritability of risk should be present. This tension leaves me feeling intellectually unsettled. I would like to find a way to explore this discrepancy more fully.

Finally, I would like to move away from risk aversion to explore the phenotype of cognition. Cognition, like risk tolerance, is a central part of life—particularly for older Americans—and yet remains poorly understood. The HRS has dedicated significant effort to carefully measuring cognition and has some of the richest longitudinal data on cognition that exists. Moreover, my GCTA analysis shows that positive heritability estimates can be gleaned from the GWAs data, which is in line with the twin studies and candidate gene literature. There are a number of potential inquiries related to this data. Of particular interest to me is the role of cognitive decline in older Americans on a variety of subsequent outcomes. While it is unlikely that there are a number of alleles with strong effects that would be highly deterministic in any model, it may be possible to create a genetic risk score that could be used in models of retirement decisions, insurance purchase, and financial decision-making.

Chapter 7

References

Anderson, L. R., & Mellor, J. M. (2008). Predicting health behaviors with an experimental measure of risk preference. *Journal of Health Economics*, 27(5), 1260-1274.

Anderson, L. R., & Mellor, J. M. (2009). Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure. *Journal of Risk and Uncertainty*, 39(2), 137-160.

Apicella, C. L., Cesarini, D., Johannesson, M., Dawes, C. T., Lichtenstein, P., Wallace, B. . . . Westberg, L. (2010). No association between oxytocin receptor (OXTR) gene polymorphisms and experimentally elicited social preferences. *PLoS ONE*, 5(6), e11153.

Apicella, C. L., Dreber, A., Campbell, B., Gray, P. B., Hoffman, M., & Little, A. C. (2008). Testosterone and financial risk preferences. *Evolution and Human Behavior*, 29(6), 384-390.

Arrondel, L. (2002). Risk management and wealth accumulation behavior in France. *Economics Letters*, 74(2), 187-194.

Barnea, A., Cronqvist, H., & Siegel, S. (2010). Nature or nurture: What determines investor behavior? *Journal of Financial Economics*, 98(3), 583-604.

Barseghyan, L., Prince, J., & Teitelbaum, J. C. (2011). Are risk preferences stable across contexts? Evidence from insurance data. *The American Economic Review*, 101(2), 591-631.

Barsky, R. B., Juster, F. T., Kimball, M., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the Health and Retirement Study. *The Quarterly Journal of Economics*, 112(2), 537-579.

Beauchamp, J., Cesarini, D., Van der Loos, M. J. H. M., Koellinger, P., Groenen, P. J.

F., Fowler, J. . . . Christakis, N. A. (2010). A genome-wide association study of educational attainment. Unpublished manuscript. Harvard University.

Beauchamp, J. P., Cesarini, D., Johannesson, M., Van der Loos, M. J. H. M., Koellinger, P., Groenen, P. J. . . . Christakis, N. A. (2011). Molecular genetics and economics. *The Journal of Economic Perspectives*, 25(4), 57-82.

Bellante, D., & Green, C. A. (2004). Relative risk aversion among the elderly. *Review of Financial Economics*, 13(3), 269-281.

Belsky, E. S., & Duda, M. (2002). Asset appreciation, timing of purchases and sales, and returns to low-income homeownership. Cambridge, MA: Joint Center For Housing Studies.

Benjamin, D. J., Cesarini, D., Chabris, C., Glaeser, E. L., & Laibson, D. I. (2012). The promises and pitfalls of geno-economics. *Annual Review of Economics*, 4, 627-664.

Benjamin, D. J., Cesarini, D., Van der Loos, M. J. H. M., Dawes, C., Koellinger, P., Magnusson, P. K. E. C. . . . Visscher, P. (2012). The genetic architecture of economic and political preferences. *Proceedings of the National Academy of Sciences*, 109(21), 8026-8031.

Boardman, J. D., Domingue, B. W., & Daw, J. (2014). What can genes tell us about the relationship between education and health?. *Social Science & Medicine*.

Bodie, Z., Merton, R. C., & Samuelson, W. F. (1992). Labor supply flexibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control*, 16(3), 427-449.

Bonin, H., Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2007). Cross-sectional earnings risk and occupational sorting: The role of risk attitudes. *Labour Economics*, 14(6), 926-937.

Brown, S., & Taylor, K. (2007). Education, risk preference, and wages. Department of Economics, University of Sheffield. Retrieved from <https://ideas.repec.org/p/shf/wpaper/2006002.html>

Brown Kruse, J., & Thompson, M. A. (2001). A comparison of salient rewards in experiments: Money and class points. *Economics Letters*, 74(1), 113-117.

Brunello, G. (2002). Absolute risk aversion and the returns to education. *Economics of Education Review*, 21(6), 635-640.

Buccioli, A., & Miniaci, R. (2011). Household portfolios and implicit risk preference. *Review of Economics and Statistics*, 93(4), 1235-1250.

Byrne, E. M., Carrillo-Roa, T., Penninx, B. W., Sallis, H. M., Viktorin, A., Chapman, B., ... & Psychiatric Genomic Consortium Major Depressive Disorder Working Group. (2014). Applying polygenic risk scores to postpartum depression. *Archives of women's mental health*, 17(6), 519-528.

Camerer, C. F., & Hogarth, R. M. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19(1-3), 7-42.

Campbell, B. C., Dreber, A., Apicella, C. L., Eisenberg, D. T., Gray, P. B., Little, A. C., . . . Lum, J. K. (2010). Testosterone exposure, dopaminergic reward, and sensation-seeking in young men. *Physiology & behavior*, 99(4), 451-456.

Catalano, R. (2009). Health, medical care, and economic crisis. *New England Journal of Medicine*, 360(8), 749-751.

Catalano, R., Golman-Mellor, S., Saxton, K., Margerison-Zilko, C., Subbaraman, M., LeWinn, K., & Anderson, E. (2011). The health effects of economic decline. *Annual Review of Public Health*, 32, 431-450.

Cesarini, D., Dawes, C. T., Johannesson, M., Lichtenstein, P., & Wallace, B. (2009). Experimental game theory and behavior genetics. *Annals of the New York Academy of Sciences*, 1167(1), 66-75.

Cesarini, D., Dawes, C. T., Johannesson, M., Lichtenstein, P., & Wallace, B. (2009). Genetic variation in preferences for giving and risk taking. *The Quarterly Journal of Economics*, 124(2), 809-842.

Cesarini, D., Johannesson, M., Lichtenstein, P., Sandewall, O., & Wallace, B. (2010). Genetic variation in financial decision-making. *The Journal of Finance*, 65(5), 1725-1754.

Chabris, C. F., Hebert, B. M., Benjamin, D., Beauchamp, J., Cesarini, D., Van der Loos, M. J. H. M., . . . Laibson, D. (2012). Most reported genetic associations with general intelligence are probably false positives. *Psychological Science*, 25(2), 1314-1323.

Chew, S., Ebstein, R., & Zhong, S. (2012). Ambiguity aversion and familiarity bias: Evidence from behavioral and gene association studies. *Journal of Risk and Uncertainty*, 44(1), 1-18.

Chew, S. H., Ebstein, R. P., & Zhong, S. (2013). Sex-hormone genes and gender difference in ultimatum game: Experimental evidence from China and Israel. *Journal of Economic*

Behavior & Organization, 90, 28-42.

Conley, D., Siegal, M. L., Domingue, B., Harris, K. M., McQueen, M., & Boardman, J. (2014). Testing the key assumption of heritability estimates based on genome-wide genetic relatedness. *Journal of Human Genetics*.

Crian, L. G., Pana, S., Vulturar, R., Heilman, R. M., Szekely, R., Druga, B. . . . Miu, A. C. (2009). Genetic contributions of the serotonin transporter to social learning of fear and economic decision making. *Social Cognitive and Affective Neuroscience*, 4(4), 399-408.

Da, Y., Wang, C., Wang, S., & Hu, G. (2014). Mixed model methods for genomic prediction and variance component estimation of additive and dominance effects using SNP markers. *PloS One*, 9(1), e87666.

Dave, D., & Saffer, H. (2007). Risk tolerance and alcohol demand among adults and older adults. The National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w13482>

De Neve, J.-E., Christakis, N., Fowler, J., & Frey, B. S. (2012). Genes, economics, and happiness. *Journal of Neuroscience, Psychology, and Economics*, 5, 193-211.

De Neve, J.-E., & Fowler, J. H. (2010). The MAOA gene predicts credit card debt. SSRN working paper 1457224.

Davies, G., Tenesa, A., Payton, A., Yang J., Harris, S.E., Liewald, D. . . . Deary, I. J. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry*, 16(10), 996-1005.

Dickie, E. W., Tahmasebi, A., French, L., Kovacevic, N., Banaschewski, T., Barker, G. J., ... & IMAGEN consortium. (2014). Global Genetic Variations Predict Brain Response to Faces. *PLoS Genetics*, 10(8), e1004523.

Ding, X., Hartog, J., & Sun, Y. (2010). Can we measure individual risk attitudes in a survey? IZA Discussion Paper No. 4807.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522-550.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *The American Economic Review*, 100(3), 1238-1260.

- Dreber, A., Apicella, C. L., Eisenberg, D. T. A., Garcia, J. R., Zamore, R., Lum, J. K., & Campbell, B. (2009). The 7R polymorphism in the dopamine receptor D4 gene (DRD4) is associated with financial risk taking in men. *Evolution and Human Behavior*, 30(2), 85-92.
- Dreber, A., Rand, D. G., Garcia, J. R., Wernerfelt, N., Lum, J. K., & Zeckhauser, R. (2010). Dopamine and risk preferences in different domains. Harvard Kennedy School Faculty Research Working Paper Series.
- Dreber, A., Rand, D. G., Wernerfelt, N., Garcia, J. R., Vilar, M. G., Lum, J. K., & Zeckhauser, R. (2011). Dopamine and risk choices in different domains: Findings among serious tournament bridge players. *Journal of Risk and Uncertainty*, 43(1), 19-38.
- Dreber, A., Rand, D., Wernerfelt, N., Montgomery, C., & Malhotra, D. K. (2012). Genetic correlates of economic and social risk taking. SSRN working paper 2141601.
- Eckel, C. C., & Grossman, P. J. (2008a). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization*, 68(1), 1-17.
- Eckel, C. C., & Grossman, P. J. (2008b). Men, women, and risk aversion: Experimental evidence. *Handbook of Experimental Economics Results*, 1(7), 1061-73.
- Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. CESifo Working Paper Series No. 2894.
- Fowler, J. H., Settle, J. E., & Christakis, N. A. (2011). Correlated genotypes in friendship networks. *Proceedings of the National Academy of Sciences*, 108(5), 1993-1997.
- Fisher, R. A. (1928). The possible modification of the response of the wild type to recurrent mutations. *American Naturalist*, 115-126.
- Fowler, J. H., & Dawes, C. T. (2013). In defense of genopolitics. *American Political Science Review*, 107(2), 362-374.
- Frydman, C., Camerer, C., Bossaerts, P., & Rangel, A. (2010). MAOA-L carriers are better at making optimal financial decisions under risk. *Proceedings of the Royal Society of Biological Sciences*. Advance Online Publication. doi:10.1098/rspb.2010.2304
- Francis Galton (1886) Anthropological Miscellanea: "Regression towards mediocrity in hereditary stature," *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15: 246-263

Ganna, A., Rivadeneira, F., Hofman, A., Uitterlinden, A. G., Magnusson, P. K., Pedersen, N. L. . . . Tiemeier, H. (2013). Genetic determinants of mortality. Can findings from genome-wide association studies explain variation in human mortality? *Human Genetics*, 132(5), 553-561.

Guiso, L., & Monica, P. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association*, 6(6), 1109-1150.

Guo, G. (2006). Genetic similarity shared by best friends among adolescents. *Twin Research and Human Genetics*, 9(1), 113-121.

Guo, G., Elder, G. H., Cai, T., & Hamilton, N. (2009). Gene-environment interactions: Peers alcohol use moderates genetic contribution to adolescent drinking behavior. *Social Science Research*, 38(1), 213-224.

Hartog, J., FerreriCarbonell, A., & Jonker, N. (2002). Linking measured risk aversion to individual characteristics. *Kyklos*, 55(1), 3-26.

Hatemi, P. K., Gillespie, N. A., Eaves, L. J., Maher, B. S., Webb, B. T., Heath, A. C. . . . Martin, N. G. (2011). A genome-wide analysis of liberal and conservative political attitudes. *The Journal of Politics*, 73(1), 271-285.

Homberg, J. R., van den Bos, R., den Heijer, E., Suer, R., & Cuppen, E. (2008). Serotonin transporter dosage modulates long-term decision-making in rat and human. *Neuropharmacology*, 55(1), 80-84.

Hurd, M. D. (1990). The joint retirement decision of husbands and wives. In D. A. Wise (Eds.), *Issues in the Economics of Aging* (pp. 231-258). Chicago, IL: University of Chicago Press.

Jung, S., & Treibich, C. (2014). Is self-reported risk aversion time varying? *Paris-Jordan Sciences Economiques*. Advance online publication.

Karatzas, I., Lehoczky, J. P., & Shreve, S. E. (1987). Optimal portfolio and consumption decisions for a small investor on a finite horizon. *SIAM Journal on Control and Optimization*, 25(6), 1557-1586.

Kan, K. (2003). Residential mobility and job changes under uncertainty. *Journal of Urban Economics*, 54(3), 566-586.

Kapteyn, A., & Teppa, F. (2002). Subjective measures of risk aversion and portfolio choice.

Kimball, M. S., Sahm, C. R., & Shapiro, M. D. (2007). Users guide for risk preference parameters. Retrieved from http://www-personal.umich.edu/shapiro/data/risk_preference/ImputationUsers

Kimball, M. S., Sahm, C. R., & Shapiro, M. D. (2009). Risk preferences in the PSID: Individual imputations and family covariation. National Bureau of Economic Research, No. w14754.

Kimball, M. S., Sahm, C. R., & Shapiro, M. D. (2008). Imputing risk tolerance from survey responses. *Journal of the American Statistical Association*, 103(483), 1028-1038.

Knowler, W. C., Williams, R. C., Pettitt, D. J., & Steinberg, A. G. (1988). Gm3; 5, 13, 14 and type 2 diabetes mellitus: An association in American Indians with genetic admixture. *American Journal of Human Genetics*, 43(4), 520.

Koellinger, P., Loos, M. H. M., Groenen P., Thurik, R., Rivadeneira, F., van Rooij, F. J. A., . . . Hofman, A. (2010). Genome-wide association studies in economics and entrepreneurship research: Promises and limitations. *Small Business Economics*, 35(1), 1-18.

Kuhnen, C. M., & Chiao, J. Y. (2009). Genetic determinants of financial risk taking. *PLoS ONE*, 4(2), e4362.

Kuhnen, C. M., Samanez-Larkin, G. R., & Knutson, B. (2011). Serotonin and risk taking: How do genes change financial choices? Available at SSRN working paper 1780552.

Lahiri, K., & Song, J. G. (2000). The effect of smoking on health using a sequential selfselection model. *Health Economic*, 9(6), 491-511.

Le, A., Miller, P. W., Slutske, W., & Martin, N. (2010). Are attitudes towards economic risk heritable? Analyses using the Australian twin study of gambling. *Twin Research on Human Genetics*, 13(4), 330-339.

Lee, J. J., & Chow, C. C. (2014). Conditions for the validity of SNP-based heritability estimation. *Human Genetics*, 1-12.

Lee, S. H., Wray, N. R., Goddard, M. E., & Visscher, P. M. (2011). Estimating missing heritability for disease from genome-wide association studies. *The American Journal of Human Genetics*, 88(3), 294-305.

Lee, S. H., DeCandi, T. R., Tang, J., Schizophrenia Psychiatric Genome-Wide Association Study Consortium, International Schizophrenia Consortium, Molecular Genetics of Schizophrenia Collaboration . . . Wray, N. R. (2012). Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature Genetics*, 44(3),

247-250.

Lusardi, A. (1998). On the importance of the precautionary saving motive. *American Economic Review*, 88(2), 449-453.

Lusk, J. L., & Coble, K. H. (2005). Risk perceptions, risk preference, and acceptance of risky food. *American Journal of Agricultural Economics*, 87(2), 393-405.

Macgregor, S., Cornes, B. K., Martin, N. G., & Visscher, P. M. (2006). Bias, precision and heritability of self-reported and clinically measured height in Australian twins. *Human Genetics*, 120(4), 571-580.

Mandal, B., & Roe, B. E. (2014). Risk tolerance among national longitudinal survey of youth participants: The effects of age and cognitive skills. *Economica*, 81(323), 522-543.

Pearson, T. A., & Manolio, T. A. (2008). How to interpret a genome-wide association study. *The Journal of the American Medical Association*, 299(11), 1335-1344.

Picone, G., Sloan, F., & Taylor Jr., D. (2004). Effects of risk and time preference and expected longevity on demand for medical tests. *Journal of Risk and Uncertainty*, 28(1), 39-53.

Powell, M., & Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology*, 18(6), 605-628.

Roe, B. E., Tilley, M. R., Gu, H. H., Beversdorf, D. Q., Sadee, W., Haab, T. C., & Paap, A. C. (2009). Financial and psychological risk attitudes associated with two single nucleotide polymorphisms in the nicotine receptor (CHRNA4) Gene. *PLoS ONE*, 4(8), e6704.

Patterson N., Price A. L., & Reich D. (2006). Population structure and eigenanalysis. *PLoS Genetics*, 2(12), e190. doi:10.1371/journal.pgen.0020190

Plomin, R., & Deary, I. J. (2014). Genetics and intelligence differences: five special findings. *Molecular Psychiatry*.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8), 904-909.

Purcell, S. M., International Schizophrenia Consortium, Wray, N. R., Stone, J. L., Visscher, P.M., ODonovan, M.C. . . . Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, 460(7256), 748-752.

- Roiser, J. (2008). The effect of serotonin depletion on the neural correlates of emotional processing in patients recovered from depression. *International Journal of Neuropsychopharmacology*, 34-35.
- Rosen, H. S., & Wu, S. (2004). Portfolio choice and health status. *Journal of Financial Economics*, 72(3), 457-484.
- Sahm, C. R. (2012). How much does risk tolerance change? *The Quarterly Journal of Finance*, 2(4), 1-38.
- Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences*, 106(36), 15268-15273.
- Schmidt, L. (2008). Risk preferences and the timing of marriage and childbearing. *Demography*, 45(2), 439-460.
- Schubert, R., Brown, M. Matthias, G., & Brachinger, H. W. (1999). Financial decision-making: Are women really more risk-averse? *American Economic Review*, 89(2), 381-385.
- Schubert, R., Gysler, M., Brown, & M. Brachinger, H. W. (2000). Gender specific attitudes towards risk and ambiguity: An experimental investigation. Institut für Wirtschaftsforschung, Eidgenössische Technische Hochschule. Retrieved from <http://www.cer.ethz.ch/research/wp017>.
- Schulhofer-Wohl, S. A. (2007). Heterogeneity, risk sharing, and the welfare costs of risk (Doctoral dissertation). University of Chicago, Chicago, IL.
- Silventoinen, K., Sammalisto, S., Perola, M., Boomsma, D. I., Cornes, B. K., Davis, C., ... & Kaprio, J. (2003). Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Research*, 6(05), 399-408.
- Sloan, F. A., & Norton, E. C. (1997). Adverse selection, bequests, crowding out, and private demand for insurance: Evidence from the long-term care insurance market. *Journal of Risk and Uncertainty*, 15(3), 201-219.
- Smith, V. K., Evans, M. F., Kim, H., & Taylor Jr., D. H. (2004). Do the near-elderly value mortality risks differently? *The Review of Economics and Statistics*, 86(1), 423-429.
- Soane, E., & Chmiel, N. (2005). Are risk preferences consistent? The influence of decision, domain, and personality. *Personality and Individual Differences*, 38(8), 1781-1791.

Spivey, C. (2010). Desperation or desire? The role of risk aversion in marriage. *Economic Inquiry*, 48(2), 499-516.

Stanton, S. J., Mullette-Gillman, O. D. A., McLaurin, R. E., Kuhn, C., LaBar, K., Platt, M., & Huette, S. (2011). Low- and high-testosterone individuals exhibit decreased aversion to economic risk. *Psychological Science*, 22(4), 447-453.

Stoltenberg, S. F., Batiem, B. D., & Birgenheir, D. G. (2008). Does gender moderate associations among impulsivity and health-risk behaviors? *Addictive behaviors*, 33(2), 252-265.

Sunden, A. E., & Surette, B. J. (1998). Gender differences in the allocation of assets in retirement savings plans. *The American Economic Review*, 88(2), 207-211.

Szpiro, G. G. (1997). The emergence of risk aversion. *Complexity*, 2(4), 31-39.

University of Washington. (2012). Quality control report for genotypic data. Retrieved from <http://hrsonline.isr.umich.edu/sitedocs/genetics/HRS_QC_RE_PO_RT_MA_R2012.pdf>

Van den Bos, R., Hartevelde, M., & Stoop, H. (2009). Stress and decision-making in humans: Performance is related to cortisol reactivity, albeit differently in men and women. *Psychoneuroendocrinology*, 34(10), 1449-1458.

Van Der Loos, M. L. M., Mulder, P., Hartong, E. G., Blom M. B., Vergouwen, A. C., van Noorden, M. S. . . . LamLit Study Group. (2010). Efficacy and safety of two treatment algorithms in bipolar depression consisting of a combination of lithium, lamotrigine or placebo and paroxetine. *Acta Psychiatrica Scandinavica*, 122(3), 246-254.

Van Praag, C. M., & Cramer, J. S. (2001). The roots of entrepreneurship and labour demand: Individual ability and low risk aversion. *Economica*, 68(269), 45-62.

Verdejo-Garca, A., Lawrence, A. J., & Clark, L. (2008). Impulsivity as a vulnerability marker for substance-use disorders: Review of findings from high-risk research, problem gamblers, and genetic association studies. *Neuroscience & Biobehavioral Reviews*, 32(4), 777-810.

Vinkhuyzen, A. A. E., Pederson, N. L., Yang, J., Lee, H. S., Magnusson, P. K. E., Iacono, W. G. . . . Wray, N. R. (2012). Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Translational Psychiatry*, 2(4), e102.

Visscher, P. M., Yang, J., & Goddard, M. E. (2010). A commentary on common SNPs explain a large proportion of the heritability for human height by Yang et al. (2010). *Twin Research and Human Genetics*, 13(06), 517-524.

Watson, J., & McNaughton, M. (2007). Gender differences in risk aversion and expected retirement benefits. *Financial Analysts Journal*, 63(4), 52-62.

Weber, E. U., Blais, A-R., & Betz, N. E. (2002). A domainspecific riskattitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, 15(4), 263-290.

Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., ... & Lim, U. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, 46(11), 1173-1186.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R. . . . Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42(7), 565-569.

Yang, J., Manolio, T. A., Pasquale, L. R., Boerwinkle, E., Caporaso, N., Cunningham, J. M. . . . Visscher, P. M. (2011). Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics*, 43(6), 519-525.

Yang, J., Lee, H. S., Goddard, M. E., Visscher, & P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, 88(1), 76-82.

Yang, S. K., Hong, M., Zhao, W., Jung, Y., Baek, J., Tayebi, N., ... & Song, K. (2014). Genome-wide association study of Crohn's disease in Koreans revealed three new susceptibility loci and common attributes of genetic susceptibility across ethnic populations. *Gut*, 63(1), 80-87.

Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., Von Schoultz, B., Hirschberg, A. L., & Johannesson, M. (2009). A randomized trial of the effect of estrogen and testosterone on economic behavior. *Proceedings of the National Academy of Sciences*, 106(16), 6535-6538.

Zhong, S., Israel, S., Shalev, I., Xue, H., Ebstein, R., & Chew, S. H. (2010). Dopamine D4 receptor gene associated with fairness preference in ultimatum game. *PLoS ONE*, 5(11), e13765.

Zhong, S., Israel, S., Xue, H., Ebstein, R., & Chew, S. H. (2009). Monoamine Oxidase A Gene (MAOA) associated with attitude towards longshot risks. *PLoS ONE*, 4(12), e8516.

Zuk, O., Schaffner, S. F., Samocha, K., Do, R., Hechter, E., Kathiresan, S., ... & Lander, E. S. (2014). Searching for missing heritability: Designing rare variant association studies.

Proceedings of the National Academy of Sciences, 111(4), E455-E464.

Zuk, O., Hechter, E., Sunyaev, S. R., & Lander, E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences*, 109(4), 1193-1198.

Zyphur, M. J., Narayanan, J., Arvey, R., & Alexander, G. (2009). The genetics of economic risk preferences. *Journal of Behavioral Decision Making*, 22(4), 367-377.