UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays on Health Economics

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

in

Economics

by

Jason T. Shafrin

Committee in charge:

Professor Julie Cullen, Chair
Professor Eli Berman
Professor John Fontanesi
Professor Roger Gordon
Professor Rick Kronick

2009

The dissertation of Jason T. Shafrin is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2009

DEDICATION

To my advisor, Julie Cullen for her suggestions, criticisms, and critiques which were always accompanied by unwavering encouragement and honesty.

To my parents, for a lifetime of support and love.

To Lisa, the love of my life.

EPIGRAPH

Energy and persistence conquer all things.

—*Benjamin Franklin*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my co-authors. Thank you to Daniel Wiesen during our international collaboration on Chapter 2,"Why Aristotle didn't get his flu shot: The impact of prudence on prevention." Thank you to John Fontanesi for his help on Chapter 3, 'Your Patients' Up-to-date Status and the CDCs Vaccine Distribution System." Finally, I would like to thank Uri Gneezy for his assistance on Chapter 4, "Why Does Getting Married Make You Fat? Incentives and Appearance Maintenance."

Thank you to *Health Economics* for their forthcoming publication of my paper "Operating on commission: How physician financial incentives affect surgery rates."

I would also like to thank my dissertation committee of Julie Cullen, Eli Berman, John Fontanesi, Roger Gordon and Rick Kronick. Their feedback and critiques have been invaluable. Thank you to Louis Eeckhoudt for his insight, time, and sense of humor. Thank you to the UCSD Economics department staff of Rafael Acevedo, Rebecca Franco, Mary Jane Hubbard, and Devaney Kerr.

On a personal note, I would like to thank my parents, Fred and Bonnie Shafrin, for their continual support and Lisa for her tireless love and affection.

VITA

2002	B. S. in Economics *cum laude*, University of Pennsylvania, Wharton School
2002	B. A. in Spanish *cum laude*, University of Pennsylvania, College of Arts and Sciences
2009	Ph. D. in Economics, University of California, San Diego
2001	Capital Budget Analyst, Government Scholars Fellow, NYC Housing Authority (New York, NY)
2002-2004	Financial Analyst, General Electric (Hartford, CT; Charlotte, NC; Barcelona, Spain)
2004-2009	Graduate Teaching Assistant, University of California, San Diego (San Diego, CA)
2006-2009	Founder and Editor of Healthcare Economist blog, (San Diego, CA)

PUBLICATIONS

Shafrin J, "Operating on commission: How physician financial incentives affect surgery rates", *Health Economics*, forthcoming.

Fontanesi J, Hirsch JD, Lorentz SM, Bowers D, Shafrin J, "Comparison of Pharmacists and Primary Care Providers as Immunizers", *American Journal of Pharmacy Benefits*, forthcoming.

ABSTRACT OF THE DISSERTATION

Essays on Health Economics

by

Jason T. Shafrin
Doctor of Philosophy in Economics
University of California San Diego, 2009
Professor Julie Cullen, Chair

This dissertation utilizes tools from the field of economics to analyze policy-relevant healthcare research questions. The four chapters in this paper can be summarized as follows:

Chapter 1. Physician compensation has a significant impact on surgery rates. When specialists are paid through a fee-for-system scheme rather than on a capitation basis, surgery rates increase 78%. The impact of primary care physician compensation on surgery rates depends on whether or not referral restrictions are present.

Chapter 2. Prudence preferences have been shown to influence precautionary savings, asset allocation, and optimal prevention levels. Using new measures of prudence, we find that 53% of survey participants are prudent and 15% are imprudent. Preferences for prudence are strong among both risk loving and risk averse individuals.

Chapter 3. Analyzing new data from Southern California providers, we find that the CDC's decision to centralize their vaccine distribution system lead to longer delivery times and increased vaccine stockouts.

Chapter 4. Married people weigh more than non-married individuals. We

suggest that exiting the dating market decreases one's incentive to maintain their appearance and leads to an increase in body weight. Evidence from a 14-year panel from the Netherlands shows that decreased incentives to maintain one's appearance after marriage partially explains this weight differential for women.

Chapter 1

"Operating on Commission: How physician financial incentives affect surgery rates"

This paper employs a nationally representative, household-based dataset in order to test how the compensation method of both specialists and primary care providers affects surgery rates. After controlling for adverse selection, I find that when specialists are paid through a fee-for-system scheme rather than on a capitation basis, surgery rates increase 78%. The impact of primary care physician compensation on surgery rates depends on whether or not referral restrictions are present.

1.1 Introduction

The question of how financial incentives affect physician decision-making has been a frequent subject of investigation in both the economics and medical fields.

Are doctors perfect agents for their patients, solely basing their medical care decisions on what is in the patient's best interest, or do physicians behave as *homo economicus*, strictly acting in a profit maximizing fashion? While many studies have focused on the relationship between financial incentives and primary care services, the treatment of specialist services has been inadequate. The topic of specialist care is particularly important as more and more graduates of U.S. medical schools are choosing to practice in specialty fields (Newton and Grayson, 2002). This paper explains how the joint financial incentives of the specialist and primary care physicians affect surgery rates.

In this study, I test two hypotheses. The first, or direct, hypothesis predicts that changing specialist compensation from capitation to fee-for-service will increase surgery rates. I find that financial incentives do significantly influence patient surgery frequencies. If a specialist is compensated via fee-for-service (FFS) as opposed to capitation, surgery rates increase by approximately 78%. In the outpatient setting, changing physician compensation increases surgery rates by 84%.

The second, or indirect, hypothesis examines the relationship between primary care compensation and surgery rates. If the quantity of primary care services influences the quantity of specialist services, one must take into account how primary care physician compensation affects the level of primary care services and, indirectly, the number of surgeries. Primary care can act as either a substitute or complement for specialist care. For example, it is possible that a drug administered to the patient within the primary care setting may be a substitute for surgery provided by a specialist. In some cases, having the primary care physician (PCP) prescribe the drug may negate the need for the specialist to perform a surgery. In other cases, primary care may be a complement to specialist care. For instance more frequent primary care visits may increase the likelihood of catching breast cancer at an early stage where surgery would be beneficial. If patients of fee-for-service PCPs have more office visits, I would observe higher rates of surgeries among patients with FFS PCPs

compared to patients with PCPs paid via capitation. Economic theory predicts that increasing primary care services should decrease (increase) specialist services if the two are net substitutes (complements).

My analysis indicates that the impact of PCP compensation on surgery depends on whether or not a patient's health insurance policy includes referral restrictions. In the absence of referral restrictions, primary care and specialist care are net substitutes. Changing the PCP's compensation from FFS to capitation increases surgery rates by 35% when the health plan does not apply referral restrictions. On the other hand, in the presence of referral restrictions, no statistically significant relationship appears between PCP compensation and surgeries. The presence of referral restriction likely tempers the substitutability of primary care and specialist medical services.

While previous studies have examined the relationship between physician compensation and specialist care, this study is unique because it uses nationally representative household data. Shrank et al. (2005) and Brinker et al. (2006) analyzed the connection between financial incentives and specialist care, but use patient visit data that is subject to bias from non-random sampling. Unlike most studies of specialist medical service provision, I use nationally representative data from the Community Tracking Study (CTS). The CTS includes not only survey response variables, but also physician compensation data collected directly from the patient's health plan. Relying on direct measures of physician compensation rather than proxies such as an HMO insurance dummy variable allows more accurate analysis of the influence of physician financial incentives. Taking into account both primary care and specialist physician compensation facilitates a more comprehensive analysis of the data.

The remainder of the paper proceeds as follows. Section 1.2 reviews the literature's findings about how financial incentives affect medical service provision. Also discussed are methods used in the literature to control for adverse selection and how

I employ these techniques in this paper. Section 1.3 develops a theoretical framework which generates two testable hypotheses. Section 1.4 describes the Community Tracking Study (CTS) 1996/1997 Restricted Use data file. Section 1.5 describes the how the hypotheses derived from the theoretical framework will be tested econometrically and section 1.6 shows the results of these tests. A concluding discussion can be found in section 1.7.

1.2 Literature Review

The manner in which financial incentives affect the provision of medical services has been studied extensively in both the economics and medical fields. The most recognized of these studies is the RAND Health Insurance Experiment (HIE). In one portion of the RAND HIE, households were randomly assigned between fee-for-service or prepaid group plans (Manning et al. 1987). Manning and co-authors concluded that patients in prepaid group practice plans had only 72% of the expenditures of those in fee-for-service plans. Other randomized trials such as Hickson, Altemeier, and Perrin (1987) and a randomized survey by Shen et al. (2004) found similar results. A review paper by Gosden et al. (2000) examined this literature in more detail.

While these randomized trials provide strong evidence that financial incentives matter to physicians, they are not without problems. The Manning and Hickson studies are over 20 years old and neither is nationally representative. The Shen paper is more recent, but the study used hypothetical survey responses rather than real-world data. Further, none of these papers examine on the connection between specialist compensation and the quantity of specialist care provided. Most importantly, these studies ignore interaction effects between PCP and specialist compensation.

Myriad studies have been conducted outside the RCT framework. Seminal work by Akerlof (1970) and Rothschild and Stiglitz (1976) suggests that when oper-

ating outside of the RCT setting, researchers must be mindful of adverse selection issues. In the medical setting, adverse selection can occur when individuals use their private information about their health risks to select physicians with varying compensation methods in ways that are correlated with this underlying risk. In the presence of adverse selection, it may appear that patients of capitation-reimbursed physicians utilize fewer medical services because of their compensation structure, when in reality, the reason for the lower patient utilization may be a healthier patient base. Evidence demonstrating the existence of adverse selection has been documented within Medicaid (Leibowitz, Buchanan, and Mann, 1992), Medicare (Dhanani et al., 2004), and employer-provided insurance plans (Cutler and Reber, 1998; Buchmueller, 2000). Since this paper relies on observed insurance type rather than random assignment, it is imperative that I address the issue of adverse selection.

There are three prominent methods used to deal with adverse selection. The first method is to control for the selection problems using observables. The use of primary diagnosis (Brinker et al. 2006), Diagnostic Cost Group (Yu, Ellis and Ash, 2001) or self reported health measures (Ettner 1997) is common. A second group of papers uses variation in plan compensation occurring due to a 'natural experiment.' Studies employing the 'natural experiment' framework have utilized changes in primary care compensation in England (Dusheiko et al., 2006), in Ireland's 'medical card' population (Madden et al., 2005), and within certain U.S. medical groups (Gosden et al., 2000). This 'quasi-experimental' approach is very appealing though it generally restricts the scope of the study to individuals affected by a government or employer policy change.

A final effective means to control for adverse selection is to examine a subpopulation of the data that has a restricted choice of insurance plans. Using the Community Tracking Study—the same dataset utilized in this paper—Polsky and Nicholson (2004) and Nicholson et al. (2004) analyzed medical expenditure data for privately insured individuals who only received one choice of health insurance from

their employer. According to Pauly and Percy (2000), nongroup insurance plans tend to have significantly higher load factors, provide less generous benefits, and are at a significant tax disadvantage compared to group policies. Consequently, for most individuals who have the option of only one type of health insurance at their place of work, this is their only *de facto* health insurance option. Adverse selection should not be a problem for individuals offered only one choice of health insurance at work.

This study significantly reduces the adverse selection problem using two of the these three methods. First, I control for observable health level using a detailed, twelve component measure of each individual's health. Second, I limit the data to a subsample of individual whose employers offer them only one choice of health insurance. Since the dataset comes from a single year, longitudinal analysis employing a 'natural experiment' policy shock is infeasible using the CTS data.

Among the few papers using household data that examine specialist service provision and adequately control for adverse selection are papers by Polsky and Nicholson (2004) and Saver et al. (2004). The Polsky and Nicholson paper analyzed how HMO-membership affects total expenditures, as well as hospital admissions. Limiting their sample to employed individuals offered only one choice of insurance from their employer, the authors found that the $188 difference between HMO and non-HMO medical expenditures per enrollee was explained by differences in physician compensation. HMOs, however, are only a proxy for physician compensation since many HMOs pay some physicians on a capitation basis and others on a FFS basis. Saver et al. (2004) compared the procedure frequencies within three different HMOs that paid some doctors on a FFS basis and others on a salaried or capitation basis. The authors found that switching from capitation to FFS compensation led to a 28% increase in the medical care provision rates and switching from salaried to FFS compensation led to a 44% increase. Unfortunately, the Saver study did not employ city or region fixed effects despite the authors' claim that variation in payment methodology within each HMO was largely due to local market considerations.

Furthermore, the paper only looked at three HMOs in the western United States, and thus the paper is not nationally representative.

1.3 Theoretical Framework

Economic theory predicts that specialist physician compensation should influence the quantity of specialist care a patient receives. Ellis and McGuire (1986) derive a model where physicians choose quantities of medical care in order to optimize a utility function that depends on patient health and the physician's own profit level. The model shows that physicians will under-supply medical care when they are paid via a prospective payment system such as capitation, but will over-supply medical care when they are paid based on a fee-for-service or cost-based paradigm. Intuitively, physicians are more likely to supply medical care when their marginal revenue is positive (i.e., when they are paid via FFS) compared to when the marginal revenue is zero or negative (i.e., when they are paid via capitation). The theoretical prediction from their model leads to my first hypothesis.

Hypothesis 1 (Direct Hypothesis)**.** *Holding all else constant, specialists paid via a fee-for-service (FFS) compensation method will perform more surgeries than those paid via capitation.*

This paper diverges from the standard literature in that it takes a comprehensive view of how medical care is provided. Not only should specialist physician compensation affect surgery rates, but primary care physician compensation should influence surgery rates as well. Under a comprehensive view of the supply of medical care, the quantity of primary care services should influence specialist care. Using similar logic as in the specialist case, PCPs will provide more medical services to their FFS patients than their capitation patients. The conclusion that primary care physicians provide more medical care to FFS patients is widely accepted in the literature

(Hickson, Altemeier, and Perrin, 1987; Krasnik et al., 1990; Polsky and Nicholson, 2004; Shen et al. 2004;) but not universally so (Davidson et al., 1992, Frank and Zeckhauser, 2007).

If one accepts that PCP compensation affects primary care levels, then one must also analyze how the quantity of primary care services affects the number of surgeries. If primary care and surgeries are substitutes, an increase in primary care services will decrease the marginal health benefit a patient receives for a given quantity of specialist services. For example, a salubrious drug may decrease the marginal benefit from surgery. When the specialist trades off profits against patient health, he will see that the marginal benefit of any specialist procedure will be lower when the patient is already receiving medication, and thus he will decrease the quantity of specialist care. On the other hand if primary care and specialist care are complements, one would find the opposite relationship. As the marginal benefit of surgery increases with additional quantities of primary care medicine, the specialist will increase the quantity of specialist care. This would be the case if visiting a PCP more frequently leads to a higher probability of an early diagnosis. The early diagnosis could increase the marginal benefit of surgery and also increase the quantity of surgeries per person. Since the quantity of primary care services influences the quantity of specialist care, primary care compensation should indirectly affect the quantity of specialist care provided.

The literature is mixed as to whether or not primary and specialist care are substitutes or complements. Using data from Italy, Atella and Deb (2008) demonstrated that primary and specialist care are in fact substitutes. In Fortney et al. (2005), primary care was found to be a complement for mental health services, a substitute for speciality care, and have a negative but statistically insignificant relationship with inpatient care. It is likely that primary care is a substitute for some forms of specialist care and a complement for other forms of specialist care. With the data available, this paper will only be able to test whether the primary care

substitution or complement effect dominates with respect to surgery rates, but can not isolate whether primary and specialist care are substitutes or complements on a procedure by procedure basis.

The above arguments are summarized in the following hypothesis:

Hypothesis 2 (Indirect Hypothesis)**.** *When primary care physicians are compensated via capitation rather than a fee-for-service mechanism, then the amount of surgeries performed will* increase (decrease) *if primary care and specialist care are net* substitutes (complements).

Further, non-financial incentives may influence how PCP care compensation affects the number of surgeries a patient receives. In the absence of referral requirements, the indirect hypothesis can be tested without complication. On the other hand, the presence of referral restrictions may reduce the substitutability between primary and specialist care. Later in this paper, I will examine in detail how referral restrictions affect the results from testing the Indirect Hypothesis.

1.4 Data

In order to test the two hypotheses outlined above, this paper employs the 1996/1997 Community Tracking Study (CTS) Household Survey Restricted Use dataset. The CTS is a nationally representative household survey of over sixty thousand individuals. The survey methodology uses stratified sampling and all subsequent coefficient estimates and standard errors derived in this paper will account for the idiosyncracies of this survey design. For more detailed information on the survey design of the CTS see: Center for Studying Health System Change (2000).

In general, household surveys have the advantage of providing a representative view of the health of a society. These surveys, however, often have inaccurate data regarding patient insurance information and diagnostic variables and health status

is often imprecisely reproted. Survey respondents generally do not know how their physician is compensated. On the other hand studies using the patient-physician encounter as the unit of observation give more detailed diagnostic information and more accurate information regarding physician compensation measures. Yet the physician-visit surveys also have their flaws. These studies often lack important demographic information such as patient income and education levels, and come from a non-representative sample. For instance, surveying individuals who visit a doctor will necessarily produce a sample that is more heavily weighted towards patients who are sick and who prefer more aggressive treatment methods. Furthermore, the number of visits may themselves be affected directly by physician compensation.

The CTS, however, draws from the best aspects of both household and physician visit data. The CTS is a nationally representative household survey. Thus, the data are not overly weighted towards sick individuals or those who treat their illnesses agressively. Furthermore by using the Restricted Use Followback Survey, I am able to observe detailed information regarding patient insurance information. In the Followback section, survey workers collected information regarding an individual's insurance coverage by directly contacting the patient's insurance company. The data collected from the insurance company include variables such as patient deductibles, copayment and coinsurance rates, referral requirements, and—most importantly for this paper—how the insurance company compensates primary care and specialist physicians.

The benefits of having observational data from the individual's insurance company include not only more information, but information less susceptible to measurement error. Using the CTS, Cunningham, Denk and Sinclair (2001) showed that only 30% of individuals were able to correctly answer four basic questions about their own insurance coverage. A study by Reschovsky et al. (2002) found that patient satisfaction with their health plan depended on whether they believed they were in an HMO, not whether they were actually in an HMO. These two studies

demonstrate the importance of using observational data when using insurance or physician compensation variables.

In the empirical analysis section of this paper, the data set is limited to only those covered by private insurance in which the Followback survey was able to match the respondent's information with the insurance companies' information. Among the 38,310 observations with private insurance, the match rate was 75%, limiting our survey to 28,578 people. On average, individuals for whom a match was not obtained were poorer, less educated, less likely to be married and more likely to be a minority than those for whom a match was possible. In the subsequent analysis, I use Followback survey weights which correct for this attrition bias.

The sample is then further pared down to include only individuals between ages 18 and 64. The new sample size is 22,958 observations. The elderly are excluded in order to avoid insurance coverage confusion after one becomes eligible for Medicare. Children are excluded because of missing values in important covariates (i.e.: health level, years of education, number of health plans offered by an employer).

1.5 Estimation Strategy

1.5.1 Creating the variables

The goal of this paper is to find how different physician compensation schemes affect surgery rates. In order to do this, I create four dummy variables: CC, FF, CF, and FC. The first letter of the two letter abbreviation determines whether the primary care physician is paid via fee-for-service (F) or capitation (C). The second letter represents the manner in which the specialist is compensated.[1] Thus, the variable CF is equal to unity when the individual's insurance plan pays the primary care provider with a capitation payment and the specialist on a fee-for-service

1. The questions used in the survey are: "What is the typical method of payment that your organization uses for primary care providers (specialists)?"

basis. In the data, I combine capitation and salaried physicians into one "capitation" category. This is done because both salaried and capitated physicians receive a non-positive pecuniary benefit from providing marginal services and thus both have have an incentive to provide less treatment than in the fee-for-service setting. All subsequent regressions were also run separating out salaried and capitation physicians. The coefficient estimates—which are not reported here—are similar to when the capitation and salaried physicians are grouped together, but the standard errors are less precise.

In the data 27.5% of patient's insurance plans pay their primary care physician via capitation compensation. These same insurance companies pay specialists via capitation compensation 16.2% of the time. Remler et al. (1997) completed a national survey of physicians in 1995 and found that the mean primary care physician received capitation payment for 18% of patients and the mean specialist received capitation payment for 10% of patients. Remler's capitation reimbursement figures are likely lower than the ones found here for three reasons. First, the Remler data is survey based, not observational, and thus there could be some response bias. Secondly if physicians base their responses on their typical patient interaction, in the presence of adverse selection the physician will interact with capitation patients less frequently if this sub-population is comprised of healthier individuals. Finally, Remler included Medicare-eligible individuals in the data. Medicare generally pays physicians on a FFS basis and this will increse the proportion of observations with fee-for-service physicians. For these reasons, the proportion of fee-for-service PCPs in the CTS data appear to be reasonable.

One complication arising in the data is that 13% of patients have physicians who are compensated through a global capitation scheme. Under this arrangement, the costs for primary, specialist and hospital care are born entirely by the primary care provider. Examples of group practices paid via global capitation include single-plan group practices such as Kaiser Permanente, individual practice associations

(IPAs), and large medical groups. [2] Under global capitation, I observe how the medical group, but not the individual physician is paid. Previous studies by Rosenthal et al. (2002) and Conrad et al. (1998) document significant variation in how medical groups compensate individual physicians under global capitation. [3] Further, neither of these studies distinguishes the manner in which specialists are paid.

In order to isolate the ambiguities regarding physician compensation under global capitation, I create a fifth physician compensation variable, G, which is equal to unity if the individual is part of a health plan with global capitation and zero otherwise.[4] A description of all five physician compensation dummy variables is given in Table 1.1.

Table 1.2 breaks down the financial compensation dummy variables by type of health plan. The series of physician compensation variables do not fall neatly into insurance product categories. For instance, assuming that HMOs pay primary care physicians under a capitation scheme would be incorrect. Approximately 32% of PCPs and 54% of specialists who work for HMOs are paid via FFS. On the other hand, one fifth of PCPs paid via capitation do not work for HMOs.

In order to analyze the effect of physician compensation on specialist medical services in the data, I assume that individual i's number of surgeries is a function of the five physician compensation dummy variables as well as a vector of covariates $\mathbf{z}_i$. The vector of all independent variables is defined to be $\mathbf{x}_i$.

2. Ogrod (1997) notes that the physician structure under global capitation "...is usually a large, multi-specialty structure spread over a large geographic area."

3. Rosenthal et al. (2002) find that Californian IPAs are more likely to compensate primary care physicians through capitation, but medical groups are more likely to utilize either FFS or salaried payments. Further, many of these practices include significant bonuses based on cost of care, quality of care, productivity or profit sharing. Using data from Washington state, Conrad et al. (1998) find that 35% of medical groups pay PCPs through a salaried or modified salaried system, while 45% of medical groups pay PCPs strictly on a production basis.

4. The variable G also includes 'full professional capitation' in which the practice is liable for paying for primary and specialist care, but do not have to pay for hospital care. Under global capitation, the practice is also responsible for hospital expenses as well as primary and specialist care.

$$Surgeries_i = f(FF_i, CC_i, CF_i, FC_i, G_i, \mathbf{z}_i) = f(\mathbf{x}_i) \tag{1.1}$$

The dependent variable that will be used in the subsequent analysis is the number of surgeries each individual received during the previous year. Using the 'number of surgeries' is chosen as the dependent variable has potential advantages and disadvantages. Patient recall of the number of surgeries undergone in the past year is likely more accurate compared to other medical utilization metrics. Further, the number of surgeries is the only variable in the data that pertains exclusively to specialists. On the other hand, because less that 5% of the sample has more than one surgery, subsequent regression estimates will be less precise.

The vector of explanatory variables, $\mathbf{z}_i$, includes individual information such as age, age-squared, gender, income, marital status and a constant term. Other covariates in $\mathbf{z}_i$ are education and race dummies, and each individual's deductible and coinsurance rate. Papers such as Saver et al. (2004) claim that much of the variation in physician compensation method is due to local market conditions and in order to control for this, dummy variables for each metropolitan statistical area (MSA) are included in all regressions. Finally, a health variable, which is proportional to the Physical Component Summary (PCS), is included. The PCS is continuous variable constructed from twelve detailed questions regarding the respondent's health status.[5] Summary information for each of these measures is shown in Table 1.3.

While some physicians may have all patients pay them in a homogenous manner, the majority of physicians treat some patients whose insurance company will pay them via capitation and other patients whose insurance will pay them via FFS. Newhouse and Marquis (1978) provide some evidence that physicians can differentiate levels of care for each patient in a mixed payment environment. The authors

5. For example, two of the twelve questions included in the PCS calculation ask whether a person's health limits them from "Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf" and "Climbing several flights of stairs?" Other questions inquire as to the emotional state of the person and whether or not they suffer significant pain.

observed that individual patient compensation arrangements do affect physician behavior in the case of hospital admissions, the length of a hospital stay and the number of office visits. On the other hand, physicians are not able to discriminate care levels by insurance type in the case of the length of an office visit (Glied and Zivin, 2002; Frank and Zeckhauser, 2007) and prescribing behavior (Frank and Zeckhauser, 2007; Hellerstein, 1998). If discriminating medical care quantities on a patient-by-patient basis is impracticable, then subsequent results should be interpreted as the effect of increasing the percentage of the physician's patient base who is paid in certain manner.

1.5.2 Regression specification

I will be using the number of surgeries an individual has had during the prior year as the dependent variable. Since the number of surgeries is restricted to be a non-negative integer, a negative binomial regression is employed. A logical starting point when working with count data is to use the Poisson model, but a Poisson regression will systematically underestimate standard errors if the dependent variable's conditional mean and variance are not equal. To test whether or not the Poisson model is appropriate, one can estimate a type I negative binomial regression (Negbin I) where $Var(S|\mathbf{x}) = (1+\alpha)E(S|\mathbf{x})$, and use a Wald test to verify whether or not $\alpha = 0$ (Cameron and Trivedi, 1986). When applied to the CTS data, the Wald test rejects the null hypothesis that $\alpha = 0$ ($p \leq 0.01$). Therefore, I employ the Negbin I model through the majority of the paper.[6]

Coefficients are reported as the marginal effect of a change in the explanatory variable on the absolute surgery rate. For the physician compensation variables, estimates are calculated as the absolute change in the surgery rate for an average

6. Although the results from the Negbin I model are presented in this paper, the analysis was also conducted using the Negbin II framework in which it is assumed that $Var(S|\mathbf{x}) = E(S|\mathbf{x})[1 + \alpha E(S|\mathbf{x})]$. Neither the qualitative or quantitative results change significantly when the Negbin II specification is used.

individual when the binary variable changes from zero to one.

The parameter estimates on the physician compensation variables are estimated in a partial equilibrium setting. These parameter estimates should be interpreted as the effect of moving a single patient from a physician compensated in one manner to a different physician compensated in a different manner. Hellinger (1996) notes that it is possible that physicians who treat patients in a less aggressive manner may migrate towards jobs with capitation payment while those with more aggressive treatment styles may choose FFS employment. With the data used in this paper, however, it is impossible to separate whether the physicians migrate to financial plans whose incentives favor their practice style or whether the financial incentives actually change the physician's practice style from an unobserved counterfactual. It may also be the case that financial incentives affect the number of specialist hours worked, and thus differences in surgery rates by physician compensation may be due to differences in patient wait times rather than an explicit physician decision. From the data used, I am unable to determine through which of these pathways does the causal mechanism operate.

1.5.3 Controlling for Adverse Selection

All non-randomized health economics studies must address the issue of adverse selection. If physicians who are compensated via capitation provide less generous medical services, but also contract with health plans with lower premiums, it is likely that healthy individuals will choose these health plans. If this was the case, individuals with capitated PCPs would have fewer surgeries not due to financial incentives, but because they are healthier patients.

To control for these selection effects, I examine a subsample of the data where individuals have no choice of health insurance. Individuals whose employers offer them only one health plan choice do not have this sorting option. Non-group in-

surance is generally a poor substitute for group insurance purchased through one's employer. Thus, individuals offered one choice of health insurance at work are *de facto* constrained to choose this plan. In fact, the main determinants of whether or not a worker is offered more than one health plan at work is the firm's size and location—not the individual's characteristics.[7] By using the subsample of individuals offered a single health insurance plan, adverse selection will not contaminate coefficient estimates from subsequent regression analysis.

Even if different types of individual are not sorting to firms offering only one health plan, there are still three confounding factors that could bias the coefficients estimated within the"no choice" subsample; these are spousal insurance, job choice, and employer agency. Individuals who have only one choice of insurance at work may not, in fact, be limited to this single option if they are married. In the data, I count individuals as having no choice of health insurance even if they take up their spouse's health insurance. The take-up of spousal insurance could lead to significant endogeneity problems since 11.1% of individuals are in households where both spouses maintain employer-provided health insurance. In order to verify whether or not having a spouse with employer-provided health insurance alters the results, I run the preferred negative binomial regression while excluding working individuals who can take-up health insurance both at their own and their spouse's place of employment. The results are quantitatively similar to those forthcoming.[8]

I also assume that an individual's place of employment is chosen without regard to physician compensation. Since Cunningham, Denk and Sinclair (2001) demonstrate that individuals have poor knowledge of their own insurance, it is un-

7. To verify whether or not this is the case, I run a probit regression (not shown) of whether or not a firm offers more than one insurance plan on a variety of individual and firm characteristics. The psedo-R^2 including all variables is 0.103. When I include only individual demographic and socioeconomic variables, the psuedo-R^2 falls to 0.018. If I include only firm size, industry, and city fixed effects variables, the pseudo-R^2 decreases only slightly to to 0.098.

8. Excluding spouses who both have employer-provided health insurance, I observe a 61.6% (75.7%) increase in total (outpatient) surgeries compared to a 77.5% (84.3%) increase in surgeries when these couples are included in my later regressions.

likely that health plans are chosen based on physician compensation method. Individuals may, however, select their employer based on whether they offer an HMO, preferred provider organization (PPO), point-of-service plan (POS) or FFS health plan. Health plan type is correlated with the manner in which physicians are compensated so any remaining bias could be driven by endogenous job choice.

Furthermore, employers could act as agents for their workers. An employer with sicker employees may elect to offer a single plan which pays physicians on a FFS basis. Another employer composed of healthier workers may choose to offer one plan which pays physicians via capitation. Thus, average employee health levels may be correlated with plan choice even if the employers only offer workers one health plan.

In order to test empirically whether or not job selection will contaminate the identification strategy, Table 1.4 shows mean covariate values by physician compensation type. The top portion of the table displays the results for individuals who are offered one choice of insurance through their employer (i.e., "no choice") and the bottom portion of the table displays the covariate means for individuals who are offered multiple health plans at work (i.e., "choice"). I test whether or not the mean covariate values for each physician compensation type are the same.

For individuals with a "choice" of health plans, there are statistically significant differences by compensation type for nearly all the covariates. Even though I cannot statistically reject the equality of means for the $Health$ variable, the p-value is 0.108. On the other hand, for individuals in the "no choice" subsample, the observations are much more homogenous across physician compensation types. I observe that the p-value for equality of means for $Health$ is 0.577 and for Age is 0.193. Since health and age are the observable variables most likely to uncover selection effects, this is strong evidence that limiting the sample to individuals offered only one choice of health insurance at work has significantly attenuated any adverse selection problems.

I do note, however, that there appears to be statistically significant differ-

ences in income, gender and minority status in the "no choice" subsample. As a robustness check, I use two regression variations to analyze the impact of income on surgery rates. First, I run the preferred regression separately for high and low income individuals; secondly, I add (*income*)(*physician compensation*) interaction terms in the regression. In both cases, the results from testing the direct and indirect hypotheses are similar to those of subsequent regressions, but the magnitude of this effect is larger for those in higher income brackets. To further test the sensitivity of my results, I run the preferred negative binomial regression separately by gender and by minority status. The results of these four regressions (i.e., Caucasian-only, minority-only, male-only, female-only) are quantitatively very similar to the results found in the rest of the paper. These results provide support that adverse selection issues are attenuated when the sample is limited to individual offered one choice of health insurance at work.

1.6 Results

Now I will test the direct and indirect hypotheses proposed in earlier sections with the CTS data.

1.6.1 'No Choice' Specification

In the first set of specifications I restrict the sample to individuals whose employers offer them only one health plan. As discussed previously, this methodology—along with the use of the PCS variable to control for health level—should largely eliminate biases resulting from adverse selection. The first three columns of Table 1.5 give the results from the negative binomial regression for individuals in the "no choice" subsample.

Looking across all three specifications, the parameter estimates from variables in the **z** vector are mostly as expected. Healthier individuals have fewer surgeries.

Age does not have any impact on surgery rates, but this is likely due to the fact that age is correlated with health. Because they are at-risk for pregnancy, women are more likely to have surgery. Individuals who are married, have more education, come from a Caucasian background and earn more income are more likely to have surgery. These patterns may be explained by differences in insurance quality, however due to data restrictions I can not empirically test this possibility. Finally, it seems that coinsurance and deductible rates have little impact on the quantity of surgeries observed.

Let us examine physician compensation dummy variables using total surgeries as the dependent variable. Since the omitted dummy variable in the regression is FF, the marginal effects should be interpreted as the absolute change in the number of surgeries per person relative to this baseline where both the PCP and specialist are compensated via FFS. The average number of total surgeries per thousand persons with FF health plans in the 'no choice' specification is 143. The CF compensation scheme leads to 12 more surgeries per thousand people than the FF group; being in the CC group decreases surgeries by 56 for every thousand people compared to the FF plans. The estimates for FC should be viewed with caution since there are only 28 individuals in the 'no choice' subsample where $FC = 1$. Global capitation payment, G, has no significant effect on surgery rates, likely due to the ambiguous nature in which the physicians are compensated.

Let us now test the direct and indirect hypotheses. The direct hypothesis predicts that FFS specialists will preform more surgeries than specialists paid via capitation. If this hypothesis is true, this would mean that holding the primary care physician's compensation constant, a change in the specialist compensation from capitation to fee-for-service will increase surgery rates. In the regression this implies that $\beta_{CF} > \beta_{CC}$. Column 1 reveals that switching specialist compensation from capitation to fee-for-service increases surgery rates by 77.5%. The prediction that surgeons paid via FFS perform more operations than surgeons paid via capitation

holds true in the data at the 5% significance level ($p \leq 0.011$).

This result is in line with some of the results found in related literature. Saver et al. (2004) found that switching from capitation to FFS increased average procedure rates by 28% and switching from salaried to FFS compensation increased average procedure rates by 44%. On the other hand, Shrank et al. (2005) found that switching from capitation to FFS increased cataract surgery rates by 94% for Medicare beneficiaries and by 123% for individuals in commercial plans. The magnitude of effect of physician compensation on specialist care found in this study falls in between the Saver and Shrank results.

The indirect hypothesis predicts that holding specialist compensation constant, changing PCP from FFS to capitation compensation will increase (decrease) surgery rates if primary care and surgeries are net substitutes (complements). With respect to the regression, the indirect hypothesis implies that $\beta_{CF} > \beta_{FF}$ if primary care and surgery are net substitutes but $\beta_{CF} < \beta_{FF}$ if they are net complements. Since FF is the omitted variable, the prediction becomes $\beta_{CF} > 0$ and $\beta_{CF} < 0$ respectively. The coefficient estimates show that switching the primary care physician from capitation to FFS increases surgery rates 9% but this estimate is not statistically different from zero. The interpretation of this coefficient is ambiguous. It could mean that quantities of primary care and surgeries are unrelated. On the other hand, the estimate could imply that primary care and specialist care are sometimes substitutes and sometimes complements, but the net effect is no relationship between the two sectors.

The results of the negative binomial regression for outpatient and inpatient surgeries are shown in columns 2 and 3 respectively. When I test the direct hypothesis, I find that outpatient surgery rates are 84.3% higher when the specialist switches from capitation to fee-for-service compensation. This result is significant at the 1% level ($p \leq 0.003$). With respect to the indirect hypothesis, changing the primary care provider from a fee-for-service to a capitation compensation basis increases surgery

rates by 15.2%, but this result is not statistically significant at even the 10% level.

The third column of Table 1.5 shows the results from the negative binomial regression using inpatient surgeries as the dependent variable. The test of the direct hypothesis is signed as predicted but statistically insignificant and of a very small magnitude. The indirect effect of the PCP compensation on surgery rates is also not statistically different from zero.

The finding that the effect of physician compensation is stronger in the outpatient than the inpatient case should not come as a surprise. Shen et al. (2004) found survey responses of treatment intensity varied greatly between fee-for-service and capitation patients in three cases where treatment was 'more elective,' but in the relatively non-elective case of the management of end-stage heart failure, the authors found no difference in the physician's intended treatment intensity for FFS and capitation patients. Shrank et al. (2005) found that cataract surgery rates were responsive to financial incentives, but ophthalmological surgery rates for non-elective procedures did not respond when physician financial incentives changed. It seems reasonable to believe that outpatient surgeries are more likely to be considered 'elective' than inpatient surgeries; thus observing that outpatient surgeries are more responsive to financial incentives than inpatient surgeries is consistent with the findings in the literature.

1.6.2 'Choice' Specification

Although Table 1.4 gives evidence that patients are not randomly distributed across physician compensation types in the "choice" specification, one may still worry whether or not subdividing the sample was a necessary step. Subdividing the sample decreases precision and thus should only be utilized if adverse selection occurs in the "choice" subsample. If I observed that $\tilde{\beta}_{CF}$ in the "choice" specification were of a smaller magnitude than the same coefficient in the "no choice" setting, this

would give some suggestive evidence that adverse selection is present. Two separate effects are operating in the "choice" specification; PCP capitation should increases surgery rates if primary care and specialist care are substitutes, but PCPs paid via capitation should attract healthier patients and thus reduce surgeries. If adverse selection is present in the "choice" specification, then the patient sorting effect should be stronger than in the "no choice" subsample. If selection is based on how the primary care physicians are paid, healthier individuals will sort into the CF physician compensation group.

In table 1.5, columns 4, 5, and 6 give the results of a negative binomial regression when individuals have a choice of insurance plans at their work. The analysis is conducted using total surgeries, outpatient surgeries and inpatient surgeries as the dependent variable. In all three specifications, the parameter estimate for $\tilde{\beta}_{CF}$ changes signs compared to the "no choice" results and becomes negative—implying $\tilde{\beta}_{CF} < \tilde{\beta}_{FF} = 0$. Assuming primary and specialist care are substitutes, these results suggest that omitted factors affecting surgery rates are correlated with a person's choice of health plan.

1.6.3 Robustness Check

To further test for the presence of adverse selection, one ideally would like to examine health behaviors unrelated to physician compensation. As a robustness check, I evaluate if physician compensation affects the probability an individual receives a flu shot. While it is possible that differences in primary care compensation could influence vaccination rates, specialist compensation should have no effect on the probability a patient receives a flu shot. Any differences in vaccination rates across the different types of specialist compensation are attributable to patient sorting.

To test this hypothesis I run a probit regression of flu shot receipt on physician compensation and all other covariates included in the preferred specification. Table

1.6 displays the results from this regression. For the "no choice" group, I fail to reject the null hypothesis that physician compensation has no effect on flu shot probabilities ($p \leq 0.199$). I also fail to reject the null hypothesis that specialist compensation has no effect on flu shot probabilities ($p \leq 0.249$). On the other hand, when individuals have a choice of health insurance, PCPs paid via capitation have patients who are more likely to get flu shots. I reject the hypothesis that physician compensation does not influence the probability of getting a flu shot ($p \leq 0.001$). Further, I find that having a specialist paid via capitation increases the probability that an individual will receive a flu shot ($p \leq 0.061$). Because specialist compensation is only strongly correlated with flu shot frequency when individuals are offered multiple health plan choices, restricting the sample to individuals with only one health plan choice has attenuated problems of adverse selection.

Another method to test for adverse selection is to investigate whether or not individual covariates affect health plan choice. Even if employees are not able to choose health plans individually, one could still observe sorting behavior if firms choose health plans based on underlying worker characteristics. In order test whether or not employee covariates predict health plan choice, I use a two stage methodology. In the first stage, I exclude the physician compensation variables and run the preferred negative binomial regression for all insured individuals (i.e., the 'choice' and 'no choice' groups combined). In the second stage, I run a probit regression of health plan choice on the predicted values from the initial regression. Table 1.7 displays these results.

For the cases of selecting an HMO or a PPO/FFS health plan, individual characteristics related to surgery are more highly correlated with health plan choice when the individual is able to choose from a menu of health plans. Individual characteristics correlated with surgery decrease the probability of choosing an HMO when the individual has a choice of plan ($p \leq 0.015$), but these same predicted values have no statistically significant relationship to HMO choice when I limit the sam-

ple to those offered only one health plan ($p \leq 0.420$). Similarly for the case of choosing a PPO or FFS plan, demographic variables predicting surgery also are correlated with an increased probability of choosing a PPO or FFS when individuals are offered multiple plans ($p \leq 0.011$), but the same information has no predictive value of a choice of a PPO or FFS plan when individual are only offered one health plan ($p \leq .870$). This provides further evidence that even if employers are selecting health plans based on average worker characteristics, this problem is fairly small in magnitude and unrelated to surgery rates.

1.6.4 Alternate Regression Specifications

Even though the negative binomial is my preferred econometric specification, it has some limitations. For instance, let us assume that person A has one surgery and person B has two. While it would be reasonable to assume that person B utilizes more medical services person than person A, she may not use twice as many services. To take this into account, columns 1 and 2 of Table 1.8 use an ordered probit regression, so that the total number of surgeries becomes a categorical variable rather than a count variable. The results shown in Table 1.8 confirm that the ordered probit regression does not materially change the results in either the "choice" or "no choice" subpopulations.

Also, since less than 5% of individuals in the Community Tracking Study have more than one surgery during the year, the surgery variable can be redefined to be a dummy variable equal to unity when an individual had one or more surgeries during the year and zero otherwise without much loss in information. A logit regression is run using this binary variable as the dependent variable. This ensures that a few outliers are not driving the results. The results from the logit regression are reported in Columns 3 and 4 of Table 1.8. In the logit, "no choice" case, individuals with CF health plans are 60.3% ($p \leq 0.045$) more likely to have surgery during the year

than than those with CC plans and 6.8% ($p \leq 0.758$) more likely to have surgery than individuals with FF plans. The similarities across parameter estimates of the negative binomial, ordered probit, and logit regressions indicate that the findings of this study are robust.

To this point in the paper, I have treated global capitation of group practices as distinct from direct physician compensation via capitation. However, Rosenthal et al. (2002) show that in California, about 50% of medical groups who receive global capitation have profit sharing. Profit sharing would create incentives for physicians with global capitation to act more like physicians paid directly via capitation. To control for this, I conduct all previous regressions combining the CC and G groups. This implicitly assumes that all specialists in global capitation groups are paid via capitation. The results combining the capitation and global capitation groups are qualitatively similar to previous results, but parameter estimates of the direct hypothesis are generally of smaller magnitude.

1.6.5 Referral Requirements

One confounding factor ignored to this point is the possibility that non-financial incentives play a role in surgical rates. Referral requirements compel patients to receive 'permission' from their primary care doctor before they visit a specialist. These referral restrictions act as a constraint on how primary care financial incentives can indirectly affect the quantity of surgeries.

A new specification to isolate the effects of referrals would be:

$$Surgeries = g(nFF, nCC, nCF, nFC, nG, rFF, rCC, rCF, rFC, rG, \mathbf{z})$$

where $rFF = ref * FF$ and $nFF = (1 - ref) * FF$, where ref is a dummy variable equal to one when a referral is required by the health plan and zero otherwise. I now test whether or not referral requirements have an impact on my tests of the direct

or indirect hypotheses. The direct hypothesis would predicts that $\beta_{nCF} > \beta_{nCC}$ and $\beta_{rCF} > \beta_{rCC}$; the indirect hypothesis predicts that $\beta_{nCF} > \beta_{nFF}$ and $\beta_{rCF} > \beta_{rFF}$ if primary and specialist care are net substitutes.

Table 1.9 displays the results of the negative binomial regression by referral requirement status for the direct and indirect hypotheses. The theoretical prediction that the direct effect of moving specialist compensation from capitation to FFS will increase surgery rates still holds. When referrals are not required, there is an 87% ($p \leq 0.112$) increase in the total surgery rate when the specialist is paid via FFS rather than capitation. When referral requirements are present, there is a 66% ($p \leq 0.074$) increase in the total surgery rate when specialist compensation changes from capitation to FFS.

For the indirect hypothesis, the net effect of changing primary care compensation rates from FFS to capitation depends on whether or not there are referral requirements. With no referral requirement in place, when the PCP is paid via capitation, surgery rates increase by 35% ($p \leq 0.082$). On the other hand, when referral restrictions are in place, having a capitation-compensated PCP does not seem to have an impact on surgery rates. The regression shows a statistically insignificant 13% decrease ($p \leq 0.477$) in surgery rates when there are referral restrictions and the PCP is paid on a capitation basis. These results suggest that primary care and surgery are net substitutes in the absence of referral restrictions, but that referral restrictions attenuate this effect. It may be the case that capitation compensated PCPs who would have referred patients to a specialist for surgery are now compelled to treat them themselves because of the referral restrictions.

1.6.6 The problem with using HMO dummy variables as a proxy for capitation payment schemes

A naive researcher may use an HMO dummy variable as a proxy for how capitation payment affects surgery rates. Yet from Table 1.2, one can clearly see that assuming that all HMO's pay their physician on a capitation basis is unwise. Over 30% of HMO primary care physicians and over 50% of HMO specialists are paid on a FFS basis. Table 1.10 shows us the results of the standard negative binomial regression using an HMO dummy variable to proxy for capitation payment.

In the first column, I examine the entire sample of individuals with employer provided insurance aged 18-64. Patients who have an HMO health plan have about 10% fewer surgeries than those with a POS, PPO or FFS health plan. These results are statistically significant at the 10% level ($p < .080$) but do not control for adverse selection. When one limits the sample to individuals with only one choice of employer-provided insurance, I see that HMO patients only have -0.4% fewer surgeries and this coefficient has a large p-value ($p \leq 0.96$). The naive researcher would conclude that capitation payment schemes do not materially affect surgery rates.

The findings using an HMO dummy variable starkly contrast with the results using direct observation of physician compensation. Using the physician compensation data, I find that the surgery rates for patients who pay their specialist on a FFS basis are 78% higher than the surgery rates for patients whose health plan uses a capitation scheme. By using more accurate data with using robust controls for adverse selection, I directly identify the impact of different physician compensation schemes without resorting to noisy measures such as an HMO health plan dummy variable.

1.7 Discussion

This study has found that switching specialist compensation from capitation to FFS increased total (outpatient) surgery rates by 78% (84%). As predicted by the literature, the results using inpatient surgeries were not statistically significant. Even when taking into account referral requirements or altering the econometric specification, it was consistently shown that FFS specialists preform more surgeries per person on average than specialists paid on a capitated basis.

The relationship between PCP compensation and surgery rates depended on the presence of referral restrictions. In the absence of referral requirements, changing PCP compensation from FFS to capitation increased total surgeries rates by 35%. This provides some evidence that primary care and surgeries are net substitutes. When referral requirements are in place, however, this relationship disappears, likely because referral requirements restrict capitation-compensated PCPs from sending patients to specialists.

This study is one of the first to use nationally representative, household data to estimate how physician financial incentives affect the provision of specialist services. Further, using the Community Tracking Study Restricted Use data set eliminates serious measurement error biases that have been shown to occur in household data. Any selection effects that may remain after restricting the sample to individuals offered one health insurance plan through their employer would bias the estimates against concluding in favor of the indirect hypothesis. The effect of any remaining adverse selection bias on direct hypothesis is unknown.

Future research should investigate how specialist financial incentives affect a broader range of specialist medical services. Also, subsequent empirical studies should model physician compensation in an even more sophisticated manner, taking into account the possibility that physicians are paid simultaneous via capitation and fee-for-service by the same patient. For instance, physicians may receive a capitation

payment, but their contract with a health plan may allow carve-outs where the physician receives FFS compensation for performing certain procedures. Finally, collecting data on patient outcomes or surgical quality would be useful in addressing the welfare implications of different physician compensation schemes.

1.8 Acknowledgements

Thank you to *Health Economics* for agreeing to publish this paper in a forthcoming edition.

1.9 Tables

Table 1.1. Physician Compensation Variables Explained

Variable	Explanation
FF	PCP paid via fee-for-service; Specialist paid via fee-for-service
CC	PCP paid via capitation; Specialist paid via capitation
CF	PCP paid via capitation; Specialist paid via fee-for-service
FC	PCP paid via fee-for-service; Specialist paid via capitation
G	Practice paid via *global* capitation, physician compensation uncertain

Table 1.2. Physician Compensation by Insurance Type

	FF	CC	CF	FC	G	Total
HMO	2726	1101	1956	78	2767	8628
	(0.173)	(0.853)	(0.682)	(0.907)	(0.924)	(0.376)
POS	2129	173	897	6	207	3412
	(0.135)	(0.134)	(0.313)	(0.070)	(0.069)	(0.149)
PPO	7698	16	15	2	20	7751
	(0.490)	(0.012)	(0.005)	(0.023)	(0.007)	(0.338)
FFS	3167	0	0	0	0	3167
	(0.201)	(0.000)	(0.000)	(0.000)	(0.000)	(0.138)
Total	15,720	1290	2868	86	2994	22,958
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Number of observations; proportion of column in parentheses

Table 1.3. Table of Means

Variable	Mean	Std. Dev.	Min	Max
$Total\ Surgeries^{\dagger}$	0.187	0.524	0	5
$Outpatient\ Surgeries^{\dagger}$	0.137	0.448	0	5
$Inpatient\ Surgeries^{\dagger}$	0.050	0.257	0	5
FF	0.721	0.448	0	1
CC	0.054	0.227	0	1
CF	0.106	0.308	0	1
FC	0.004	0.064	0	1
G	0.114	0.318	0	1
$Health$	5.126	0.844	1.1	6.8
$Male$	0.506	0.500	0	1
Age	40.0	11.98	18	64
$Yrs.\ of\ Edu.$	13.5	2.40	6	19
$Married$	0.713	0.453	0	1
$Caucasian$	0.778	0.415	0	1
$African-American$	0.101	0.302	0	1
$Asian-American$	0.039	0.193	0	1
$Latino$	0.082	0.274	0	1
$Coinsurance$ (%)	5.2	8.6	0	50
$Deductible$ ($)	125.7	174.9	0	3000
$Income$ $(\$10,000s)^{\dagger}$	5.38	3.37	0	15
$_n$	22958			

† Surgeries top-coded at 5; Income top-coded at 15 (i.e.:$150,000)

Table 1.4. Variable Means by Physician Compensation

No Choice

	FF	CC	CF	FC	G	Mean	P(equal means)
Health	5.19	5.27	5.19	5.22	5.17	5.19	0.577
Male	0.583	0.570	0.508	0.747	0.550	0.573	**0.003**
Age	39.0	37.6	38.4	39.9	38.4	38.9	0.193
Yrs. of Edu.	13.4	13.9	13.6	13.5	13.4	13.4	0.050
Married	0.699	0.686	0.652	0.717	0.692	0.693	0.272
Minority	0.166	0.425	0.317	0.464	0.293	0.203	**0.000**
Income ($10,000s)	5.07	5.42	5.36	6.71	5.18	5.12	**0.006**

Choice

	FF	CC	CF	FC	G	Mean	P(equal means)
Health	5.08	5.12	5.13	5.23	5.10	5.09	0.108
Male	0.469	0.470	0.468	0.453	0.487	0.472	0.641
Age	41.0	40.1	39.6	40.7	39.5	40.6	**0.001**
Yrs. of Edu.	13.5	13.7	13.8	13.3	13.7	13.6	**0.001**
Married	0.738	0.658	0.688	0.848	0.698	0.720	**0.001**
Minority	0.196	0.387	0.295	0.516	0.285	0.236	**0.001**
Income ($10,000s)	5.49	5.77	5.58	6.16	5.38	5.51	**0.014**

Table 1.5. Results from the Negative Binomial Regression

	No Choice			Choice		
Variable	Total	Outpatient	Inpatient	Total	Outpatient	Inpatient
CC	-0.056*	-0.035*	-0.004	-0.025	-0.024	0.000
	(0.031)	(0.020)	(0.004)	(0.023)	(0.026)	(0.009)
CF	0.012	0.014	-0.002	-0.029**	-0.030**	0.000
	(0.031)	(0.027)	(0.004)	(0.012)	(0.012)	(0.004)
FC	0.056	0.042	0.003	-0.071***	-0.072*	0.000
	(0.071)	(0.047)	(0.004)	(0.017)	(0.038)	(0.029)
G	0.020	0.028	-0.004*	-0.024*	-0.020**	-0.005
	(0.023)	(0.022)	(0.002)	(0.013)	(0.010)	(0.005)
$Health$	-0.048***	-0.024***	-0.005***	-0.082***	-0.054***	-0.023***
	(0.008)	(0.005)	(0.002)	(0.008)	(0.006)	(0.002)
$Male$	-0.049***	-0.017**	-0.010***	-0.033***	-0.008	-0.022***
	(0.009)	(0.008)	(0.003)	(0.007)	(0.007)	(0.004)
Age	-0.002	0.001	-0.001	0.001	0.002	-0.001
	(0.003)	(0.003)	(0.000)	(0.002)	(0.002)	(0.001)
$Married$	0.018	0.018*	-0.001	-0.033***	-0.050***	0.008**
	(0.015)	(0.010)	(0.003)	(0.012)	(0.011)	(0.003)
$Coinsurance$	0.000	-0.001	0.000**	0.000	0.000***	0.000
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
$Deductible$	0.000	0.000	0.000	0.000*	0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$Income$	0.001	0.000	0.001	0.006**	0.006**	0.000
	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.000)
CF vs. CC	77.5%**	84.3%***	2.5%	-2.8%	-4.7%	0.6%
CF vs. FF	8.6%	15.2%	-2.0%	-16.0%**	-20.8%**	0.7%
$P(CF = CC)$	0.011	0.003	0.677	0.835	0.818	0.982
$P(CF = 0)$	0.692	0.601	0.621	0.018	0.014	0.944

Age^2; Education, Race and City Dummies included in regression but not shown.

***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$

Table 1.6. Probability of Receiving a Flu Shot

	No Choice	Choice
CC	0.005	0.049***
	(0.048)	(0.011)
CF	-0.039	0.029**
	(0.027)	(0.015)
FC	-0.007	-0.004
	(0.143)	(0.018)
G	0.049*	0.0173
	(0.028)	(0.016)
P(All coeff. equal 0)	0.199	0.001
$P(CF = CC)$	0.249	0.061

All covariates from Table 5 included in regression but not shown
Coefficients reported as percentage point marginal effects

Table 1.7. Covariate Index and health plan choice

	HMO		PPO/FFS	
Variable	No Choice	Choice	No Choice	Choice
Index	-0.221	-0.469**	0.046	0.471**
	(0.275)	(0.193)	(0.283)	(0.185)
Constant	-0.643***	-0.238*	0.319	-0.086
	(0.195)	(0.143)	(0.208)	(0.160)

***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$

Table 1.8. Alternative Econometric Specifications

Dependent Variable: Total Number of Surgeries

	Ordered Probit		Logit	
Variable	No Choice	Choice	No Choice	Choice
CC	-0.2240	-0.0712	-0.0384	-0.0140
	(0.1739)	(0.0855)	(0.0283)	(0.0200)
CF	0.0579	-0.0944***	0.0078	-0.0192***
	(0.1179)	(0.0343)	(0.0254)	(0.0065)
FC	0.0258	-0.2833***	-0.0282	-0.0642***
	(0.2111)	(0.0780)	(0.0295)	(0.0168)
G	0.0757	-0.0638	0.0130	-0.0113
	(0.0771)	(0.0446)	(0.0160)	(0.0105)
$Health$	-0.2007***	-0.2767***	-0.0363***	-0.0605***
	(0.0273)	(0.0158)	(0.0053)	(0.0034)
$Male$	-0.1915***	-0.0988***	-0.0453***	-0.0224***
	(0.0303)	(0.0201)	(0.0074)	(0.0054)
% Change CF vs. CC			60.3%**	-3.8%
% Change CF vs. FF			6.8%	-12.8%***
$P(CF = CC)$	0.023	0.778	0.045	0.786
$P(CF = 0)$	0.623	0.006	0.758	0.003

All variables from Table 5 included in the regression but not shown.

***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$

Table 1.9. Referral Requirements (No Choice)

Dependent Variable: Total Number of Surgeries

	No Referral Requirement		Referral Requirement	
Hypothesis Test	% Change	p-value	% Change	p-value
Direct (CF vs. CC)	87.0%	0.112	66.0%*	0.074
Indirect (CF vs. FF)	35.3%*	0.082	-12.5%	0.477

All covariates from Table 5 included in regression but not shown

***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$

Table 1.10. HMO Naive Regression

	Full Sample	No Choice
HMO	-0.018*	-0.001
	(0.010)	(0.015)
Health	-0.074***	-0.048***
	(0.007)	(0.007)
Male	-0.039***	-0.050***
	(0.005)	(0.009)
Predicted Surgeries/person	0.173	0.145
HMO (% *Change*)	-10.2%*	-0.4%

All covariates from Table 5 included in regression but not shown
***: $p < 0.01$; **: $p < 0.05$; *: $p < 0.10$

Chapter 2

"Why Aristotle didn't get his flu shot: Experimental measures of Prudence and Prevention" with Daniel Wiesen

Prudence preferences have been shown to influence precautionary savings, asset allocation, and optimal prevention levels. In this paper, we will collect data from an experimental setting to measure prudence non-parametrically over both gains and losses. We find that 53.1% of individuals are prudent and 15.0% are imprudent. These preferences are constant regardless of the individual's level of risk aversion. Contrary to previous work in the literature, our results demonstrate that estimating prudence using parametric assumptions on the utility function often incorrectly categorizes risk lovers as imprudent. We also provide suggestive evidence that individuals classified as prudent are more likely to choose lower levels of prevention.

2.1 Introduction

Throughout history, prudence has been seen as one of mankind's ultimate virtues. Aristotle defined prudence or *phronesis* ($\phi\rho\acute{o}\eta\sigma\iota\varsigma$) as man's ability to think well about the nature of the world. He also stated that "the full performance of man's functions depends upon a combination of prudence and moral virtue; virtue ensures the correctness of the end at which we aim, and prudence that of the means towards it." Thomas Aquinas believed prudence was one of man's four principal virtues–along with temperance, justice, and fortitude.

In the field of economics, prudence is defined less poetically but perhaps more concisely. Seminal work by Kimball (1990) claims that an individual is prudent only if the third derivative of an agent's utility function is positive. Given lotteries with the same expected value and variance, prudent individuals prefer lotteries with more upside risk (i.e., right-skewed) to those with more downside risk (i.e., left-skewed). Prudence has been shown to be an important factor in areas such as intertemporal savings decisions (Kimball, 1990), asset allocation (Gomes and Michaelides, 2005) to preventive care decisions (Courbage and Rey, 2006).

In this paper, we use a new non-parametric methodology to measure prudence preferences. Dat was collected from 113 individuals using an online survey. We find that 53.1% of individuals are prudent and 15.0% are imprudent. Kahneman and Tversky (1979) suggest that risk preferences may vary significantly depending on whether a outcomes are perceived as gains or losses from a status quo. To test this point, we repeat our tests over both over gains and losses. We find that 56.6% of individuals are prudent over gains, but only 45.1% are prudent over losses. On the other hand, 17.7% of individuals are imprudent over gains, while 19.5% of individuals are imprudent over losses. Moreover, we found that both risk averse and risk loving individuals evince a strong preferences for prudence. This gives rise to the complementarity of both concepts Eckhoudt and Schlesinger (2006).

Our research also questions the usefulness of employing parametric utility function assumptions to capture prudence preferences. We fit constant relative risk aversion (CRRA) utility function parameters to individual responses regarding certainty equivalents. Parameter estimates from the CRRA utility function predict that 53 of 133 individuals should be imprudent. Of these 53 individuals, however, only 17.0% were imprudent and 56.6% were prudent. This suggests that functional form assumptions may be a poor method to estimate prudence preferences.

In order to verify that our experimental design truly does test for prudence, we present risky alternatives in two different ways. One method asks individuals to choose between two lotteries which are mean-variance preserving transformations of one another. Probabilities and outcomes are presented as "ballot boxes." In a second "coin flip" specification, we present individuals with a single 50:50 lottery. We then ask individuals to attach a zero mean risk to either the 'good' or the 'bad' outcome of the lottery. Preferences for prudence are consistent across both specifications.

Prudent preferences also can affect optimal levels of prevention. Because prudent individuals prefer to avoid very 'bad outcomes,' electing higher levels of prevention involves a chance of a worse outcome. For instance, if you decide to get a flu shot, this reduces the probability that you get the flu. However, you still have some probability—albeit smaller—of getting sick. Getting the flu *and* paying for a flu shot is worse than just getting the flu. Hence, a prudent individual such as Aristotle would prefer lower levels of prevention and may be less likely to receive a flu shot.

In this paper, we use our data to test whether or not prudent individuals select lower levels of prevention than imprudent individuals. Prevention is tested in a neutral framing. Selecting a higher level of prevention involves a larger up-front payment, but it also reduces the probability of a larger loss. In general, we find suggestive evidence that higher levels of prudence are associated with lower levels of prevention.

The paper proceeds as follows. Section 2 will give a brief literature review. Section 3 will define prudence in more detail and outline the expected utility theory of why prudence will affect risk aversion. In Section 4 we explain our experimental methodology and section 5 we give our preliminary results. We conclude in section 6.

2.2 Related literature

2.2.1 Prudence

Before we discuss our methodology, we must first rigorously define what is meant by the terms "prudence" and "prevention." Eeckhoudt and Schlesinger (2006) provided a definition of prudence outside of expected utility theory. If $\tilde{\epsilon}$ is a non-degenerate, zero mean random variable and k is a positive constant, than a prudent person will always have the following preferences: $[-k, \frac{1}{2}; \tilde{\epsilon}, \frac{1}{2}] \succeq [0, \frac{1}{2}; \tilde{\epsilon} - k, \frac{1}{2}]$. Further, if we define a risk premium as $w_r(x) = Eu(x + \tilde{\epsilon}) - u(x)$, then we can similarly define a prudence premium (i.e.: the change in the risk premium as wealth changes) by $w_p(x) = w_r'(x) = Eu'(x + \tilde{\epsilon}) - u'(x)$. Using Jensen's inequality, the prudence premium is positive if and only if $u''' > 0$ (i.e.: u' is convex). Kimball (1990) also demonstrated that an individual is prudent when the third derivative of their utility function is positive (i.e.: $u''' > 0$) and imprudent when this term is negative (i.e.: $u''' < 0$). The Kimball paper is of particular note since the research demonstrated that prudence will lead to "precautionary savings" in a two period model where income is uncertain in the second period. Menezes, Geiss and Tressler (1980) defined prudence as "downside risk aversion." Critical to our empirical work, this paper proved that a prudent person prefers lottery A to lottery B if lottery A was a particular mean-variance preserving transformation (MVPT) of lottery B. This MVPT must include a mean preserving contraction at the left tail of the distribution and a mean

preserving spread at the right tail of the distribution. By moving from lottery A to lottery B, we switch from a lottery with more downside risk (lottery A) to one with more upside risk (lottery B). In other words, comparing two random variables whose distributions have the same first two moments, a prudent person prefers the lottery whose the distribution is right skewed (positive skewness) compared to the lottery that is left skewed (negative skewness). Several empirical studies attempt to test for prudence as defined within the EUT-framework. In his seminal article Mao (1970) elicited preferences about uncertain investment opportunities varying in their third moment. Executives of leading US-companies participated in his (non-incentivized) survey study. He found that 4 out of 8 queried executives exhibited preferences for the lottery with the positive skew. More recently Unser (2000) analyzes individual's risk perception by applying the concept of lower partial moments. Within a complex financial decision context students were asked in a (non-incentivized) classroom experiment to state their preferences over several different risky assets. Using data from an Italian household survey, Eisenhauer and Ventura (2003) measure absolute and relative prudence for a broad cross-section of Italian household concerning financial investments. They found evidence for decreasing absolute and increasing relative prudence. A first attempt to test for prudence using the means of experimental economics was put forward by Gomez (2003). She determines prudence parameters by eliciting individuals' preferences over lottery pairs. The work allows individuals to attach a mean preserving increase in risk to either the best lottery outcomes (the prudent choice) or the worst outcomes (the imprudent choice). She found that more than 60% of the individuals to exhibit a prudent behavior. An experimental study by Deck and Schlesinger (2008) tests for higher order risks, i.e. prudence and temperance, outside the expected utility theory framework. Participating subjects are asked to make eight decisions between lottery pairs of the form introduced by Eeckhoudt and Schlesinger (2006). The authors find some evidence for prudence but no evidence for temperance.

2.3 Related literature

2.3.1 Prudence

Before we discuss our methodology, we must first rigorously define what is meant by the terms "prudence" and "prevention." Eeckhoudt and Schlesinger (2006) provided a definition of prudence outside of expected utility theory. If $\tilde{\epsilon}$ is a non-degenerate, zero mean random variable and k is a positive constant, than a prudent person will always have the following preferences: $[-k, \frac{1}{2}; \tilde{\epsilon}, \frac{1}{2}] \succeq [0, \frac{1}{2}; \tilde{\epsilon} - k, \frac{1}{2}]$. Further, if we define a risk premium as $w_r(x) = Eu(x + \tilde{\epsilon}) - u(x)$, then we can similarly define a prudence premium (i.e.: the change in the risk premium as wealth changes) by $w_p(x) = w_r'(x) = Eu'(x + \tilde{\epsilon}) - u'(x)$. Using Jensen's inequality, the prudence premium is positive if and only if $u''' > 0$ (i.e.: u' is convex). Kimball (1990) also demonstrated that an individual is prudent when the third derivative of their utility function is positive (i.e.: $u''' > 0$) and imprudent when this term is negative (i.e.: $u''' < 0$). The Kimball paper is of particular note since the research demonstrated that prudence will lead to "precautionary savings" in a two period model where income is uncertain in the second period. Menezes, Geiss and Tressler (1980) defined prudence as "downside risk aversion." Critical to our empirical work, this paper proved that a prudent person prefers lottery A to lottery B if lottery A was a particular mean-variance preserving transformation (MVPT) of lottery B. This MVPT must include a mean preserving contraction at the left tail of the distribution and a mean preserving spread at the right tail of the distribution. By moving from lottery A to lottery B, we switch from a lottery with more downside risk (lottery A) to one with more upside risk (lottery B). In other words, comparing two random variables whose distributions have the same first two moments, a prudent person prefers the lottery whose the distribution is right skewed (positive skewness) compared to the lottery that is left skewed (negative skewness). Several empirical studies attempt to test for prudence as defined within the EUT-framework. In his seminal article Mao (1970)

elicited preferences about uncertain investment opportunities varying in their third moment. Executives of leading US-companies participated in his (non-incentivized) survey study. He found that 4 out of 8 queried executives exhibited preferences for the lottery with the positive skew. More recently Unser (2000) analyzes individual's risk perception by applying the concept of lower partial moments. Within a complex financial decision context students were asked in a (non-incentivized) classroom experiment to state their preferences over several different risky assets. Using data from an Italian household survey, Eisenhauer and Ventura (2003) measure absolute and relative prudence for a broad cross-section of Italian household concerning financial investments. They found evidence for decreasing absolute and increasing relative prudence. A first attempt to test for prudence using the means of experimental economics was put forward by Gomez (2003). She determines prudence parameters by eliciting individuals' preferences over lottery pairs. The work allows individuals to attach a mean preserving increase in risk to either the best lottery outcomes (the prudent choice) or the worst outcomes (the imprudent choice). She found that more than 60% of the individuals to exhibit a prudent behavior. An experimental study by Deck and Schlesinger (2008) tests for higher order risks, i.e. prudence and temperance, outside the expected utility theory framework. Participating subjects are asked to make eight decisions between lottery pairs of the form introduced by Eeckhoudt and Schlesinger (2006). The authors find some evidence for prudence but no evidence for temperance.

2.3.2 Prevention

Prevention is defined in a very strict mathematical sense. Let the variable e represent the amount of preventive effort and let $p(e)$ represent the probability of a loss. Within the expected utility framework, we can calculate the individual's

decision function as follows:

$$e^* \equiv max_e \quad p(e)u(w - L - e) + [1 - p(e)]u(w - e). \tag{2.1}$$

The loss L could be a financial loss, a decrease in health, or some other type of disutility. Since our empirical specification uses monetary losses, we will assume L and e are all measured in dollars. The variable w is the individual's current wealth level.

Preventive effort, e, reduces the probability of a loss (i.e.: $p'(e) \leq 0$). It is important, to distinguish this from the case of self-insurance. Under the self-insurance framework, effort does not affect the probability of a loss, but increasing effort decreases the quantity of a loss if it occurs. In this paper, we only examine prevention as defined in equation (2.1). We abstract from any issues of how preventive effort affects the loss level. A real-world example of pure prevention is flu shots. An influenza vaccine or 'flu shot' immunizes the patient against a handful strains of the influenza virus, thus reducing the chance the patient will be stricken with influenza. The flu shot, however, does not protect the individual against *all* influenza virus strains. If you do fall ill from an influenza virus strain which was not included in the vaccine, you will experience the same health loss regardless of whether you received the flu shot or not. Thus, the flu shot decreases your probability of getting sick, but does not alter the degree or magnitude of the illness if one does contract a flu strain not included in the vaccine.

2.4 Theoretical predictions: The effect of Prudence on Prevention

Our hypothesis in this paper is that prudent agents will choose a prevention effort level which is *lower* than that chosen by imprudent agents. This hypothesis

stems from the theoretical work of Eeckhoudt and Gollier (2005). Let us assume a similar decision structure as in equation (2.1)

$$V(e) \equiv p(e)u(w - L - e) + [1 - p(e)]u(w - e). \tag{2.2}$$

Let e_n be the optimal level of preventive effort chosen by a risk neutral agent

$$e_n \equiv max_e \quad p(e)(w - L - e) + [1 - p(e)](w - e). \tag{2.3}$$

If this is the case, then we know that an individual exerts more effort if and only if

$$\begin{aligned} V'(e) &= -\{p(e_n)u'(w - L - e_n) + [1 - p(e_n)u'(w - e_n)]\} \\ &- p'(e_n)[u(w - e_n) - u(w - L - e_n)] \geq 0. \end{aligned} \tag{2.4}$$

Let e_p^* be the optimal level of prevention chosen by a prudent person and e_i^* be the optimal level of prevention chosen by a imprudent individual. Eeckhoudt and Gollier showed that if a risk neutral agent selects an optimal level of preventive effort such that the probability of a loss at this optimum is equal to $\frac{1}{2}$, (i.e.: $p(e_n) = \frac{1}{2}$), then all prudent agents will choose a preventive level less than the risk neutral agent, $e_p^* < e_n$ and all imprudent agents will choose a level of effort greater than the risk neutral agent, $e_i^* > e_n$. The intuition here is that at $p(e_n) = \frac{1}{2}$, a marginal change in e will not influence preferences over the second moment and thus only the agents prudence preferences will affect their choice of e^*. Any differences in risk preferences are irrelevant at $p(e_n) = \frac{1}{2}$.

In our experiment, however, we investigate prevention choices even when $p(e_n) \neq \frac{1}{2}$. In order to do this, we must take into account that increasing e from e_1 to e_2 will decrease $p(e)$. That is, if $e_1 < e_2$, $p(e_1) > p(e_2)$. If $p(e_1), p(e_2) > \frac{1}{2}$, then we know that moving from e_1 to e_2 will increase the variance of the lottery in equation (2.1).

If $p(e_1), p(e_2) < \frac{1}{2}$, then moving from e_1 to e_2 will decrease the variance.

To summarize, if $p(e_n) > \frac{1}{2}$, and an individual is risk averse and prudent, then the effect of risk aversion and prudence operate in the same direction, compelling the agent to choose $e_p^* < e_n$. If $p(e_n) < \frac{1}{2}$, then a risk averse agent will prefer a higher level of effort due to risk considerations, but if they are prudent they will prefer a lower level of effort due to the changing skewness of the distribution. Thus we have opposing effects.

The analysis for a risk loving, prudent person is very similar. If $p(e_n) < \frac{1}{2}$, risk and prudence preferences will induce the agent to decrease e_p^*. If $p(e_n) > \frac{1}{2}$, however, then risk preferences urge the agent to increase e_p^*, but prudence preferences pull them towards decreasing e_p^*. Here again the prediction is ambiguous. Table 2.1 provides a summary of hypotheses.

Table 2.1. Hypotheses

If			Then	Where
Risk	Prudence	$p(e_n)$		
Averse	Prudent	$> \frac{1}{2}$	$e^* < e_n$	
Loving	Prudent	$< \frac{1}{2}$	$e^* < e_n$	
Averse	Imprudent	$< \frac{1}{2}$	$e^* > e_n$	
Loving	Imprudent	$> \frac{1}{2}$	$e^* > e_n$	
r	p	$<> \frac{1}{2}$	$e^*_{r,p_i} < e^*_{r,p_j}$	$p_i > p_j$

2.5 Experimental design and procedure

In our experiment, we measure individual risk preferences, prudence preferences, and levels of preventive effort. Below is a table of our experimental stages.

The experiment was conducted online without any actual payoffs. We offered the chance of a $25 gift certificate to Amazon to induce participation. One hundred

Table 2.2. Experimental stages

Stage	Short description
RIAV	20 paired lotteries testing for risk aversion
PRUDA	'Ballot Box': 20 paired lotteries testing for prudence*
PRUDB	'Coin Flip': Where to place 0 mean risk to test for prudence**
PREV	10 Decsions on effort to reduce probability of loss
BACK	12 questions on the individuals background characterisitcs

* 10 decision over gains and 10 over losses
** 1 decision over gains and 1 over losses

twenty one individuals participated in the survey. We dropped 8 individuals from the survey because their responses revealed that the did not fully understand the questionnaire.[1] Thus, we were left with 113 observations.

The average age of the participants was 32.3 years old and average income was $63,450. Sixty four percent of the participants were male and 30% of the survey participants had some education in either economics or business.

2.5.1 Risk Aversion Questions (RIAV)

There are several accepted ways one can measure risk aversion suitable for a laboratory setting. Examples include eliciting the certainty-equivalent for a given lottery using Vickrey autions or the Becker-DeGroot-Marschak procedure Becker, DeGroot and Marschak (1964), or observing subjects' choices over lotteries varying in prizes offered for given probabilities Binswanger (1980). We implement a version of the method developed by Holt and Laury (2002). Because the Holt and Laury framework is widely implemented in laboratory experiments and the decicison task involved is relatively transparent, we apply it to evaluate risk aversion.

When applying Holt and Laury's method in RIAV, subjects make 20 choices

1. 5 individuals were dropped who choose weakly dominated choices [i.e., ($0, .5; -$40, .5)¿$0]. 3 more individuals were dropped who had multiple crossover points in the RIAV section.

between the paired lotteries. Our method applied slightly differs from Holt and Laury's. Instead of using positive payoff, subjects decide between lotteries with negative payoffs (losses). It is meaningful to use these kind of lotteries as subjects may exhibit different preferences towards losses compared with gains (Kahneman and Tversky, 1979). We use losses because in the section where individuals choose optimal prevention levels, their prevention level choice reduces the probability of a loss. Thus, estimates of risk preferences over losses are most relevant to this section.

Table 2.6 in the appendix shows the paired lotteries. Option A gives sure payoffs of \$0 in the first question and decreases in two dollar increments until it reaches a loss of \$38. For option B, participants face the following lottery in each question: (\$0, 50%; -\$40, 50%). Thus, individuals who switch to the risky option before option A equals \$20 are risk loving, and those who switch to the risky option when Option A equals \$20 or more are risk averse. Through this experimental design, the participants reveal their certainty equivalent over this lottery.

2.5.2 Prudence Questions (PRUDA, PRUDB)

In part A of the prudence portion of the survey, subjects are presented with two lotteries, denoted Option A and Option B. Both lotteries have the same mean and variance, but Option A is a mean-variance preserving transformation (MVPT) of Option B (see Table 2.7). In other words, the lotteries are identical in terms of their first two moments, but one is positively skewed while the other is negatively skewed. For instance, in question 1 on Table 2.7, the participant must choose between the lotteries (-\$10, 75%; -\$30, 25%) or (\$0, 25%; -\$20, 75%). Both lotteries have the same mean (-\$15) and variance (\$75), but only option B is positively skewed. In our experimental design, each positively skewed lottery is a MVPT of the negatively skewed lottery. According to Menezes, Geiss and Tressler (1980), only individuals with $u''' > 0$ will prefer the positively skewed lottery. It is not always the case that

a prudent person prefers positively skewed lotteries to negatively skewed lotteries; only when the positively skewed lottery is a MVPT of the other do we know that the individual's lottery choice will reveal their preference for prudence.

We use a ballot box format to visualize lotteries on subjects' decision screens. For an example see Figure 2.1. For the lottery, (-$10, 75%; -$30, 25%) in question 1 of PRUDA, the ballot box has 75 balls marked with a -$10 and 25 balls marked with a -$30. Questions 1-10 of the PRUDA section ask participants to choose between two lotteries involving losses; questions 11-20 ask participants to choose between lotteries involving gains.

Because we are dealing with preferences based on the third moment of an individual's utility function, one would expect there to be some error. Even very prudent people may not choose the prudent option every time. In order to ensure that our experimental design is robust, we include another prudence section: the 'coin flip' or PRUDB section.

In the coin flip section, individuals are presented with the lottery (-$10, 50%; -$20, 50%). Participants are faced with a choice over what to do with a second risky lottery, x, where x =($10, 50%; -$10, 50%). Would they rather play x when they lose 10 in the first lottery or when they lose 20 in the first lottery? A prudent person will play x when they lose only $10 in order to avoid the worst possible out come of -$30. An imprudent individual will play x only when they lose $20. The second question of PRUDB ask the same question except that the initial lottery is over gains, i.e., (+$10, 50%; +$20, 50%).

Mathematically, the two questions in coin flip section are exactly the same as two of the questions in the ballot box test. Thus, we can test whether or not our presentation methodology is influencing individual choices..

2.5.3 Prevention Questions (PREV)

When we test for prevention subjects are asked to choose an optimal level of effort in order to influence the probability of facing a loss, L. Here, subjects can choose an effort level from an interval between 0 and 9, i.e. $e_{ij} \in \{0, 1, ..., 9\}$. An effort level of 0 costs \$0, and effort level of 1 costs \$1, and so on. The specific questions can be found in Table 2.8.

When subject i chooses a higher level of prevention in decision j, this will reduce the probability of incurring the loss L (i.e., $p'(e_{ij}) \leq 0 \ \forall j$). However, increasing the effort level involves a higher upfront payment. Further, a higher effort level magnifies the level of the worse possible outcome, $-(L + e_{ij})$. Since prudent individual try to avoid left skewed distributions with large losses, we predict that prudent individuals will choose lower levels of e_{ij}. Since we have evaluated the risk and prudence preferences of our subjects, we can predict how much prevention they will choose.

In many classical utility functions (e.g.: CRRA) people who are risk averse are prudent and those who are risk loving are generally imprudent. Yet this need not be the case. Menezes, Geiss and Tressler state that "...both risk averters and risk preferrers can be downside risk averse. Similarly, individuals with increasing or decreasing utility functions can be averse to a leftward shift of risk. Thus $u'''(x) > 0$ is the only property of $u(x)$ necessary to define the set of downside risk averters." For instance, an individual with the utility function $u(x) = x - x - x^3$ will be risk averse, but imprudent when $x \in (0, \frac{1}{3})$. When an individual has the utility function $u(x) = x^3$, they will be risk loving and prudent .

To verify whether or not prudent individuals choose lower levels of prevention, we examine whether or not prudent individuals do indeed choose lower levels of prevention. Further, we will attempt to control for risk preferences, and then measure within each risk preference category whether or not prudent individuals choose lower levels of prevention than other individuals.

We also test whether or not preferences for prudence (imprudence) lead to less (more) prevention in the case where risk and prudence preference are allowed to have opposing predicted effects on prevention levels. To do this, we group together individuals with the same risk aversion levels as calculated from the RIAV section of the questionnaire. For those with the same risk aversion levels, we test whether or not prudent people select lower prevention levels than imprudent individuals. A regression framework can be implemented in the two prevention scenarios.

$$e_{ij} = \beta_0 + \beta_1 \mathbf{r}_i + \beta_2 \mathbf{r}_i p_i + \epsilon_{ij} \tag{2.5}$$

In equation (2.5), the vector $\mathbf{r}$ is a set of dummy variables indicating the individuals risk aversion score and p gives the probability that the individual is prudent. The key parameter is the vector β_2. If $\beta_2 < 0$, than higher prudence levels–conditional on risk preferences–lead to lower levels of prevention. Going forward $\mathbf{r}$ will include 22 dummy variables indicating risk averse behavior and the other indicates risk loving behavior. Using a vector of $\mathbf{r}$ with each individual's particular risk aversion score produces similar qualitative results.

2.5.4 Background questions (BACK)

At the end of the survey, we also asked the individuals basic questions about their background, such as their age, income, gender and education. Additional questions also asked about their general attitudes towards risk.

2.6 Results

2.6.1 Risk Aversion

Figure 2.2 displays the results from the Risk Aversion (RIAV) portion of the experiment. Risk averse individuals will choose the sure option 11 or more times. A risk loving person will choose the sure option less than 11 times. [2]

In our data, 69.0% of the sample is risk loving over losses and 31.0% of the sample is risk averse over losses. Evidence from the prospect theory literature has shown that individuals are generally risk loving over losses. In fact, Tversky and Kahneman (1992) sample 25 students and find that 87% of the students are risk loving over losses.[3]

2.6.2 Prudence

The results from the Prudence section are displayed in Table 2.3.

In the full sample of the PRUDA section, the average person answered 63.6% of the questions prudently. There are differences between average levels of prudence over gains and losses. The average person answered 61.3% of questions prudently when considering lotteries over losses, but for gains, the average person answers 65.5% of questions prudently. A t-test reveals that this 4.2% difference, however, is not statistically significant at conventional levels ($p < 0.173$).

The average number of questions individuals answer prudently may be a poor proxy for what proportion of the population is truly prudent. Instead, we define a person to be prudent if we can reject the null of them being being 'prudent neutral'

2. On the 11th question, an individual is asked whether the prefer (-$20) vs. ($0, 50%; -$40, 50%). We do not allow for indifference and thus a person who prefers -$20 is considered risk averse and the person who prefers the risky lottery with the same expected value is considered risk loving.

3. The 87% number corresponds to the case where the probability of the loss is greater than one half.

Table 2.3. Prudence Results by Risk Level

		Total	Loss	Gains
Full Sample	Avg	0.636	0.613	0.655
	SD	0.493	0.278	0.310
	% Prudent	0.531	0.451	0.566
	% Imprudent	0.150	0.195	0.177
Risk Loving	Avg	0.636	0.633	0.638
	SD	0.251	0.278	0.319
	% Prudent	0.551	0.487	0.551
	% Imprudent	0.167	0.192	0.192
Risk Averse	Avg	0.630	0.569	0.691
	SD	0.239	0.277	0.291
	% Prudent	0.486	0.371	0.600
	% Imprudent	0.114	0.200	0.143

Prudent individuals answer prudently 13/20 questions (for the total sample) or 7/10 questions (for gain/loss sections).

at an α of 0.132.[4] In our definition of prudence, 53.1% of the sample is prudent and 15.0% of the sample is imprudent.

The finding that most individuals are prudent holds both for risk loving and risk averse individuals. There is no statistically significant difference between the average number of questions answered prudently by risk averters compared to risk lovers ($p < 0.907$). We do note that for individuals with both types of risk preference have stronger prudence preferences for gains than losses.

In order to confirm our results from the 'ballot box' section, we compare individual responses in this section to those in the 'coin flip' section. The results are displayed in Table 2.4. We find similar results across both the ballot box and coin flip sections, however the average level of prudence is higher when we use the coin

4. A person is considered prudent if they answer 13 or more questions prudently out of 20. When analyzing prudence separately over gains and losses, individuals must answer at least 7 of 10 questions prudently to be considered prudent, which corresponds to an α of 0.172. For imprudent individuals, the cutoffs are ≤ 7 and ≤ 3 questions respectively.)

flip.

Table 2.4. Prudence Results for the Ballot Box and Coin Flip Sections

Type of Question	Total	Risk Loving	Risk Averse	CRRA: $(-1 < \alpha < 0)$
Losses(Ballot Box)	0.613	0.633	0.569	0.626
Losses(Coin Flip)	0.796	0.795	0.800	0.849
Gains(Ballot Box)	0.655	0.638	0.691	0.626
Gains(Coin Flip)	0.814	0.795	0.857	0.792
Total (Ballot Box)	0.634	0.636	0.630	0.626
% Prudent (Ballot Box)	0.531	0.551	0.486	0.566
% Imprudent (Ballot Box)	0.150	0.167	0.114	0.170

Two questions in the ballot box section were mathematically identical to those in the coin flip section. When we test whether or not the answers given by individuals are consistent across the two sections, we reject the null that the answers in the coin flip section are uncorrelated with their mathematical equivalents in the ballot box section ($p < 0.001$).

To further test the consistency of our results, two questions in the ballot box section were identical in both the gain and the losses section. The only change was that we switched the order in which the lotteries were presented. For losses, 62.9% of individuals answered the identical questions the same; for the gains section this figure was 69.0%. It should not be surprising that these figures are not at 100% or even 90%. Our experiment asks individuals to compare lotteries with the same mean and variance. Thus, the experimenter must expect a some degree of error on the part of the participants. Nevertheless, we strongly reject the null that individuals are answering these questions randomly by conducting a t-test for both losses ($p < 0.003$) and gains ($p < 0.001$).

We also test whether or not having an economics or business background affects prudence preferences. Those who majored in economics were more likely to

answer questions prudently for losses ($p < 0.010$) but were no more likely to answer prudently over gains ($p < 0.312$). This could be explained by underlying differences between the two groups. It could also be the case that non-economics majors had a steeper learning curve to understanding these questions since the prudence questions with losses were presented first.

In summary, we have found that 53.1% of individuals are prudent while 15.0% of individuals are not. The proportion of prudent individuals is similar across risk lovers and risk averters. Responses to the prudence questions are robust to a change in the framing of the question (i.e., ballot box vs. coin flip presentation) as well as the order in which the lotteries are presented.

Other papers have used certainty equivalents in order to measure prudence. For instance,Eisenhauer and Ventura (2003) measure absolute and relative prudence by fitting a constant relative risk aversion function (CRRA) to a single question about the willingness to pay for a 50/50 gamble.[5] If we assumed a utility function of the form $u(x) = \frac{x^{1-\theta}}{1-\theta}$, then $u''' = \frac{\theta(1+\theta)}{x^{\theta+2}}$. If $x > 0$, then an individual is imprudent when $\theta \in (-1, 0)$ and prudent otherwise.

In the RIAV section, we examine individuals for whom the fitted value of θ is between -1 and 0. Under expected utility theory, we would predict that these individuals should act imprudently and should be more likely to choose the negatively skewed lotteries in our prudence sections. The result of this analysis are shown in Table 2.4.

We see that the majority of individuals whom EUT predicts would be imprudent under CRRA are actually prudent in our experiment. Individuals who are risk loving with $\theta \in (-1, 0)$ answer 62.6% of questions prudently. Further, we classify only 17.0% of these individuals as imprudent compared to 56.6% who are classified as prudent. From these results, we conclude that simply fitting a CRRA function to

5. The actual question asked is "You are offered the opportunity of acquiring a security permitting you, with the same probabilities, either to gain 10 million lire or to lose all the capital invested. What is the most you are prepared to pay for this security?"

a certainty equivalents questions does not accurately characterize an individuals prudence preferences. We propose that our non-parametric measure of prudence better characterizes how individuals view skewed lotteries.

2.6.3 Prevention

One question remains: do prudent individuals select lower levels of prevention? In order to test this hypothesis, we first regress the level of prevention chosen in each question on the level of prudence. The results are in Table 2.5. We use the number of questions an individuals answered prudently as a proxy for their level of prudence. Since the prevention section involves making payments to reduce the probability of a loss, we only take into account prudence questions over losses. Using categorical dummy variables for prudent, imprudent, and prudent neutral individuals gives similar qualitative results.

We first examine the constant term. The average prevention level selected was greater than the one that maximized the expected value for a risk neutral person (i.e., $\frac{1}{n}\Sigma_i e^*_{ij} > e_n \ \forall j$). This is likely due to the experimental design. For questions 1-4, e_n=3; for questions 5-6 e_n= 2 or e_n=3; and for questions 7-8, e_n=4. If an survey participant wanted to choose the median prevention level, they would select e^*=4 or e^*5. We do see individual choices congregating around these values. For questions 5-6, however, survey respondents generally choose e^* to be lower than for the other questions.

Now we turn to the coefficient on *Prudence*. Across all 8 prevention questions, we find a similar trend. Prudence decreases prevention levels in all 8 regressions but not in a statistically significant manner. On average, the Prudence coefficient is -0.043. This means that on a scale from \$0 to \$9, individuals who prefer prudent lotteries every time contribute \$0.43 less towards prevention than an individual who always prefers imprudent choices. We cannot reject the null hypothesis that *Prudence*

has no impact on prevention even across all 8 prevention questions ($p < 0.592$). Still, this consistency of the negative coefficient across all 8 specification provides some suggestive evidence that higher levels of prudence lead to less prevention, even if this effect is small in magnitude.

Section 3 of the paper noted that both risk and prudence preferences should influence the preferred level of prevention e^*. Higher levels of risk aversion should decrease e^* when $p(e_n) > \frac{1}{2}$ and increase e^* when $p(e_n) < \frac{1}{2}$. In two of four prevention questions where $p(e_n) > \frac{1}{2}$, higher levels if risk aversion does decrease e^*; in zero of four prevention questions where $p(e_n) < \frac{1}{2}$, higher levels of risk aversion increases e^* . Further, none of these results are statistically different from 0 at the 5% level. When we examine the interaction of the prudence and the dummy variables for risk aversion and risk loving, we find very little predictive power. None of the interaction coefficients were statistically different from 0 at the 5% level.

Overall, we did find some suggestive evidence that higher levels of prudence may decrease prevention, but these results are far from conclusive.

2.7 Conclusion

Prudence has been shown theoretically to be an important determinant of precautionary savings, asset allocation and optimal prevention levels. However, empirical measures of of prudence have been scare and generally identify prudence through assumption with respect to the functional form of the utility function.

This paper takes a non-parametric approach to reveal a more comprehensive understanding of individual's preferences for prudence. We find that on average 53.1% of individuals are prudent and 15% of individuals are imprudent. Risk averse or risk loving behavior does not seem to impact the probability that an individuals will be prudent. Although participants respond more prudently when questions were asked using the coin flip presentation as opposed to ballet box presentation,

participants did answer questions consistently across specifications.

In order to calculate prudence, previous research has fit standard utility functions to participant certainty equivalent responses responses. Our results contradict the validity of this methodology. When we isolate imprudent individuals as predicted by CRRA certainty equivalents, we find only 17.0% are actually imprudent compared to 56.6% who are prudent. We propose that our non-parametric measure of prudence better characterizes the manner in which individuals view skewed lotteries.

We did find some suggestive evidence that prudent individuals choose lower levels of prevention. On a scale from \$0 to \$9, prudent individuals choose to \$0.43 less of prevention than imprudent individuals. Although this result was small in magnitude and not statistically significant, in all 8 questions more prudent individuals choose less prevention.

One limitation to this approach is that we do not measure the strength of an individuals' preferences for prudence. Because we measure preferences over the third derivative of the utility function, precisely estimating the magnitude of prudence is difficult. In the future we aim to estimate a prudence premium over skewed lotteries. Future research should also try to take these findings out of the experimental setting. Utilizing the prudence measures in this paper, researchers should be able to determine whether or not prudent individuals are more likely to get a flu shot or select different investment allocations. This paper gives suggestive evidence that prudent individuals like Aristotle may prefer lower levels of prevention and thus may be less likely to get flu shots.

2.8 Acknowledgements

Thank you to my co-author Daniel Wiesen during our international collaboration on Chapter 2,"Why Aristotle didn't get his flu shot: The impact of prudence on prevention." Also thank you to Louis Eeckhoudt for his insight, time, and sense

of humor.

2.9 Tables

Table 2.5. Prevention and Prudence Regressions

		$p(e_n) > .5$				$p(e_n) < .5$			
	Variable	1	3	5	7	2	4	6	8
1	Prud	-0.054	-0.049	-0.066	-0.044	-0.051	-0.038	-0.037	-0.043
		(0.084)	(0.077)	(0.081)	(0.084)	(0.083)	(0.071)	(0.081)	(0.079)
	Cons.	5.640	6.019	4.575	4.463	4.554	5.030	3.767	4.281
		(0.567)	(0.518)	(0.548)	(0.563)	(0.559)	(0.478)	(0.545)	(0.533)
2	(RA)(Prud)	0.009	-0.020	0.160	0.187	0.070	0.060	0.193	0.145
		(0.154)	(0.141)	(0.146)	(0.151)	(0.152)	(0.130)	(0.146)	(0.144)
	(RL)(Prud)	-0.069	-0.053	-0.149	-0.150	-0.107	-0.089	-0.144	-0.127
		(0.102)	(0.094)	(0.097)	(0.100)	(0.101)	(0.086)	(0.097)	(0.095)
	RA	0.019	0.148	-1.191	-2.172	-1.129	-1.188	-2.210	-1.657
		(1.202)	(1.100)	(1.141)	(1.179)	(1.185)	(1.010)	(1.139)	(1.122)
	Cons.	5.590	5.939	4.908	5.194	4.934	5.448	4.514	4.831
		(0.706)	(0.646)	(0.670)	(0.693)	(0.696)	(0.593)	(0.669)	(0.659)

Table 2.6. Paired lotteries in stage RIAV

Question	Option A	Option B	E(A)-E(B)
1	\$0	(\$0, 50%; -\$40, 50%)	\$20
2	-\$2	(\$0, 50%; -\$40, 50%)	\$18
3	-\$4	(\$0, 50%; -\$40, 50%)	\$16
4	-\$6	(\$0, 50%; -\$40, 50%)	\$14
5	-\$8	(\$0, 50%; -\$40, 50%)	\$12
6	-\$10	(\$0, 50%; -\$40, 50%)	\$10
7	-\$12	(\$0, 50%; -\$40, 50%)	\$8
8	-\$14	(\$0, 50%; -\$40, 50%)	\$6
9	-\$16	(\$0, 50%; -\$40, 50%)	\$4
10	-\$18	(\$0, 50%; -\$40, 50%)	\$2
11	-\$20	(\$0, 50%; -\$40, 50%)	\$0
12	-\$22	(\$0, 50%; -\$40, 50%)	-\$2
13	-\$24	(\$0, 50%; -\$40, 50%)	-\$4
14	-\$26	(\$0, 50%; -\$40, 50%)	-\$6
15	-\$28	(\$0, 50%; -\$40, 50%)	-\$8
16	-\$30	(\$0, 50%; -\$40, 50%)	-\$10
17	-\$32	(\$0, 50%; -\$40, 50%)	-\$12
18	-\$34	(\$0, 50%; -\$40, 50%)	-\$14
19	-\$36	(\$0, 50%; -\$40, 50%)	-\$16
20	-\$38	(\$0, 50%; -\$40, 50%)	-\$18

Table 2.7. Paired lotteries in stage PRUDA

	Option A				Option A				Prudent
Question	P(X_1)	X_1	P(X_2)	X_2	P(X_1)	X_1	P(X_2)	X_2	Option
1	75%	-$10	25%	-$30	25%	$0	75%	-$20	B
2	25%	-$5	75%	-$25	75%	-$15	25%	-$35	A
3	90%	-$19	10%	-$29	10%	-$11	90%	-$21	B
4	20%	$0	80%	-$20	80%	-$12	20%	-$32	A
5	5%	-$1	95%	-$21	95%	-$19	5%	-$39	A
6	80%	-$17	20%	-$42	20%	-$2	80%	-$27	B
7	90%	-$13	10%	-$23	10%	-$5	90%	-$15	B
8	90%	-$21	10%	-$1	10%	-$37	90%	-$17	A
9	75%	-$15	25%	-$35	25%	-$5	75%	-$25	B
10	90%	-$11	10%	-$21	10%	-$3	90%	-$13	B
11	10%	$11	90%	$21	90%	$19	10%	$29	B
12	95%	$19	5%	$39	5%	$1	95%	$21	A
13	10%	$3	90%	$13	90%	$11	10%	$21	B
14	10%	$37	90%	$17	90%	$21	10%	$1	A
15	80%	$12	20%	$32	20%	$0	80%	$20	A
16	25%	$5	75%	$25	75%	$15	25%	$35	B
17	10%	$5	90%	$15	90%	$13	10%	$23	B
18	75%	$15	25%	$35	25%	$5	75%	$25	A
19	20%	$2	80%	$27	80%	$17	20%	$42	B
20	25%	$0	75%	$20	75%	$10	25%	$30	B

Table 2.8. Effort choices and corresponding lotteries in stage PREV

Q1	e	0	1	2	3	4	5	6	7	8	9
	P(Lose \$30)	0.99	0.80	0.70	0.65	0.63	0.62	0.61	0.60	0.60	0.60
Q2	e	0	1	2	3	4	5	6	7	8	9
	P(Lose \$30)	0.49	0.31	0.20	0.15	0.13	0.12	0.11	0.10	0.10	0.10
Q3	e	0	1	2	3	4	5	6	7	8	9
	P(Lose \$40)	0.99	0.80	0.70	0.65	0.63	0.62	0.61	0.60	0.60	0.60
Q4	e	0	1	2	3	4	5	6	7	8	9
	P(Lose \$40)	0.49	0.31	0.20	0.15	0.13	0.12	0.11	0.10	0.10	0.10
Q5	e	0	1	2	3	4	5	6	7	8	9
	P(Lose \$20)	0.99	0.80	0.70	0.65	0.63	0.62	0.61	0.60	0.60	0.60
Q6	e	0	1	2	3	4	5	6	7	8	9
	P(Lose \$20)	0.49	0.31	0.20	0.15	0.13	0.12	0.11	0.10	0.10	0.10
Q7	e	0	1	2	3	4	5	6	7	8	9
	P(Lose \$30)	0.94	0.80	0.73	0.65	0.61	0.58	0.56	0.55	0.54	0.54
Q8	e	0	1	2	3	4	5	6	7	8	9
	P(Lose \$30)	0.44	0.30	0.23	0.15	0.11	0.08	0.06	0.05	0.04	0.04

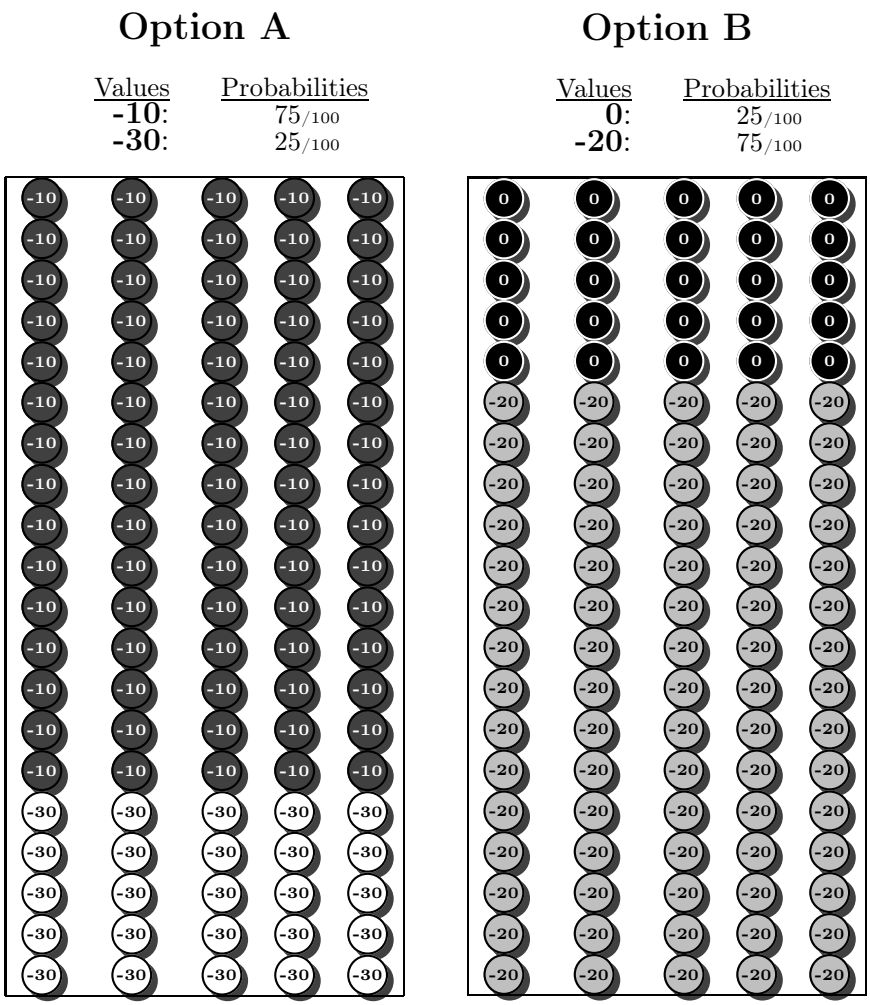

Figure 2.1. Sample Questions in Stage PRUD

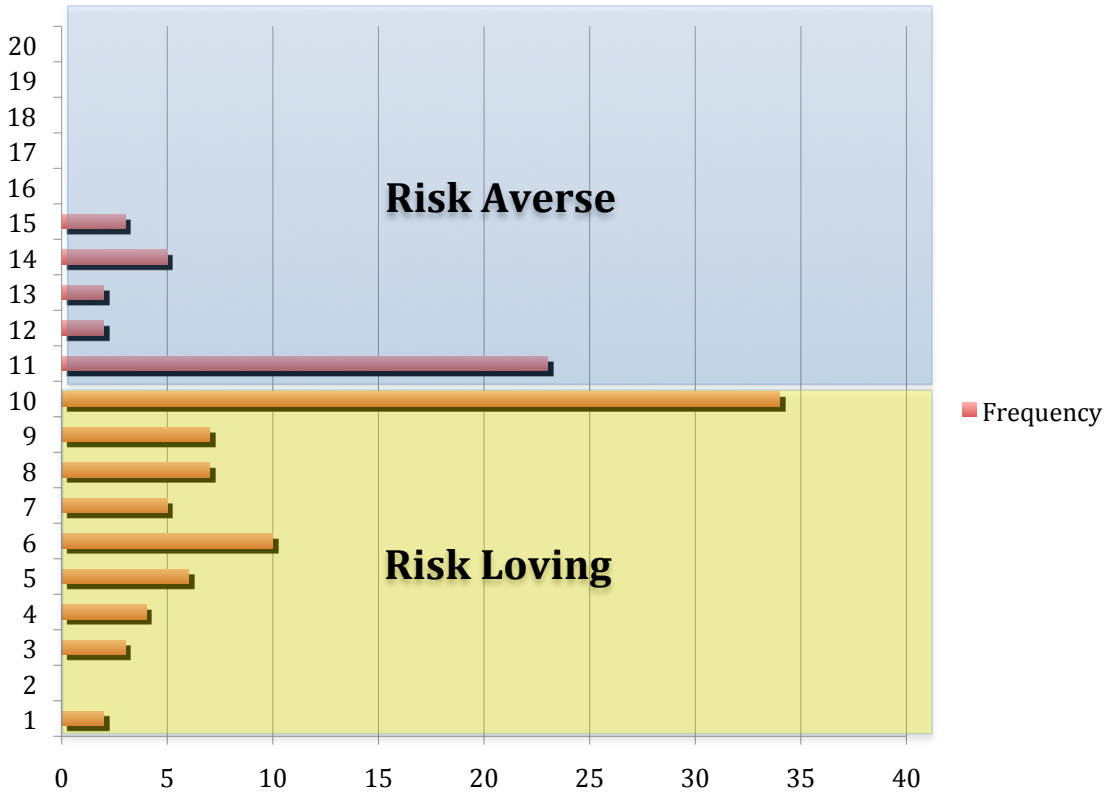

Figure 2.2. Risk Aversion Levels

Chapter 3

"Your Patients' Up-to-date Status and the CDCs Vaccine Distribution System" with John Fontanesi

Objective: To estimate the initial impact of the CDCs new centralized distribution system on vaccine availability.

Study Design: In March 2007, California providers participating in the Vaccines for Children (VFC) program were compelled to begin placing vaccine orders under the Vaccine Management Business Improvement Project (VMBIP) centralized logistical system. Vaccine ordering, usage, inventory, and delivery data was collected from two large Southern California providers.

Methods: The change in the average vaccine delivery delay is measured before and after VBMIP. The data underwent simulation to estimate three other metrics important to the CDCs vaccination goals: 1) the

probability a provider will run out of vaccines before the new shipment arrives, 2) the number of days per year a provider will have zero vaccine inventory, and 3) change in the up-to-date status of 12 month old children.

Results: After VMBIP was implemented, delivery times increased from 1.6 to 12 business days ($p \leq 0.001$). For 7 of the 11 vaccines investigated, there was a strong, statistically significant increase ($p \leq 0.01$) in the number of days a provider would be without vaccine. The simulation predicts that VMBIP will cause a decrease in the up-to-date status for 12-month-old children.

Conclusion: VMBIP was implemented in order to save costs through the centralization of vaccine ordering and delivery. Although there may have been a decrease in costs to the centralized distributor, during VMBIPs initial implementation timeline that savings has decreased the likelihood of immunization in the clinics. The decrease in up-to-date status imposes hidden costs on provider operation, as well as on the patients.

One of the great public health achievements of the 20th century was the near eradication of vaccine preventable disease through public health vaccination programs. Collectively, the programs are estimated to create a net societal savings of up to $43 billion (CDC MMWR 1991, Zhou et al. 2004). In the early 1990s, responsibility for vaccinating children shifted from the public health sector to the medical home. This became possible by the inclusion of vaccinations in both commercial and government-sponsored health insurances. According to DeNavas-Walt, Proctor, and Smith (2008), over 11% of the children (8.9 million) are uninsured. Further an additional 29% of children are insured under SCHIP/Medicaid programs. In order to ensure that these under-served populations were vaccinated, the Federal government established the Vaccines for Children (VFC) program.

Enacted in 1993 under Title XIX of the Social Security Act and placed under the administrative control of the Centers for Disease Control and Prevention (CDC) the VFC program is the largest supplier of childhood vaccines. With an estimated 2009 budget of $2.8 billion, VFC provides approximately 43% of all routinely recommended childhood vaccines (see CDC FY2009 Budget Request Summary, CDC Program Brief 2007). Therefore, any major changes to the VFC logistical system will directly affect vaccine availability at the provider level.

One such change in vaccine delivery logistics was the creation of the Vaccine Management Business Improvement Project (VMBIP). On a pilot basis, the CDC began using electronic order entry and a centralized distribution system in the hopes of decreasing costs and reducing waste and fraud (Bell et al. 2001).

VMBIP was a dramatic shift in vaccine ordering and fulfillment processes compared to the prior system. Originally, VFC providers placed orders through their state/local Department of Health. Vaccines would then be distributed to the provider from either a local vaccine depot, or a contracted distributor. Health care organizations, including integrated managed care organizations, with multiple sites could have vaccine delivered to a central location and distributed internally on an as-

needed basis. Providers who unexpectedly ran out of VFC-provided vaccines could petition the local vaccine depot and receive emergency supplies in very short order.

Under VMBIP, vaccines are shipped directly from a centralized warehouse to individual provider sites. Redistribution of VFC-provided vaccines within a health care organization has been disallowed, and each clinical site is responsible for ordering and receiving vaccine for their site solely. Providers now order vaccine on a tiered order frequency schedule known as Economic Order Quantity (EOQ). A providers reported vaccine usage and the eligible patient population they serve determine how often and in what amounts a provider may order vaccine. According to the CDC Operations Guide (2007), what has not changed is the admonition: Order vaccine in accordance with actual vaccine need; avoid stockpiling or build-up of excess vaccine inventory which generally means ordering with one months lead time.

This paper evaluates the new, centralized distribution systems ability to supply vaccines to providers in a timely manner. Specifically, VMBIPs distribution efficiency is measured by changes in: 1) vaccine delivery times, 2) the probability that a provider will run out of vaccines before a new order arrives, 3) the number of business days per year a provider will run out of any given vaccine, and 4) the up-to-date status of 12-month-old children specifically due to out-of-stock vaccine inventory.

3.1 Methods

3.1.1 Data Source

The data for this study were collected from a convenience sample of the vaccine ordering behavior of two large Southern California providers in San Diego and Orange Counties. Variables collected include quarterly vaccine usage, vaccine ordering dates and quantities, the dates the vaccine shipments were received and the

vaccine inventory levels at the times orders were placed. Between January 2005 and June 2008, we collected data for 11 vaccines . These vaccines were: Diphtheria, tetanus, and pertussis (DTaP); DTap/Hep B/IPV; Haemophilus influenzae (Hib); Hepatitis A; Hepatitis B; Measles, Mumps and Rubella (MMR); Meningococcal Vaccines (MCV); Pneumococcal conjugate vaccine (PCV); Polio (IPV); Tetanus and Diphtheria (Td); Tetanus, Diphtheria and Pertussis (Tdap). Data for Varicella vaccines was collected but not included due to missing data. Although VMBIP completed its nationwide rollout in June 2008, California was one of four pilot states that began rollout in February 2007 (California DHS Internal Memo 2007a). Californias actual go-live date was in March 2007 (California DHS Internal Memo 2007b). Because of the early enactment date, we have 27 months of data prior to the VMBIP start date and 15 months of data after VMBIP was implemented.

3.1.2 Statistical Analysis

Analysis began with determining the average duration between the date when a vaccine order was placed and when it was received under the old and new distribution systems. A standard t-test was used to test for equality of means between the two samples.

To further understand the effect of changing delivery times, a simulation methodology was employed using the additional vaccine data collected. The first two simulated statistics were 1) the probability a provider will run out of vaccines before a new order arrives and 2) the number of business days per year a provider will have zero vaccine inventory on hand. A provider will run out of vaccines if the number of doses used between the vaccine order and delivery dates is higher than the vaccine inventory levels at the time an order is placed. The inventory levels at the time of each order are known from the data collected; however, vaccine usage data was only available on a quarterly basis.

In order to estimate daily vaccine usage, we developed a simulation in Matlab. The simulation assumed daily vaccine usage was distributed according to a negative binomial function. The only exception is that for the Dtap/Hep B/ IPV, we assumed a normal distribution since it fit the data better. For both distributions, the parameters of the distributions were set so that the simulated datas expected mean quarterly usage matched those of the actual data collected. The negative binomial distribution incorporates the fact that daily vaccine usage is a count variable that suffers from overdispersion. The simulated number of vaccines used per day was drawn from this distribution.

Using historical ordering schedules and actual inventory levels at the time an order was placed, we estimated the two simulation statistics. Vaccine usage was randomly drawn from the simulated negative binomial distribution described above. We bootstrapped actual delivery delay data separately before and after VMBIP to create simulated delivery delay data. In order to calculate the probability a provider will run out of vaccines and the number of business days per year a provider will have zero vaccine inventory, we simulated 40 months of vaccine ordering and delivery data 10,000 times.

After the number of days a provider will be without a vaccine is estimated, one can also estimate how VMBIP will affect a facilitys cohort of childrens up-to-date vaccination status. If the provider is out of stock of a given vaccine, then it is assumed that the child is not vaccinated. If the child was not vaccinated, we assume that they may be vaccinated at their next well-child visit. In the simulation, children visit the doctor for well-child visits according to the CDCs standard well-child visit schedule (CDC Immunization Schedule 2008). In the first year, the CDC recommends well-child visits at 1, 2, 4, 6 and 12 months. Because physician visits are so frequent during the first year, it is unlikely that a parent will bring their child to the physician for an additional well-child visit if they miss an immunization. By standardizing patient-provider visit rates, one can isolate what portion of the

decreased up-to-date status is specifically due to changes in the probability that a provider has zero vaccine inventory.

3.2 Results

3.2.1 Average Delay Time

Figure 3.1 plots the vaccine delivery duration over time. The vertical black line demarcates the date VMBIP was implemented in California. Delivery times, under VMBIP are significantly longer than was previously the case. In 2008, this effect seems to have attenuated somewhat, but delay levels at the end of the sample are still far above those at the beginning of the sample.

Table 3.1 displays summary statistics for vaccine delivery times. After VMBIP was implemented in California, the average vaccine delivery time increased from 1.6 to 12 business days. The probability that a vaccine delivery arrived more than one week after an order was placed jumped from 7% to 89%. A t-test for equality of means is strongly rejected ($p \leq 0.001$). The evidence shows that during the initial phases of implementation, VMBIP worsened vaccine delivery times.

3.2.2 Simulation Results

Clinic data clearly demonstrates that delivery times increased significantly under VMBIP. Using delivery delay times from before and after the VMBIP implementation, Figure 3.2 displays the simulated probability that a provider will run out of MMR and Hib vaccines before a new shipment arrives. For both MMR and Hibas well as for the majority of the vaccines under considerationthe longer vaccine delivery times resulted in a significantly increased risk of running out of vaccines before an order arrives. For instance, if a clinic had 100 doses of MMR vaccine and 100 doses of Hib vaccine in inventory when placing a new order, under the old system there was

less than a 1% chance that the clinic would run out of MMR and a 7% chance they would run out of Hib. Under VMBIP, these probabilities increase to 41% and 76%, respectively. Figure 3.2 shows that VMBIPs longer delays significantly increased the probability of running out of vaccines before a new shipment arrives.

A more important issue, however, is analyzing how vaccine delivery times will affect vaccine availability under typical inventory levels . Using actual ordering dates, actual inventory levels when placing an order, and simulated daily vaccine usage data, a second simulation in Table 3.2 compares the number of days per year a provider will be left without vaccines before and after the VMBIP implementation. In the case of Hep A, Hib, PCV, and Td vaccines under the VMBIP simulation, vaccine stock-outs in the clinic will occur during more than two months of the year. Vaccine availability is much higher under the old system and these results are statistically significant at the 1% level. For Dtap/Hep B/IPV, IPV, MMR, the decrease in vaccine availability under VMBIP is also statistically significant at conventional levels. On the other hand, for Dtap, Hep B, MCV, and Tdap, VMBIPs longer delays seem to have little impact on vaccine availability. This is due to the fact that in our sample, these vaccines were used less frequently and thus stock-outs are less sensitive to delivery delays.

The effect of longer delivery times on up-to-date immunization status of children on their first birthdays is shown in Table 3.3. The simulation results show dramatic decreases in the number of children who are up-to-date on Dtap/Hep B/IPV, Hib, Hep A, and PCV vaccines, and a fairly large decrease in the number of children who are up-to-date on MMR. On the other hand, for the Dtap, Hep B, IPV, there is a minimal decrease in the number of children who are up-to-date. A larger effect on up-to-date status was found for vaccines with higher daily usage rates and for vaccines with more doses required in each series.

3.3 Conclusion

There are a number of implications issuing from this study. The first and most obvious is that VMBIP increased delays in vaccine delivery resulting in an increased number of days where vaccines were unavailable for patients. The effect is particularly pronounced when vaccine usage is high or when inventory levels are low at the time an order is placed.

The second issue concerns VMBIPs original goal of reducing waste. Not only was there a dramatic increase in the mean time to vaccine delivery, there was a parallel increase in the variability for vaccine delivery (see Table 3.1). Receiving a vaccine shipment too early is almost as bad as too late as providers did not have space available from proper storage and some of these vaccines had to be discardedclearly running counter to the goal of reducing wastage.

The third issue has to do with this Countrys present capacity to meet unexpected demand under either pandemic or bio-terrorist scenarios.

The forth issue has to do with the impact of vaccine shortages and missed opportunities to vaccinate on pay-for-performance and other contractual relationships between payors and providers. A decline in California for reimbursable immunization has direct financial consequences to both health plans and providers operating under the various pay-for-performance schemas. Hopefully, one lesson learned from the VMBIP pilot is that a change on one side of a systems processes deserves to have its outcomes actually measured. There is much that can be done to achieve the original goals and objectives of the program. Distribution centers for biological products typically have an order fill time of 24 hours with an additional dock to stock time of 8-24 hours with no more than 0.5% pick/pack error rate (Frazelle and Hackman 1994, Hackman et al. 2001).

Though it is quite possible to spend several hundreds of thousands of dollars on consultants, a novel approach presently used by many industry leaders is a self-

assessment tool available online from the Georgia Institute of Technology, School of Industrial & Systems Engineering (McGinnis et al. 2006). This online product uses Data Envelopment Analysis to determine how a user compares to their industry segment leaders in a way that highlights potential areas of improvement.

3.3.1 Limitations of the Study

Data was collected from a convenience sample of Southern California providers and thus may not be nationally representative. Further, the data for the post-VMBIP period were collected during the start-up phase of VMBIP and prior to the initiation of a new ordering system. It is possible that VMBIPs long delays are caused by a poor initial implementation rather than long-term issues with the centralized distribution system. Regardless of start-up problems, the goals of providing safer vaccine more quickly to poor children at less cost to the tax-payer is one that continues to need to be pursued. It is hoped this study will be used to achieve those goals.

3.4 Take-Away Points

In March 2007, the CDCs Vaccines for Children (VFC) program begin a pilot program placing vaccine orders under the Vaccine Management Business Improvement Project (VMBIP) centralized logistical system. This paper evaluates the effect of VMBIP on provider access to vaccines. The results of this paper reveal:

- Vaccine delivery times increased from 1.6 days at baseline to 12 business days after VBMIP was implemented.
- A significant increase in the number of days per year providers will be left with zero vaccine doses in inventory.
- A potential worsening of up-to-date status of children if the delays caused by VMBIP continue.

3.5 Acknowledgements

Thank you to co-author John Fontanesi for his help on Chapter 3, 'Your Patients' Up-to-date Status and the CDCs Vaccine Distribution System." We would like to thank Nancy Fasano of the Centers for Disease Control and Prevention and Dr. Howard Backer of the California Department of Public Health for their support and review of this study.

3.6 Tables and Figures

Table 3.1. Vaccine Delivery Duration (in weekdays)

	After VMBIP	Before VMBIP	Δ	p-value
Average Delivery time	12.3	1.6	10.7	0.0001
Median Delivery time	13	1	12	0.0001
Standard Deviation of delivery time	6.3	2.4	3.9	0.0002
P(Delivery time > 1 week)	89%	7%	82%	0.0001

Table 3.2. Simulated Days per year with zero vaccines

		Days/year with 0 vaccines			
Vaccine	Vaccines/Month	After VMBIP	Baseline	Δ	p-value
Dtap	182.1	0.7	0.0	0.7	0.458
Dtap/Hep B/IPV	198.6	31.3	0.0	31.3	0.001
Hep A	162.6	95.7	2.3	93.4	0.001
Hep B	141.8	3.1	0.0	3.1	0.074
Hib	318.3	122.1	4.8	117.3	0.001
IPV	134.3	6.1	0.0	6.1	0.004
MCV	6.0	0.0	0.0	0.0	0.999
MMR	156.6	18.5	0.0	18.5	0.001
PCV	327.7	137.8	6.8	130.9	0.001
Td	54.7	62.7	1.6	61.1	0.001
Tdap	5.4	0.0	0.0	0.0	0.999

Table 3.3. One-year olds not UTD due to vaccine stockout (millions)

Vaccine	After VMBIP	Before VMBIP	Δ
Dtap/Hep B/IPV	12.9	0	12.9
Dtap, Hep B, IPV (separate vaccines)	0.8	0	0.8
Hep A	11.2	0.3	11
Hib	34.4	2.2	32.2
MMR	2.2	0	2.2
PCV	36.4	3.1	33.3

Assumes a birth cohort of 42.85 million (U.S. Census Bureau, 2004)

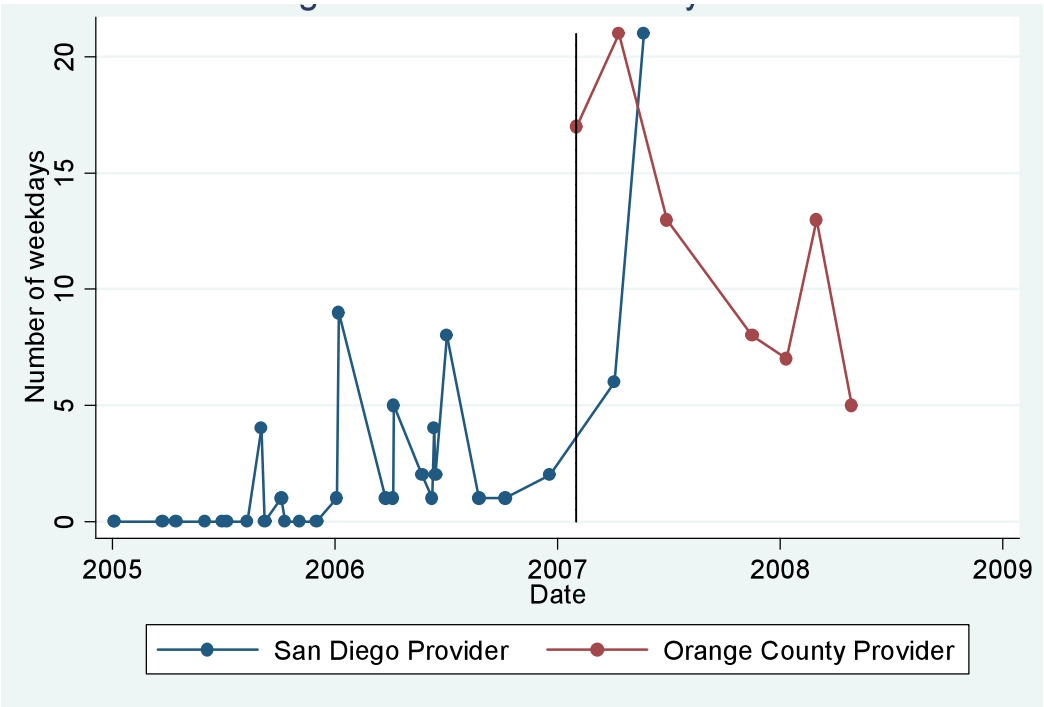

Figure 3.1. Vaccine Delivery Duration

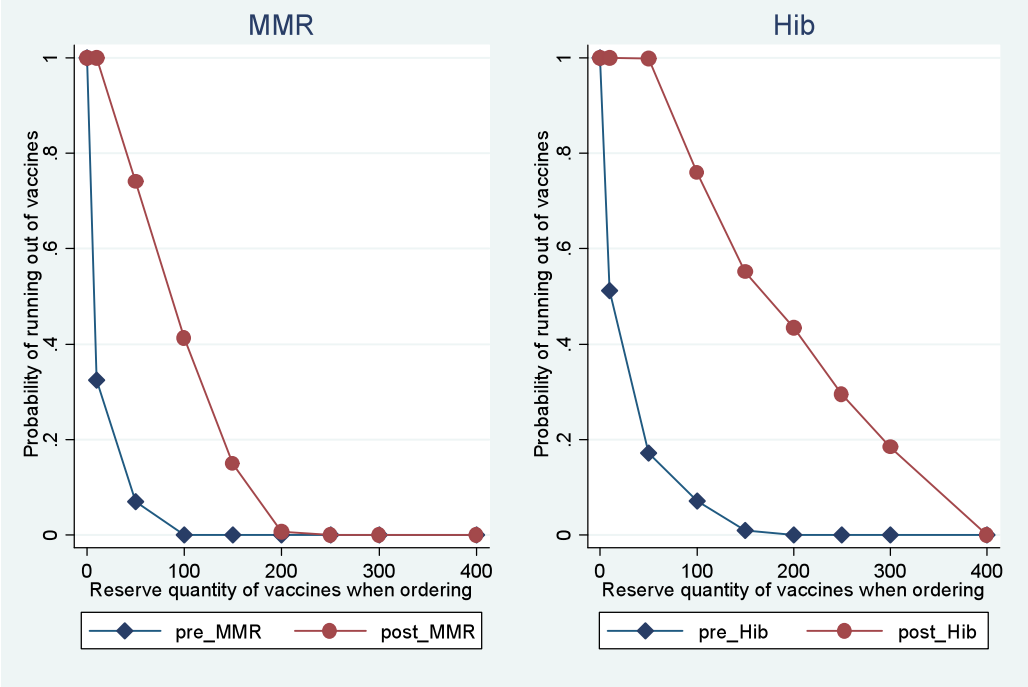

Figure 3.2. Vaccine stockout probability before and after VMBIP

Chapter 4

"Why Does Getting Married Make You Fat? Incentives and Appearance Maintenance" with Uri Gneezy

Married individuals weigh more on average than non-married individuals. We suggest that exiting the dating market decreases ones incentive to maintain their appearance and leads to an increase in body weight. We hypothesize that it is most difficult for individuals to exit a traditional marriage, and easiest for individuals to exit if the couple is cohabitating but not legally married. Using a 14-year panel data set, we test whether or not the ease of exiting a domestic relationship affects weight gain. For men, we find that the type of domestic relationship has little impact on weight gain. For women, however, marriage leads to a 2.4 kg weight gain compared to cohabitating.

4.1 Introduction

Married individuals are more likely to be overweight than non-married individuals (Sobal, Rauschenbach and Frongillo, 1992; Hahn, 1993). One reason for this increase in body weight after entering into marriage: decreased incentives to maintain your appearance. Individuals engage in costly behavior to maintain their appearance such as purchasing make-up, name-brand clothes or expensive cars. In modern western societies, an important component of maintaining ones appearance includes having a healthy body weight. In order to achieve an ideal body weight, many people engage in costly time-consuming activities such as exercising and preparing healthy meals at home.

The maintaining appearances hypothesis claims that upon entering into any monogamist domestic relationship, ones incentive to maintain an attractive body weight decreases. Entering into a domestic relationship decreases the probability of being left without a mate in subsequent periods. Barriers to exit in domestic relationships include the legal costs of involved with divorce, psychic costs of separation, the cost to divide assets between two individuals, and fixed costs associated with finding and sharing a residence. These factors all contribute to the persistence of domestic relationships over time. If the probability of re-entering the dating market drops after establishing a shared-residence domestic relationship, an individuals should experience a decreased incentive to maintain their appearance and body weight. In summary, the maintaining appearances hypothesis holds that the decrease in the probability of being left without a mate after one enters a domestic relationship will lead to an increase in body weight.

In this paper we test the maintaining appearances hypothesis empirically using a 14 year panel data set from the Netherlands. We hypothesize that it is most difficult for individuals to exit a traditional marriage or a marriage with a prenuptial agreement, and easiest for individuals to exit if the couple is cohabitating

but not legally married. We exploit variation in relationship status over time in order to determine whether individuals entering domestic relationships with a higher termination probability gain less weight than those entering relationships with a lower probability of termination.

We find that cohabitators are much more likely to terminate their relationship than couples in either a traditional marriage or those who are married with a prenuptial agreement. We also find evidence that when individuals begin to cohabitate, they will gain less weight than if they had gotten married. For the average Dutch male men, getting married increases weight by 0.6 kg. There is no statistically significant difference in weight gain when a man enters into a cohabitating relationship compared to marriage. The probability that a man will separate from his significant other has no effect on weight gain. For Dutch women, marriage causes a significant increase in weight. Women entering into traditional marriages gain 3.0 kg while those entering into marriages with a prenuptial agreement gain 2.9 kg. For women who enter in cohabitation relationships outside of marriage, weight gain is only gain 0.6 kg. This difference in weight gain between cohabitators and women in traditional marriages ($p \leq .001$) as well as the difference between cohabitators and women in marriages with prenuptial agreements ($p \leq 0.001$) is strongly statistically significant. For women, we do find evidence that getting married not only decreases separation probabilities, but also increases weight gain.

4.2 Background

Married individuals are more likely to be overweight than their non-married counterparts (Sobal, Rauschenbach and Frongillo, 1992; Hahn, 1993). This result persists after controlling for age and other covariates. Interpreting this empirical finding is not trivial. Cross-sectional analyses do not take into account the fact that marriage and body mass interact through two different mechanisms: marital

selection and marital causation (Sobal, Rauschenbach and Frongillo, 1992). Marital selection refers to the phenomenon that overweight or obese individuals are less likely to date or marry than their healthy-weighted peers (Averett and Korenman, 1999; Gortmaker et al., 1993; Cawley, Joyner and Sobal, 2006). Marital causation claims that something about being married directly affects an individuals weight.

It has been reported in the medical literature that individuals who enter marriage gain weight and those who exit marriage lose weight. These findings hold despite the fact that married individuals engage in healthier behaviors and have lower mortality rates than their non-married peers (Umberson, 1992).

What specific marriage-related factors lead to weight gain? Craig and Truswell (1988) claim that marriage may lead to higher food intake which will cause more weight gain. Marriage may alter activity levels as well. Verhoef, Love and Rose (1992) and Myers, Weigel, Holliday (1989) observe that married individuals are less likely to be active or exercise, but this finding is not universal (King et al. 1998). Marriage is correlated with a decline in smoking rates and smoking cessation can induce weight gain (Waldron and Lye, 1989; Wee et al., 2001). The closest explanation to the maintaining appearances hypothesis that we test can be found in Sobal (1984). Sobal proposes the possibility that marriage may reduce the incentive to maintain an attractive appearance since mate attraction is not a priority.

Empirically testing this litany of theories is not trivial. When individuals wed, a variety of changes happen all at once. To solve this problem, we utilize variation in the "types of marriages or domestic relationships. To be specific, this paper examines three different domestic relationships. Individuals can either be married under the traditional common property law, married with a prenuptial agreement, or they can live together but not be officially married (i.e., cohabitation). We hypothesize that it is most difficult for individuals to exit a traditional marriage, but easiest for individuals to exit if the couple is cohabitating but not legally married.

In order for our hypothesis to be valid, the probability of actually terminating

the relationship must be correlated with our predictions. We hypothesize that individuals who enter a traditional marriage will gain the most weight after entering into a domestic relationship because there is a lower probability that their marriage will dissolve. The likelihood that those in a traditional marriage will re-enter the dating market is low and thus individuals in this type of domestic relationship have less motivation to maintain their appearance or keep a healthy body weight. Those who cohabitate but are not married will gain the least amount of because the probability that the relationship will dissolve is the highest. Thus, a higher probability that the relationship will terminate will compel individuals to maintain their appearance and body weight.

4.3 Data

We use the DNB Household Survey (DHS) from the CentERdata of the Netherlands. The data set is a panel with data collected between 1995 and 2008. The dependent variable of interest is an individuals body mass index (BMI). Marriage status in the DHS is defined into one of six categories: 1) married under traditional, common property law, 2) married with a prenuptial agreement, 3) cohabiting, 4) divorced, 5) widowed, or 6) never married. The DNB Household Survey contains 70,879 observations. Constraining the sample to individuals aged between 18 and 54 years old leaves 38,151 observations. Deleting observations where height, weight, or marital status data is missing reduces the sample size to 21,184. We omit extreme value of height, weight, BMI or income reduced the sample by 132 observations. These omissions do not materially alter the results. We also drop 3231 observations in years when women were pregnant. Thus, the sample that will be used in the empirical portion of this paper has 17,821 observations for 7,027 unique individuals.

Table 4.1 shows the sample means for the variables used in subsequent regressions. In order to test our hypothesis we need variation in Dutch communal living

arrangements. Traditional marriages make up 55.2% of the sample, marriages with prenuptial agreements make up 8.8%, while 10.5% of Dutch individuals cohabitate without being married. Legally if not culturally, cohabitation in the Netherlands is more of a substitute for marriage than would be the case in the United States. Cohabitation acquired virtually equal status with marriage under Dutch law in the 1980s. Unlike in the United States, Dutch cohabitators are able to file joint tax returns (Hantrais 2004).

Figure 4.1 displays the average BMI levels across marital status. Individuals in traditional marriages have a higher BMI than individuals in marriages with a prenuptial agreement, but individuals in either type of marriage have a higher BMI than cohabitators. Single individuals have the lowest levels of BMI. Figure 4.1 also examines BMI by age. A clear trend emerges that BMI increases with age.

4.4 Methods: Variables affecting the probability a relationship dissolves

We propose that individuals who enter into traditional marriages will gain more weight than cohabitators because the probability that married individuals will re-enter the dating market is lower than is the case for cohabitating individuals. In order to test this hypothesis, the type of domestic relationship must influence the probability a couple will dissolve their relationship. Figure refObesityF2 shows the average separation probabilities in the data set. The separation rate each year is 1.2%, but the separation rate after five years is 4.0%. Using a probit regression controlling for age, income and education, we investigate if the probability an individual will be single or divorced subsequent years is related to their current type of domestic relationship.

$$P(Separated_{i,t+1} = 1|\mathbf{W}_{i,t}) = \beta_0 + \beta_1 Trad_{it} + \beta_2 Prenup_{it} + \beta_3 Cohab_{it} \quad (4.1)$$
$$+ \beta_4 homeowner_{it} + \beta_5 Kids_{it} + \beta_6(\mathbf{X}_{it}) + \epsilon_{it}$$

The dependent variable, *Separated* equals unity if the individual has become single or divorces one year in the future. The variable $\mathbf{W}_{it}$ is a vector of the independent variables, where i and t are indices for the individual and year respectively. $Trad_{it}$, $Prenup_{it}$ and $Cohab_{it}$ are variables dummy variables representing (Trad)itional Marriage, Marriage with a (Prenup)tial Agreement, and (Cohab)itiation respectively. $\mathbf{X}_{it}$ is a vector of the other explanatory variables including age, income and education.

Table 4.2 shows the results of these tests separately by gender. Men and women who are married are much less likely to be single or divorced after one year than all other individuals ($p \leq 0.001$). This is true regardless of whether the individual was involved in a "traditional marriage or a marriage with a prenuptial agreement. Individuals cohabitating are less likely to be single the next year than current singles ($p \leq 0.001$). On the other hand, cohabitators are more likely terminate their relationship than either type of married individuals ($p \leq 0.001$). The probability of separation for married individuals with a prenuptial agreement is not statistically different from people with a traditional marriage for both men and women. We also note that having a child decreases separation probabilities

The last two columns of Table 4.2 repeat the exercise above, but instead use two year lead in the dependent variable. The results are very similar to those with a one year lead. Based on this analysis, we predict that cohabitators will gain less weight after moving in together than individuals who get married with or without a prenuptial agreement.

Homeownership may also serve as predictor of relationship stability. Conditional on income, couples who decide to buy a home may expect to have longer-term relationships than others. If the act of buying a home is orthogonal to weight gain, the purchase of a home could be used as a proxy for relationship strength. We also include a homeownership variable in the probit regressions above. Table 4.2 shows that in fact homeownership is not related to the probability a couple separates.

We have now derived two empirically testable hypotheses:

1. Individuals who enter into a cohabitation arrangement should gain less weight than individuals who enter into a traditional marriage.

2. The weight gain differences between cohabitators and those in traditional marriage should be of larger magnitude for women than men.

The subsequent sections will test these hypotheses.

4.5 Results: Does the type of domestic relationship affect body weight?

4.5.1 OLS and Fixed Effects

In an ideal scientific experiment, the researcher would randomly assign individuals to different domestic relationship statuses. In the case of random assignment, the researcher has eliminated any sorting biases. If this were the case, we could run the following OLS regression.

$$BMI_{it} = \beta_0 + \beta_1 Trad_{it} + \beta_2 Prenup_{it} + \beta_3 Cohab_{it} + \beta_4 \mathbf{X}_{it} + \epsilon_{it} \qquad (4.2)$$

The results from this regression are displayed in Table 4.3 separately for men

and women. Other covariates in X include age, age-squared, age-cubed, education, income, smoking status, the whether or not the individual has a child, as well as year and region dummy variables. For men we observe a monotonic increase in weight from single individuals (the omitted variable) to cohabitators to married individuals with a prenuptial agreement to individuals in a traditional marriage (i.e., $0 < \beta_{Cohab} < \beta_{Prenup} < \beta_{Trad}$). For females, cohabitators are actually thinner than single women. As predicted, both types of married women weigh more than cohabitators. For both men and women, our hypothesis that married women will weigh more than cohabitators holds.

However, in order for OLS to be valid, an individuals choice of domestic relationship must be unrelated to their underlying BMI. OLS will overestimate weight gain from marriage if heavier people are choosing to be married, and will underestimate weight gain from marriage if thinner people are selecting into marriage. Thus, we must take advantage of the panel nature of the data in our regression specification. The Hausman test rejects (p¡0.001) the null hypothesis that the random effects specification produces consistent coefficient estimates.

Thus, in order to mitigate any sorting bias across the different types of domestic relationships, we utilize a fixed effects regression. The fixed effects specification isolates how entering a domestic relationship changes body weight. Identification comes from within-person changes in BMI over time.

$$\Delta BMI_{it} = \beta_0 + \beta_1 \Delta Trad_{it} + \beta_2 \Delta Prenup_{it} + \beta_3 \Delta Cohab_{it} + \beta_4 \Delta \mathbf{X}_{it} + \Delta \epsilon_{it} \quad (4.3)$$

In the equation above, $\Delta z_{it} = z_{it} t_i^{-1} \Sigma z_{it}$. The fixed effects regression is valid if the change in marital status is uncorrelated with unobserved changes that also affect weight. This strong condition is unlikely to hold. Since our major research question is to test how weight changes across domestic relationships, however, we only need for

unobserved changes affecting weight to be similar for cohabitators and both groups of married individuals. In other words, we assume that sharing a residence influences the evolution of BMI levels similarly across all three types of domestic relationships. We attribute any additional differences in weight gain across the three groups to differences in the probability a relationship will dissolve.

The assumption that shared residence has a homogeneous affect on BMI level could be violated if weight gain trends differ across relationship types for reasons unrelated to the probability a relationship is terminated. For instance, a secular increase in weight over time could account for some of these differences. To control for this possibility, we include year dummy variables. Thus, the coefficients of interest represent changes in BMI above or below a secular time trend.

Columns 1 of Table 4.4 shows the results of our preferred fixed effects regression separately for males. Men increase their BMI when they enter into any domestic relationship but this increase is not statistically significant. Further there are negligible differences in the change in a males weight gain across traditional marriage, marriage with a prenuptial agreement and cohabitation . A males weight gain is 0.186 BMI in a traditional marriage and 0.267 BMI in marriages with a prenuptial agreement. These differences are not statistically different from one another ($p \leq 0.422$). For the average Dutch male, this corresponds to weight increases of 0.61 kg and 0.87 kg respectively. Dutch male cohabitators gain 0.195 BMI (0.63 kg), but this gain is not statistically different from the weight gain observed in traditional marriages ($p \leq 0.939$) or marriages with prenuptial agreements ($p \leq 0.597$).

Column 2 of Table 4.4 shows the results for women. Women gain more weight after marriage than men. Females in traditional marriage gain 1.069 BMI (3.0kg) and those in marriage with a prenuptial agreement gain 1.028 BMI (2.9 kg). Unlike for men, we observe significant differences between how entering into a domestic relationship affects weight. Women who begin to cohabitate gain only 0.203 BMI (0.57 kg); this is a significantly smaller weight gain than women in a traditional marriage

($p \leq 0.001$) or married women with prenuptial agreement ($p \leq 0.002$). Cohabitators have a significantly higher probability that their relationship will dissolve than either of the married groups. Thus, cohabitators have a strong incentive to maintain their weight because the probability they re-enter the dating market is high.

It is possible that married women are more likely to have children and thus the increased weight gain upon getting married may be caused entirely by pregnancy. To control for this, years in which a woman was pregnant were dropped from the sample. Nevertheless, giving birth could lead to a permanent increase in weight even months or years after giving birth. To control for this, we include a dummy variable indicating whether the mother had a child. The length of a relationship may also affect weight gain. To control for this, we include a variable measuring the length of a couples relationship (not shown). The length of the couples relationship has no effect on a males weight. For women, longer relationships tend to decrease weight (-0.42 BMI) compared to the secular trend. This finding indicates that the majority of weight gain occurs during the first few years after marriage.

While the maintaining appearances hypothesis could explain a sudden weight gain for women after marriage, women dieting before their wedding day could also explain this phenomenon. Pre-wedding dieting could lead to overestimates in the impact of marriage on weight gain. The post-wedding weight gain for women may simply be a regression to the mean and not a true effect from marriage market incentives. Employing 14 years of data and a fixed-effects regression framework should help to attenuate this problem. To further address this problem, we run our preferred fixed effects regression with lagged marital status. Columns 3 and 4 of Table 4.4 display the results of this regression.

We see that although men gain some weight after entering into any domestic relationship, there are no statistically significant differences in weight change across the three types of domestic relationships. For women, we still see significant weight gain for both types of marriage, but we actually observe a decrease in weight for

cohabitating individuals. Cohabitating women gain significantly less weight than married women ($p \leq 0.014$) or married women with a prenuptial agreement ($p \leq 0.002$). Because the results using lagged marital status are similar to our preferred regression, it does not seem to be the case that dieting before marriage is driving these results. Table 4.5 compares our results to those of other papers. Jeffrey and Rick (2002) found that males who enter into marriage gain 0.70 BMI compared to their baseline, whereas our results show that weight gain was only 0.19 BMI. The Jeffrey and Rick paper, however, does not control for income, education, or smoking status as our specification does. On the other hand, for women, Jeffrey and Ricks estimate of female weight gain from marriage is very similar to ours.

Table 4.5 also compares our results to those of Averett, Sikora and Argys (2008), who use ln(BMI) as the dependent variable. In order to make our results comparable, we conduct our preferred fixed effects regression but substitute ln(BMI) for BMI as the dependent variable. Compared to the Averett paper, our estimates of weight gain caused by marriage are of a smaller magnitude for men and of larger magnitude for women. The magnitude of the weight gain due to cohabitation is nearly identical between the Averett paper and the results presented here.

4.5.2 Two Stage Specification

Let us assume that type of domestic relationship an individual enters into is correlated with the probability that a relationship dissolves, but is not correlated with factors influencing weight gain outside of a shared living arrangement. If this is the case, we can use the fitted values of a probit regressions as a proxy for the strength of the marriage.

$$P(Sep_{it+1}) = \gamma_0 + \gamma_1 Trad_{it} + \gamma_2 Prenup_{it} + \gamma_3 Cohab_{it} + \gamma_4 \mathbf{X}_{it} + u_{it} \qquad (4.4)$$

$$\Delta BMI_{it} = \beta_0 + \beta_1 \Delta Dom_{it} + \beta_2 (\Delta Dom_{it})(\Delta \widehat{Sep}_{it}) + \beta_3 \Delta \mathbf{X}_{it} + \Delta \epsilon_{it} \quad (4.5)$$

In equation 4.4, the variable Sep is equal to unity if the individual separated from their significant other one year into the future and Dom is a variable equal to unity if the individual is any type of domestic relationship. The variable $\widehat{Sep}$ represents the fitted values of the probit regression in equation 4.4. Thus, $\widehat{Sep}$ variable is a proxy for the probability of being single or divorced one year in the future. In the data, we do not observe when singled, divorced, or widowed individuals separate from the person they are dating. Because the Sep variable equals zero for all individuals outside of domestic relationships, we interact the $\widehat{Sep}$ variable with Dom in the second stage. Also included in the vector $\mathbf{X}_{it}$ are dummy variables indicating whether a person is a divorcée or widow.

The exclusion restriction here is that the type of domestic relationship and homeownership status need to be correlated with the probability of the breakup, but not correlated with the error term from . This means that different types of domestic relationships effect weight only through relationship termination probabilities or through factors shared across all types of domestic relationships (e.g., shared housing, more domesticated lifestyle). Table 4.6 shows the results from this regression.

As expected, we see that entering into a domestic relationship leads to an increase in BMI. Estimates of the direct impact of being in a domestic relationship on weight gain are similar to those found in our preferred fixed effects specification shown in Table 4.4.

The probability an individual will separate from their mate has a significant impact on weight gain. A 5% increase in the predicted separation probability leads to a 0.139 decrease in BMI for men and a 0.597 decrease in BMI for women. This corresponds to a 0.6% decrease in body mass for men and a 2.4% decrease in body mass for women.

4.6 Conclusion

It is documented that married individuals weigh more than their non-married peers. In this paper, we test whether the prospect of separating from your significant other and reentering the dating market can effect body mass. For men, we find little evidence that separation probabilities affect weight gain. For women, however, separation probabilities have a large effect on weight gain. Moving from a cohabitation relationship to a married relationship leads to an increase in weight of 2.4 kg. Further, our two-stage regression also finds that predicted separation probabilities have a significant impact on weight gain.

Understanding the mechanisms through which weight gain occurs is important for policy-makers. However, one should not interpret this paper as opprobrium against marriage. First, weight gain caused specifically by the maintaining appearances hypothesis is modest for women (2.4 kg) and not statistically significant for men. In our data, marriage is also correlated with decreased rates of smoking. Marriage may have difficult to measure psychic benefits as well. Nevertheless, this paper has shown that withdrawal from the dating market does cause weight gain. Extending this paper to find practical policies that could maintain marriage stability and decrease obesity rates would be fruitful.

4.7 Acknowledgements

I would like to thank my co-author Uri Gneezy for his assistance on Chapter 4, "Why Does Getting Married Make You Fat? Incentives and Appearance Maintenance."

4.8 Tables and Figures

Table 4.1. Table of Means

	Variable	Mean	Std. Dev.	Min	Max
	BMI	24.78	3.98	10.4	59.5
Marital Status	Single	0.213		0	1
	Tradiational Marriage	0.552		0	1
	Married, Prenutpial Agreeme	0.088		0	1
	Cohabitating	0.105		0	1
	Divorced	0.038		0	1
	Widowed	0.005		0	1
	Kids	1.24	1.22	0	7
	Homeowner	0.271	0.444	0	1
	Smoker	0.338	0.473	0	1
	Gender	0.602	0.490	0	1
Education	Primary Education or lower	0.090	0.286	0	1
	Pre-vocational	0.144	0.352	0	1
	Pre-university	0.110	0.313	0	1
	Vocational	0.486	0.500	0	1
	University	0.157	0.364	0	1
	Other	0.013	0.112	0	1
Region	Three largest cities	0.163	0.369	0	1
	Other West	0.288	0.453	0	1
	North	0.110	0.313	0	1
	East	0.203	0.402	0	1
	South	0.236	0.425	0	1
	Income	57.81	50.98	0	492.8
	Year	####	4.1	1995	2008
	n	*17,821*			

Table 4.2. Probit Regression: Probability of Being without a Mate

	P(No mate in 1 year)		P(No mate in 2 years)	
	Male	Female	Male	Female
Married	-3.503	-3.397	-3.046	-3.069
	(0.098)***	(0.106)***	(0.099)***	(0.116)***
Prenup	-3.624	-3.829	-3.225	-2.991
	(0.212)***	(0.319)***	(0.221)***	(0.205)***
Cohab	-2.769	-2.522	-2.379	-2.103
	(0.091)***	(0.096)***	(0.099)***	(0.105)***
Div	-0.314	0.163	-0.179	0.223
	(0.115)***	(0.111)	(0.132)	(0.117)*
Homeowner	-0.003	0.034	0.026	0.166
	(0.069)	(0.075)	(0.073)	(0.080)**
Kids	-0.258	-0.057	-0.361	-0.143
	(0.078)***	(0.075)	(0.082)***	(0.082)*
Age	-0.096	0.016	0.005	-0.001
	(0.158)	(0.028)	(0.004)	(0.000)*
Age-squared	0.004	0.000	0.000	0.000
	(0.004)	(0.000)	(0.000)*	(0.000)*
Age-cubed	0.000	0.000	-0.001	-0.003
	(0.000)	(0.000)	(0.001)	(0.001)***
Income	0.000	-0.001	-0.247	0.031
	(0.001)	(0.001)*	(0.167)	(0.186)
Constant	1.557	0.297	1.572	-0.976
	(1.876)	(0.521)	(1.700)	(0.620)
P(Trad=Prenup)	0.5691	0.1768	0.4213	0.7116
P(Trad=Cohab)	0.0001	0.0001	0.0001	0.0001
P(Prenup=Cohab)	0.0001	0.0001	0.0003	0.0001

*Education level dummies included but not shown, ***: p<.01, **: p<.05, *: p<.10*

Table 4.3. OLS, Dependent Variable: BMI

	Male	Female
Trad	0.738***	0.140
	(0.109)	(0.196)
Div	0.631***	-0.133
	(0.205)	(0.290)
Wid	0.158	-0.138
	(0.491)	(0.766)
Prenup	0.462***	0.131
	(0.145)	(0.250)
Cohab	0.365***	-0.245
	(0.121)	(0.219)
Smoker	-0.466***	-1.037***
	(0.066)	(0.116)
Age	1.360***	1.360***
	(0.149)	(0.251)
Age-squared	-0.032***	-0.032***
	(0.004)	(0.007)
Age-cubed	0.00026***	0.00025***
	(0.00004)	(0.0001)
Income	0.0009	-0.0057***
	(0.0007)	(0.0012)
Kids	-0.117	-0.181
	(0.080)	(0.133)
Constant	4.919**	6.382***
	(1.711)	(2.860)
P(Trad=Prenup)	0.016	0.965
P(Trad=Cohab)	0.002	0.066
P(Prenup=Cohab)	0.520	0.149

*Education dummies included but not shown, ***: p<.01, **: p<.05, *: p<.10*

Table 4.4. Fixed Effects, Dependent Variable: BMI

	(1) Male	(2) Female	(3) Male	(4) Female
Trad. Marr	0.186	1.077***	0.127	0.162
	(0.147)	(0.287)	(0.079)	(0.143)
Div	-0.377*	-0.729*	0.272	-0.058
	(0.209)	(0.394)	(0.201)	(0.351)
Wid	0.952	1.979***	0.546	-0.199
	(0.599)	(0.756)	(0.806)	(0.852)
Prenup	0.267	1.028***	0.091	0.497**
	(0.165)	(0.317)	(0.122)	(0.234)
Cohab	0.195	0.203	0.186*	-0.381*
	(0.128)	(0.214)	(0.107)	(0.203)
Kids	-0.182*	-0.296*	-0.158	-0.418*
	(0.096)	(0.175)	(0.111)	(0.222)
Smoker	-0.335***	-0.782***	-0.320***	-0.734***
	(0.075)	(0.132)	(0.088)	(0.177)
Income	0.001	-0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Trad=Prenu	0.4219	0.7761	0.7478	0.1308
Trad=Comn	0.9393	0.0001	0.6092	0.0144
Prenup=Co	0.5972	0.0015	0.5075	0.0021

*Year dummies included but not shown, ***: p<.01, **: p<.05, *: p<.10*

Table 4.5. Comparing Results against the Literature)

	Dependent Variable: BMI				Dependent Variable: Log(BMI)			
	Gneezy & Shafrin		**Jeffrey & Rick (2002)**		**Gneezy & Shafrin**		**Averett et al. (2008)**	
	Men	Women	Men	Women	Men	Women	Men	Women
Trad	0.19	1.08***	0.70***	0.96***	0.008	0.034***	0.015***	0.017***
Prenup	0.27	1.03***			0.011*	0.035***		
Cohab	0.19	0.20			0.009*	0.006	0.010***	0.006**
Div	-0.38*	-0.73*	-0.27	-0.63***	-0.013*	-0.029**	-0.001	-0.007***
Wid	0.95	1.98***			0.030	0.069***		

Table 4.6. Two Stage Regression, Dependent Variable: BMI

	Males	Females
Dom	0.323	1.064
	(0.151)**	(0.295)***
Šeparation	-2.77	-11.94
	(2.01)	(3.31)***
Div	-0.335	-0.693
	(0.208)	(0.398)*
Wid	1.010	1.976
	(0.598)*	(0.758)***
Kids	-0.207	-0.287
	(0.097)**	(0.175)
Smoker	-0.352	-0.761
	(0.075)***	(0.133)***
Income	0.000	-0.002
	(0.001)	(0.001)*
Constant	24.240	23.765
	(0.124)***	(0.244)***

*Year dummies included but not shown, ***: p<.01, **: p<.05, *: p<.10*

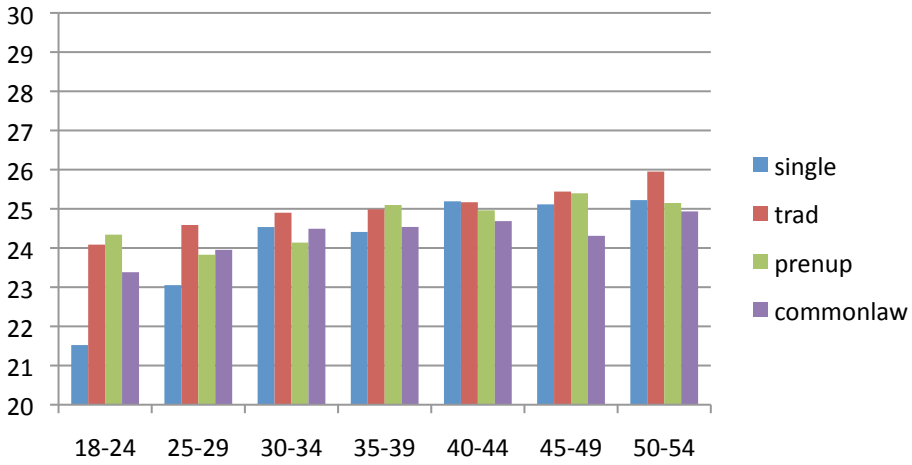

Figure 4.1. Average BMI by Age and Marital Status

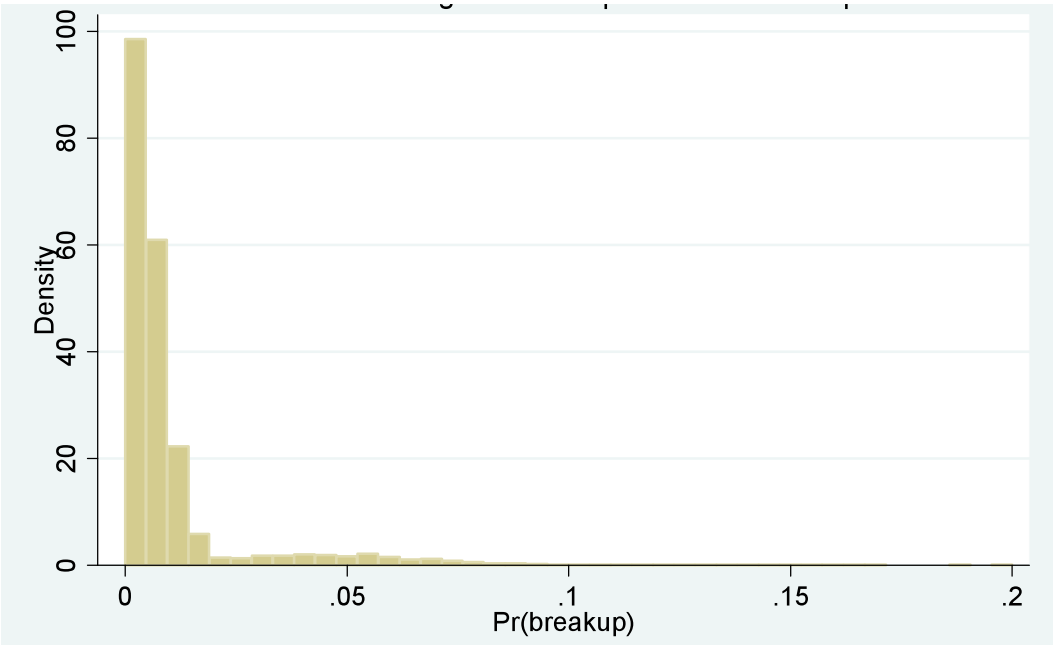

Figure 4.2. Probability of Separation:

Bibliography

42 U.S.C. 1396s. Program for distribution of pediatric vaccines (parallel cite Section 1928 of the Social Security Act).

Akerlof GA. 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics* **84**(3): 488-500.

Atella V, Deb P. 2008. Are primary care physicians, public and private sector specialists substitutes or complements? Evidence from a simultaneous equations model for count data. *Journal of Health Economics* **27**(3): 770-785.

Averett S and Korenman S. 1999. "Black-White Differences in Social and Economic Consequences of Obesity" *International Journal of obesity and related metabolic disorders: Journal of the International Association for the Study of Obesity*, **23**(2): 166-173.

Averett SL, Sikora A, Argys LM. 2008. "For better or worse: Relationship status and body mass index ." *Economics and Human Biology*, **6**(3):330-349

Becker G, DeGroote M, Marschak J. (1970) 'Measuring Utility by a Single-Response Sequential Method', *Behavioral Science* **9**: 226,232.

Bell KN, Hogue CJR, Manning C, Kendal AP. Risk Factors for Improper Vaccine Storage and Handling in Private Provider Offices. *Pediatrics*. 2001; 107(6):e100.

Binswanger, H. 'Attitudes Towards Risk: Experimental Measurement in Rural India' *American Journal of Agricultural Economics* **62**: 395-407.

Brinker MR, O'Connor DP, Pierce P, Spears JW. 2006. Payer type has little effect on operative rate and surgeons' work intensity. *Clinical Orthopaedics and Related Research* **451**:257-262.

Buchmueller T. 2000. The Health Plan Choices of Retirees under Managed Competition. *Health Services Research* **35**(5): 949.

Burkhauser, Richard V. and John Cawley. 2008. "Beyond BMI: The Value of More Accurate Measures of Fatness and Obesity in Social Science Research" *Journal of Health Economics*, **27**(2): 519-529.

California Department of Health Services Immunization Branch, Internal Memo. Program in Brief: Vaccine Management Business Improvement Project. February, 2007. http://www.docstoc.com/docs/DownloadDoc.aspx?doc_id=565104

California Department of Health Services Immunization Branch, Internal Memo. VFC Prepares to Transition to National Vaccine Distribution System. March 15, 2007. http://www.cdph.ca.gov/programs/immunize/Documents/TX%20to%20Centralized%20Ditribution%20FAX.doc

Cameron AC, Trivedi PK. 1986. Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests. *Journal of Applied Econometrics* **1**(1): 29-53.

Cawley, John, Kara Joyner, and Jeffery Sobal. 2006. "Size Matters: The Influence of Adolescents' Weight and Height on Dating and Sex" *Rationality and Society*, **18**(1): 67-94.

Center for Studying Health System Change. 2000. Community Tracking Study Household Survey, 1996-1997, and Followback Survey, 1997-1998: Restricted Use User Guide. ICPSR Study No. 2524, Washington, DC, 2nd ICPSR version.

Centers for Disease Control and Prevention. Recommended Immunization Schedule for Persons Aged 0-6 Years–United States, 2008.
http://www.cdc.gov/vaccines/recs/schedules/downloads/
child/2008/08_0-6yrs_schedule_bw_pr.pdf

Centers for Disease Control and Prevention. Vaccines for Children Operations Guide. August, 2007;M-6:10.
http://www.cdc.gov/vaccines/programs/vfc/downloads/vfc-op-guide
/vfc-op-guide-all-chaptr-files.pdf

Centers for Disease Control and Prevention. FY 2009 Budget Request Summary. Accessed December 22, 2008 at
http://www.cdc.gov/fmo/PDFs/FY09budgetreqsummary.pdf

Centers for Disease Control and Prevention. Ten great public health achievements-United States, 1900-1999. MMWR Morb Mortal Wkly Rep. 1999;48:241-243.

Centers for Disease Control and Prevention. Vaccines for Children Program - Program Brief. February 2007. Accessed December 22, 2008 at
www.cdc.gov/NCIRD/progbriefs/downloads/vfc.pdf

Conrad DA, Maynard C, Cheadle A, et al. 1998. Primary Care Physician Compensation Method in Medical Groups: Does It Influence the Use and Cost of Health Services for Enrollees in Managed Care Organizations? *JAMA* **279**(11): 853-858.

Courage C, Rey B. (2006), 'Prudence and optimal Prevention for Health Risks, *Health Economics* **15**: 1323-1327.

Craig, P. L. and A. S. Truswell. 1988. "Changes in Food Habits when People Get Married: Analysis of Food Frequencies." In *Food Habits in Australia*, ed. A. S. Truswell and M. L. Wahlqvist, 151-159. Melbourne.

Cunningham PJ, Denk C, Sinclair M. 2001. Do consumers know how their health plan works? *Health Affairs* **20**(2): 159-166.

Cutler DM, Reber SJ. 1998. Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection. *The Quarterly Journal of Economics* **113**(2): 433-466.

Davidson SM, Manheim LM, Hohlen MM, Werner SM, Yudkowsky BK, Fleming GV. 1992. Prepayment with Office-Based Physicians in Publicly Funded Programs: Results from the Children's Medicaid Program. *Pediatrics* **89**(4): 761-767.

Deck C, Schlesinger H. (2008) 'Prudence and Temperance: Exploring Higher Order Risk Effects in the Laboratory', Working Paper, University of Alabama.

DeNavas-Walt C, Proctor BD, Smith JC. Income, Poverty, and Health Insurance Coverage in the United States: 2007. Current Population Reports. U.S. Census Bureau., August 2008. Access January 15, 2009 at `http://www.census.gov/prod/2008pubs/p60-235.pdf`.

Dhanani N, O'Leary JF, Keeler E, Bamezai A, Melnick G. 2004. The effect of HMOs on the inpatient utilization of medicare beneficiaries. *Health Services Research* **39**(5): 1607-1627.

Dusheiko M, Gravelle H, Jacobs R, Smith P. 2006. The Effect of Financial Incentives on Gatekeeping Doctors: Evidence from a Natural Experiment. *Journal of Health Economics* **25**(3): 449-478.

Eeckhoudt L, Gollier C. (2005) 'The Impact of Prudence on Optimal Prevention', *Economic Theory* **26**: 989994.

Eeckhoudt L, Schlesinger H. (2006) 'Putting Risk in Its Proper Place, *American Economic Review* **96**: 280289.

Eisenhauer J, Ventura L. (2003) 'Survey Measures of Risk Aversion and Prudence, *Applied Economics* **35**: 14771484.

Elder, Glen H.,Jr. 1969. "Appearance and Education in Marriage Mobility" *American Sociological Review*, **34**(4): 519-533.

Ellis RP, McGuire TG. 1986. Provider Behavior under Prospective Reimbursement. *Journal of Health Economics* **5**(2): 129-151.

Ettner SL. 1997. Adverse selection and the purchase of Medigap insurance by the elderly. *Journal of Health Economics,* **16**(5): 543-562.

Finkelstein, Eric A., Ian C. Fiebelkorn, and Guijing Wang. 2003. "National Medical Spending Attributable to Overweight and Obesity: How Much, and Who's Paying?" *Health Affairs*, Web Exclusive.

Fortney JC, Steffick DE, Burgess JF, Maciejewski ML, Petersen LA. 2005. Are Primary Care Services a Substitute or Complement for Specialty and Inpatient Services? *Health Services Research* **40**(5): 1422-1442.

Frank RG, Zeckhauser RJ. 2007. Custom-made versus ready-to-wear treatments: Behavioral propensities in physicians choices. *Journal of Health Economics,* **26**(6): 1101-1127.

Frazelle EH, Hackman ST. The Warehouse Performance Index: A single Point Metric for Benchmarking Warehouse Performance. Material Handling Research Center Technical Report. TR-93-14, Georgia Tech, Atlanta, GA.

Glied S, Zivin JG. 2002. How do doctors behave when some (but not all) of their patients are in managed care? *Journal of Health Economics* **21**(2): 337-353.

Gomes F, Michaelides A. (2005) 'Optimal life-cycle asset allocation: Understanding the empirical evidence, *The Journal of Finance* **60**(2): 869904.

Gomez M. (2003) 'Are Individuals Prudent? An Experimental Approach Using Lottery Choices', Working Paper, Copenhagen Business School.

Gortmaker, Steven L., Aviva Must, James M. Perrin, Arthur M. Sobol, and William H. Dietz. 1993. "Social and Economic Consequences of Overweight in Adolescence and Young Adulthood" *The New England Journal of Medicine*, **329**(14): 1008-1012.

Gosden T, Forland F, Kristiansen IS, Sutton M, Leese B, Giuffrida A, Sergison M, Pedersen L. 2000. Capitation, salary, fee-for-service and mixed systems of payment: effects on the behaviour of primary care physicians. *Cochrane Database of Systematic Reviews (Online)* **(3)**(3): CD002215.

Hackman ST, Frazelle EH, Griffin PM, Griffin SO, Vlatsa DA. Benchmarking Warehousing and Distribution Operations: An Input-Output Approach. Journal of Productivity Analysis. 2001;16(1):79-100.

Hahn, Beth A. 1993. "Marital Status and Women's Health: The Effect of Economic Marital Acquisitions" *Journal of Marriage and the Family*, **55**(2): 495-504.

Hantrais, Linda. 2004. *Family Policy Matters: Responding to Family Change in Europe*, Policy Press, 246 pages.

Hellerstein JK. 1998. The Importance of the Physician in the Generic versus Trade-Name Prescription Decision. *The Rand Journal of Economics* **29**(1): 108-136.

Hellinger FJ. 1996. The impact of financial incentives on physician behavior in managed care plans: a review of the evidence. *Medical Care Research and Review* **53**(3): 294-314.

Hickson GB, Altemeier WA, Perrin JM. 1987. Physician reimbursement by salary or fee-for-service: effect on physician practice behavior in a randomized prospective study. *Pediatrics* **80**(3): 344-350.

Holt C, Laury S. (2002), 'Risk Aversion and Incentive Effects, *American Economic Review* **92**: 16441655.

Jeffrey Robert W. and Rick, Allison M. 2002. "Cross-Sectional and Longitudinal Associations between Body Mass index and Marriage-Related Factors. *Obesity Research*, **10**(8):809-815.

Kahn, H. S. and D. F. Williamson. 1990. "The Contributions of Income, Education and Changing Marital Status to Weight Change among US Men" *International Journal of Obesity*, **14**(12): 1057-1068.

Kahn, H. S. and D. F. Williamson. 1991. "Is Race Associated with Weight Change in US Adults After Adjustment for Income, Education, and Marital Factors?" *The American Journal of Clinical Nutrition*, **53**(6 Suppl): 1566S-1570S.

Kahn, H. S., D. F. Williamson, and J. A. Stevens. 1991. "Race and Weight Change in US Women: The Roles of Socioeconomic and Marital Status" *American Journal of Public Health*, 81(3): 319-323.

Kahneman D, Tversky A. (1979) 'Prospect Theory: An Analysis of Decision under Risk, *Econometrica* **47**: 263292.

Kimball M. (1990), 'Precautionary Savings in the Small and in the Large, *Econometrica* 58, 5373.

Krasnik A, Groenewegen PP, Pedersen PA, von Scholten P, Mooney G, Gottschau A, Flierman HA, and Damsgaard MT. 1990. Changing remuneration systems: effects on activity in general practice. *British Medical Journal* **300**(6741): 1698-1701.

Leibowitz A, Buchanan JL, Mann J. 1992. A randomized trial to evaluate the effectiveness of a Medicaid HMO. *Journal of Health Economics* **11**(3): 235-257.

Madden D, Nolan A, Nolan B. 2005. GP Reimbursement and Visiting Behaviour in Ireland. *Health Economics* **14**(10): 1047-1060.

Manning WG, Newhouse JP, Duan N, Keeler EB, Leibowitz A. 1987. Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. *The American Economic Review* **77**(3): 251-277.

Mao J. (1970) 'Survey of Capital Budgeting: Theory and Practice, *Journal of Finance* **25**: 349369.

McGinnis LF, Chen WC, Griffin P, Sharp G, Govindaraj T, Bodner D. Benchmarking Warehouse Performance: Initial Results for Internet-based Data Envelopment Analysis for Warehousing. W. M. Keck Virtual Factory Lab. 2006; (iDEAs-W 1.0).

Menezes C, Geiss C, Tressler J. (1980) 'Increasing Downside Risk, *American Economic Review* **70**: 921932.

Mokdad, Ali H., Mary K. Serdula, William H. Dietz, Barbara A. Bowman, James S. Marks, and Jeffrey P. Koplan. 1999. "The Spread of the Obesity Epidemic in the United States, 1991-1998" *JAMA: The Journal of the American Medical Association*, **282**(16): 1519-1522.

Myers, A. M., C. Weigel, and P. J. Holliday. 1989. "Sex- and Age-Linked Determinants of Physical Activity in Adulthood" *Canadian journal of public health.Revue canadienne de sante publique*, **80**(4): 256-260.

Newhouse JP, Marquis MS. 1978. The Norms Hypothesis and the Demand for Medical Care. *The Journal of Human Resources* **13**(Supplement: National Bureau of

Economic Research Conference on the Economics of Physician and Patient Behavior): 159-182.

Newton DA, Grayson MS. 2003. Trends in Career Choice by US Medical School Graduates. *JAMA: The Journal of the American Medical Association* **290**(9): 1179-1182.

Nicholson S, Bundorf K, Stein RM, Polsky D. 2004. The magnitude and nature of risk selection in employer-sponsored health plans. *Health Services Research* **39**(6): 1817-1838.

Nyholm, Maria, Bo Gullberg, Juan Merlo, Cristina Lundqvist-Persson, Lennart Rastam, and Ulf Lindblad. 2007. "The Validity of Obesity Based on Self-Reported Weight and Height: Implications for Population Studies" *Obesity*, **15**(1): 197-208.

Ogrod ES. 1997. Perspective: Compensation And Quality: A Physician's View. *Health Affairs* **16**(3): 82-86.

Pauly M, Percy A. 2000. Cost and performance: a comparison of the individual and group health insurance markets. *Journal of Health Politics, Policy and Law* **25**(1): 9-26.

Polsky D, Nicholson S. 2004. Why Are Managed Care Plans Less Expensive: Risk Selection, Utilization, or Reimbursement? *Journal of Risk and Insurance* **71**(1): 21-40.

Rand, C. S., J. M. Kuldau, and L. Robbins. 1982. "Surgery for Obesity and Marriage Quality" *JAMA: The Journal of the American Medical Association*, **247**(10): 1419-1422.

Remler DK, Donelan K, Blendon RJ, Lundberg GD, Leape LL, Calkins DR, Binns K, Newhouse JP. 1997. What do managed care plans do to affect care? Results from a survey of physicians. *Inquiry* **34**(3): 196-204.

Reschovsky JD, Hargraves JL, Smith AF. 2002. Consumer Beliefs and Health Plan Performance: It's Not Whether You Are in an HMO But Whether You Think You Are. *Journal of Health Politics, Policy and Law* **27**(3): 353-377.

Rissanen, A. M., M. Heliovaara, P. Knekt, A. Reunanen, and A. Aromaa. 1991. "Determinants of Weight Gain and Overweight in Adult Finns" *European Journal of Clinical Nutrition*, **45**(9): 419-430.

Rosenthal MB, Frank RG, Buchanan JL, Epstein AM. 2002. Transmission Of Financial Incentives To Physicians By Intermediary Organizations In California. *Health Aff* **21**(4): 197-205.

Rothschild M, Stiglitz J. 1976. Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics* **90**(4): 629-649.

Saver BG, Ritzwoller DP, Maciosek M, Goodman MJ, Conrad DA, Finkelstein E, Haase M, Barrett P, Cain KC. 2004. Does payment drive procedures? Payment for specialty services and procedure rate variations in 3 HMOs. *The American Journal of Managed Healthcare* **10**(3): 229-237.

Shen J, Andersen R, Brook R, Kominski G, Albert P, Wenger N. 2004. The effects of payment method on clinical decision-making: physician responses to clinical scenarios. *Medical Care* **42**(3): 297.

Shrank W, Ettner SL, Slavin PH, Kaplan HJ. 2005. Effect of physician reimbursement methodology on the rate and cost of cataract surgery. *Archives of Ophthalmology* **123**(12): 1733-1738.

Sobal, J., B. S. Rauschenbach, and E. A. Frongillo Jr. 1992. "Marital Status, Fatness and Obesity" *Social science and Medicine*, **35**(7): 915-923.

Sobal, Jeffery, Barbara Rauschenbach, and Edward A. Frongillo. 2003. "Marital Status Changes and Body Weight Changes: A US Longitudinal Analysis" *Social Science and Medicine*, **56**(7): 1543-1555.

Sturm, Roland. 2003. "Increases in Clinically Severe Obesity in the United States, 1986-2000" *Archives of Internal Medicine*, **163**(18): 2146-2148.

Sui, Xuemei, Michael J. LaMonte, James N. Laditka, James W. Hardin, Nancy Chase, Steven P. Hooker, and Steven N. Blair. 2007. "Cardiorespiratory Fitness and Adiposity as Mortality Predictors in Older Adults" *JAMA: The Journal of the American Medical Association*, **298**(21): 2507-2516.

Tversky A, Kahneman D. (1992) 'Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty* **5**:297323.

U.S. Census Bureau. U.S. Interim Projections by Age, Sex, Race, and Hispanic Origin. March 18, 2004.
`http://www.census.gov/ipc/www/usinterimproj/`

Umberson, D. 1992. "Gender, Marital Status and the Social Control of Health Behavior" *Social Science and Medicine*, **34**(8): 907-917.

Unser M. (2000) 'Lower Partial Moments as Measures of perceived risk: An Experimental Study, *Journal of Economic Psychology* **21**: 253280.

van Baal,Pieter H. M., Johan J. Polder, G. A. de Wit, Rudolf T. Hoogenveen, Talitha L. Feenstra, Hendriek C. Boshuizen, Peter M. Engelfriet, and Werner B. F. Brouwer. 2008. "Lifetime Medical Costs of Obesity: Prevention no

Cure for Increasing Health Expenditure" *PLoS Medicine*, **5**(2): e29.

Vasan, Ramachandran S., Michael J. Pencina, Mark Cobain, Matthew S. Freiberg, and Ralph B. D'Agostino. 2005. "Estimated Risks for Developing Obesity in the Framingham Heart Study" *Annals of Internal Medicine*, **143**(7):

473-480.

Verhoef, M. J., E. J. Love, and M. S. Rose. 1992. "Women's Social Roles and their Exercise Participation" *Women and Health*, **19**(4): 15-29.

Waldron, I. and D. Lye. 1989. "Family Roles and Smoking" *American Journal of Preventive Medicine*, **5**(3): 136-141.

Wee, Christina C., Nancy A. Rigotti, Roger B. Davis, and Russell S. Phillips. 2001. "Relationship between Smoking and Weight Control Efforts among Adults in the United States" *Archives of Internal Medicine*, **161**(4): 546-550.

Wolfers, Justin and Betsey Stevenson. 2007. "Marriage and Divorce: Changes and their Driving Forces" *Journal of Economic Perspectives*, **21**(2): 27-52.

Wooldridge, Jeffrey M. 2001. *Econometric Analysis of Cross Section and Panel Data*, First ed: The MIT Press.

Yang, Zhou and Allyson G. Hall. 2008. "The Financial Burden of Overweight and Obesity among Elderly Americans: The Dynamics of Weight, Longevity, and Health Care Cost" *Health Services Research*, 10.1111/j.1475-6773.2007.00801.x.

Yu W, Ellis RP, Ash A. 2001. Risk Selection in the Massachusetts State Employee Health Insurance Program. *Health Care Management Science* **4**(4): 281.

Zelfgerapporteerde medische consumptie, gezondheid en leefstijl, Central Bureau of Statistics, March 18, 2008 accessed March 19, 2009:
`http://statline.cbs.nl/StatWeb/publication/?VW=T&DM=SLNL&PA=03799&D1=242&D2=0-2&D3=a&D4=a&HD=081122-0545&HDR=G2,T,G3&STB=G1`

Zhou F, Santoli J, Messonnier ML, Yusuf HR, Shefer A, Chu SY, Rodewald L, Harpaz R. Economic Evaluation of the 7-Vaccine Routine Childhood Immunization Schedule in the United States, 2001. *Arch Pediatr Adolesc Med.* 2005; **159**(12):1136-1144.