Empirical Essays on Health Insurance Demand

Inaugural-Dissertation

zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München,
Volkswirtschaftliche Fakultät

2016

vorgelegt von

Maximiliane Hörl

Referent:	Prof. Dr. Joachim K. Winter
Korreferent:	Prof. Dr. Amelie C. Wuppermann
Promotionsabschlussberatung:	16.11.2016

Acknowledgements

First and foremost I would like to thank my supervisor Joachim Winter. I am grateful for his great support, constructive remarks and general advice during my dissertation. I would further like to thank Amelie Wuppermann for her guidance and encouraging research discussions during the past years. I also thank Derya Uysal who kindly agreed to serve as the third reviewer on my dissertation committee.

In addition, I experienced a lot of support writing this dissertation. In particular, I would like to thank my further co-authors James Smith, Iris Kesternich, Katherine Carman, Silvia Barcellos, and Sebastian Bauhoff for their critical remarks and many comments.

My dissertation has largely benefited and would not have been the same without my stay at the Leonard D. Schaeffer Center for Health Policy & Economics and my frequent visits at the RAND Corporation. A special word of thanks to Dana Goldman and Julie Zissimopoulos at the Schaeffer Center and James Smith and Katherine Carman at RAND. This time has been very inspiring for my research. I would further like to thank Patricia St. Clair, Jillian Wallis, Bryan Tysinger and Dirk Rösing for support related to data access.

It was a great experience being part of the first cohort of the international structured doctoral program "Evidence-Based Economics" with great colleagues and exposure to inspiring research through challenging courses and frequent visits of international high-profile guest lecturers. I have gratefully benefited from financial support from the Elitenetwork of Bavaria.

Finally, I would especially like to thank my family for their continuous support.

Contents

List of Figures

List of Tables

Preface

Traditional economic models were based on the *homo oeconomicus* – a rational decision maker with complete information who arrives at decisions by maximizing utility. Many anomalies that could not be explained by rational considerations lead to the behavioral revolution which challenged these traditional models. Individuals appear to struggle expecially with decisions that involve uncertainty, complex choice environments and consequences to current actions that lie in the future – elements that are combined in health insurance markets (Baicker et al. 2012; Liebmann and Zeckhauser 2008).

The discussion of choice in insurance markets was traditionally dominated by *adverse selection* models. In the context of health insurance, these models assume an ideal world in which health and associated future medical expenditures are perfectly observable to the individual. The rational individual compares easily understandable health plans and chooses the optimal plan that covers his needs while minimizing cost. As a consequence, those expecting higher health expenditures choose to obtain (more) insurance (for instance, Akerlof 1970; Rothschild and Stiglitz 1976). If there are information asymmetries between beneficiary and insurer regarding the beneficiary's risk type, this can lead to an unravelling of insurance markets. This is characterized in the disappearance of insurance options as a reaction to self-selection of risks types into plans as shown in the case of Harvard university (Cutler and Reber 1998).

Moreover, the finding that those with (more) insurance coverage are typically associated with higher health expenditures is also consistent with *moral hazard.* Those with (more) insurance coverage face lower marginal cost when seeking medical care since expenditures are partly or fully covered by a third party. Therefore, these beneficiaries have incentives to spend more than if they would bear the cost themselves (for instance, Arrow 1963; Pauly 1968). For the non-elderly, this strain of research was largely influenced by the RAND health insurance experiment (Aron-Dine et al. 2013; Manning et al. 1987) which randomly assigned families to plans with different degrees of cost-sharing. Beneficiaries with higher cost-sharing were found to have a significantly lower health care utilization in each category of care. Similar demand responsiveness to prices was reported for the elderly (Chandra et al. 2010).

Recently, research has discussed several barriers individuals face during the health insurance choice and utilization process (see, for instance, Baicker et al. 2012; Liebmann and Zeckhauser 2008). First, when choosing health insurance the individual has to assess the own risk. Individuals differ in their access to and evaluation of health information (Cutler and Lleras-Muney 2006) and can make systematic errors when assessing probabilities (Kahneman and Tversky 1979). Research has further shown that insurance decisions are sensitive to perceptions of risk (Johnson et al. 1993) and that the degree of coverage is likely influenced by the individual risk aversion (de Meza and Webb 2001; Finkelstein and McGarry 2006). Second, the individual faces a complex choice environment. The ability to understand health insurance concepts – a prerequisite for evaluating different options – varies widely between individuals (Barcellos et al. 2014; Loewenstein et al. 2013; Long et al. 2014b). Limitations can constitute a barrier to even choosing a health insurance plan at all (Heiss et al. 2006) or can lead to bad health insurance choices (Bhargava et al. 2015). Third, being better able to understand the terms of the own health insurance contract might imply a better ability to make use of the insurance coverage. Beneficiaries are often poorly informed about the plan attributes of their own insurance plans – which can be an important barrier when seeking medical care (Cunningham et al. 2001).

This dissertation deals with the demand for health insurance of a specific group: the U.S. population. A large fraction of this population is covered by employer-sponsored health insurance (ESI, about 50% in 2013, see KFF 2013a). Moreover, many are covered by public, socialized health insurance. Those with very low incomes can obtain health insurance via Medicaid, whereas the population 65+ and the disabled are primarily covered by Medicare (about 18% and 13% in 2013, respectively). However, a part of the U.S. population lacks having any health insurance coverage (13% in 2013). Traditionally, those with coverage did not have much choice. For instance, 85% of firms offered only one health plan in 2014, 12% offered two health plans (KFF 2014a).

Relying on consumer choice, recent health reforms in the U.S. aimed at expanding health insurance coverage in several dimensions. The Medicare Modernization Act of 2003 included the introduction of Medicare Part D which aimed at decreasing financial risk of prescription drug expenditures by establishing the possibility to add subsidized prescription drug coverage to traditional Medicare (Part A and B). The Patient Protection and Affordable Care Act of 2010 (ACA) aimed at decreasing the number of uninsured, inter alia, by subsidies, the individual mandate, and the states' option to expand Medicaid. Both reforms rely on individual decisions which should ultimately result in lower premia due to increased competition. Individuals are exposed to complex decisions and face many options. For instance, to obtain Medicare Part D coverage as a supplement to traditional Medicare, individuals can choose between 19 to 29 plans (KFF 2015d). Moreover, on the newly established Marketplaces under the ACA, individuals have to weight benefits and cost of typically more than 40 health plans (Bhargava et al. 2015). These examples illustrate the general trend towards relying on consumer choice for health insurance decisions.

This dissertation contains three empirical essays with different perspectives on health insurance demand exploring heterogeneity related to knowledge and preferences and considering recent developments in the U.S. health insurance markets. Chapter 1 investigates education-induced heterogeneity in health insurance demand for the Medicare population. In Chapter 2, I explore for the non-elderly in how far financial and health insurance literacy as well as risk preferences are related to the choice between health insurance plans. Chapter 3 provides insights into how these factors as well as political preferences can constitute barriers to health insurance coverage in context of the ACA. In addition, Chapter 4 sheds light on the formation of preferences investigating the impact of a shock in childhood.

In Chapter 1, I test traditional asymmetric information models which predict a positive coverage-risk correlation and explore potential education-related heterogeneity for the Medigap market. Medigap policies are heavily regulated health plans sold by private insurers to cover gaps arising in coverage from traditional Medicare which typically includes substantial cost-sharing. Using the representative Medicare Current Beneficiary Survey (MCBS), I analyze the education-dependent association between Medigap coverage and realized health expenditures for more than 1,700 beneficiaries aged 65+ comparing those without supplemental insurance to those with Medigap coverage but no other sources of supplemental coverage. The MCBS presents an extraordinary rich data source combining administrative claims (which contain detailed payment information) with survey instruments (which contain individual characteristics such as educational attainment).

Consistent with the predictions of traditional asymmetric information models and in contrast to recent empirical research, I find a strong significant positive correlation between Medigap coverage and health expenditures. Those beneficiaries with Medigap coverage spend on average about $4,000 more on health care than those without. This correlation does not change including a rich set of health measures which suggests that the positive correlation is rather driven by moral hazard than adverse selection.

Investigating heterogeneity, I find that the coverage-risk correlation is significantly larger for those with completed college degree compared to those with some college and less. While those with some college and less spend about $3,000 more when they have Medigap coverage, those with completed college degree spend about $7,100 more. Conditioning on health, the difference in the coverage-risk correlation between both groups decreases by a third and – while still being sizable – is not statistically significant anymore. Overall, my findings provide suggestive evidence for a substantial degree of moral hazard and additional adverse selection for those with completed college degree.

My findings have two important policy implications: First, this study provides support that the Medicare Access and CHIP Reauthorization Act of 2015 aiming to increase cost-sharing for a reduction of moral hazard presents an important milestone to limit federal spending. Second, heterogeneity in the coverage-risk correlation between beneficiaries with different levels of educational attainment might entail serious consequences. Since insurers do not consider educational attainment when setting prices, potential increases will affect all beneficiaries. In addition to paying mark-ups for a higher health care utilization of those with higher educational attainment, those with lower educational attainment could lose their policies related to affordability issues.

The next two chapters investigate health insurance choices in context of the ACA. Both chapters draw on data collected by my coauthors of Chapter 3 using the nationally representative RAND American Life Panel (ALP). The ALP has two major advantages: First, it includes more information on potential behavioral barriers to health insurance choices than included in earlier studies. It contains standard measures of *financial literacy* testing the individual's ability to understand the calculation of interest rates (numeracy), the change in the value of money due to inflation (inflation) as well as the difference between a single stock and a mutual fund (risk). Furthermore, it contains measures of *health insurance literacy* testing the individual's knowledge about deductibles, co-pays, coinsurance, networks, and prescription drug pricing (generic versus brand name). As a measure for the individual's *risk aversion*, participants further have to indicate their willingness to take risks. Moreover, it contains self-stated *political preferences* (Democrat, Republican, Independent/Other). Second, the longitudinal design of the ALP allows to track the health insurance coverage of the same individual over time which is important to identify prior uninsured individuals in Chapter 3.

In Chapter 2, I explore heterogeneity in the willingness to pay (WTP) across health insurance plans focusing on the role of financial and health insurance literacy as well as risk preferences. This chapter draws on a health insurance experiment in the ALP that was collected in the context of the ACA. In this experiment, participants have to choose between two health insurance plans while the price spread between the plans is varied randomly. The individuals can choose between a catastrophic type plan (1) where they have to pay the first $1,500 of their annual health expenditures out-of-pocket (OOP) and beyond this sum, all further expenditures are covered by insurance. Choosing the alternative plan (2), the individuals have to pay $200 OOP and then face copayments up to a total annual OOP maximum of $3,500, above which all further expenditures are covered by insurance. Both plans yield equivalent OOP cost for total annual health expenditures up to \$200 and Plan 2 yields lower OOP cost thereafter unless the individuals have extraordinary high health expenditures. In this case, Plan 1 shields them from additional expenditures. Translating the copayments into a coinsurance rate of 20% which is common for health insurance plans, about 90% of comparable individuals in another representative survey would benefit from choosing Plan 2, and hence most individuals should be willing to pay mark-ups for this plan.

To investigate the heterogeneity in the WTP across both health plans, I implement nonparametric and semiparametric estimators proposed by Lewbel et al. (2011) in Stata. These estimators can be applied to studies using referendum format elicitation where individuals are asked whether they are willing to pay a specified amount for a good while this amount is varied randomly across individuals. Given the choice of the individuals, the proposed value and covariates, the distribution of the unobserved WTP across individuals can be estimated. Using the proposed estimators, I investigate the heterogeneity in the price differential at which individuals switch from the first to the second health plan. I estimate univariate differences in the WTP across health plans using nonparametric estimators and multivariate differences using semiparametric estimators.

I find that the WTP does not differ largely between the plans. Considering typical health expenditures for this population, this cannot be explained by health and expenditure risk considerations alone. My nonparametric analysis indicates an increasing, almost linear trend in the valuation for Plan 2 with higher levels of financial and health insurance literacy. This is confirmed in the semiparametric analysis where I find that those with high levels of financial and health insurance literacy have a significantly higher WTP for Plan 2 of about \$18 which translates into a mark-up of 2-5 percentage points depending on the age-dependent premia. In contrast to prior research, I observe a less clear picture for risk attitudes in the nonparametric analysis and do not find significant differences for the WTP in the semiparametric analysis.

My results indicate that a better understanding of the contract terms facilitates identifying a potentially suitable health plan. This implies serious consequences for the low income population. Considering that this knowledge is correlated with income, my findings show that the most vulnerable part of the population can leave money on the table when selecting their health insurance plans.

Chapter 3 of this dissertation draws on joint work with Silvia Barcellos, Sebastian Bauhoff, Katherine Carman, Joachim Winter, and Amelie Wuppermann. In the context of the first two enrollment periods of the ACA and remaining high rates of uninsurance by spring 2015, we investigate in how far financial and health insurance literacy, as well as attitudes towards risk and political affiliation constitute barriers to insurance coverage. We make use of the longitudinal design of the ALP and track insurance choices for more than 2,500 individuals aged 18-64 in 2013 and 2015. For 525 individuals uninsured in 2013, we investigate the importance of the aforementioned factors for the probability of obtaining coverage by 2015. In addition, we identify which types of insurance individuals had in 2015, both for the previously uninsured and the general population.

Our results indicate that of those uninsured in 2013, about 60% had insurance by spring 2015. Among the uninsured in 2013, higher financial and health insurance literacy and greater risk aversion were associated with a greater probability of being insured in 2015. Uninsured Republicans were both less likely than Democrats to obtain insurance and to obtain it via Medicaid or the Marketplaces. For the general population, those with high financial and health insurance literacy were more likely to obtain insurance via Medicaid or the Marketplaces compared to being uninsured. Republicans were less likely than Democrats to be covered via Medicaid or the Marketplaces. Strikingly, the magnitude of the coefficients for our novel predictors was of the same order as that of the more traditional covariates like demographic characteristics including education and employment.

Our results have important policy implications. We stress that a lack of understanding of health insurance is likely to be an important barrier to becoming insured. Individual preferences, such as risk aversion and political leanings, may also constitute barriers to obtaining health insurance. Outreach and consumer-education programs designed to increase coverage should consider these barriers.

The last chapter with focus on the formation of preferences draws on joint work with Iris Kesternich, James Smith, and Joachim Winter. Specifically, we investigate in how far a major shock experienced in childhood can permanently shape trust as an adult. We consider a severe hunger episode in Germany after WWII and construct a measure of exposure to hunger from official data on caloric rations that were set monthly by the occupying forces in the four occupation zones thus providing both regional and temporal variation in the exposure to hunger. We combine these exposure to hunger measures with valid measures of trust that are available for a nationally representative sample of the German population in the German Socio-Economic Panel (SOEP). We focus on the population born between 1929 and 1955 and investigate differential impacts of exposure to hunger at ages 0-3 (infant), 4-7 (child), 8-16 (youth) as well as overall at ages 0-16.

We show that individuals who were exposed to low caloric rations in their childhood and youth have significantly less trust as adults many years after the hunger exposure. Strikingly, we find that the hunger effects are largest for the oldest age group. Exploring known trust differences in East and West Germany, we show that about one-fifth of lower levels of adult trust among those who lived as children in what became Eastern Germany were due to differences in caloric rations during their childhood. However, we did not find any support that food insecurity and unpredictability due to varying caloric rations played a role for the formation of trust.

Our findings have important implications. We show that experiences during childhood can have a permanent impact on individual preferences which can play an important role for human interactions and economic decisions. Our results provide a new explanation for the variation in preferences across individuals.

1 Heterogeneity in asymmetric information: Evidence from the Medigap market

Abstract. *Using the representative Medicare Current Beneficiary Survey, I investigate traditional predictions for the coverage-risk correlation and potential education-related heterogeneity for the Medigap market. Consistent with traditional models and in contrast to recent research, I find that Medigap coverage is associated with higher health expenditures. I further find that this coverage-risk correlation is higher for those with high educational attainment which is consistent with recent research showing differences in the ability to understand health insurance contracts. I provide suggestive evidence that these correlations are rather driven by moral hazard than adverse selection. Moreover, my results suggest some degree of adverse selection for those with high educational attainment. My results are robust controlling for further background characteristics such as income.*[1]

1.1 Introduction

Asymmetric information has served as a prominent explanation for failure of insurance markets. One problem can be adverse selection. Those individuals with private information on high expected claims have incentives to obtain (more) insurance coverage (see, for instance Akerlof 1970; Rothschild and Stiglitz 1976). Moreover, there can be moral hazard. Those facing lower cost have incentives to use more insurance than if they would bear the cost themselves (Arrow 1963; Pauly 1968). Both theories predict a positive correlation between insurance coverage and the realized loss (coverage-risk correlation).[2] One implicit assumption of the traditional models is that individuals are in general equally capable of choosing and using insurance. However, recent evidence has shown that individuals are quiet heterogeneous. For instance, they differ in their knowledge of insurance terms (Barcellos et al. 2014; Loewenstein et al. 2013) or characteristics of their own insurance plan (Cunningham et al. 2001). This variation should result in differences in the coverage-risk correlation between alternative groups of insurance holders.

[1]This paper benefited from helpful remarks of Raphael Guber and further colleagues.

[2]In this paper, I use the term coverage-risk correlation. *Coverage* refers to insurance coverage in this context, *risk* to the realization of risk from an insurer's perspective which can be both due to a higher initial risk of a loss (size and/or probability) or a higher risk of moral hazard (or both).

As in other insurance markets, the empirical evidence in health insurance markets has been mixed, providing evidence for a positive, negative or no coverage-risk correlation in different insurance contexts (Baicker and Goldman 2011; Cohen and Siegelman 2010; Cutler and Zeckhauser 2000). The puzzling finding of a zero or negative correlation is often explained by offsetting factors that are positively correlated with coverage and negatively correlated with risk (for instance risk aversion, see de Meza and Webb 2001; Finkelstein and McGarry 2006). Cohen and Siegelman (2010) provide an overview of alternative explanations for this puzzle. First, different results regarding the coverage-risk correlation can be obtained by analyzing different groups of policyholders with a varying degree of private information. Second, even if policyholders have private information on their risk type, they do not necessarily adjust their decisions, and thus might not use the private information they have. In a similar vein, this paper investigates heterogeneity in asymmetric information related to differences in educational attainment. Higher educated individuals might be better informed about their risk or better able to identify a suitable health insurance plan that fits their needs – implying differences in adverse selection. In addition, they might be more knowledgeable of what is covered under their health insurance plan when they seek medical care – implying differences in moral hazard.

In this paper, I test traditional asymmetric information models which predict a positive coverage-risk correlation and explore potential heterogeneity depending on the level of educational attainment for the Medigap market. Medigap policies are heavily regulated health insurance plans sold by private insurers to cover gaps arising in coverage from traditional Medicare which typically includes substantial cost-sharing requirements. More than 80% of the beneficiaries have some form of supplemental insurance. Using the representative Medicare Current Beneficiary Survey (MCBS), I analyze the education-dependent correlation between Medigap coverage and realized health expenditures for more than 1700 beneficiaries aged 65+. I compare those with Medigap coverage but no other sources of supplemental coverage to those without supplemental coverage.

The prerequisite for investigating heterogeneity in the coverage-risk correlation is a data set that combines detailed information on health insurance coverage, health expenditures and background characteristics, most importantly educational attainment. Most administrative or survey data sets do not contain all measures. The MCBS presents an extraordinary rich data source, combining administrative claims with survey instruments, and thus containing the required information.

Consistent with the predictions of traditional asymmetric information models and in contrast to recent empirical research, I find a strong positive and significant correlation between Medigap coverage and realized health expenditures. Those beneficiaries with Medigap coverage spend on average about $4,000 more on health. The correlation does not change substantially when including a rich set of health measures. Assuming valid health measures, this finding provides evidence that the strong correlation is rather driven by moral hazard than adverse selection.

I propose that the coverage-risk correlation should be higher for those with high educational attainment – potentially related to both adverse selection and moral hazard. Compared to those with low educational attainment, I find that the coverage-risk correlation is significantly higher for those with high educational attainment. While those with low educational attainment spend about $3,000 more, those with high educational attainment spend about $7,100 more when they have Medigap coverage compared to those without Medigap coverage. Controlling for health, I find a higher coverage-risk correlation for the group with the higher level of educational attainment. However, compared to the findings without health controls, the difference in the coverage-risk correlation for those with high compared to those with low educational attainment is reduced by a third and gets insignificant. This provides suggestive evidence for adverse selection in the group with high educational attainment, in addition to a higher degree of moral hazard compared to those with low educational attainment. To disentangle the exact proportions of adverse selection and moral hazard is beyond the scope of this paper.

This study is closely related to Fang et al. (2008). The authors provide evidence for advantageous selection in the Medigap market and investigate its channels. Conditioning on pricing characteristics, they show that those with Medigap have on average about $4,000 lower health expenditures than those without. However, as soon as they condition on health, they find that those with Medigap spend on average about $2,000 more – evidence for some degree of moral hazard. The authors argue that these findings provide evidence for multidimensional private information driving selection into Medigap and show that the sources of this advantageous selection are education, longevity expectations, financial planning horizons and cognitive abilities.

This paper investigates asymmetric information in the Medigap market as well as education-induced heterogeneity and makes some important contributions to the literature: First, I present evidence for a positive coverage-risk correlation in the Medigap market which is consistent with adverse selection and moral hazard. My findings differ from Fang et al. (2008) primarily because I exclude those beneficiaries living in long-term care facilities, considering that their incentives to obtain Medigap are rather limited. Second, rather than thinking of educational attainment as an offsetting factor, I provide evidence for heterogeneity in asymmetric information depending on educational attainment. This finding is in line with the reasoning in Cohen (2005) and Cohen and Siegelman (2010). If insurance markets differ in their perceived complexity of insurance terms or if insurance markets target specific groups of the population with varying abilities to understand and use their insurance plans, this could provide an alternative explanation why asymmetric information can be found in some insurance markets, but not in others.

The remainder of this paper is structured as follows. Section 1.2 presents theoretic predictions and Section 1.3 the related literature. In Section 1.4 and 1.5, I describe the institutional background, the data and empirical methods. Section 1.6 presents the empirical results and Section 1.7 concludes.

1.2 Predictions

Coverage-risk correlation. *Adverse selection* is present in insurance markets if there are informational asymmetries with respect to the risk type that is known by the individual but either not known or cannot be used by the insurer.[3] The literature has been largely influenced by the seminal work by Akerlof (1970) and Rothschild and Stiglitz (1976) and later work by others (Miyazaki 1977; Riley 1979; Spence 1978; Wilson 1977). The prediction that high risk types have incentives to obtain (more) insurance translates into a positive correlation between risk and coverage. A high risk type implies, for instance, to expect health problems because of an own bad health history or hereditary diseases within the family. This so called "coverage-risk correlation" has been used to test empirically for the existence of adverse selection (Chiappori and Salanié 2000).

However, Chiappori and Salanié (2000) stress that this test is only a necessary but not sufficient condition for adverse selection considering that it is also consistent with moral hazard. *Moral hazard* can be a problem if the individual faces lower cost when seeking medical care since expenditures are partly or fully covered by a third party – and the third party can not observe the true health status of the individual. In this case, the individual will spend more than if he would bear the cost himself (Arrow 1963; Pauly 1968; Zeckhauser 1970). Higher expenditures can be both related to lower incentives to prevent bad states (for instance, through smoking, referred to as "ex ante moral hazard") or getting more (or more expensive, referred to as "ex post moral hazard") treatments once seeking care (Cutler

[3] For instance, the individual can be either better informed about the probability, the size of the loss or both (Cohen and Siegelman 2010). Another source of asymmetric information can be regulatory frameworks that prevent insurance companies to use some known individual characteristics for pricing insurance policies (Cohen and Siegelman 2010; Finkelstein and Poterba 2014).

and Zeckhauser 2000). Both adverse selection and moral hazard should lead to a positive correlation between insurance coverage and the realized loss, in this case the associated health expenditures:

Hypothesis 1 *Individuals choosing more coverage are associated with higher health expenditures (positive correlation property).*

Heterogeneity in coverage-risk correlation. In recent research, the prominent explanations for varying results with respect to asymmetric information in insurance markets have been off-setting factors – such as risk aversion – being positively correlated with insurance coverage and negatively correlated with risk (see Cutler et al. 2008; de Meza and Webb 2001; Finkelstein and McGarry 2006). Cohen and Siegelman (2010) recently discuss alternative explanations. First, there can be "private information in the possession of some but not all policyholders". Evidence for this explanation has been provided by Cohen (2005) who finds that the coverage-risk correlation is present for experienced, but not for new policyholders in the automobile market. The author explains this finding by learning of policyholders about their risk type. Second, Cohen and Siegelman (2010) stress the possibility of "failure of policyholders to use the private information that they have". Even if policyholders have private information about their risk type, they do not necessarily adjust their decisions (for instance, Hurd et al. 2004).[4]

From an adverse selection perspective, those with higher educational attainment might be better able to identify their risk type and to evaluate different health insurance terms with respect to their health needs when selecting health insurance plans. Prior research discussed that higher educated people have better access to health information and are better able to evaluate them (Cutler and Lleras-Muney 2006).[5] Being better informed about health risks should simplify the evaluation of the individual's own health status. Moreover, the individual faces a complex choice environment. Prior research has shown that the ability to understand health insurance concepts – a prerequisite to evaluate the different options – varies widely between individuals (Barcellos et al. 2014; Loewenstein et al. 2013; Long et al. 2014b). Limitations can constitute a barrier to even choosing a health insurance plan at all (Heiss et al. 2006) or can lead to bad health insurance choices (Bhargava et al. 2015).[6] Health

[4] Hurd et al. (2004) show for the elderly that despite of being able to predict mortality well, this knowledge is not related to actual retirement behavior.

[5] For instance, smoking declined earlier among the higher educated after 1950 when the information about dangers associated with smoking started to diffuse (de Walque 2010). Other studies show that a part of the positive correlation between education and health is due to differences in health knowledge (Kenkel 1991; Nayga 2000).

[6] Bhargava et al. (2015) show that health insurance literacy is correlated with the ability to identify bad health insurance options (so called "dominated plans").

insurance literacy has been shown to be strongly correlated with educational attainment (Loewenstein et al. 2013).

Taking a moral hazard perspective, being better able to understand the terms of the own health insurance contract might imply a better ability to use insurance coverage. Research has shown that beneficiaries are often poorly informed about the plan attributes of their own insurance plans (Cunningham et al. 2001). The degree of knowledge has been shown to be correlated with educational attainment. Lacking knowledge of what the own health insurance plan covers can constitute a barrier when seeking medical care. Considering that those with higher educational attainment both might be in the possession of more information and/or better able to use the information they have, I propose:

Hypothesis 2 *Individuals with higher educational attainment are associated with a stronger correlation betweeen coverage and health expenditures.*

1.3 Related literature

Prior evidence for asymmetric information. The empirical evidence for *adverse selection* in different insurance markets has been mixed.[7] There is a significant amount of empirical research in health insurance markets that is discussed in detail by Cutler and Zeckhauser (2000). Most of the reviewed studies find some degree of adverse selection. A well-known example is the study by Cutler and Reber (1998) analyzing health insurance choices of Harvard University employees in response to a health insurance pricing policy change moving from the subsidization of generous insurance to a voucher-type system which pays a fixed subsidy independent of plan choice. As a reaction to facing higher prices for the comprehensive policies, better risks moved out of the more generous plans. Within three years, the market for more generous health insurance was eliminated entirely (known as an "unraveling of the markets").

The main focus of this study is on the Medigap insurance market.[8] The empirical evidence for adverse selection in this market has been quite mixed. While early research reports

[7]For instance, evidence for adverse selection has been provided for the annuity (Finkelstein and Poterba 2004), but not for the long-term care market (Finkelstein and McGarry 2006) while evidence in the life insurance market (Cawley and Philipson 1999; He 2009) and automobile insurance market (Chiappori et al. 2006; Cohen 2005; Puelz and Snow 1994) has been mixed. For an exellent overview on the adverse selection literature in different markets, see Cohen and Siegelman (2010).

[8]Major parts of the health expenditures of Medicare beneficiaries are already covered by free or heavily subsidized parts of traditional Medicare. Hence, adverse selection for the Medicare population should primarily play a role for the supplemental insurance decision.

some degree of adverse selection into supplemental insurance based on high prior health expenditures (but not self-rated health, see Wolfe and Goddeeris 1991), later work does not find consistent evidence for differences in the probability to obtain Medigap based on the health status (Ettner 1997; Hurd and McGarry 1997). This early literature has some limitations. First, to study adverse selection, it is important to have a good measure of health risk. Hence, studies based on survey responses can be limited in their analysis (Hurd and McGarry 1997) or include measures subject to potential measurement error (Wolfe and Goddeeris 1991). Second, private supplemental insurance should be especially important for specific parts of the population without access to other forms of supplemental coverage (e.g. through Medicaid or a former employer). Hence, studies not distinguishing between sources of private insurance (e.g. via employer vs. self-purchased, see Hurd and McGarry 1997) or comparing the decision to obtain Medigap with the decision to obtain other sources of supplemental coverage (e.g. Medicaid, see Ettner 1997) might provide limited insights.

My study is most closely related to Fang et al. (2008). The authors rely on detailed claim information with respect to health expenditures and exclude individuals with access to free or heavily subsidized supplemental coverage provided by a former employer, Medicaid or some other government agency. They show that beneficiaries with Medigap coverage generally have lower health expenditures.[9] The authors interpret this as evidence for advantageous selection into Medigap. In my analysis, I closely follow their empirical strategy. In addition to excluding beneficiaries with other sources of coverage, I also take into account that the incentives to obtain Medigap might be limited for some beneficiaries.

Moral hazard has been shown to play an important role in insurance markets (Cohen and Siegelman 2010). With respect to health insurance markets, most studies find some degree of moral hazard. Excellent overviews are provided by Cutler and Zeckhauser (2000) and Baicker and Goldman (2011). A prominent investigation is the RAND health insurance experiment initiated by the U.S. government in the 1970s which randomly assigned families to plans with different degrees of cost-sharing.[10] Beneficiaries with higher cost-sharing are shown to have a significantly lower health care utilization in essentially each category of care with an elasticity around 0.2 (Manning et al. 1987). Although the analysis of the RAND health insurance experiment has been criticized (Aron-Dine et al. 2013), later studies often arrive at remarkably similar price sensitivity estimates (Baicker and Goldman 2011). Moreover,

[9] However, once conditioning on health, they find higher health expenditures for those with Medigap coverage which they interpret as evidence for moral hazard.

[10] The random assignment to plans with different degrees of generosity provides exogenous variation that facilitates identifying the effect of cost sharing on health care utilization in isolation from potential selection effects.

responsiveness to cost-sharing has been also reported for the poor (Finkelstein et al. 2012) and the elderly (Chandra et al. 2010).[11]

In the context of Medigap, the empirical evidence has been limited. Some studies provide evidence that those with more insurance coverage use more health services (Christensen and Shinogle 1997; Hurd and McGarry 1997). However, these studies concentrate primarily on the probability of a doctor visit or on the probability and length of a hospital stay. Moreover, the studies have been criticized potentially overestimating moral hazard by controlling insufficiently for care providers and the underlying health status (Lemieux et al. 2008). More recent evidence is provided by Keane and Stavrunova (2014). Using detailed health expenditure data, the authors provide evidence for substantial moral hazard, increasing health expenditures by about 24%. They further investigate heterogeneous effects and find that health care demand is more elastic for healthier beneficiaries. The authors do not analyze heterogeneity with respect to educational attainment.

Heterogeneity related to educational attainment. There has been limited empirical work on education-induced heterogeneity in asymmetric information, likely for the following reasons. First, most data sets including health expenditure data typically lack information on individual background characteristics such as education. Second, most research has focused on homogeneous spending elasticities and selection mechanisms. One important exception is the aforementioned study by Cohen (2005) showing that the coverage-risk correlation is present for experienced, but not for new policyholders in the automobile market.

However, research has shown that insurance knowledge might be limited among Medicare beneficiaries. Evidence has pointed towards seniors having generally very limited knowledge about the Medicare and Medigap regulation which is correlated with educational attainment (McCall et al. 1986). This lack of knowledge is found to have serious effects on the health insurance choices of the elderly. Heiss et al. (2006) show in the context of Medicare Part D that many of those who – in contrast to their immediate self-interest – remain without insurance coverage do not have more than a high school degree. At the same time, many of those who do not enroll into Medicare Part D do not understand the program or the enrollment process. This implies that asymmetric information should be less pronounced among the less educated. Most closely related to this paper concerning education-induced heterogeneity, Davidson et al. (1992) investigate the effect of the interaction between health insurance knowledge and health status on the decision to purchase insurance coverage supplementing Medicare. They find that the higher the level of Medicare knowledge, sicker beneficiaries are

[11] Chandra et al. (2010) analyze a policy change that raised patient cost sharing of a supplemental insurance for retired employees in California. They find that physician office visits and prescription drug utilization are price sensitive with elasticities similar to the RAND estimates.

more likely to have Medigap or HMO coverage. However, the authors do not have detailed information on realized health expenditures. Hence, their analysis is restricted to measures of the perceived health status.

1.4 Institutional background

Medicare. Medicare is a federal health insurance program established in 1965 targeted at the population aged 65+ and those with permanent disabilities.[12] In 2013, it covered 13% in the U.S. population and about 14% of the federal budget was spent for these beneficiaries. To become eligible for Medicare, an individual or the spouse has to have made payroll tax contributions for 10 years and more. Medicare consists of multiple components covering different health services. Table 1.1 provides an overview on these components. Part A mainly covers inpatient and Part B outpatient care. Part C, also known as Medicare Advantage, is organized via private health plans that provide Part A and B benefits (and in addition mostly also prescription drug benefits). Premiums in Medicare Part C are generally low and not related to age or health history, at the same time, access to care is usually organized via health maintenance organizations (HMOs) or preferred provider organizations (PPOs). Since 2006, beneficiaries enrolled in Medicare Part A and/or Part B can choose to obtain Medicare Part D to cover prescription drugs expenditures through private plans.

This paper focuses on those beneficiaries with traditional Medicare (Part A/B) and not covered by Medicare Advantage.[13] Traditional Medicare has high deductibles and cost sharing requirements and places no limit on the beneficiaries' out-of-pocket spending. Even if beneficiaries choose to obtain Medicare Part D to cover prescription drug cost, they face substantially higher cost when they reach the drug benefit's coverage gap (the so called "doughnut hole") until the catastrophe coverage limit. To cover the gaps, more than 80% of beneficiaries have some form of supplemental insurance such as employer-sponsored retiree health plans, Medicaid for the low-income population or Medigap which is the focus of this study.

Medigap. In 2010, about 23% of the Medicare population (9.3 million) chose a Medigap policy as supplemental coverage. Almost 40% of these beneficiaries combined Medigap with another source of supplemental coverage. Medigap enrollees are typically individuals that do not have access to employer- or union-sponsored retiree health plans and are not poor enough

[12] Unless noted otherwise, this section is based on information retrieved from KFF (2013a, 2014b,c, 2015e), https://www.medicare.gov/.

[13] Beneficiaries with Medicare Advantage face different access and cost conditions. In addition, beneficiaries can not combine Medicare Advantage and Medigap policies. It is actually illegal for insurers to sell a Medigap policy to a beneficiary with Medicare Advantage. Beneficiaries with Medigap policies obtained at an earlier point of time than joining Medicare Advantage are advised to drop their policy (https://www.medicare.gov/).

to qualify for Medicaid. Shielding beneficiaries from unexpected medical out-of-pocket cost, Medigap helps individuals to budget their health expenditures and in addition, it minimizes paperwork. The Medigap market has become heavily regulated over time. Beneficiaries can enroll into different plan types. All plans of the same type have to provide the same benefits. Some of these plan types cover the Part B deductible and thus provide first-dollar coverage which has been subject to substantial discussions recently.[14]

Table 1.1: Medicare overview

Medicare Part A	
Coverage	Inpatient hospital stays, skilled nursing facility stays, some home health visits, and hospice care
Cost-sharing	No premium, but deductible ($1,100 for benefit period in 2010) and coinsurance
Enrollment	Automatic when beneficiary turns 65
Medicare Part B	
Coverage	Physician visits, outpatient services, preventive services, some home health visits
Cost-sharing	Monthly premium ($96.40/month in 2010, higher for those with higher incomes), deductible ($155 in 2010) and coinsurance (20% for most services)
Enrollment	Voluntary
Medicare Part C	
Coverage	Enrollment in private health plans (HMO or PPO), includes all Medicare covered Part A and B benefits and typically Part D benefits.
Cost-sharing	Depends on terms
Enrollment	Voluntary, yearly open enrollment period
Medicare Part D	
Coverage	Outpatient prescription drug coverage through private plans that contract with Medicare
Cost-sharing	Monthly premiums ($37.25 weighted average in 2010), deductible ($310 in 2010) and cost-sharing, doughnut hole
Enrollment	Voluntary

Sources: KFF (2014b,c), https://www.medicare.gov/.

The best time for beneficiaries to enroll into individual Medigap policies is during their 6-month open enrollment period which starts the month the beneficiary turns 65 *and* is enrolled in Medicare Part B.[15] During this period, insurers cannot deny coverage (so called "guaranteed issue") and are not allowed to underwrite based on past or present health conditions. After this period, beneficiaries may not be able to buy a Medigap policy or it will cost more. All Medigap policies are guaranteed renewable at the end of the coverage period even if beneficiaries have health problems. Being enrolled in Medigap, beneficiaries have to pay a

[14] To limit health care expenditures and reduce federal spending, the Medicare Access and CHIP Reauthorization Act that prohibits Medigap policies to cover the Part B deductible was signed into law in April 2015.

[15] The specific regulations for beneficiaries under the age of 65 vary by state. In general, insurers are not required by federal law to sell any of their Medigap policies to beneficiaries under 65 (KFF 2013b).

monthly premium in addition to their Medicare premiums (Part B and D, if applicable). In 2010, the average Medigap premium was $183 per month. However, there are variations across states and depending on characteristics of the beneficiary. I follow Fang et al. (2008) and use age, gender and state of residence to control for variations in Medigap pricing.[16]

1.5 Data and methods

1.5.1 Data

MCBS. The Medicare Current Beneficiary Survey (MCBS) is a continuous survey of a representative sample of the Medicare population conducted since September 1991 by the Centers for Medicare and Medicaid Services (CMS) in the Department of Health and Human Services of the U.S. Government (MCBS 2010).[17] The population contains individuals aged 65+ or being disabled, residing in households or long-term care facilities. Beneficiaries are sampled from Medicare enrollment files. Each respondent is interviewed three times a year over a period of four years, using mostly computer-assisted personal interviews.[18] The sample is annually supplemented during the September to November interview to account for attrition (e.g. refusal or death) and newly enrolled beneficiaries.

The central goal of the MCBS is to determine all sources of payments for all services used by Medicare beneficiaries, going beyond expenditures covered by Medicare and including also co-payments, deductibles and non-covered services. Thereby, the MCBS aims to obtain a more complete picture of health expenditures (Adler 1994). Combining survey and administrative data, the MCBS includes detailed information on claims and total cost, but also cost by service[19] and cost by payment source.[20] The survey data further provide rich information on subjective health measures and background characteristics, importantly educational attainment. Information on health insurance coverage such as Medicare Part A and/or B as well as supplemental coverage is provided as well.

Sample selection and variables of interest. This study is based on 10,741 beneficiaries participating in the MCBS in 2010. Most beneficiaries are not only covered by traditional

[16] In some states, Medigap premia also depend on smoking status. The inclusion of this characteristic does not affect my results.

[17] Refer to http://www.cms.hhs.gov/mcbs/ for further information.

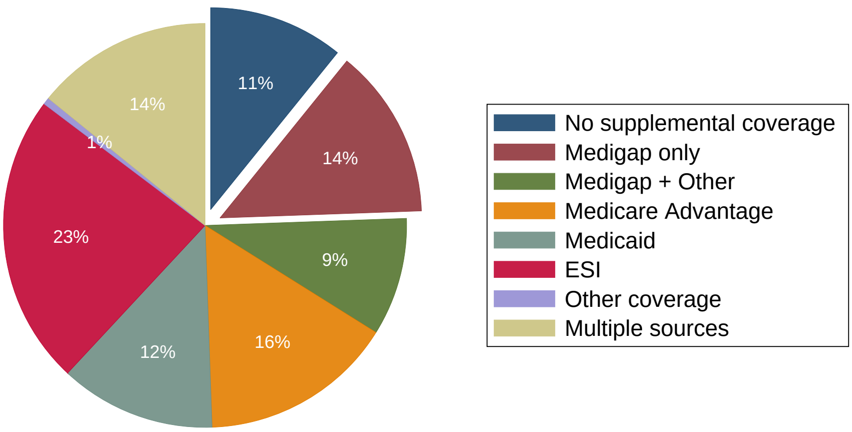
[18] For individuals living in long-term care facilities, the responses are collected from staff members designated by the facility director and not from the sample person or family members (about 8% of beneficiaries).

[19] E.g. Inpatient, outpatient, prescription drug, facility, dental care, medical services and goods.

[20] E.g. Medicare, Medicaid, Medicare HMO, private HMO, out-of-pocket, employer-sponsored insurance (ESI), individually purchased insurance, unknown private insurance, Veterans Affairs (VA), other payments, uncollected liability.

Medicare, but often have supplemental health insurance coverage. Following Fang et al. (2008), I define Medigap coverage as an indicator equal to one if the beneficiary has another form of self-purchased private health insurance that is secondary to Medicare. Beneficiaries with Medigap coverage can have other sources of supplemental coverage as well. Figure 1.1 provides an overview on the sources of supplemental coverage for the full Medicare population in the MCBS 2010 which are similar to numbers provided elsewhere (KFF 2013b). About 11% of the 2010 sample does not have any supplemental health insurance coverage. Moreover, about 14% has Medigap coverage only, another 9% has Medigap plus other sources of supplemental coverage.

Figure 1.1: Sources of supplemental coverage among Medicare beneficiaries in 2010



Source: Medicare Current Beneficiary Survey 2010, full sample. ***Notes:*** Figure displays sources of supplemental coverage among Medicare beneficiaries, weighted ($N = 10,741$).

I focus on beneficiaries with no other sources of supplemental coverage or Medigap coverage only. Beneficiaries with other or additional sources of supplemental coverage (such as Medicaid, employer-sponsored insurance) may anticipate ("selection") and face ("moral hazard") different cost structures. In addition, beneficiaries with Medicare Advantage are prohibited to use and advised to drop Medigap. Furthermore, beneficiaries can only enroll in Medigap if they are enrolled in Part A and B. Hence, I restrict the sample to beneficiaries enrolled in both parts. I further restrict the sample to beneficiaries aged 65+ considering that insurers are not required by federal law to sell Medigap policies to beneficiaries under 65. With respect to extraordinary high cost in the last year of life (estimated to be six times higher than for survivors, see Hogan et al. 2001), I exclude beneficiaries who die during the covered year.[21] I also focus on beneficiaries living in one of the U.S. states plus the District of Columbia.

[21] If the beneficiary died during the covered year and did not name a proxy respondent, the MCBS also uses imputation methods to determine health expenditures for the period before the death of the beneficiary (MCBS 2010).

Importantly, I exclude all beneficiaries living in a facility for the following reasons. First, Medigap policies generally do not cover long-term care. If beneficiaries can not expect their expenditures to be at least partially covered, they have weak incentives to obtain this form of supplemental health insurance. Beneficiaries living in facilities are typically covered by Medicaid, multiple sources of supplemental insurance (importantly, Medicare Advantage) or do not have any supplemental coverage at all (refer to Figure 1.5 in Appendix 1.A). Only 4% of these beneficiaries are covered by Medigap. This illustrates that beneficiaries in facilities do likely not belong to the target group of Medigap. Second, beneficiaries living in facilities have health expenditures which are on average five times as high as those of the community population (refer to Table 1.8 in the Appendix 1.A). Hence, including these beneficiaries could have an impact on the investigated coverage-risk correlations. Third, responses for these beneficiaries are not obtained in personal interviews but from several proxy individuals at the facility. Thus, they might be subject to measurement error. In 2010, about 8% of the beneficiaries lived in facilities.

To obtain a measure of health expenditure risk, I follow Fang et al. (2008) and use total health expenditures incurred within the covered year which is a combination of survey responses and administrative claims for those covered by traditional Medicare. Considering potential effects from the possibility of Medicare Part D coverage for prescription drugs, I use non-drug related total health expenditures in a robustness check. The MCBS further provides information on attributes used for rating (gender, age, state of residence) and on a set of self-reported health measures related to general health, diagnoses, limitations to activities of daily living and to instrumental activities of daily living. Moreover, the data set contains detailed information on the attained school degree[22] which I summarize in the two categories "some college and less" and "completed college degree".[23] Generally, I follow Fang et al. (2008) closely in the selection and construction of variables. Appendix 1.A provides a comprehensive overview on the construction of variables. The selected sample comprises $N = 1,744$ beneficiaries.

Descriptive statistics. Table 1.2 displays the descriptive statistics. Compared with the full sample which is representative of the general Medicare population, the selected sample has on average lower total health expenditures (about $15,400 vs. $12,200). A similar picture emerges for the non-drug related health expenditures. While about 23% have Medigap

[22] The schooling information is collected in the categories: No schooling, nursery to 8th grade, 9th to 12th grade (no diploma), high school graduate, vocation/technical/business/etc., some college (no degree), associate's degree, bachelor's degree, post graduate degree.

[23] Splitting the sample by educational attainment into those with high school and less, some college, and completed college degree instead, I do not find significantly different results between those with high school and less and those with some college.

coverage in the overall Medicare population, this is true for 71% in the selected sample.[24] This illustrates that Medigap presents an important possibility to cover cost-sharing requirements for beneficiaries with no other source of supplemental insurance. The selected sample shows similar characteristics with respect to gender, educational attainment and marital status, but differs to a certain extent with respect to race, ethnicity and income – certainly also related to the exclusion of the Medicaid population.

Concentrating on the selected sample, those with Medigap coverage have on average significantly higher total health expenditures which suggests the existence of asymmetric information. Those with Medigap are further significantly more likely to be female which is likely related to potential discounts for women offered by some insurers. With respect to socioeconomic background, an interesting pattern emerges. Compared to those without Medigap coverage, those with Medigap are less likely to be non-White or Hispanic and have higher levels of income. Interestingly, there are no significant differences with respect to the level of educational attainment. Table 1.9 in Appendix 1.A displays the descriptive statistics of the health characteristics for the selected sample. For most health measures, there are no significant differences between those without and those with Medigap coverage. However, there are exceptions. Those with Medigap coverage are more likely diagnosed with (non-skin) cancer and arthritis. It is more likely that they had a cataract operation. Moreover, they are less likely to smoke currently.

[24] Considering those with no supplemental coverage and those with Medigap coverage only, about 56% is covered by Medigap. The seemingly extraordinary high Medigap coverage of 71% in my selected sample is mainly driven by excluding those beneficiaries who would not be able to enroll into Medigap: (1) Beneficiaries not enrolled both in Part A and Part B which is required for enrollment in Medigap. (2) Beneficiaries under the age of 65. Only about 6% of those age 64 and younger have Medigap policies which is related to regulatory requirements.

Table 1.2: Descriptive statistics

Characteristics	Full sample	Selected sample			
		All	No Medigap	Medigap	t-test
	[N=10,741]	[N=1,744]	[N=477]	[N=1,267]	p-value
Health expenditures					
Total health expenditures (in $)	15420.87	12240.13	8892.55	13607.43	0.000
	(26570.11)	(19664.12)	(16635.69)	(20625.04)	
Non-drug related health expenditures (in $)	12552.79	9999.62	7046.23	11205.91	0.000
	(25603.26)	(17946.72)	(15496.14)	(18728.67)	
Insurance status					
Medigap coverage	0.23	0.71	0.00	1.00	1.000
	(0.42)	(0.45)	(0.00)	(0.00)	
Rating variables					
Female	0.55	0.56	0.43	0.61	0.000
	(0.50)	(0.50)	(0.50)	(0.49)	
Age (in years)	71.44	74.81	73.23	75.46	0.000
	(12.08)	(7.48)	(7.38)	(7.43)	
Socioeconomic background					
Non-White	0.10	0.06	0.15	0.03	0.000
	(0.30)	(0.24)	(0.35)	(0.16)	
Hispanic	0.09	0.04	0.06	0.02	0.002
	(0.28)	(0.18)	(0.24)	(0.16)	
Some college and less	0.76	0.74	0.77	0.74	0.362
	(0.42)	(0.44)	(0.42)	(0.44)	
Completed college	0.24	0.26	0.23	0.26	0.362
	(0.42)	(0.44)	(0.42)	(0.44)	
Income: 0-10k	0.12	0.05	0.08	0.04	0.000
	(0.33)	(0.21)	(0.27)	(0.19)	
Income: 10-20k	0.26	0.24	0.28	0.22	0.036
	(0.44)	(0.43)	(0.45)	(0.42)	
Income: 20-30k	0.20	0.26	0.27	0.25	0.746
	(0.40)	(0.44)	(0.44)	(0.43)	
Income: 30-40k	0.12	0.13	0.12	0.14	0.716
	(0.32)	(0.34)	(0.33)	(0.35)	
Income: 40k and more	0.30	0.32	0.25	0.35	0.000
	(0.46)	(0.47)	(0.43)	(0.48)	
Married	0.51	0.54	0.51	0.55	0.403
	(0.50)	(0.50)	(0.50)	(0.50)	

Data sources: Medicare Current Beneficiary Survey 2010. ***Notes:*** Table displays weighted averages (standard deviations). Medigap coverage refers to the beneficiary having self-purchased private supplemental coverage. This does not exclude other sources of supplemental coverage for the full sample. The last column shows results of t-tests for differences between those without and those with Medigap coverage (selected sample). Sample sizes for the full sample differ for the variables Non-White, Hispanic and marital status due to missing information (<0.5% of the full sample).

1.5.2 Empirical approach

The standard adverse selection and moral hazard models predict that those with (more) insurance coverage should have a higher realization of risk which suggests to a positive coverage-risk correlation (Chiappori and Salanié 2000). In the Medigap context, this implies a positive correlation between Medigap coverage and the associated health expenditures conditional on facing the same market environment. This is achieved conditioning on observable characteristics of the beneficiary that are used by the insurer for pricing. I estimate the following standard model (see Cohen and Siegelman 2010) to investigate the coverage-risk correlation:

$$Risk_i = \alpha + \beta Coverage_i + \gamma X_i + \epsilon_i \tag{1.1}$$

where $Risk_i$ presents the realization of beneficiary i's risk. $Coverage_i$ indicates whether the beneficiary is covered by Medigap. The vector X_i contains the beneficiary's characteristics known by the insurer and used for pricing: gender, a third-order polynomial of age,[25] and state of residence. I further use heteroscedasticity-robust standard errors clustered at the state of residence level. To investigate the education-induced heterogeneity in the coverage-risk correlation, I further estimate:

$$Risk_i = \alpha + \beta_{ncc} Coverage_i + \beta_{cc} Coverage_i \times E_i + \lambda E_i + \gamma X_i + \epsilon_i \tag{1.2}$$

where E_i indicates the educational attainment of the beneficiary. The proposed heterogeneity would be consistent with a positive and significant coefficient β_{cc} (cc = completed college) for the interaction term $Coverage_i \times E_i$. However, it should be stressed that the positive correlation between coverage and risk is consistent with adverse selection and moral hazard which point in different directions. Hence, this test only indicates the existence of asymmetric information, it does not distinguish between adverse selection and moral hazard. The estimated coefficients do not have a causal interpretation. To provide some insights into alternative explanations, I include a rich set of subjective health measures in further regressions. Moreover, I check the robustness of my specifications including demographic characteristics, excluding drug expenditures and excluding beneficiaries with the top 1% and top 5% total health expenditures. Finally, I explore different channels estimating the regressions for different expenditure types and splitting the sample by gender, age and marital status, respectively.

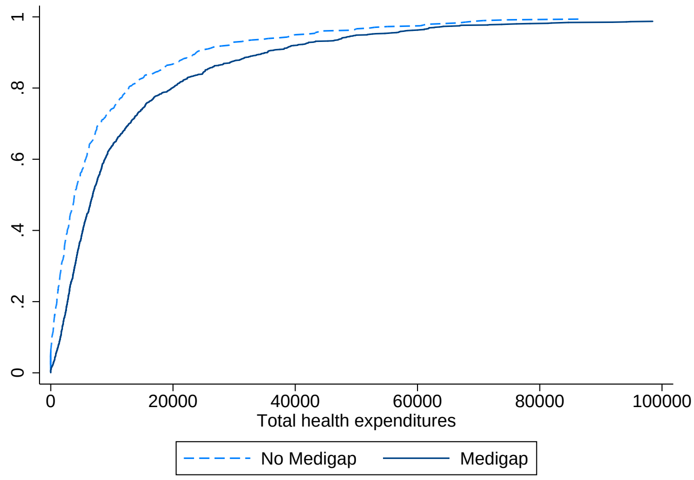
[25] The third-order polynomial of age is jointly significant at the 1% level.

1.6 Empirical evidence

1.6.1 Evidence for asymmetric information

Figure 1.2 illustrates the cumulative distributions of total health expenditures by Medigap coverage status. Compared to those without Medigap, those with Medigap generally have higher health expenditures. This observation is consistent with the positive correlation property, and hence provides a first indication of asymmetric information.

Figure 1.2: Cumulative distribution of health expenditures by Medigap coverage status



Source: Medicare Current Beneficiary Survey 2010, selected sample.
Notes: Figure displays the cumulative distribution of total health expenditures by Medigap coverage status. Beneficiaries with expenditures larger than \$100,000 are excluded (about 1% of the sample, $N = 1,725$).

In Table 1.3, I present regression results investigating the positive correlation property, following the strategy of Fang et al. (2008) closely. Not controlling for individual health characteristics (column 1), I find that those with Medigap coverage spend on average about \$4,000 more than those without (significant at the 1% level). Once controlling for the rich set of health measures (column 2), this association does not change much in terms of size and significance. However, the adjusted R-Squared increases substantially suggesting that the health measures are important predictors of health expenditures. My findings suggest that the positive correlation between Medigap coverage and health expenditures is primarily driven by moral hazard. Comparing beneficiaries with the same health status facing the same market conditions, those with Medigap coverage spend significantly more on health.

This interpretation rests on the assumption of valid health measures. Considering that the health measures are collected in the same year as the other information,[26] this assumption might be problematic. In the further discussion of results, I will assume valid health measures, but will also discuss potential changes in the limitations of this paper.

My findings stand in contrast to the results of Fang et al. (2008) who report a negative and significant coverage-risk correlation not controlling for health and a positive and significant correlation once including the health measures. These differences are primarily driven by the application of a different sample definition in this study. Importantly, I exclude beneficiaries living in a facility which typically have high health expenditures but a low degree of Medigap coverage since they generally can not expect their long-term care expenditures to be covered by Medigap. The inclusion of these beneficiaries leads to a negative coverage-risk correlation.

Table 1.3: Evidence for coverage-risk correlation

Variable	**Dependent variable: Total health expenditures**	
	Without health controls	**With health controls**
	(1)	(2)
Coverage-risk correlation		
Medigap coverage	3996.9***	3951.2***
	(789.7)	(870.7)
Rating variables		
Female	–455.6	–569.4
	(915.6)	(1241.5)
Age	7465.1	16796.6
	(14250.4)	(12833.3)
Age^2	–91.3	–213.3
	(179.6)	(161.7)
Age^3	.4	.9
	(.8)	(.7)
State-fixed effects	X	X
Health variables		X
Adj. R-Squared	.014	.188
N	1744	1744

Data sources: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Table displays coefficients (standard errors) from OLS regressions using heteroscedasticity-robust standard errors clustered at the state of residence level. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

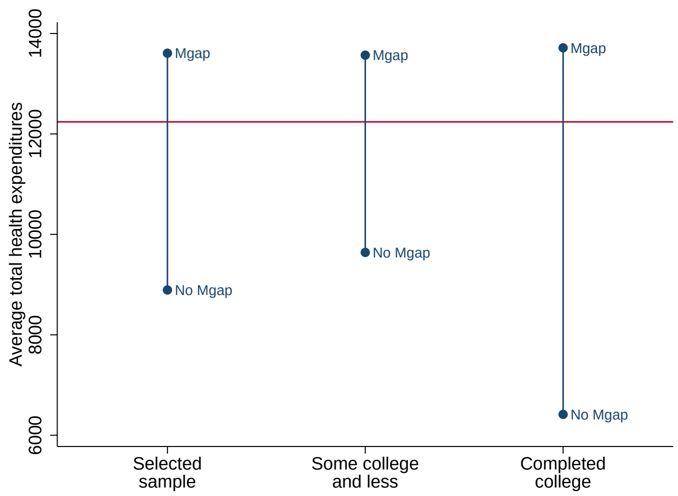
1.6.2 Evidence for heterogeneity

Figure 1.3 provides an overview on the average health expenditures by Medigap coverage status for the selected sample and for sub-samples with different levels of educational attainment. Irrespective of the level of educational attainment, those with Medigap coverage

[26] The health questions are collected in the fall round of the year (September through December).

have on average higher health expenditures which is consistent with the positive correlation property. The spread between those with and those without Medigap coverage is higher for the group with completed college degree compared to the group with some college and less. Hence, an investigation of the raw data already provides evidence for heterogeneity in the positive correlation property related to different levels of educational attainment. Interestingly, the average health expenditures for those with Medigap coverage do not differ between the groups with different levels of educational attainment.

Figure 1.3: Average health expenditures by Medigap coverage status and educational attainment



Source: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Figure displays the weighted averages of total health expenditures by Medigap coverage status and educational attainment ($N = 1,744$).

Table 1.4 presents regression results including educational attainment and its interaction with Medigap coverage. In columns 1-2, I present results without, in columns 3-4 with health controls. The results of Table 1.3 do not change substantially when including educational attainment (columns 1 and 3). Including the interaction term, I find that the coverage-risk correlation decreases to about \$3,000 for those with some college and less (difference between Medigap coverage coefficients in columns 1 and 2 statistically significant at 5% level). Moreover, I find a significantly higher coverage-risk correlation for those with completed college degree of about \$7,100 (\$3,000 + \$4,100; interaction term statistically significant at 5% level). When they are covered by Medigap, those with completed college degree spend significantly more compared to those with some college and less. When they are not covered by Medigap, they spend significantly less than the comparison group.

Once controlling for individual health characteristics, there are no differences in health expenditures between the two groups without Medigap coverage anymore, suggesting that the prior differences are driven by differences in health status. Moreover, the coverage-risk correlation for those with some college and less does not change much. However, the coefficient for the differences in the coverage-risk correlation for those with completed college degree in comparison to those with some college and less reduces – while still being sizable – by a third and is not statistically significant anymore. The described pattern is illustrated in Figure 1.6 of Appendix 1.A displaying the marginal effects related to the regressions in columns (2) and (4) of Table 1.4. The average health expenditures are predicted assuming that all beneficiaries in the sample have each level of educational attainment and each Medigap coverage status, one combination at a time.

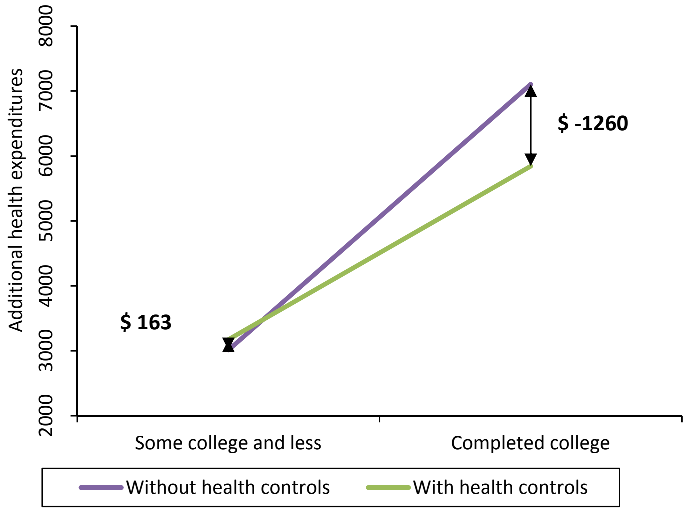
These differences in additional health expenditures related to Medigap coverage by level of educational attainment are illustrated in Figure 1.4. Those with some college and less consistently spend about \$3,000 more when they have Medigap coverage, irrespective of their health status. Hence, additional expenditures are likely not driven by selection into Medigap, but primarily by moral hazard. In contrast, those with completed college degree spend about \$7,100 more when they are covered by Medigap which decreases to about \$5,850 considering their health. These findings suggest a substantial degree of moral hazard – even larger than for those with lower educational attainment – but in addition also some degree of adverse selection into Medigap.

Table 1.4: Evidence for heterogeneity in coverage-risk correlation

Variable	**Dependent variable: Total health expenditures**			
	Without health controls		**With health controls**	
	(1)	(2)	(3)	(4)
Coverage-risk correlation				
Medigap coverage	3993.4***	3010.4***	3815.0***	3172.1***
	(780.0)	(964.1)	(875.7)	(1063.3)
Medigap x Completed college		4093.5**		2681.5
		(1885.8)		(2118.8)
Education				
Completed college	38.5	–2926.6**	1920.6**	–25.7
	(1029.3)	(1210.9)	(817.0)	(1312.0)
Rating variables	X	X	X	X
Health variables			X	X
Adj. R-Squared	.013	.015	.190	.190
N	1744	1744	1744	1744

Data sources: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Table displays coefficients (standard errors) from OLS regressions using heteroscedasticity-robust standard errors clustered at the state of residence level. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure 1.4: Differences in additional Medigap-related health expenditures



Source: Medicare Current Beneficiary Survey 2010, selected sample.
Notes: Figure displays the additional health expenditures for those with Medigap coverage compared to those without Medigap coverage by level of educational attainment, without and with health controls ($N = 1,744$).

1.6.3 Robustness and channels

In this section, I report results of robustness checks and explore potential channels. First, I perform robustness checks for the specifications of the regressions. Results are provided for the coverage-risk correlation and its heterogeneity. Second, I explore channels analyzing the heterogeneity in asymmetric information by expenditure type and for subgroups.

Robustness checks. Table 1.5 presents results of specification checks for regressions without and with health controls. Panel A presents robustness checks for the coverage-risk correlation, Panel B for heterogeneity by educational attainment. The findings could be driven by correlations between educational attainment and other socioeconomic background characteristics (such as income) which could also affect the level of health expenditures. In columns 2 and 7, I include additional controls for income, race, ethnicity, and marital status. Moreover, beneficiaries can obtain Medicare Part D to cover prescription drug expenditures. Drug expenditures account for about 20% of the health expenditures. Individual characteristics related to Part D coverage that have an impact on the decision to obtain Medigap and on the overall health expenditures could bias my findings. In columns 3 and 8, I therefore exclude prescription drug related expenditures in the dependent variable. In columns 4, 5, 9, and 10, I further trim the total health expenditure distribution excluding those with the top 1% and 5% expenditures, respectively, to verify that my findings are not driven by outliers.

Overall, the results remain remarkably robust. For the positive correlation property, I find that those with Medigap coverage spend on average \$3,000–\$4,000 more which is likely related to moral hazard considering that the correlation does not change substantially when including health. For the heterogeneity in the coverage-risk correlation, I find high and significant differences for those with completed college degree compared to those with some college and less. The group with higher educational attainment spends about \$2,600–\$4,200 more when having Medigap coverage in contrast to the group with lower educational attainment when having Medigap coverage. Importantly, this difference in the coverage-risk correlation between those with completed college and those with some college and less decreases consistently once controlling for health. This suggests some selection effects in addition to moral hazard for those with higher educational attainment.

Channels. Table 1.6 presents results examining potential channels related to different expenditure types (inpatient, outpatient, medical provider, dental, prescription drugs). I find a positive and mostly significant correlation between Medigap coverage and the realization of risk for all expenditure types which does not change much when including health controls. Significant heterogeneity in the coverage-risk correlation is present for inpatient and dental expenditures. While this difference in the coverage-risk correlation decreases for inpatient expenditures when including health controls, it does not change for dental expenditures. These findings suggest the existence of moral hazard for all expenditures types, irrespective of the level of educational attainment. Moreover, they suggest a higher degree of moral hazard for those with higher educational attainment when having inpatient and dental health events and additional selection into Medigap when expecting hospital visits.

In Tables 1.10, 1.11, and 1.12 of Appendix 1.A, I explore further channels and investigate differences in the results when splitting the selected sample by gender, age and marital status. With respect to gender, differences in the coverage-risk correlation by educational attainment seem to be driven by females. For males, I do not find a significant heterogeneity but generally a higher correlation between Medigap coverage and realized health expenditures. In light of the financial literacy literature showing that women are generally less financially literate than men (Lusardi and Mitchell 2011b), this finding illustrates the importance of educational attainment when facing decisions which are perceived to be complex. Moreover, the described heterogeneity appears to be driven by the population age 78 and younger. This finding is consistent with the literature showing that literacy decays with age (Lusardi and Mitchell 2011b). Turning to marital status, heterogeneity in the coverage-risk correlation is present for those not married, but not for those married. Hence, the own cognitive abilities appear to play a larger role for the health insurance demand of those not having a partner who can support their decisions.

Table 1.5: Robustness: Alternative specifications

Variable	Dependent variable: Total health expenditures									
	Without health controls					With health controls				
	Baseline	Include demographics	Exclude drug expenditures	Exclude top 1% expenditures	Exclude top 5% expenditures	Baseline	Include demographics	Exclude drug expenditures	Exclude top 1% expenditures	Exclude top 5% expenditures
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Positive correlation property										
Coverage-risk correlation										
Medigap coverage	3996.9***	3642.3***	3535.5***	3605.5***	3131.1***	3951.2***	3541.6***	3431.9***	3642.5***	3046.4***
	(789.7)	(741.1)	(707.3)	(757.1)	(666.4)	(870.7)	(929.9)	(759.1)	(701.3)	(594.8)
Adj. R-Squared	.014	.013	.013	.026	.033	.188	.189	.174	.217	.197
Panel B: Heterogeneity										
Coverage-risk correlation										
Medigap coverage	3010.4***	2668.1***	2509.3***	2688.7***	2350.5***	3172.1***	2849.7**	2629.1***	2899.2***	2392.3***
	(964.1)	(956.8)	(870.9)	(763.3)	(656.3)	(1063.3)	(1119.5)	(929.1)	(761.8)	(623.9)
Medigap x Completed college	4093.5**	3985.4**	4248.1**	3531.9**	2592.7**	2681.5	2582.0	2888.8	2449.0	1997.2
	(1885.8)	(1948.6)	(1669.5)	(1426.9)	(1124.8)	(2118.8)	(2177.9)	(1909.8)	(1726.1)	(1314.0)
Education										
Completed college	–2926.6**	–2987.2**	–3010.6***	–1827.7*	–266.8	–25.7	–241.1	–544.2	416.5	977.0
	(1210.9)	(1371.9)	(1037.3)	(1026.9)	(890.2)	(1312.0)	(1407.7)	(1115.2)	(1219.4)	(963.1)
Adj. R-Squared	.015	.013	.014	.028	.039	.190	.189	.175	.221	.208
Rating variables	X	X	X	X	X	X	X	X	X	X
Health variables					X	X	X	X	X	X
N	1744	1744	1744	1726	1656	1744	1744	1744	1726	1656

Data sources: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Panel A and B display coefficients (standard errors) from separate OLS regressions using heteroscedasticity-robust standard errors clustered at the state of residence level. Demographic controls contain race and ethnicity, income, and marital status. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 1.6: Channels: Expenditure types

Variable	**Dependent variable: Health expenditure type**									
	Without health controls					**With health controls**				
	Inpatient	**Out-patient**	**Medical provider**	**Dental**	**Prescrip-tion drugs**	**Inpatient**	**Out-patient**	**Medical provider**	**Dental**	**Prescrip-tion drugs**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coverage-risk correlation										
Medigap coverage	269.9	652.4**	1600.4***	249.5***	501.1**	459.2	514.1*	1532.2***	232.7**	542.9**
	(498.5)	(282.7)	(344.4)	(88.3)	(207.1)	(497.0)	(255.3)	(365.5)	(95.4)	(235.8)
Medigap x Completed college	1703.3***	8.2	942.5	1112.8***	–154.7	1031.4	–260.3	723.3	1115.8***	–207.4
	(583.6)	(431.8)	(732.9)	(344.4)	(412.6)	(661.4)	(455.5)	(798.5)	(349.1)	(398.0)
Education										
Completed college	–1874.1***	–334.7	–161.9	60.9	84.1	–850.5	82.2	413.6	–16.3	518.5
	(452.0)	(286.8)	(433.8)	(129.3)	(336.8)	(540.9)	(263.3)	(455.9)	(141.8)	(352.8)
Rating variables	X	X	X	X	X	X	X	X	X	X
Health variables						X	X	X	X	X
Adj. R-Squared	.004	.009	.023	.061	–.005	.105	.049	.120	.049	.084
N	1744	1744	1744	1744	1744	1744	1744	1744	1744	1744

Data sources: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Table displays coefficients (standard errors) from OLS regressions using heteroscedasticity-robust standard errors clustered at the state of residence level. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

1.7 Conclusion

In this paper, I test predictions of the positive correlation property as well as potential education-related heterogeneity using the representative Medicare Current Beneficiary Survey. This data set combines administrative claims with survey instruments, and hence provides extraordinary rich data which is necessary for this study. In my analysis, I compare those beneficiaries without supplemental health insurance to those with Medigap coverage but no other sources of supplemental health insurance. Importantly, I exclude beneficiaries living in facilities considering that this part of the Medicare population should have weak incentives to obtain Medigap, which is related to restrictions with respect to long-term care reimbursement.

Consistent with predictions of traditional theoretical models and in contrast to recent empirical research, I find a significant and strong positive correlation between Medigap coverage and health expenditures. This correlation is likely driven by some degree of moral hazard since results do not change when including a rich set of health measures.

Moreover, I find a consistent pattern that the coverage-risk correlation is significantly higher for those with high educational attainment. Including health controls, this difference is reduced by a third and loses significance. These findings suggest some degree of moral hazard for those with high educational attainment – even larger than for those with low educational attainment – but in addition some adverse selection into Medigap. My findings are consistent with recent research showing that beneficiaries differ in their abilities to understand health insurance policies (related to adverse selection) and in their knowledge of the terms of their own contracts (related to moral hazard) which have been both shown to be correlated with educational attainment.

This paper has some limitations. First, my sample size is rather small, leading to low power. Hence, economically significant effect sizes might not be detectable as statistically significant at conventional levels. In spite of this limitation, I find significant and robust heterogeneity in the coverage-risk correlation. Second, the health measures were collected in the same year as the other information. The health measures thus might be subject to measurement error or related to Medigap coverage: (A) Those without Medigap coverage might have undiagnosed illnesses. In this case, my findings provide even more support for substantial moral hazard comparing those with Medigap to beneficiaries without Medigap and potentially worse health. (B) Those with Medigap might feel more positive about their health due to improved access to the health care system, without their health having actually improved (similar patterns were observed in the Oregon health insurance experiment, see Finkelstein et al. 2012). If those with Medigap overreport their health, this could serve as an explanation for higher health expenditures in this group. Observing however a higher coverage-risk correlation for those with high educational attainment, this pattern is unlikely driven by increased overreporting considering their better comprehension of health information. (C) Those with Medigap might have an improved health status relative to when they obtained Medigap coverage. In this case, my findings controlling for health can be explained by a combination of moral hazard as adverse selection. To disentangle these factors is beyond the scope of this paper and left for future research. Third, my findings may not be generalizable to other populations. I investigate a specific part of the Medicare population which is characterized by not being covered by other sources of supplemental health insurance. However, the selected sample is quite comparable to the general Medicare population with somewhat higher incomes.

My findings provide two important policy implications: First, I have shown that differences in spending between those with and those without Medigap coverage are likely driven by some degree of moral hazard. In the context of the U.S. national debt reduction debate, there have been several proposals to restrict Medigap coverage in order to limit Medicare spending (see Cassidy 2011; KFF 2014c). The Medicare Access and CHIP Reauthorization Act signed into law in April 2015 prohibits Medigap policies to cover the Part B deductible for policies issued on or after January 1, 2020. Hence, this study provides support that this reform may be an important milestone to limit federal spending.

Second, heterogeneity in the coverage-risk correlation between groups of beneficiaries with different levels of educational attainment might entail serious consequences. I have shown that the combination of moral hazard and adverse selection is more pronounced for the group with higher educational attainment. However, since insurers do not distinguish between beneficiaries with different levels of educational attainment in the Medigap underwriting process, potential increases in prices will equally affect all parts of the Medicare population. This could have consequences for those with lower educational attainment (thus also lower income levels). In addition to paying mark-ups for higher health expenditures of those with higher educational attainment, this group could lose their policies related to affordability issues. Hence, these heterogeneous information asymmetries could have serious redistributional effects.

In the spirit of Cohen (2005), this paper highlights the importance to investigate heterogeneity in asymmetric information between different groups of policy holders. I provide evidence for heterogeneity by educational attainment and suggestive evidence for both factors moral hazard and adverse selection. To disentangle the exact proportions is beyond the scope of this paper and remains for future research.

1.A Appendix Figures and Tables

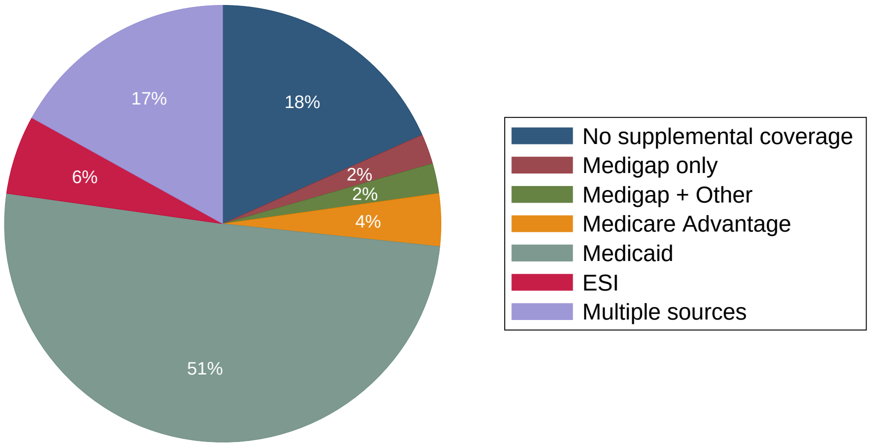
Table 1.7: Data description

Variable	**Description**
Realization of health expenditure risk	
Total health expenditures	Total annual health expenditures across all services by person for the survey year 2010. Expenditures are based on data from Medicare administrative and survey responses.
Total health expenditures w/o drug expenditures	Total annual health expenditures including all services excluding prescription drug expenditures.
Insurance coverage	
Medigap coverage	Indicator whether respondent has any self-purchased private health insurance coverage.
Heterogeneity	
Education	Indicators for the highest attained school degree: Some college and less, completed college degree.
Rating variables	
Female	Indicator for self-reported gender.
Age	Age of the respondent (as of December 31). Respondents under 65 are excluded from the sample.
State of residence	Indicators for state of residence from administrative data. Individuals with missing state or living in Puerto Rico are excluded from the sample.
Health variables	
Self-rated health	Indicators for respondent's self-reported general health compared to others at the same age collected in categories: Excellent, very good, good, fair, poor. Non-responses are marked by indicators.
Self-rated health compared to prior year	Indicators for respondent's self-reported general health compared to prior year collected in categories: Much better, somewhat better, about the same, somewhat worse, much worse. Non-responses are marked by indicators.
Height	Variables for respondent's self-reported height and height squared (in inches).
BMI	Respondent's body mass index calculated from self-reported weight and height (weight (kg)/height (m)^2).
Hearing	Indicator for respondent having at least some trouble with hearing. Non-responses are marked by indicators.
Cataract surgery	Indicator for respondent ever having a cataract surgery. Non-responses are marked by indicators.
Diagnoses	Indicators for respondent being ever told by a doctor to have arthritis, high blood pressure, diabetes, (non-skin) cancer, lung disease, heart attack, chronic heart disease, stroke, psychiatric illness, Alzheimer's disease, and broken hip. These questions reflect whether respondent has at some time been diagnosed with the condition, even if the condition has been corrected. Non-responses are marked by indicators.
Difficulties with (instrumental) activities of daily living (ADL)	Indicators respondent having at least some difficulty with walking 1/4 miles or 2-3 blocks, stooping/crouching/kneeling, reaching overhead, lifting 10 pounds, dressing, walking at all, bathing/showering, eating, getting in/out of a bed/chair, using the toilet, preparing meals, shopping, using the telephone, managing money and bills. These questions reflect whether respondent usually has trouble with these tasks. Non-responses are marked by indicators.

Currently smokes	Indicator for whether the respondent smoked a cigarette within the current month and not necessarily whether the respondent had a cigarette, cigar or pipe tobacco on the day of the interview. Since this included many missings, respondents indicating that they never smoked before are considered not to smoke currently. Non-responses are marked by indicators.
Further demographics	
Race and ethnicity	Indicators for respondent's background: Non-White, Hispanic.
Income	Indicators for respondent's self-reported income in ranges: 0-10k, 10-20k, 20-30k, 30-40k, 40k and more.
Marital status	Indicator for respondent's marital status.

Source: Medicare Current Beneficiary Survey 2010, following Fang et al. (2008) closely.

Figure 1.5: Sources of supplemental coverage among Medicare beneficiaries living in facilities in 2010



Source: Medicare Current Beneficiary Survey 2010, beneficiaries living in facilities. ***Notes:*** Figure displays sources of supplemental coverage among Medicare beneficiaries living in facilities, weighted ($N = 823$).

Table 1.8: Characteristics of community vs. facility beneficiaries

	Full sample **[N=10,741]**	**Community** **[N=9,918]**	**Facility** **[N=823]**
Total health expenditures	15420.87	12852.74	65493.73
	(26570.11)	(22539.98)	(43944.10)
Medigap coverage	0.23	0.24	0.04
	(0.42)	(0.43)	(0.21)

Data sources: Medicare Current Beneficiary Survey 2010, full sample.
Notes: Table displays weighted averages.

Table 1.9: Descriptive statistics: Health measures

Characteristics	**Selected sample**			
	All [N=1,744]	**No Medigap** [N=477]	**Medigap** [N=1,267]	**t-test** p-value
SRH: Excellent	0.21	0.20	0.21	0.999
	(0.40)	(0.40)	(0.41)	
SRH: Very good	0.33	0.33	0.33	0.480
	(0.47)	(0.47)	(0.47)	
SRH: Good	0.30	0.29	0.30	0.515
	(0.46)	(0.46)	(0.46)	
SRH: Fair	0.12	0.13	0.11	0.135
	(0.32)	(0.34)	(0.32)	
SRH: Poor	0.03	0.03	0.04	0.476
	(0.18)	(0.18)	(0.18)	
Change in SRH: Much better	0.04	0.05	0.04	0.235
	(0.21)	(0.22)	(0.20)	
Change in SRH: Somewhat better	0.11	0.10	0.11	0.561
	(0.31)	(0.30)	(0.31)	
Change in SRH: About the same	0.66	0.67	0.66	0.973
	(0.47)	(0.47)	(0.47)	
Change in SRH: Somewhat worse	0.17	0.16	0.17	0.967
	(0.37)	(0.37)	(0.37)	
Change in SRH: Much worse	0.02	0.02	0.02	0.870
	(0.13)	(0.13)	(0.13)	
Height (in inches)	66.22	66.73	66.02	0.001
	(3.97)	(3.96)	(3.96)	
BMI (kg/m^2)	27.21	27.44	27.11	0.191
	(5.26)	(5.40)	(5.20)	
Diagnosis: High blood pressure	0.67	0.64	0.68	0.126
	(0.47)	(0.48)	(0.47)	
Diagnosis: Diabetes	0.20	0.22	0.19	0.228
	(0.40)	(0.41)	(0.39)	
Diagnosis: (Non-skin) cancer	0.21	0.17	0.22	0.012
	(0.40)	(0.37)	(0.42)	
Diagnosis: Lung disease	0.16	0.17	0.16	0.974
	(0.37)	(0.37)	(0.37)	
Diagnosis: Heart attack	0.11	0.10	0.11	0.931
	(0.31)	(0.30)	(0.32)	
Diagnosis: Chronic heart disease	0.10	0.08	0.12	0.319
	(0.31)	(0.27)	(0.32)	
Diagnosis: Stroke	0.09	0.09	0.09	0.610
	(0.29)	(0.29)	(0.28)	
Diagnosis: Psychiatric illness	0.03	0.04	0.03	0.216
	(0.18)	(0.21)	(0.17)	
Diagnosis: Alzheimer's disease	0.02	0.02	0.01	0.294
	(0.13)	(0.14)	(0.12)	
Diagnosis: Broken hip	0.03	0.03	0.03	0.902
	(0.17)	(0.18)	(0.17)	
Diagnosis: Arthritis	0.58	0.51	0.60	0.001
	(0.49)	(0.50)	(0.49)	
Cataract operation	0.38	0.30	0.41	0.000
	(0.48)	(0.46)	(0.49)	
Hearing aid	0.12	0.11	0.12	0.831
	(0.32)	(0.31)	(0.33)	
Difficulties: Walking 2-3 blocks	0.29	0.31	0.28	0.093
	(0.45)	(0.46)	(0.45)	

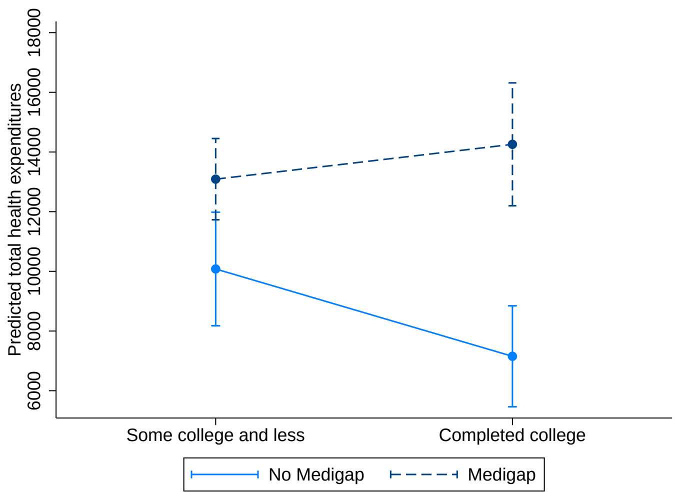
Continued on next page.

Characteristics	Selected sample			
	All [N=1,744]	No Medigap [N=477]	Medigap [N=1,267]	t-test p-value
Difficulties: Stooping	0.42	0.39	0.44	0.318
	(0.49)	(0.49)	(0.50)	
Difficulties: Reaching overhead	0.11	0.11	0.11	0.823
	(0.32)	(0.31)	(0.32)	
Difficulties: Lifting 10 pounds	0.18	0.16	0.19	0.486
	(0.38)	(0.37)	(0.39)	
Difficulties: Dressing	0.04	0.04	0.04	0.914
	(0.20)	(0.19)	(0.21)	
Difficulties: Walking at all	0.22	0.23	0.22	0.190
	(0.41)	(0.42)	(0.41)	
Difficulties: Bathing	0.08	0.09	0.07	0.109
	(0.26)	(0.29)	(0.25)	
Difficulties: Eating	0.02	0.03	0.02	0.234
	(0.14)	(0.16)	(0.13)	
Difficulties: Getting in/out of bed/chair	0.09	0.10	0.09	0.310
	(0.29)	(0.30)	(0.29)	
Difficulties: Using the toilet	0.03	0.03	0.03	0.972
	(0.18)	(0.18)	(0.18)	
Difficulties: Preparing meals	0.11	0.10	0.11	0.919
	(0.31)	(0.30)	(0.31)	
Difficulties: Shopping	0.10	0.11	0.10	0.452
	(0.30)	(0.31)	(0.30)	
Difficulties: Using the telephone	0.06	0.07	0.05	0.122
	(0.23)	(0.25)	(0.22)	
Difficulties: Managing money and bills	0.08	0.10	0.08	0.462
	(0.28)	(0.29)	(0.27)	
Currently smoking	0.10	0.17	0.07	0.000
	(0.30)	(0.38)	(0.26)	

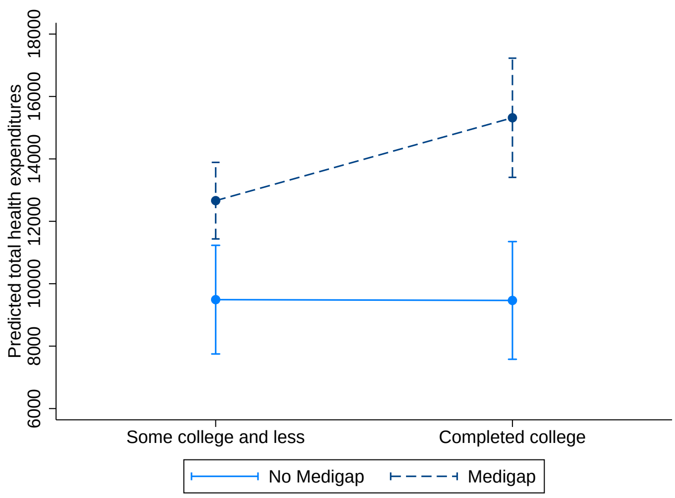
Data sources: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Table displays weighted averages (standard deviations).

Figure 1.6: Marginal effects: Health expenditures by Medigap coverage status and educational attainment

(a) Without health controls



(b) With health controls



Source: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Figure displays predicted total health expenditures for the corresponding Medigap coverage status by educational attainment along with their 95% confidence intervals ($N = 1,744$).

Table 1.10: Channels: Gender

Variable	**Dependent variable: Total health expenditures**					
	Without health controls			**With health controls**		
	All	**Female**	**Male**	**All**	**Female**	**Male**
	(1)	(2)	(3)	(4)	(5)	(6)
Coverage-risk correlation						
Medigap coverage	3010.4***	766.9	5469.4**	3172.1***	794.5	4931.9***
	(964.1)	(1520.6)	(2024.6)	(1063.3)	(1857.5)	(1511.4)
Medigap x Completed college	4093.5**	4839.5*	3159.9	2681.5	6677.8**	1145.6
	(1885.8)	(2765.2)	(3149.1)	(2118.8)	(2848.5)	(2847.1)
Education						
Completed college	–2926.6**	–4286.8*	–1658.5	–25.7	–3724.0*	637.0
	(1210.9)	(2149.3)	(1314.1)	(1312.0)	(2165.9)	(1628.9)
Rating variables	X	X	X	X	X	X
Health variables				X	X	X
Adj. R-Squared	.015	.002	.050	.190	.215	.196
N	1744	975	769	1744	975	769

Data sources: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Table displays coefficients (standard errors) from OLS regressions using heteroscedasticity-robust standard errors clustered at the state of residence level. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 1.11: Channels: Age

Variable	**Dependent variable: Total health expenditures**					
	Without health controls			**With health controls**		
	All	**Age ≤78**	**Age >78**	**All**	**Age ≤78**	**Age >78**
	(1)	(2)	(3)	(4)	(5)	(6)
Coverage-risk correlation						
Medigap coverage	3010.4***	2705.1*	2374.4	3172.1***	3029.8*	3516.6
	(964.1)	(1369.6)	(2468.7)	(1063.3)	(1497.9)	(2112.9)
Medigap x Completed college	4093.5**	4951.0**	1443.2	2681.5	2822.9	–296.1
	(1885.8)	(2125.5)	(2967.1)	(2118.8)	(2133.4)	(3503.4)
Education						
Completed college	–2926.6**	–2369.8*	–3733.3	–25.7	464.8	–582.3
	(1210.9)	(1301.9)	(2446.5)	(1312.0)	(1360.6)	(3012.5)
Rating variables	X	X	X	X	X	X
Health variables				X	X	X
Adj. R-Squared	.015	.013	–.021	.190	.180	.249
N	1744	1052	692	1744	1052	692

Data sources: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Table displays coefficients (standard errors) from OLS regressions using heteroscedasticity-robust standard errors clustered at the state of residence level. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 1.12: Channels: Marital status

Variable	**Dependent variable: Total health expenditures**					
	Without health controls			**With health controls**		
	All	**Not married**	**Married**	**All**	**Not married**	**Married**
	(1)	(2)	(3)	(4)	(5)	(6)
Coverage-risk correlation						
Medigap coverage	3010.4***	2949.9***	3296.9*	3172.1***	2652.4**	4406.0***
	(964.1)	(921.2)	(1771.8)	(1063.3)	(1147.6)	(1429.9)
Medigap x Completed college	4093.5**	4821.5*	3435.1	2681.5	3569.9	1355.8
	(1885.8)	(2832.4)	(2732.0)	(2118.8)	(2181.9)	(2661.3)
Education						
Completed college	–2926.6**	–4014.8*	–2549.9	–25.7	–1852.3	964.4
	(1210.9)	(2174.7)	(1563.9)	(1312.0)	(1319.5)	(2118.1)
Rating variables	X	X	X	X	X	X
Health variables				X	X	X
Adj. R-Squared	.015	.025	.005	.190	.315	.147
N	1744	849	895	1744	849	895

Data sources: Medicare Current Beneficiary Survey 2010, selected sample. ***Notes:*** Table displays coefficients (standard errors) from OLS regressions using heteroscedasticity-robust standard errors clustered at the state of residence level. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

2 Willingness to pay for health insurance across plans: Evidence from hypothetical choices

Abstract. *As recent health reforms in the U.S. rely more on consumer choice, it is important to understand determinants of preferences across health plans. Individuals differ in their preferences because of varying needs and less explored factors such as knowledge and risk attitudes. They reveal these preferences through their choices. Using a nationally representative survey, I investigate choices in a hypothetical health insurance experiment in which individuals have to choose between two plans while the price difference varies randomly. I implement and use nonparametric and semiparametric estimators developed by Lewbel et al. (2011) to investigate heterogeneity in the differential willingness to pay across plans. Both types of estimators identify meaningful patterns in the data. I provide robust evidence that those with higher knowledge have a significantly higher willingness to pay for the plan that potentially leads to lower annual out-of-pocket expenditures. The findings regarding risk aversion are less conclusive.*[1]

2.1 Introduction

Relying on consumer choice, recent health reforms in the U.S., including the Patient Protection and Affordable Care Act (ACA) and the introduction of Medicare Part D, aim to increase consumer welfare. Individuals have more options to find a suitable plan and should face lower prices as a result of increased provider competition. As a prerequisite, individuals have to make sensible health insurance choices. This requires, inter alia, that they understand which plan yields the lowest cost according to their needs and are risk averse enough to care about their insurance decisions.

Creating a rather complex choice environment, it is important for policy makers to understand determinants of preferences across health plans. For instance, how does the ability to understand the contract terms influences choices? In how far are individual risk attitudes important for choices across plans? While prior research suggests that these factors matter for health insurance decisions (see, for instance, Bhargava et al. 2015; Cutler et al. 2008;

[1]This paper benefited from helpful remarks of seminar participants in the joint internal seminar of the Chair of Empirical Economic Research and the Chair of Economic History at the University of Munich.

Finkelstein and McGarry 2006; Heiss et al. 2006), researchers lack clear understanding how they affect preferences across health plans. In this paper, I address these questions using novel estimators proposed by Lewbel et al. (2011) that allow to investigate the heterogeneity in preferences across health plans. I analyze choices in a hypothetical health insurance experiment. Using nonparametric estimators, I first explore how preferences across plans correlate with these factors. Using semiparametric estimators in a second step, I investigate how results change controlling for a rich set of background characteristics.

Preferences across different products – also expressed in the differential willingness to pay (WTP) – can be inferred either from the consumers' actual choices ("revealed preferences") or using "stated preference" methods (see Kesternich et al. 2013, for a detailed discussion). Preferences can be stated in different ways including direct (e.g. open-ended questions) and indirect methods. The latter methods include hypothetical choice experiments in which respondents have to choose among alternatives which contain different bundles of characteristics. The approach has three strengths: First, choice-based methods were shown to lead to less bias than open-ended WTP elicitation methods (Murphy et al. 2005). Second, hypothetical choice experiments are considered cognitively less challenging which is important in the insurance setting considering complex and infrequent choices (Braun et al. 2016). Third, research has shown that hypothetical health insurance experiments provide a useful tool to predict actual health insurance demand (Kesternich et al. 2013). However, the WTP is unobservable and needs to be inferred from the consumers' choices.

The hypothetical experiment I analyze is part of the RAND American Life Panel (ALP), a nationally representative online survey. Survey respondents have to choose between two insurance contracts that differ in their level of coverage. They can choose between Plan 1, a catastrophic type plan that covers expenditures above \$1,500, and Plan 2, including a deductible of \$200 and copayments up to an out-of-pocket (OOP) maximum of \$3,500. Respondents should benefit from choosing Plan 2 through lower OOP cost unless they have extraordinary high health expenditures (expected for only about 10% of the sample). Hence, most respondents should be willing to pay mark-ups for Plan 2. The price difference between the two contracts varies randomly which facilitates the identification of the WTP across health plans. The data set further includes detailed background characteristics, including information on financial and health insurance literacy as well as risk attitudes, typically rarely available in surveys containing health insurance choices.

To investigate the heterogeneity in the WTP across health plans, I implement and use nonparametric and semiparametric estimators proposed by Lewbel et al. (2011). These estimators facilitate the identification of features of the distribution of an unobserved random variable such as the latent WTP. Individuals are asked whether they are willing to pay a specified amount for a product while this amount varies randomly across them. Given the choice of the individuals, the proposed value and covariates, features of the distribution of the unobserved variable can be estimated. This technique stands in contrast to traditional approaches that make parametric assumptions for the functional form of the WTP and the distribution of the errors terms and estimate the model via maximum likelihood (Crooker and Herriges 2004; Lewbel et al. 2011). These earlier approaches were criticized to be potentially sensitive to model specifications (Crooker and Herriges 2004). Moreover, empirical studies using parametric approaches often concentrate on homogeneous WTP estimates by plan attribute that are derived trading off the attribute with necessary price changes (see, for instance, Kesternich et al. 2013; van den Berg et al. 2008)

I find that about half of the respondents choose Plan 2 and that their WTP does not differ much between the contracts. Considering typical health expenditures for this population, this finding cannot be explained completely by health and related expenditure risk considerations alone. Using nonparametric estimators, I explore how preferences across plans correlate with background characteristics. In a further step, I investigate using semiparametric estimators how results change controlling for a rich set of background characteristics. Both types of estimators identify meaningful patterns in the data. My nonparametric analysis indicates a clear increasing, almost linear trend in the preference for Plan 2 with higher levels of financial and health insurance literacy. This is confirmed in the semiparametric analysis. I find that those with high levels of financial and health insurance literacy have a significantly higher WTP for Plan 2 of about \$18 compared with Plan 1, even after controlling for education and income. This higher WTP translates into potential price mark-ups of 2-5 percentage points depending on the age-related premiums. In contrast to prior research, I observe a less clear picture for risk aversion in the nonparametric analysis and do not find significant differences in the WTP depending on risk aversion in the semiparametric analysis.

This paper makes two important contributions to the literature: First, I show that the estimators proposed by Lewbel et al. (2011) provide easy to use tools to investigate features of the distribution of unobserved random variables such as the WTP. These estimators can be useful in other contexts as well including, for instance, bioessays, medical dose-response studies and materials testing (Lewbel et al. 2011). Second, I show that financial and health insurance literacy facilitates the optimization of complex choices. These findings have important policy implications and give reason for concern that the designed market environment under the ACA might not match the capabilities of the population.

The remainder of the paper is structured as follows. Section 2.2 provides an overview on the related literature, Section 2.3 presents the data and survey design. In Section 2.4, I explain the key ideas of the implemented nonparametric and semiparametric estimators. Section 2.5 presents the empirical results and Section 2.6 concludes.

2.2 Related literature

This section reviews the literature in several research areas related to this study: First, I discuss the role of hypothetical choices for investigating health insurance decisions. Second, I relate to the literature on selection and barriers that can affect the underlying preferences for specific health plans.

Hypothetical health insurance choices. Individual preferences across different products can be inferred either from their actual choices ("revealed preferences") or using "stated preference" methods (see Kesternich et al. 2013, for a detailed discussion). Actual choices in real markets have the advantage to be consequential to the individual by affecting their budgets. However, there are important limitations such a lack of variation in the attributes of products or being only able to observe choices for certain individuals. Therefore, individuals are asked to state their preferences in an environment where the researcher can control the presented alternatives. Moreover, the sample design can take the interest in specific subpopulations into account (e.g. the uninsured). While direct stated preference methods (such as open-ended questions) have been favored in earlier research (see Green et al. 1998, for an overview on the historic development), these methods were crowded out by choice-based methods and were especially considered inappropriate in the insurance context with respect to complex product designs and infrequent purchases (Braun et al. 2016). Selecting between different options is considered less cognitively challenging (Braun et al. 2016) and has been shown to lead to less bias than open-ended WTP elicitation methods (Murphy et al. 2005). These so called indirect stated preference methods include hypothetical choice experiments in which respondents have to choose among alternatives containing different bundles of characteristics. Prior research comparing hypothetical health insurance choices and real decisions in the context of Medicare Part D has shown that hypothetical experiments provide a useful tool to predict actual health insurance demand (Kesternich et al. 2013). However, the underlying preferences across products are unobservable in this process and need to be inferred.

Selection and barriers. The WTP captures information about the individual's underlying preferences for a specific health plan. In a perfect world, health status and associated future medical cost would be perfectly observable to the individual. The rational decision maker compares easily understandable health plans and picks the plan that covers his needs while minimizing cost. Those with higher expected health expenditures are likely to purchase more insurance, leading to adverse selection if there are information asymmetries between insurer and beneficiary regarding the risk type (Rothschild and Stiglitz 1976). However, in the real world the individual is subject to certain barriers (among others, discussed by Baicker et al. 2012; Liebmann and Zeckhauser 2008). First, the individual faces a complex choice environment. Prior research has show that the ability to understand health insurance concepts – a prerequisite for evaluating different options – varies widely between individuals (Barcellos et al. 2014; Cunningham et al. 2001; Loewenstein et al. 2013; Long et al. 2014b). It can constitute a barrier to even choosing a health insurance plan at all. For instance, Heiss et al. (2006) show in the context with the introduction of Medicare Part D that some individuals, contrary to their immediate self-interest, fail to enroll in insurance. Moreover and most closely related to this study, Bhargava et al. (2015) show that the ability to understand and apply basic health insurance concepts (such as deductibles or coinsurance rates) plays a significant role for choice quality – in their scenario translating into a better ability to detect dominated plans. Second, the individual faces uncertainty regarding his future health status. He does not know the exact probabilities of having adverse health events in the future. In addition to making systematic errors in assessing probabilities such as placing too much weight on small probabilities and to little on high probabilities (Kahneman and Tversky 1979), he does in general not like to think about adverse health events. Moreover, the individual does not know the size of expenditures once he becomes sick (Liebmann and Zeckhauser 2008). Research has shown that hypothetical insurance decisions are sensitive to perceptions of risk (Johnson et al. 1993) and that the degree of coverage is likely to be influenced by the degree of the individual's risk aversion (Cutler et al. 2008; de Meza and Webb 2001; Finkelstein and McGarry 2006).

2.3 Data and survey design

In this section, I present the data for the empirical analysis. I first describe the American Life Panel (ALP) and then the hypothetical health insurance experiment. Furthermore, I discuss considerations with respect to the design of the experiment.

American Life Panel. This paper is based on a health insurance experiment in the RAND American Life Panel (ALP) which was conducted in the context of a number of surveys on the ACA.[2] The ALP is a nationally representative internet panel that covers individuals aged 18+

[2]Further information is available at https://alpdata.rand.org/index.php?page=data.

who agreed to participate in occasional online surveys. Participants without internet access are provided with computers and internet, eliminating the potential bias arising if this part of the population is excluded from the analysis. My analysis is based on $2,513$ participants aged 18 to 64 who participate in the experiment and have non-missing information in all relevant dimensions.[3] The survey containing the experiment was conducted from September 20, 2013 until March 5, 2014. I add measures on health and health insurance literacy that were collected in a separate survey between August 23, 2013 and February 14, 2014.

I focus primarily on literacy and risk attitudes as barriers to choices across insurance plans. The data contain measures of financial and health insurance literacy as well as risk attitudes which are typically rarely available in surveys containing health insurance choices. I observe the participants' financial literacy which is based on standard questions testing fundamental conceps of economics and finance including numeracy (ability to compound interest rates) as well as the ability to understand the concepts of inflation (changes in value of money) and risk (difference between single stock and mutual fund).[4] Prior research has shown that financial literacy is relevant for financial decisions such as borrowing and retirement planning (Lusardi and de Bassa Scheresberg 2013; Lusardi and Mitchell 2011a). Financially literate individuals should be more motivated and more capable to optimize decisions that imply financial consequences. Health insurance decisions are relevant in this respect helping individuals to manage expenditure risk from medical care. Furthermore, I observe the participants' health insurance literacy which is defined as the ability to understand key health insurance concepts.[5] Health insurance literacy (HI literacy) in the ALP is measured collecting information on the participants' ability to understand deductibles, co-pays, coinsurance, networks, and (generic vs. brand) prescription drug pricing (see Barcellos et al. 2014, for further information). These questions reflect the individual knowledge of health insurance concepts, which are the factors that distinguish health insurance plans, and may be particularly important for making sensible health insurance decisions when having to choose between many options in the Marketplaces. Individuals with low financial or health insurance literacy may have difficulties to understand the alternative health plans or to determine their optimal choices. We have shown in prior joint work (refer to Chapter 3) that both measures are predictive of having health insurance and being covered via Medicaid or the Marketplaces compared to being uninsured. For the semiparametric analysis, I construct a single indicator and classify individuals to have

[3]The survey also included individuals that did not answer the hypothetical choice questions. There are no significant differences between this control group and my sample in terms of background characteristics and how pleasant they find the interview.

[4]For an overview, see Lusardi and Mitchell (2014).

[5]The Health Insurance Literacy Expert Roundtable Report defines health insurance literacy as the "degree to which individuals have the knowledge, ability, and confidence to find and evaluate information about health plans, select the best plan for their own (or their family's) financial and health circumstances, and use the plan once enrolled (Quincy 2012).

high literacy if they answer all financial literacy questions (following Lusardi and Mitchell 2011a) and at least 5 out of 7 health insurance literacy questions correctly. This presents a higher bar (compared with prior work, refer to Chapter 3) and shall identify individuals of the general population with confidence making health insurance decisions considering that they know most of the key health insurance concepts.

I further observe the participants' risk attitudes through their stated willingness to take risks (collected on scale from 0 for 'not at all willing to take risks' to 10 for 'very willing to take risks'). Prior research has shown the behavioral relevance of this measure using paid lottery choices (Dohmen et al. 2011). Risk averse individuals should be interested in demanding more coverage (Cutler et al. 2008; de Meza and Webb 2001; Finkelstein and McGarry 2006). I reverse the scale so that a higher value indicates a higher degree of risk aversion. For the semiparametric analysis, I classify individuals to have high risk aversion, in comparison to low or medium risk aversion, if they indicate a value of 7 or more. Appendix 2.D provides an overview on the questions in the ALP.

Moreover, the data include rich background characteristics containing gender, age, race and ethnicity, education, family income, household size, employment status, health insurance type and health. Table 2.1 displays the descriptive statistics for the analytic sample in comparison to the March 2013 Current Population Survey (CPS), a national representative survey conducted by the US Census Bureau (King et al. 2015). The statistics are weighted using ALP sample weights. The statistics in my estimation sample track the distribution of key covariates in the CPS sample, including variables that were not used to construct the statistical weights.[6]

Table 2.1: Descriptive statistics

Characteristics	**ALP** [N=2,513]	**CPS** [N=122,316]
Literacy		
Low financial and HI literacy	61%	*a*
High financial and HI literacy	39%	*a*
Risk attitudes		
Low and medium risk aversion	61%	*a*
High risk aversion	39%	*a*
Gender		
Male	50%	49%
Female	50%	51%
Age		
18-25	16%	18%
26-34	19%	19%
35-44	21%	21%
45-54	22%	22%
55-64	22%	20%

Continued on next page

[6]Sample weights are based on variables gender, age, ethnicity, education, income, and householde size.

Characteristics	**ALP** [N=2,513]	**CPS** [N=122,316]
Race and ethnicity		
Non-Hispanic White	63%	63%
Non-Hispanic Non-White	15%	20%
Hispanic	22%	17%
Education		
High school and less	41%	39%
Some college	19%	21%
Completed college degree	40%	40%
Family income		
Medicaid eligible ($\leq$ 138% of FPL)	23%	23%
Subsidy eligible (138-400% of FPL)	42%	40%
Medium income (400-700% of FPL)	23%	24%
Top income ($>$ 700% of FPL)	12%	14%
Household size		
Household $<$ 2 members	41%	43%
Household $\geq$ 2 members	59%	57%
Employment status		
Not working	30%	31%
Working	70%	69%
Health insurance type		
Uninsured	19%	21%
Medicaid	6%	9%
Employer-sponsored insurance	61%	60%
All other	13%	12%
Health		
Excellent/very good/good self-rated health	86%	89%
Fair/poor self-rated health	14%	11%
Low expenditures ($<$ \$100)	61%	[a]
Medium expenditures (\$100 $-$ 500)	28%	[a]
High expenditures ($\geq$ \$500)	11%	[a]

Data sources: Fall/Winter 2013 RAND American Life Panel and March 2013 Current Population Survey, individuals aged 18-64. ***Notes:*** Tables displays weighted averages for individuals younger than 65 with non-missing values in the relevant dimensions. [a] No comparable information available in CPS.

Hypothetical choice experiment. Figure 2.1 displays the hypothetical health insurance experiment that I use to investigate the heterogeneity in the WTP across health plans. Survey respondents have to choose between two contracts that differ in their level of coverage including different annual deductibles and annual out-of-pocket (OOP) maximums (henceforth, I omit annual for simplicity). The deductible refers to the amount the beneficiary has to pay before health insurance starts to cover any health expenditures while the OOP maximum refers to the maximum amount that the beneficiary has to pay during the insured period before health insurance starts to cover 100% of health expenditures. The plan with the lower deductible is associated with the higher OOP maximum. The first plan ("Plan 1") includes a deductible of \$1,500 and an OOP maximum of \$1,500. This implies that for each dollar spent on health up to \$1,500, the beneficiary has to pay the dollar himself. Above \$1,500, all health expenditures are covered by health insurance.[7]

[7]Technically speaking, the displayed doctor visit copay and generic medicine copay are irrelevant for this

Figure 2.1: Hypothetical health insurance experiment in the RAND ALP

Now imagine you only had a choice between the two following health insurance plans and you cannot choose to stay without coverage. Please check the plan that you would pick.

Monthly premium	Annual deductible	Annual out-of-pocket maximum	Doctor visit copay	Generic medicine copay	Your choice
$c	$1,500	$1,500	$30	$20	☐
$c +/- $v	$200	$3,500	$30	$20	☐

If first option is chosen. You chose the first plan. At what monthly premium for the second plan would you choose this plan instead of the first plan?

Monthly premium	Annual deductible	Annual out-of-pocket maximum	Doctor visit copay	Generic medicine copay
$c	$1,500	$1,500	$30	$20
$[ ]	$200	$3,500	$30	$20

If second option is chosen. You chose the second plan. At what monthly premium for the first plan would you choose this plan instead of the second plan?

Monthly premium	Annual deductible	Annual out-of-pocket maximum	Doctor visit copay	Generic medicine copay
$[ ]	$1,500	$1,500	$30	$20
$c +/- $v	$200	$3,500	$30	$20

Source: RAND American Life Panel. ***Notes:*** The monthly premium $c depends on age: Individuals aged 18-34 have to pay $350, aged 35-44 have to pay $450, aged 45-54 have to pay $700 and aged 55-64 have to pay $950. The bid values contain $v \in \{1, 2, ..., 200\}$.

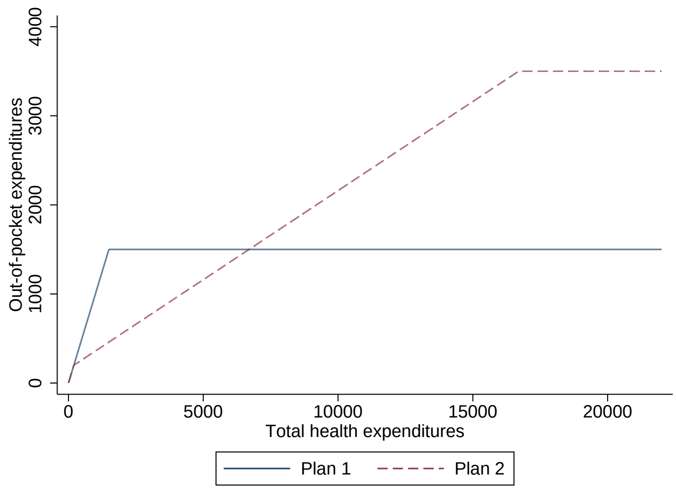
The second plan ("Plan 2") includes a deductible of $200 and an OOP maximum of $3,500. Up to $200 health expenditures, the beneficiary has to cover the cost himself. Between $200 and $3,500 OOP expenditures, the beneficiary has to cover a part of the cost (in this context, $30 per doctor visit and $20 per generic medicine prescription). Above $3,500 OOP expenditures, health insurance will cover 100% of the health expenditures. Figure 2.2 illustrates the differences between both health plans assuming that the fixed copays of Plan 2 translate into a coinsurance rate of 20%.[8] Up to $200 total health expenditures, both plans yield the same OOP cost. Between $200 and $6,500 total health expenditures, Plan 2 with the low deductible dominates Plan 1 with the high deductible by yielding lower OOP cost. Above $6,500, the OOP maximum of Plan 1 kicks in, and hence Plan 1 dominates Plan 2. When deciding which plan is more valuable for the individual, it depends both

type of contract. This contract shows similarities to "catastrophic health insurance plans" that usually have low(er) monthly premiums and high deductibles. These plans shall protect the beneficiary from worst-case scenarios while routine medical expenditures are paid by the beneficiary.

[8] A coinsurance of 20% is a common rate. Other designs include coinsurance rates of 10% or 30%. The 2014 Employer Health Benefits Survey indicates that the average coinsurance for office visits for covered workers with coinsurance is 18% for primary care and 19% for specialty care (KFF 2014a).

on the expectations regarding total health expenditures as well as the individual ability to understand the architecture of these health plans and the individual risk aversion.

Figure 2.2: Illustration of health insurance plans



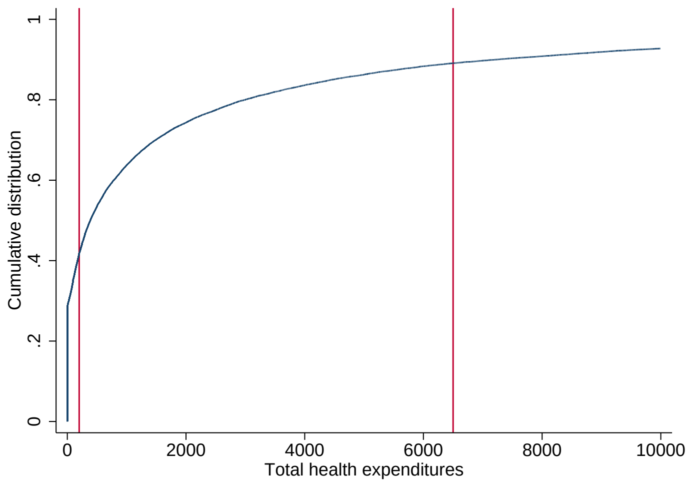
Notes: Own illustration of health plans in hypothetical health insurance experiment contained in the RAND American Life Panel assuming a coinsurance rate of 20%.

In the experiment, the premium c for Plan 1 differs between individuals depending on age with age groups 18-34, 35-44, 45-54, and 55-64. The premium for Plan 2 depends on the premium for Plan 1 but differs by a specified amount $v \in \{1, ..., 199, 200\}$ that varies randomly and can be both positive or negative. Once the respondents have chosen their preferred alternative, they have to indicate at which value they would have chosen the other plan. Furthermore, individuals can check the definitions of “monthly premium”, “annual deductible”, “annual out-of-pocket maximum”, “copay” and “generic medicine” via hyperlinks.[9]

Considerations. In how far can individuals expect to fall into the different ranges of total health expenditures? I obtained individual level data from the 2012 Medical Expenditure Panel Survey (MEPS), a representative sample of the civilian non-institutionalized population. Figure 2.3 displays the cumulative distribution of the total health expenditures for the population aged 18-64. The median annual health expenditures are about $787, the mean about $4,090. Moreover, about 42% of this population can expect to have expenditures below $200, almost 90% to have expenditures below $6,500. Hence, Plan 2 with low deductible and higher OOP maximum should yield lower total OOP expenditures for the majority which should be valuable for them – and should translate into a higher WTP for Plan 2.

[9]The health insurance literacy measures are collected in a separate survey.

Figure 2.3: Illustration of typical annual health expenditures



Data source: 2012 Medical Expenditure Panel Survey, civilian non-institutionalized population, aged 18-64. ***Notes:*** Figure displays the cumulative distribution of total health expenditures ($N = 23,653$).

2.4 Empirical framework

This section presents the key ideas of the nonparametric and semiparametric estimators used for the empirical analysis. Furthermore, I discuss the general applicability of these methods to the health insurance experiment contained in the ALP. More details on the estimators are presented in Appendix 2.A.

Econometric methods. Being interested in the heterogeneity of the unobserved differential WTP, I use both nonparametric and semiparametric estimators proposed by Lewbel et al. (2011) that estimate, conditional on observables, moments of the distribution of an unobserved random variable.[10] The estimators are presented for the case when the researcher can only observe a bid value V (but not the latent WTP W) and the individual response $Y = \mathbb{1}(W > V)$ with $Y = 1$ if the WTP is greater than the proposed value, and $Y = 0$ otherwise. In addition, the researcher can observe the covariates X (which can contain discrete and continuous variables). The estimators shall provide features of the distribution of W across individuals conditional on covariates $\mu_r(x) = E[r(W, X)|X = x]$ – in my case the conditional expected WTP $E(W|X = x)$. The conditional moments could be calculated with an estimate of the conditional density of the latent WTP $g(w|x)$: $\mu(x) = \int wg(w|x)dw$. However, the researcher can only observe the conditional survival curve

[10] The authors discuss their proposed estimator in the context of a contingent valuation study using referendum format elicitation. Further examples include bioessays, dose-response studies and materials testing.

of bid values $G(v|x) = Pr(W > v|X = x) = E[Y|V = v, X = x]$. Therefore, Lewbel et al. (2011) develop both nonparametric $[\hat{\mu}_1(x)]$ and semiparametric $[\hat{\mu}_3(x), \hat{\mu}_4(x)]$ estimators that utilize an estimate of $G(v|x)$ to identify the conditional moments of the latent variable. The key idea of the nonparametric estimators is to use kernel density estimation to obtain an estimate on the likelihood that individuals with similar characteristics accept the proposed bid $G(v|x)$. In contrast, the semiparametric estimators impose more structure and assume a functional form of the WTP where an initial estimate of the WTP is used in a multivariate regression to obtain coefficient estimates. The predicted WTP is thereafter corrected for deviators to obtain the conditional moments. They propose estimators for cases when the distribution of the bid values is either known $[\hat{\mu}_1(x), \hat{\mu}_3(x)]$ or unknown $[\hat{\mu}_4(x)]$. I implement these estimators in Stata. Appendix 2.B presents my replication of the authors' empirical application estimating the WTP for protecting wetland habitats and wildlife in California's San Joaquin Valley.

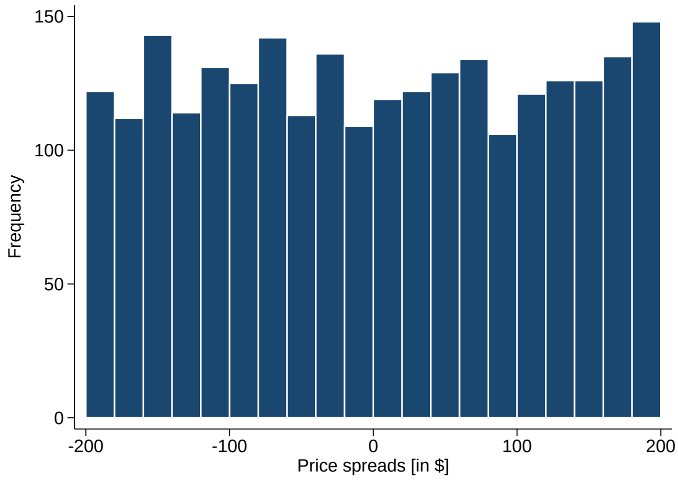
Applicability. Analyzing the experiment contained in the ALP, I am interested in preferences across health plans which are expressed in differences in the individual willingness to pay (W) for either health plan. However, I can only observe which contract an individual is willing to chose (Y) under the presented price spread (V). Hence, the experimental setting corresponds to the situation considered by Lewbel et al. (2011). The authors further discuss a set of experimental prerequisites: (1) The bid values should have at least a fair number of mass points – ideally being continuous. I have 400 mass points with positive and negative price spreads of size $v \in \{1, ..., 199, 200\}$. (2) The bid values V should be conditionally independent of the WTP W. In the experiment, the bid values are assigned randomly. (3) The support of the bid values should contain the support of the WTP which should be compact. In the experiment, respondents have to indicate in a second step at which value they would have chosen the other plan. Investigating these choices, I observe that this requirement might be problematic and account for this in additional robustness checks.[11] (4) Conditional on observable covariates, the bid values should also have a strictly positive density. This holds by construction assuming the distribution of bid values h to be uniform over the range of bid values (and considering random assignment of V).

2.5 Empirical results

In the hypothetical health insurance experiment in the ALP, I define the bid values V as the amount the premium for Plan 2 differs from the premium for Plan 1. I further define the

[11] Only about 58% of the sample indicate to be indifferent between both plans within a price spread of \$200. I perform additional robustness checks based on a restricted sample.

Figure 2.4: Distribution of price spreads



Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Figure displays distribution of bid price spreads between Plan 1 and Plan 2 in the hypothetical health insurance experiment contained in the RAND American Life Panel ($N = 2,513$).

individual choice Y in terms of Plan 2: Choosing Plan 1 is coded as "rejection of Plan 2" ($Y = 0$), choosing Plan 2 as "acceptance of Plan 2" ($Y = 1$). Regarding the first choice, 48.69% choose Plan 2. Figure 2.4 illustrates the distribution of bid values which approximately follows a uniform distribution over the range of bid values. The average proposed bid value is $\$ - 1.28$. Most individuals choose the option for which they have a higher valuation. They request (sometimes substantial) discounts to switch to the alternative option (refer to Appendix Figure 2.10). In what follows, I focus on the first choice.

I use nonparametric estimators to explore how preferences across plans correlate with background characteristics focusing especially on literacy and risk attitudes. In a second step, I investigate using semiparametric estimators how results change controlling for a rich set of background characteristics. Considering the distribution of bid values, I assume a uniform distribution for those estimators that require a known density of bid values ($\hat{\mu}_1, \hat{\mu}_3$). For the nonparametric estimator and the semiparametric estimator with unknown density of bid values ($\hat{\mu}_4$), I follow Lewbel et al. (2011) and use product Gaussian kernels and Silverman's rule of thumb bandwidths with $\lambda = 1.06sn^{-1/5}$, where s refers to the sample standard deviation and n to the sample size. I further choose a linear parameterization of the WTP for the semiparametric estimators: $W = X_i^T\theta - \epsilon$. With this specification, I investigate the conditional average differential WTP across plans $\mu(x) = E(W|X = x)$ using two different semiparametric estimators. This allows to investigate preferences in favor of the respective plans for specific subpopulations. For instance, it is possible to investigate whether individuals with high financial and health insurance literacy favor Plan 2 which would be expressed in a higher WTP for Plan 2 compared with Plan 1.

2.5.1 Nonparametric results

Figure 2.5 displays the nonparametric results for the heterogeneity in WTP across health plans for the key covariates using estimator $\hat{\mu}_1$ and looking at the continuous covariates, one at a time. The estimates are displayed along with their pointwise 95% paired bootstrap confidence interval.[12] The displayed estimates refer to the price difference between Plan 1 and Plan 2. A positive value indicates a higher valuation of Plan 2 while a negative value indicates a higher valuation of Plan 1.

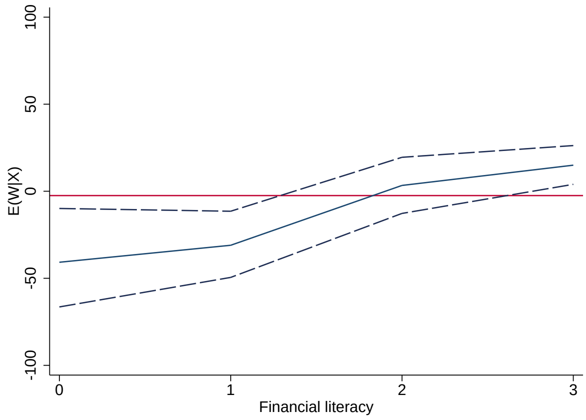
I find a meaningful heterogeneity in the WTP across health plans by financial and health insurance literacy. Both measures show an upward sloping, almost linear association between literacy and preferences across plans. While those with low levels of financial and health insurance literacy indicate a significant preference for Plan 1, those with higher levels are willing to pay mark-ups to obtain Plan 2. The findings for the impact of risk attitudes are less conclusive. Prior research has stressed that the highly risk averse are expected to demand more health insurance. This could have pointed towards preferences for Plan 1 which shields beneficiaries from extraordinary high expenditures. However, I observe a less clear and insignificant, almost hump-shaped relationship between risk aversion and preferences across plans.

Figure 2.9 in Appendix 2.C shows the nonparametric results for further background characteristics. For those with lower levels of self-rated health and those with high medical expenditures ($\geq$ \$500), I find somewhat stronger preferences towards Plan 1 suggesting some degree of adverse selection. The Medicaid eligible have a significantly stronger preference for Plan 1 which reduces catastrophic risks while those with the highest incomes have somewhat stronger preferences for Plan 2. I observe a similar pattern and find increasing preferences towards Plan 2 with higher levels of education. To account for potential correlations between my key covariates and these background factors, I investigate the multivariate associations in the next subsection.

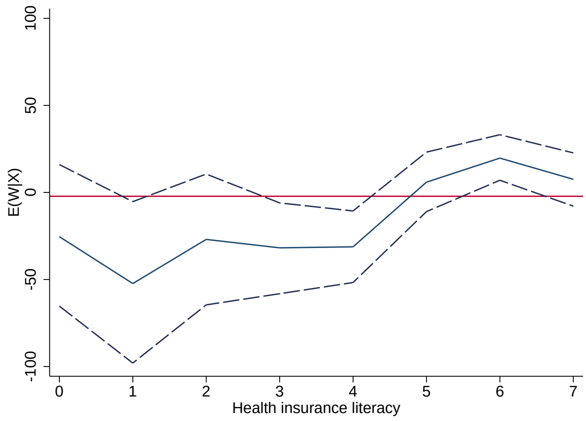
[12] The confidence intervals are based on bootstraps using 1,000 iterations. Results are similar when using 200 iterations.

Figure 2.5: Nonparametric results: Heterogeneity in WTP across health plans by knowledge and risk attitudes

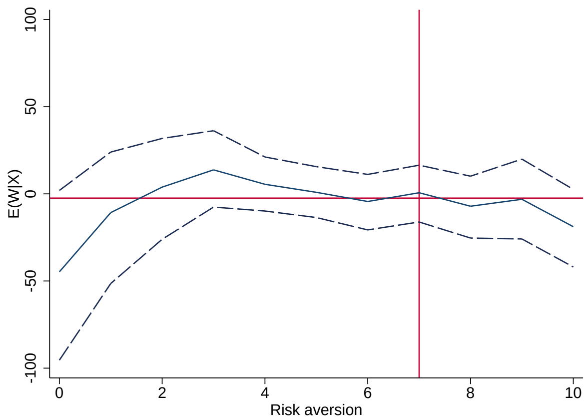
(a) Financial literacy



(b) Health insurance literacy



(c) Risk attitudes



Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Figure displays marginal smooths $\hat{\mu}_1(X_i)$ along with their pointwise 95% paired bootstrap confidence intervals (1,000 iterations) and estimated unconditional mean ($N = 2,513$).

2.5.2 Semiparametric results

Table 2.2 displays the parameter estimates for the linear specification of the WTP using 90% paired bootstrap confidence intervals.[13] In addition, I report sample averages of the conditional mean WTP $\hat{\mu}_j$ with $j = 3, 4$ along with their respective paired bootstrap confidence intervals. Using categorical variables, the parameters can be interpreted as the average price mark-up for Plan 2 in comparison to Plan 1 that the underlying linear model predicts for an individual in the respective category in comparison to the price mark-up of an individual in the base category. Without calculating the average price mark-up of an individual in the base category, it is however not possible to determine whether individuals in the respective category have a higher valuation of Plan 1 or Plan 2.[14] Hence, I present sample averages of the conditional mean WTP for certain subpopulations focusing on my main key covariates in Table 2.3.

Table 2.2: Semiparametric results: Heterogeneity in the WTP across health plans

	90% CI	90% CI	90% CI	90% CI
	(1)	(2)	(3)	(4)
Literacy				
High literacy	34.57 [20.91; 48.24]*		34.30 [20.05; 48.00]*	17.83 [3.20; 34.26]*
Risk attitudes				
High risk aversion		-6.70 [-19.57; 7.20]	-4.78 [-18.79; 8.33]	-6.95 [-20.06; 7.41]
Covariates				X
$\bar{\hat{\mu}}_3$	-4.32 [-11.55; 2.51]	-3.61 [-11.04; 4.10]	-5.19 [-11.85; 2.67]	-5.79 [-12.27; 1.63]
$\bar{\hat{\mu}}_4$	-2.01 [-8.82; 4.11]	-2.84 [-9.9; 4.56]	-2.86 [-9.28; 4.96]	2.38 [-6.23; 13.23]
N	2,513	2,513	2,513	2,513

Data sources: RAND American Life Panel (ALP), individuals aged 18-64. ***Notes:*** Estimates are presented along with their 90% paired bootstrap confidence intervals (1,000 iterations, $N = 2,513$). The star $*$ marks significance.

The semiparametric analysis allows to account for background factors that might be correlated with levels of literacy and risk attitudes. Comparing column (1) which shows raw correlations and column (4) including further covariates of Table 2.2, the indicated mark-up for Plan 2 of \$35 (that those with high levels of literacy are willing to pay more in comparison to those with low levels of literacy) reduces almost by half to about \$18. At the same time, I find that those factors likely to be correlated with literacy play a role for the choice across plans. For instance, those with higher levels of education and those with health insurance via

[13] The confidence intervals are based on bootstraps using 1,000 iterations. Results are similar using 200 iterations.

[14] For example, all else equal, Hispanics are willing to pay about \$33 less for Plan 2 in comparison to Plan 1 than Non-Hispanic Whites. If Non-Hispanic Whites would have a valuation for Plan 2 which is about \$50 lower than for Plan 1, Hispanics would have a WTP for the Plan 2 which is \$83 lower than for Plan 1. In contrast, if Non-Hispanic Whites would have a valuation of Plan 2 which is \$50 higher than for Plan 1, Hispanics would have a WTP for Plan 2 which is \$17 higher than for Plan 1. In the first case Hispanics would have a preference for Plan 1, in the second for Plan 2.

ESI (compared with the uninsured) indicate higher mark-ups for Plan 2 (refer to Table 2.7 in Appendix 2.C displaying the full set of covariates).[15] Hence, the importance of literacy for choices across health plans is found over and above these other factors which also stresses the importance of literacy in financial and insurance matters in addition to a general importance of, for instance, cognitive abilities and experience with health insurance decisions. The average predicted conditional mean WTP for individuals with low financial and health insurance literacy is negative indicating preferences towards Plan 1 while the opposite is true for those with high financial and health insurance literacy indicating preferences towards Plan 2 (refer to Table 2.3).

Even though not significant, those with high in comparison with those with low and medium risk aversion have a lower valuation of Plan 2 which is also observable comparing the average predicted conditional mean WTP for different levels of risk aversion (refer to Table 2.3). However, it is not clear in how far there are differences in plan choice for those with high compared to those with low and medium levels of risk aversion.

Table 2.3: Sample averages of conditional mean WTP by literacy and risk attitudes

	$\hat{\mu_3}(x)$	$\hat{\mu_4}(x)$	**N**
	(1)	**(2)**	**(3)**
Literacy			
Low financial and HI literacy	-19.41***	-11.23***	1,523
High financial and HI literacy	15.16***	34.33***	990
Risk attitudes			
Low and medium risk aversion	-3.13***	5.04***	1,516
High risk aversion	-9.83***	-1.66*	997

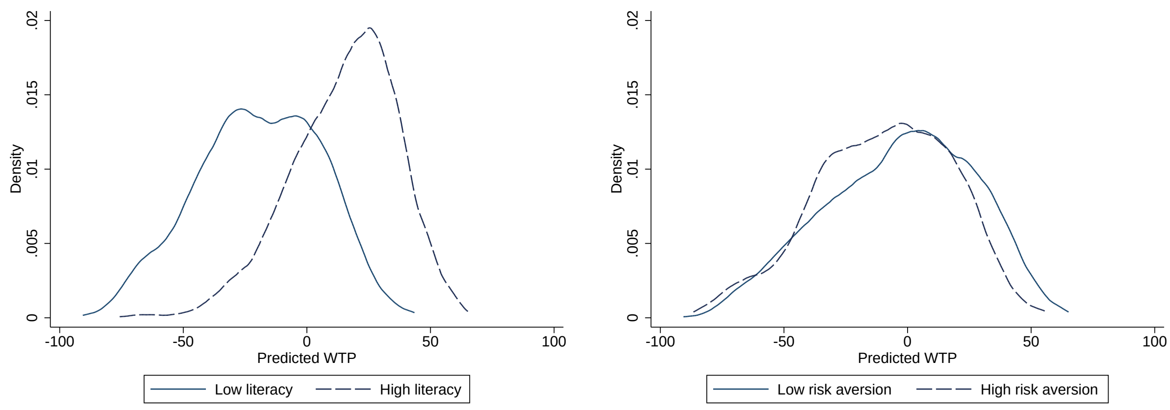
Data sources: Fall 2013 RAND American Life Panel (ALP), individuals aged 18-64. ***Notes:*** Tables displays sample averages of the conditional mean WTP using different semiparametric estimators. Stars indicate statistical difference from zero using one-sample t-tests. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

In Figure 2.6, I show kernel densities of the predicted conditional mean WTP across health plans by level of literacy and risk attitudes. Figure 2.6a further illustrates the discussed strong pattern for the influence of financial and health insurance literacy. While the majority of those with low literacy prefer Plan 1, the majority of those with high literacy are willing to pay mark-ups to obtain Plan 2. The findings are less conclusive investigating the densities by degree of risk aversion (Figure 2.6b).

[15] This could be related to better cognitive abilities in general and experience with choosing between health plans and using the health care system.

Figure 2.6: Density of conditional mean WTP by knowledge and risk attitudes

(a) Literacy **(b)** Risk attitudes



Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Figure displays kernel densities of the conditional mean WTP $\hat{\mu}_3$ by literacy and risk attitudes ($N = 2,513$).

All in all, the semiparametric multivariate analysis confirms the previous nonparametric univariate findings for my main variables of interest: In contrast to those with low literacy who favor Plan 1, those understanding all financial and most health insurance concepts are willing to pay mark-ups to obtain Plan 2 potentially leading to lower OOP cost. The results for the importance of risk attitudes are less conclusive.

These findings raise further questions: First, in how far does experience with health insurance play a role? More specifically, are there differences in preferences across health plans with respect to literacy between those insured and those uninsured? Second, which of the different concepts in the measures of financial and health insurance literacy are important? To address the first question, I calculate the bivariate sample averages of the predicted conditional mean WTP by literacy and insurance status (refer to Table 2.4). I find higher valuations of Plan 2 both for those with high literacy and those insured (differences significant at the 1% and 10% level, respectively). For the uninsured, this increase is however not sufficient to shift respondents towards preferences for Plan 2. While those uninsured with low literacy strongly favor Plan 1, those uninsured with high literacy are indifferent between both plans. For the insured, the pattern is more consistent with the prior findings indicating preferences for Plan 1 with low levels and for Plan 2 with high levels of literacy. My findings have two implications: First, both experience with the health system and health insurance literacy help individuals with their decisions. Second, even for those with experience, financial and health insurance literacy has a strong impact on preferences across health plans.

Table 2.4: Sample averages of conditional mean WTP: Literacy vs. insurance status

	Insurance status		
	Uninsured	Insured	Bootstrapped difference (p-value)
Literacy			
Low financial and HI literacy	-34.76***	-14.58***	0.068
	[N=364]	[N=1,159]	
High financial and HI literacy	-1.86	17.16***	0.092
	[N=104]	[N=886]	
Bootstrapped difference (p-value)	0.000	0.000	

Data sources: Fall 2013 RAND American Life Panel (ALP), individuals aged 18-64. ***Notes:*** Tables displays sample averages of the conditional mean WTP using the semi-parametric estimator $\hat{\mu_3}$ by literacy and insurance status. Stars indicate statistical difference from zero using one-sample t-tests. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

To address the second question, I investigate the importance of understanding the different financial and health insurance concepts. In Table 2.5, I present the estimates along with their 90% paired bootstrap confidence intervals. In column 1, 3 and 5, I include the individual literacy measures without additional covariates which are added in columns 2, 4 and 6. In terms of the size of coefficients and the level of significance, several concepts stand out: (1) The ability to understand that the visit of a doctor who is not part of the insurer's network will cost more. (2) The concept of a deductible (not significant when adding covariates). (3) Understanding the concept of risk diversification. Those participants knowing these concepts have significantly stronger preferences towards Plan 2. If I include all measures at the same time and additional controls, the coefficients of the first and the latter concept remain large and significant. Interestingly, both concepts relate to rather specific knowledge. Understanding the cost structure of insurer networks implies that the individual has already put some effort into how to optimize the utilization of health insurance. At the same time, understanding the concept of risk diversification shows that the individual has spent time on managing his savings.[16]

[16] In contrast to the concepts of numeracy and inflation which are widely understood (85 and 73%, respectively), only 62% understand the concept of risk diversification (see Barcellos et al. 2014).

Table 2.5: Semiparametric results: Importance of literacy measures

	90% CI **(1)**	**90% CI** **(2)**	**90% CI** **(3)**	**90% CI** **(4)**	**90% CI** **(5)**	**90% CI** **(6)**
HI literacy						
Deductible/Premium	2.65 [-14.59; 19.43]	-5.26 [-21.84; 12.20]			-6.61 [-23.98; 12.16]	-9.65 [-27.84; 8.78]
Networks	25.55 [6.52; 45.15]*	23.15 [2.76; 42.33]*			22.04 [2.90; 40.39]*	20.94 [1.52; 40.96]*
Generic/branded drugs	5.63 [-17.44; 30.86]	0.61 [-21.54; 23.35]			1.19 [-23.03; 25.21]	-1.68 [-25.38; 23.01]
HMO restrictions	7.10 [-6.88; 21.32]	3.02 [-11.73; 18.15]			3.97 [-9.87; 17.11]	1.99 [-11.72; 16.95]
Deductible description	19.07 [3.21; 34.80]*	12.47 [-3.06; 27.67]			15.38 [-0.08; 30.59]	10.87 [-4.58; 27.37]
Understands coinsurance	-7.84 [-33.69; 19.52]	-8.86 [-37.86; 19.16]			-12.10 [-41.03; 15.95]	-10.94 [-39.60; 16.22]
Understands co-pay	7.06 [-21.79; 33.89]	2.67 [-23.43; 29.08]			5.25 [-20.66; 32.52]	2.07 [-24.38; 27.24]
Financial literacy						
Numeracy			13.93 [-7.16; 34.27]	8.05 [-11.32; 28.13]	9.71 [-10.86; 31.83]	6.82 [-13.87; 27.71]
Inflation			10.19 [-6.70; 28.55]	5.33 [-12.98; 24.45]	5.66 [-13.24; 24.29]	4.53 [-14.78; 23.33]
Risk diversification			34.70 [18.89; 50.07]*	26.99 [11.13; 42.79]*	30.38 [15.14; 45.29]*	25.67 [9.45; 42.19]*
Covariates		X		X		X
$\bar{\hat{\mu}}_3$	-4.17 [-11.06; 2.71]	-4.16 [-12.86; 1.65]	-2.97 [-10.13; 3.00]	-4.90 [-11.56; 2.21]	-4.32 [-10.71; 2.89]	-5.72 [-11.22; 2.22]
$\bar{\hat{\mu}}_4$	1.13 [-6.26; 9.11]	6.49 [-5.54; 20.83]	-1.68 [-8.36; 4.73]	-0.00 [-6.86; 10.00]	-0.61 [-6.93; 7.97]	1.09 [-6.26; 13.70]
N	2,513	2,513	2,513	2,513	2,513	2,513

Data source: RAND American Life Panel (ALP), individuals aged 18-64. ***Notes:*** Estimates are presented along with their 90% paired bootstrap confidence intervals. *Covariates* contain further variables gender, age, race and ethnicity, education, family income, household size, employment status, health insurance type and health ($N = 2,513$). The star $*$ marks significance at the 10% level.

2.5.3 Robustness checks

I have discussed the prerequisite that the support of the bid values should contain the support of the latent WTP. This requirement might be violated considering the indicated price spreads in the second choice data. Therefore, I perform robustness checks excluding all individuals indicating a price spread for their indifference outside of the support of bid values. Figure 2.10 in Appendix 2.C displays the nonparametric estimation results excluding this part of the sample. The figure shows similar patterns as discussed before. Table 2.8 in Appendix 2.C displays the robustness check for the semiparametric estimation. I find some variability and significant negative sample averages of the estimates of the conditional mean WTP. Importantly however, the results for financial and health insurance literacy remain robust and significant.

2.6 Conclusion

I investigate heterogeneity in the WTP across health insurance plans using a hypothetical experiment contained in the representative RAND American Life Panel. In the experiment, survey respondents can choose between Plan 1, a catastrophic type plan that covers expenditures above \$1,500, and Plan 2, including a deductible of \$200 and copayments up to an OOP maximum of \$3,500. In an ideal world, the advantageousness of either plan depends on expected health expenditures. However, other factors such as literacy and risk attitudes can influence the individual perceived utility to choose either plan. Investigating health expenditures for a comparable representative sample, I show that about 90% could potentially save money by picking the low deductible, high OOP maximum option which should translate into a higher WTP for Plan 2.

Implementing and using nonparametric and semiparametric estimators proposed by Lewbel et al. (2011), I do not find significant differences in the WTP across plans which cannot be entirely explained by health expenditure considerations. Both types of estimators identify meaningful patterns in the data. I find a higher willingness to pay mark-ups for a plan with potentially lower OOP cost with higher levels of financial and health insurance literacy. This effect is robust and occurs over and above general educational attainment and income. Hence, my findings show the importance of financial and health insurance literacy in contrast to factors such as general cognitive abilities and affordability issues. While experience with the health care system appears to matter, I find that the degree of literacy has a strong impact on the insurance decisions of both the uninsured and the insured. Moreover, this pattern emerges despite of the possibility to check the required definitions such as deductibles, copays, etc. during the experiment. This decision aid might have been especially valuable for those with low levels of literacy. In spite of this, my findings suggest that those with low levels of literacy struggle to make sensible health insurance choices which might be even more problematic when support is not provided.

Furthermore, the catastrophe type plan should be especially valuable for those with high risk aversion by reducing the risk of extraordinary high health expenditures. In contrast to prior research indicating the importance of risk aversion for health insurance choices (see, for instance, Finkelstein and McGarry 2006), I do not find conclusive evidence for differences in the WTP across plans depending on risk aversion.

Other issues such as heuristics that simplify choices might play a role.[17] Previous research has found different forms of heuristics, such as choosing the cheapest plan available without evaluating the plan attributes (Ericson and Starc 2012) or picking the plan with the lowest deductible (Sydnor 2010). Despite of the fact that the plan with the lower deductible should yield lower future health expenditures for most individuals, only about half of the sample choose this alternative.[18] In contrast, about 73% choose the cheaper plan. However, this fraction is actually somewhat higher for those with high financial and health insurance literacy (78%) compared to those with low financial and health insurance literacy (70%). These findings suggest that my results are not only driven by the use of heuristics.

My study has some limitations. First, the hypothetical choice experiment is conducted in the context of collecting other information on health insurance status, type and knowledge of ACA features which could have changed the awareness of the importance of health insurance choices. Despite of this, my results imply a substantial degree of suboptimal behavior. Second, the choices are hypothetical and may differ from actual choices that affect the individual budget. However, prior research has shown that real and hypothetical choices match well (Kesternich et al. 2013). Third, the choices of the hypothetical choice experiment are limited to two health plans without the option to remain without health insurance. A sequential experiment including this option might provide additional insights. This is left to future research.

My findings have some important policy implications. Under the Affordable Care Act, individuals can compare and shop for health insurance via the Marketplaces. They typically face more than 40 different plans (Bhargava et al. 2015). Having to choose only between two plans, my results indicate that a better understanding of the contract terms facilitates identifying

[17] This should be especially relevant for those with a poor ability to understand the contract terms.
[18] The fraction is somewhat higher for those with high (52%) compared to those with low literacy (47%).

a potentially suitable health plan. Prior research has further indicated that choice quality should even decrease with a higher number of alternatives (Johnson et al. 2013; Schram and Sonnemans 2011). Policies promoting consumer choice in health insurance – considering also to increase competition to lower prices – should take this into account and simplify the choice environment. Moreover, my findings imply serious consequences for the low income population. Considering that literacy is correlated with income, my findings imply that the most vulnerable part of the population can leave money on the table when selecting their health insurance plans.

Finally, this paper shows that the estimators proposed by Lewbel et al. (2011) provide easy to use tools to investigate features of the distribution of an unobserved random variable. This paper presents a highly policy relevant application of these estimators. I can identify meaningful and robust patterns in the heterogeneity in the WTP across health insurance plans.

2.A Econometric methods

The general idea is to rewrite the conditional moment[19]

$$\begin{aligned}\mu(x) &= \int_{\rho_0(x)}^{\rho_1(x)} wg(w|x)dw \\ &= \rho_0(x) + \int_{\rho_0(x)}^{\rho_1(x)} 1\, G(v|x)dv \\ &= \kappa(x) + \int_{\rho_0(x)}^{\rho_1(x)} \left[G(v|x) - \mathbb{1}(v < \kappa(x)\right] dv \end{aligned} \tag{2.1}$$

where integration by parts yields the second line. In the third line, the lower bound is replaced by $\rho_0(x) < \kappa(x) < \rho_1(x)$, so any value contained in the support of the WTP, which can be thought of as integrating around a midpoint in the support of the WTP to obtain the conditional moment.

Nonparametric estimator. In the nonparametric case, kernel density estimation is used to estimate the survival curve of the bid values conditional on observables $G(v|x)$. This survival curve estimates the fraction of individuals with background characteristics similar to x accepting a specific bid value v. Lewbel et al. (2011) substitute $G(v|x)$ in equation 2.1, rewrite and obtain their proposed nonparametric estimator $\hat{\mu}_1$ (density of bid values is known). Being interested in how the average WTP depends on one continuous variable, I simplify as follows:

$$\hat{\mu}_1(x) = \kappa(x) + \frac{1}{\hat{C}(x)} \underbrace{E_n\left[\frac{Y - \mathbb{1}(V < \kappa(X))}{h(V|X)} \lambda^{-1} K\left(\frac{X-x}{\lambda}\right)\right]}_{\hat{I}(x)} \tag{2.2}$$

where $\hat{C}(x) = E_n\left[\lambda^{-1}K\left(\frac{X-x}{\lambda}\right)\right]$, $K(\cdot)$ is a symmetric continuously differentiable kernel function with compact support and λ denotes the bandwidth parameter. In this context, $\hat{I}(x)$ indicates whether the proposed latent WTP $\kappa(x)$ needs to be adjusted. If an individual behaves consistently with the proposed WTP, $Y - \mathbb{1}(V < \kappa(X))$ equals zero; it deviates to 1 if the individual accepts bid values larger than $\kappa(X)$ and to -1 if the individual does not accept bid values smaller than $\kappa(X)$. In a next step, the expected probability $\hat{I}(x)$ is calculated indicating whether and how individuals with characteristics similar to x deviate from the proposed latent WTP $\kappa(X)$, which is weighted by the conditional density of the proposed bid values $h(V|X)$. Finally, this proportion is set in relation to the general likelihood to have

[19] I simplify this equation being interested in the conditional mean WTP.

similar characteristics $\hat{C}(x)$ and the proposed latent WTP $\kappa(x)$ is adjusted up- or downwards accordingly.

Semiparametric estimators. Since nonparametric moments can be restricted with respect to the number of variables used, Lewbel et al. (2011) propose semiparametric estimators with W satisfying

$$W = \Lambda[m(X, \theta_0) - \epsilon] \tag{2.3}$$

where $m(.)$ and $\Lambda[.]$ are known functions, for instance, $W = X'\theta_0 - \epsilon$ or $log(W) = X'\theta_0 - \epsilon$ where Λ is an identity or exponential function (linear and log-linear specification) and m is linear in parameters. I concentrate on the linear specification.[20] The estimation of the conditional WTP contains several steps:

1. Each individual is assigned an initial estimate of the WTP $s(x, v, y)$ satisfying

$$s(x, v, y) = \kappa(x) + \frac{y - \mathbb{1}(v < \kappa(x))}{h(v|x)} \tag{2.4}$$

which proposes a WTP $\kappa(x)$ which is individually adjusted depending on whether the respective individual decides in a with the proposed value consistent way, weighted by the conditional density to observe bid value v for individuals with characteristics x. Hence, the initial estimate is corrected more for individuals deviating for less likely bid values.

2. This initial estimate of the WTP is used to obtain an estimate of θ minimizing the least squares criterion

$$(\hat{\theta}, \hat{\alpha}) = arg \min_{\theta,\alpha} \frac{1}{n} \sum_{i=1}^{n} \left[s(X_i, V_i, Y_i) - \alpha - X_i'\theta\right]^2 \tag{2.5}$$

where α is an arbitrary location constant.

3. In a third step, the researcher can obtain an estimate on the conditional WTP in $\hat{\mu}_3(x)$ and $\hat{\mu}_4(x)$, respectively:

$$\hat{\mu}_3(x) = x'\theta + \underbrace{\frac{1}{n} \sum_{i=1}^{n} \frac{Y_i - \mathbb{1}(U_i > 0)}{\hat{\phi}(U_i)}}_{\hat{J}} \tag{2.6}$$

with $U_i = X_i\theta - V_i$. Here, U_i is larger than 0 if the individual is assigned a WTP which is larger than the proposed bid V_i. In this case, the individual should accept the bid and $Y_i - \mathbb{1}(U_i > 0)$ should equal 0. It deviates to -1 if the individual does not accept

[20] Further information on the log-linear case are provided in the original paper by Lewbel et al. (2011).

bids although he should have and to 1 if he accepts bids although he should not have accepted them according to the assigned WTP. Thereafter, the assigned WTP $x'\theta$ is corrected for these deviators that are weighted by $\hat{\phi}(U_i)$, the parametric density of U_i (h known). Furthermore,

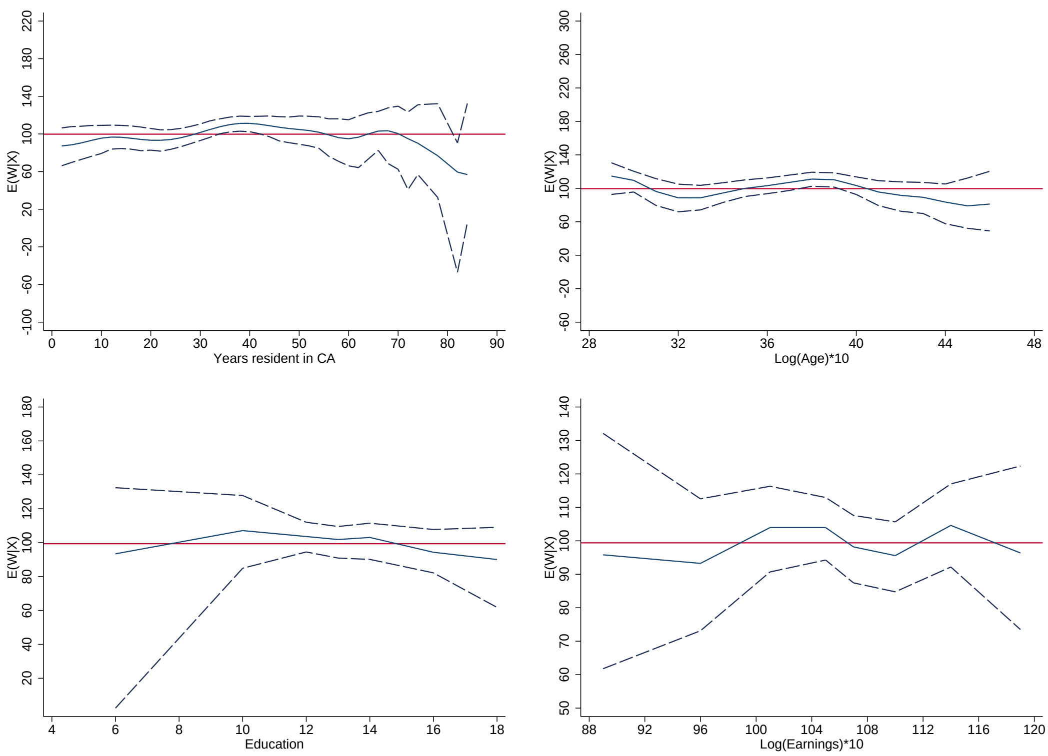
$$\hat{\mu}_4(x) = x'\theta + \underbrace{\frac{1}{n}\sum_{i=1}^{n}\frac{Y_i - \mathbb{1}(U_i > 0)}{\tilde{\phi}(U_i)}}_{\tilde{J}} \tag{2.7}$$

where $\tilde{\phi}(U_i)$ is the semiparametric density of U_i. In this context, $\hat{J}$ and $\tilde{J}$ present adjustments of the assigned WTP $x'\theta$ depending on how well the estimated WTPs correspond to the actual decisions in the experiment.

2.B Replication

Lewbel et al. (2011) provide data for their empirical application to estimate the WTP for protecting wetland habitats and wildlife in California's San Joaquin Valley as well as code in R and GAUSS for their semiparametric estimators (refer to http://www.oliver linton.me.uk/research/software). I implement their non- and semiparametric estimators $\hat{\mu}_1, \hat{\mu}_3$ and $\hat{\mu}_4$ in Stata. Figure 2.7 provides an overview of the replicated figures that are shown in Figure 2 in Lewbel et al. (2011), Table 2.6 an overview of the replicated estimates in comparison with estimates shown in Tables 7 and 8 in Lewbel et al. (2011). I replicate the results for the empirical application assuming both a linear $[W = X_i^T\theta - \epsilon]$ and log-linear $[log(W) = X_i^T\theta - \epsilon]$ specification of the WTP. My replication reproduces their findings almost perfectly.

Figure 2.7: Replication: Nonparametric estimates



Data sources: Data retrived from Oliver Linton's webpage (http://www.oliverlinton .me.uk/research/software) ***Notes:*** Figures show replication of Figure 2 in Lewbel et al. (2011). Estimates are presented along with their 95% paired bootstrap confidence intervals (200 iterations) which can deviate due to sampling differences.

Table 2.6: Replication: Semiparametric estimates

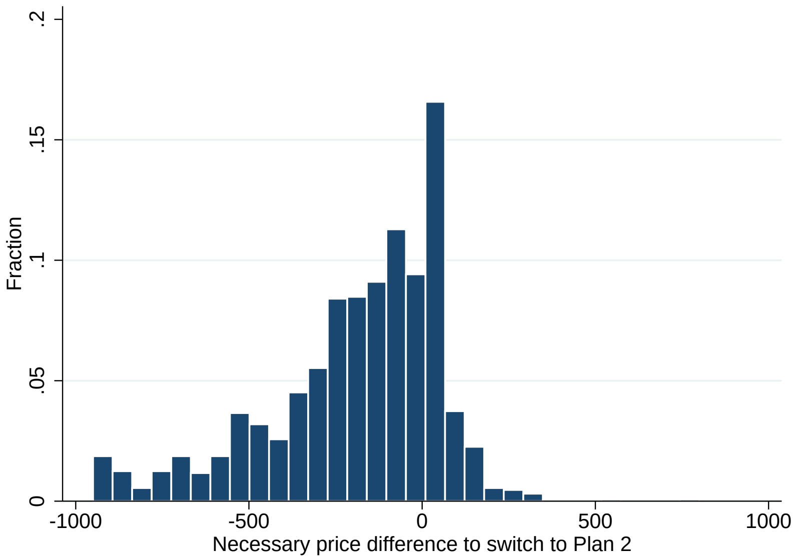
	Bid 1		Bid 2	
	Lewbel et al. (2011) (1)	**Replication** (2)	**Lewbel et al. (2011)** (3)	**Replication** (4)
Panel A: Linear specification				
Parameter estimates				
Years resident in CA	0.382 [-0.118; 0.935]	0.382 [-0.176; 0.985]	0.538 [-1.240; 1.970]	0.538 [-0.974; 2.331]
Female	0.556 [-12.540; 11.750]	0.556 [-12.865; 13.064]	28.290 [-10.500; 70.800]	28.290 [-16.831; 67.109]
log(Age)	-13.591 [-37.310; 8.780]	-13.591 [-35.725; 9.069]	-31.714 [-98.900; 32.700]	-31.713 [-95.964; 35.294]
Education	-2.024 [-4.980; 0.720]	-2.024 [-5.316; 1.087]	1.292 [-10.660; 10.100]	1.292 [-10.182; 10.671]
White	6.238 [-9.360; 26.070]	6.238 [-8.877; 23.593]	60.210 [3.000; 112.000]*	60.210 [7.082; 110.185]*
Membership envir. organization	1.968 [-15.070; 16.490]	1.968 [-15.673; 19.739]	34.860 [-23.900; 88.700]	34.860 [-25.448; 89.839]
log(Earnings)	2.378 [-9.700; 12.210]	2.378 [-8.143; 11.551]	40.414 [6.090; 68.930]*	40.414 [3.479; 74.349]*
Estimates of WTP				
$\bar{\hat{\mu}}_3$	110.480 [101.300; 126.000]*	110.480 [101.831; 123.425]*	172.838 [154.300; 202.300]*	172.838 [154.348; 198.175]*
$\bar{\hat{\mu}}_4$	105.611 [97.800; 118.100]*	105.611 [97.463; 116.906]*	246.059 [196.800; 294.500]*	246.059 [191.039; 311.432]*
Panel B: Log-linear specification				
Parameter estimates				
Years resident in CA	0.005 [-0.001; 0.010]	0.005 [-0.002; 0.011]	0.010 [-0.008; 0.030]	0.010 [-0.012; 0.034]
Female	0.011 [-0.140; 0.150]	0.011 [-0.147; 0.161]	0.503 [0.073; 0.980]*	0.503 [0.031; 0.982]*
log(Age)	-0.159 [-0.500; 0.120]	-0.159 [-0.439; 0.119]	-0.608 [-1.730; 0.330]	-0.608 [-1.583; 0.236]
Education	-0.027 [-0.067; 10.010]	-0.027 [-0.060; 0.007]	0.056 [-0.050; 0.150]	0.056 [-0.056; 0.152]
White	0.021 [-0.170; 0.210]	0.021 [-0.145; 0.208]	0.521 [0.040; 1.080]*	0.521 [-0.002; 1.101]
Membership envir. organization	0.042 [-0.170; 0.200]	0.042 [-0.168; 0.271]	0.093 [-0.560; 0.660]	0.093 [-0.619; 0.752]
log(Earnings)	0.046 [-0.090; 0.170]	0.046 [-0.081; 0.153]	0.277 [-0.150; 0.560]	0.277 [-0.107; 0.736]
Estimates of WTP				
$\bar{\hat{\mu}}_3$	112.676 [106.400; 126.000]*	112.676 [106.533; 125.662]*	356.771 [317.300; 631.300]*	356.771 [313.046; 630.149]*
$\bar{\hat{\mu}}_4$	104.674 [99.700; 115.000]*	104.674 [99.982; 115.058]*	715.210 [380.800; 1,810.400]*	715.210 [373.774; 1,739.754]*

Data sources: Data retrived from Oliver Linton's webpage (http://www.oliverlinton.me.uk/research/software) ***Notes:*** Estimates in column 1 and 3 refer to estimation results presented in Table 7 and 8 in Lewbel et al. (2011). Estimates are presented along with their 95% paired bootstrap confidence intervals (200 iterations) which can deviate because of sampling differences. Significance levels: * $p < 0.05$.

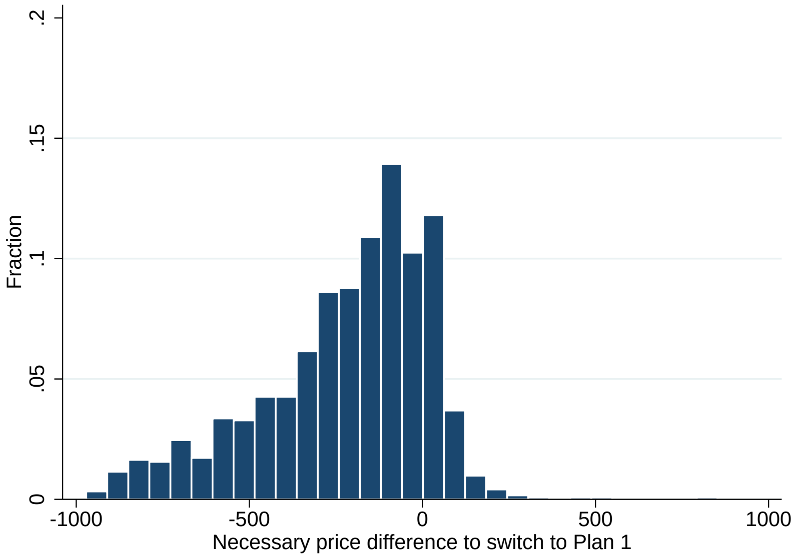
2.C Appendix Figures and Tables

Figure 2.8: Necessary price spread to choose alternative option

(a) First choice: Plan 1



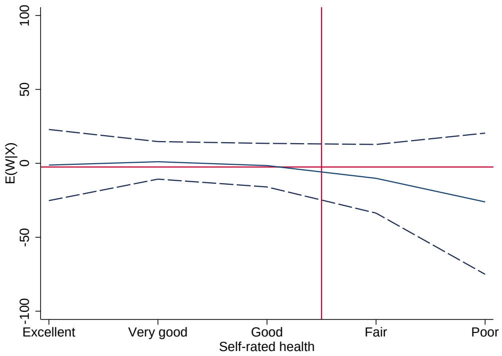
(b) First choice: Plan 2



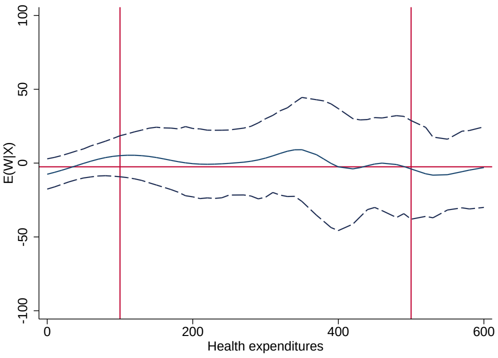
Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Figure displays necessary price spread to choose alternative option. For Figure (a), this indicates the price of Plan 2 relative to the price of Plan 1. For Figure (b), this indicates the price of Plan 1 relative to the price of Plan 2. I restrict the figures to the range [-1000;1000], excluding less than 1% of the observations. Figures (a) and (b) are based on 1,286 and 1,227 observations, respectively.

Figure 2.9: Nonparametric results: Heterogeneity in WTP across health plans by further background characteristics

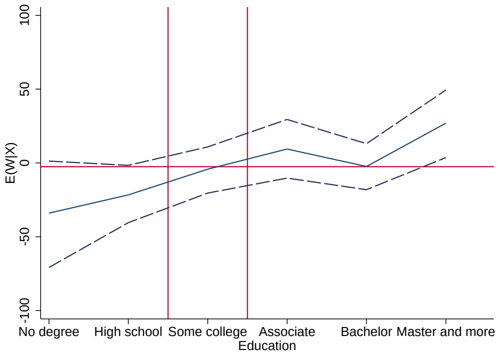
(a) Self-rated health



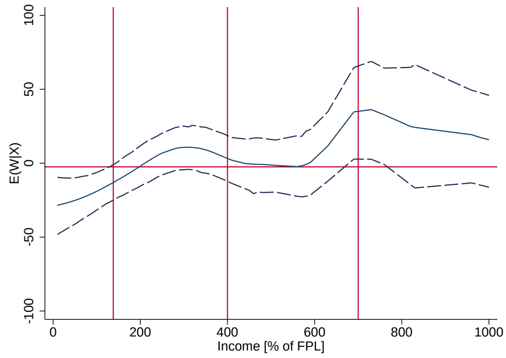
(b) Health expenditures



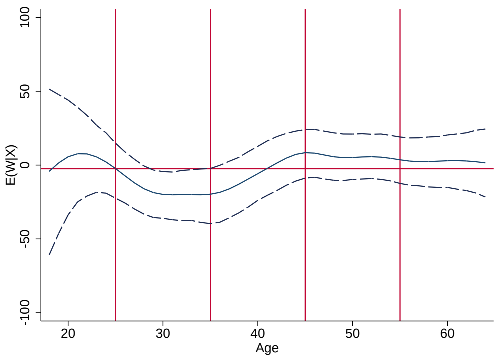
(c) Education



(d) Income



(e) Age



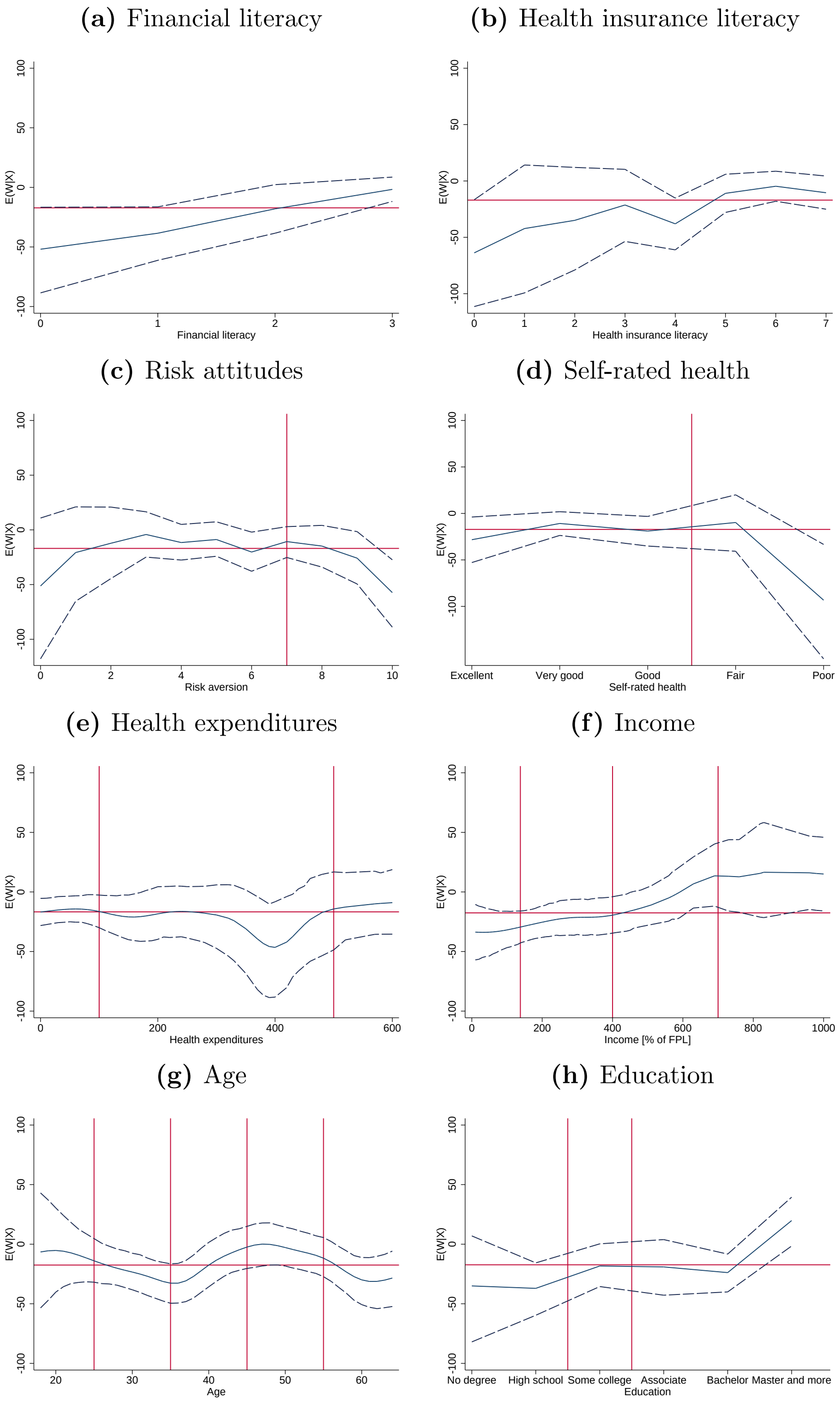
Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Figure displays marginal smooths $\hat{\mu}_1(X_i)$ along with their pointwise 95% paired bootstrap confidence intervals (1,000 iterations) and estimated unconditional mean ($N = 2,513$). The density for those with high medical expenditures or with high incomes is very low at some points of the distribution. I censor medical expenditures at \$600 to limit the impact of outliers (about 8% of the sample). I further censor income at 1000% of FPL (about 5% of the sample). This has a decreasing impact on the bandwidth and leads to an improvement with respect to meaningful cell sizes. Results on the lower tails of the distributions of health expenditures and income do not change much.

Table 2.7: Semiparametric results: Heterogeneity in the WTP across health plans (full regression)

	Coefficient [90% CI]	
Literacy		
High financial and HI literacy	17.83	[3.20; 34.26]*
Risk attitudes		
High risk aversion	-6.95	[-20.06; 7.41]
Gender		
Female	-2.04	[-15.62; 12.56]
Age		
Age 18-25	33.96	[2.62; 65.48]*
Age 35-44	10.58	[-10.26; 31.92]
Age 45-54	21.90	[2.67; 41.49]*
Age 55-64	12.08	[-8.19; 35.89]
Race and ethnicity		
Non-Hispanic non-White	-9.09	[-28.57; 10.12]
Hispanic	-32.76	[-51.49; -11.73]*
Education		
Some college	11.38	[-9.59; 30.27]
Completed college degree	16.90	[-1.74; 34.59]
Family income		
Medicaid eligible	11.39	[-14.44; 36.83]
Subsidy eligible	20.09	[0.79; 38.94]*
Top household income	19.08	[-7.15; 43.09]
Household size		
Household $\geq$ 2 members	8.33	[-6.80; 22.46]
Employment status		
Working	4.13	[-13.09; 20.55]
Health insurance type		
Medicaid	7.93	[-20.71; 36.92]
Employer-sponsored insurance	21.91	[1.89; 43.28]*
All other	-6.37	[-29.37; 18.68]
Health		
Fair/poor self-rated health	5.82	[-15.63; 27.46]
Medium health expenditures	-1.74	[-18.28; 14.57]
High health expenditures	-13.02	[-36.06; 10.20]
Constant	-52.96	[-91.34; -17.58]*
Estimates of the WTP		
$\bar{\hat{\mu}}_3$	-5.79	[-12.27; 1.63]
$\bar{\hat{\mu}}_4$	2.38	[-6.23; 13.23]
N	2,513	

Data sources: RAND American Life Panel (ALP), individuals aged 18-64. ***Notes:*** Estimates are presented along with their 90% paired bootstrap confidence intervals (1,000 iterations, $N = 2,513$). The star $*$ indicates significance.

Figure 2.10: Nonparametric robustness check: Excluding individuals indicating price spreads outside the support of bid values



Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Figure displays marginal smooths $\hat{\mu}_1(X_i)$ along with their pointwise 95% paired bootstrap confidence intervals (1,000 iterations) with estimated unconditional mean ($N = 1,465$). Income and health expenditures variables are censored at 1000% of FPL and $600, respectively.

Table 2.8: Semiparametric robustness check: Excluding individuals indicating price spreads outside the support of bid values

	Coefficient	[90% CI]
Literacy		
High financial and HI literacy	18.568	[1.968; 34.880]*
Risk attitudes		
High risk aversion	-8.915	[-24.689; 6.380]
Gender		
Female	10.222	[-4.820; 25.976]
Age		
Age 18-25	19.530	[-11.989; 53.055]
Age 35-44	-12.424	[-33.938; 8.112]
Age 45-54	5.238	[-16.667; 27.972]
Age 55-64	-15.341	[-40.555; 9.110]
Race and ethnicity		
Non-Hispanic non-White	7.919	[-13.136; 29.048]
Hispanic	-33.237	[-55.124; -10.164]*
Education		
Some college	9.250	[-15.943; 33.517]
Completed college degree	6.054	[-15.258; 29.201]
Family income		
Medicaid eligible	-14.803	[-41.696; 13.402]
Subsidy eligible	-11.229	[-32.043; 8.115]
Top income	17.441	[-7.670; 43.804]
Household size		
Household $\geq$ 2 members	7.752	[-8.878; 23.409]
Employment status		
Working	6.637	[-12.569; 24.269]
Health insurance type		
Medicaid	16.547	[-20.430; 51.132]
Employer-sponsored insurance	12.758	[-11.222; 34.974]
All other	-9.333	[-40.524; 20.025]
Health		
Poor self-rated health	12.847	[-11.984; 39.035]
Medium health expenditures	-17.027	[-34.760; 2.374]
High health expenditures	-3.232	[-28.452; 20.632]
Constant	-33.048	[-73.595; 8.166]
Estimates of the WTP		
$\bar{\tilde{\mu}}_3$	-17.032	[-25.773; -8.902]*
$\bar{\tilde{\mu}}_4$	-15.172	[-24.098; -8.104]*
N	1,465	

Data sources: RAND American Life Panel (ALP), individuals aged 18-64. ***Notes:*** Estimates are presented along with their 90% paired bootstrap confidence intervals (1,000 iterations, $N = 1,465$). The star $*$ indicates significance.

2.D Questions in the ALP

Financial literacy questions

Correct answers are marked with a star.

1. (*Numeracy*) Suppose you had $100 in a savings account and the interest rate was 2 percent per year. After 5 years, how much do you think you would have in the account if you left the money to grow: more than $102, exactly $102, or less than $102?
   - More than $102*
   - Exactly $102
   - Less than $102
   - Don't know
2. (*Inflation*) Imagine that the interest rate on your savings account was 1 percent per year and inflation was 2 percent per year. After 1 year, would you be able to buy more than, exactly the same as, or less than today with the money in this account?
   - More than today
   - Exactly the same as today
   - Less than today*
   - Don't know
3. (*Risk*) Do you think that the following statement is true or false? "Buying a single company stock usually provides a safer return than a stock mutual fund."
   - True
   - False*
   - Don't know

Health insurance literacy questions

Correct answers are marked with a star.

1. If an insurance policy has a higher deductible the premium should be lower, everything else equal.

☐ True*

☐ False

☐ Don't know

2. If you visit a doctor who is not part of your insurer's network you will have to pay more.

   ☐ True*

   ☐ False

   ☐ Don't know

3. Generic prescription drugs cost the patient more than brand name drugs.

   ☐ True

   ☐ False*

   ☐ Don't know

4. Which type of insurer places greater restrictions on the patient's choices of the providers they see?

   ☐ HMO*

   ☐ PPO

   ☐ They are the same (HMO equals PPO in terms of choice of providers)

   ☐ Don't know

5. Which of the following best describes a deductible?

   ☐ A small amount that patients must pay each time they visit a doctor

   ☐ The amount patients must pay during a year before their insurance will pay for care*

   ☐ The price policy holders must pay for insurance

   ☐ Don't know

6. You go to the doctor and the bill for your visit is $100. You have to pay a co-insurance of 20% for all doctor visits. How much will you be expected to pay for this visit?

   ☐ $0

   ☐ $20*

   ☐ $80

- ☐ Don't know

7. (...) You have a co-pay of $15 for all doctor visits. How much will you be expected to pay for this visit?
   - ☐ $0
   - ☐ $15*
   - ☐ $85
   - ☐ Don't know

Risk attitudes

How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid risks? Please rate your willingness to take risks on the scale below where the value 0 means: "not at all willing to take risks" and the value 10 means: "very willing to take risks". With values between 0 and 10 you can express where you lie between these two extremes.

- ☐ 0 – Not at all willing to take risks
- ☐ 1
- ☐ (...)
- ☐ 9
- ☐ 10 – Very willing to take risks

3 Knowledge and preferences as barriers to health insurance coverage under the Affordable Care Act

The Affordable Care Act established new policy mechanisms to increase health insurance coverage in the United States. While the rate of insurance coverage among adults has increased, estimates suggest that 10 to 15% of the U.S. population remains uninsured. Using the nationally representative RAND American Life Panel, we investigate whether financial literacy, health insurance literacy, attitudes toward risk, and political affiliation constitute barriers to coverage. Among the uninsured in 2013, higher financial and health insurance literacy and greater risk aversion were associated with a greater probability of being insured in 2015. Uninsured Republicans were both less likely than Democrats to obtain insurance and to obtain it via Medicaid or the Marketplaces. For the general population, those with high financial and health insurance literacy were more likely to obtain insurance via Medicaid or the Marketplaces compared to being uninsured. Republicans were less likely than Democrats to be covered via Medicaid or the Marketplaces. The magnitude of the coefficients for our novel predictors was of the same order as that of the more traditional covariates.[1]

3.1 Introduction

The Affordable Care Act (ACA) of 2010 introduced multiple policy mechanisms designed to increase health insurance coverage in the United States. First, it offered incentives for states to expand Medicaid to low-income groups. Second, it introduced health insurance Marketplaces to facilitate enrollment and create competition across insurers. Third, it subsidized the purchase of insurance in the Marketplaces for those with incomes between 100% (effectively 138% in Medicaid expansion states) and 400% of the Federal Poverty Level (FPL). Finally, it

[1] This paper is joint work with Amelie Wuppermann, Department of Economics, University of Munich, Ludwigstr. 33, D-80539 Munich, Germany; Silvia H. Barcellos, Center for Economic and Social Research, University of Southern California, 635 Downey Way, Los Angeles, CA 90089-3332; Sebastian Bauhoff, Center for Global Development, 2055 L Street NW, Fifth Floor, Washington DC 20036; Joachim K. Winter, Department of Economics, University of Munich, Ludwigstr. 33, D-80539 Munich, Germany; and Katherine G. Carman, The RAND Corporation, 1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138. This paper benefited from helpful comments of partipants in the Evidence Based Economics Research Strategy Seminar, the joint internal seminar of the Chair of Empirical Economic Research and the Chair of Economic History at the University of Munich, the brown bag seminar at the Leonard D. Schaeffer Center for Health Policy and Economics in Los Angeles, the World Risk and Insurance Economics Congress 2015 in Munich and the 2015 Doctoral Workshop SSPH+ in Luzern.

established a penalty for those who opt to remain uninsured, with some exemptions. While surveys estimate that the rate of insurance coverage among adults age 18-64 increased by 4 to 8 percentage points since 2013, they also estimate that 10 to 15% of the U.S. population remains uninsured (Carman et al. 2015; Collins et al. 2014; Long et al. 2014a; Smith and Medalia 2015; Sommers et al. 2015a, 2014; Vistnes and Cohen 2015; Witters 2015). This lingering gap suggests that important barriers to enrollment persist.

Previous research on these insurance choices following the first ACA open enrollment has focused on socio-demographic characteristics, finding, for example, that lower educational and income levels and being Hispanic are associated with being uninsured, and that people who live in states that did not expand Medicaid are also likely to remain uninsured (Collins et al. 2014; Long et al. 2014a; Smith and Medalia 2015; Sommers et al. 2015a, 2014; Vistnes and Cohen 2015; Witters 2015). However, there are potentially many other unrecognized factors that may influence choices about health insurance. Financial literacy (defined as knowledge about financial concepts) and attitudes towards risk, for instance, have been found to play an important role in other financial decisions (Cutler et al. 2008; Lusardi and de Bassa Scheresberg 2013; Lusardi and Mitchell 2011a). It seems likely that they would influence decisions about health insurance as well. Furthermore, given the highly politicized nature of the debate about the ACA, political affiliation could also play a part.

In this paper, we turn the lens onto these potential barriers, investigating whether health insurance transitions are associated with health insurance literacy (defined as knowledge of health insurance concepts), financial literacy, levels of risk aversion, and political leanings. Unlike most previous research on this topic, we made use of rich longitudinal data, acquired from the nationally representative RAND American Life Panel (ALP), that offer information on more of the potential barriers to enrollment than datasets used in earlier studies. In particular we made use of a new measure of health insurance literacy documented in previous work (Barcellos et al. 2014) and assessed whether knowledge is associated with health insurance coverage. With access to these data, we could analyze the role of these variables, along with those of traditional socio-demographic insurance predictors. A key strength of this research design is the use of longitudinal data enabling us to investigate how these four factors were related to individual-level changes in insurance coverage over time. Notably, the data allowed us to measure these characteristics prior to the Marketplace's first open enrollment period in 2013, ensuring that enrollment decisions made after the rollout of the Marketplaces did not impact our measures of knowledge or political leanings. Furthermore, by collecting data about respondents' current insurance status, we avoided recall bias. Other studies assessing these effects used cross sectional data, lacked the rich set of covariates, or focused on samples not representative of the general population (KFF 2015a).

3.2 Methods

Data sources and study sample. The ALP is a nationally representative internet panel of respondents who participate in occasional online surveys. Panel members without internet access are provided with a computer and an internet connection, eliminating the bias found in many internet surveys, which include only existing computer users.[2] We calculated sample weights to make the distributions of age, gender, ethnicity, education, income, and household size approximate the distributions in the Census Bureaus Current Population Survey (CPS, see King et al. 2015). Panel members have been recruited over time, via several different mechanisms. The cumulative response rate across all recruitment waves is approximately 9%. While this is low, it is similar to other nationally representative surveys conducted by non-governmental organizations. Furthermore, ALP estimates of the uninsured rate among working-age adults in 2013 and 2014 (2013: 20.3%; 2014: 13.7%, refer to Carman and Eibner 2015) are similar to those from the Current Population Survey (2013: 18.5%; 2014 14.3%, refer to Smith and Medalia 2015). The ALP also provided accurate predictions of the 2012 Presidential Election (Gutsche et al. 2014).

Our analyses focused on 2,742 respondents age 18-64 who completed surveys in fall 2013 (baseline) and spring 2015 (follow-up). We conducted our baseline survey from August 23 to September 30, 2013, just before the first open-enrollment period for the Marketplaces. We fielded our follow-up survey from March 1 to May 26, 2015. 84% of those answering the baseline survey also responded to the follow-up. Both surveys included questions about health insurance literacy and coverage. Measures of risk aversion and financial literacy were collected between September 20, 2013 and March 5, 2014. Information about political affiliation was collected in earlier ALP surveys fielded by other researchers before our baseline survey. This information was missing for approximately 28% of respondents who were not previously asked about their political affiliation. We focused on respondents who answered all questions included in our analysis constituting 92% of those who completed both surveys (except those about political affiliation where we included an indicator if this information was missing).

Outcome measures. At both baseline and follow-up, we collected data about whether the respondent currently had health insurance, and if so, what type – employer sponsored insurance (ESI), Medicaid, insurance through a Marketplace, or insurance through other sources (e.g., privately purchased non-group insurance, Medicare, Military or Veterans insurance, insurance from other small state programs). Employer sponsored insurance includes coverage

[2]More information is available at https://alpdata.rand.org/.

obtained through COBRA, the survey respondent's employer, or the employer of the respondent's spouse or parent. The rates of uninsurance, enrollment in Medicaid, and enrollment via the Marketplaces that we estimated from these data were similar to rates measured in other studies (Carman et al. 2015; Collins et al. 2014; Long et al. 2014a; Sommers et al. 2014), as well as to estimates from the Centers for Medicare and Medicaid Services (CMS), the Department of Health and Human Services (HHS), and the Census Bureau (CMS 2015; HHS 2015; Smith and Medalia 2015).

Independent variables of interest. We focused on four barriers to insurance coverage, all measured at baseline, or before in the case of political affiliation. First, we collected standard measures of financial literacy covering fundamental concepts of economics and finance which have been shown to affect complex financial decisions, for instance, retirement planning and stock market participation (Lusardi and Mitchell 2011a; van Rooij et al. 2011). The questions measure basic knowledge of interest rates (numeracy), inflation, and the safety of stocks versus bonds (risk). We classified individuals as having high financial literacy if they correctly answered all three questions (Lusardi and de Bassa Scheresberg 2013; Lusardi and Mitchell 2011a). Individuals who are less financially literate may have more difficulties or feel less confident making decisions about health insurance, which aims to shield individuals from uncertain, potentially large financial consequences related to medical expenditures. Second, we measured health insurance literacy (7 questions about deductibles, co-pays, coinsurance, networks, and prescription drug pricing (generic versus brand name), refer to Barcellos et al. 2014). These questions reflect individual knowledge of health insurance concepts, which are the factors that distinguish health insurance plans, and may be particularly important for those who have many plans to compare in the Marketplaces. Knowledge of these concepts is thus likely to be important for choosing health insurance (Heiss et al. 2006) and making good choices (Bhargava et al. 2015). We classified individuals as having low health insurance literacy if they correctly answered fewer than 3 out of the 7 questions. This presents a rather low bar and captures individuals with potentially limited experience with health plans which is likely to be the case for our main population of interest: the uninsured. Individuals with lower health insurance or financial literacy may be less likely to purchase insurance, because of difficulties in fully considering cost and benefits of different products when comparing insurance products in the complex environment (Baicker et al. 2012; Liebmann and Zeckhauser 2008). Third, we measured risk aversion, defined as one's willingness to take risks (measured on a 0 to 10 scale). We classified individuals as having either a high (less willing to take risks) or low (more willing to take risks) aversion to risk by splitting them at the median. Prior research has shown the behavioral validity of this measure using paid lottery choices (Dohmen et al. 2011). Those who are more risk averse may be more likely to demand health insurance (de Meza and Webb 2001; Finkelstein and McGarry 2006). Finally, we included political affiliation prior to baseline. Public opinion of the ACA is divided strongly along party lines, with Republicans often opposing the ACA (Blendon and Benson 2014; KFF 2015b,c) and showing lower intentions to comply with the individual mandate (Gallup 2014). Republicans may consequently be less likely to take advantage of Medicaid or the Marketplaces, leading to lower levels of insurance coverage. However, previous research has been inconclusive on this issue (Sommers et al. 2015b).

Covariates. We controlled for demographic characteristics collected at baseline, such as gender, age, race and ethnicity, education, household income, employment status, family status (single, married), and health. We also included an indicator for whether or not the respondent's state of residence expanded Medicaid. We obtained similar results from alternative specifications controlling for whether the states had state-based Marketplaces, the states' leaning in the 2012 presidential election or using state-fixed effects.

Statistical analysis. We analyzed two populations: respondents uninsured at baseline and all respondents, regardless of initial insurance status. For those uninsured at baseline, we calculated the fraction who obtained insurance by spring 2015 by various demographic characteristics. For both populations, we separately estimated logit regressions to model the predicted probability of having insurance at follow up (i.e., the choice to become insured), and multinomial logit regressions to model the predicted probability to select different types of insurance (i.e., Medicaid, the Marketplaces, or through their employers) in comparison to being uninsured. We report odds ratios (OR) for logit regressions and relative risk ratios (RRR) for multinomial logit regressions. We used heteroscedasticity-adjusted standard errors clustered at the state of residence. We conducted our analyses using Stata 14 (StataCorp 2015). In the multinomial logit regressions for the full sample, we controlled for the type of insurance individuals had at baseline, as individuals who had insurance in 2013 were likely to have the same type of insurance at follow-up.

Descriptive statistics. Columns 1, 3, and 5 of Table 3.1 provide summary statistics for the full sample and those uninsured and insured at baseline while columns 2, 4, and 6 provide them for the March 2013 Current Population Survey (CPS). Our weighted ALP survey sample tracks the distribution of key covariates in the broader, nationally representative CPS sample, including for variables that were not used to construct the statistical weights. 54% of our sample had low financial literacy; 12%, low health insurance literacy. Both numbers were remarkably higher among those uninsured compared to those insured at baseline. The correlation between the two variables is 0.271 among the general population and 0.268 among the uninsured (both $p<0.001$). By construction, approximately half of the sample was not highly risk averse. Those uninsured at baseline are less likely to be Republican. Table 3.4 in Appendix 3.A shows summary statistics by source of insurance.

Table 3.1: Descriptive statistics

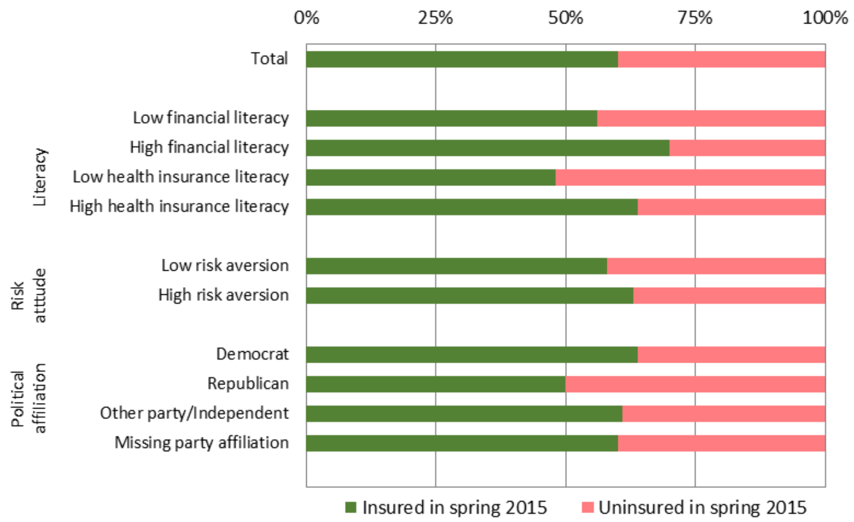
Characteristics	Full sample		Uninsured		Insured	
	ALP	CPS	ALP	CPS	ALP	CPS
	[n=2,742]	[n=122,316]	[n=525]	[n=24,523]	[n=2,217]	[n=97,793]
Health insurance in fall 2013						
Uninsured	18%	21%	100%	100%	0%	0%
Medicaid	6%	9%	0%	0%	7%	11%
ESI	60%	60%	0%	0%	73%	74%
All other	17%	12%	0%	0%	20%	15%
Literacy						
Low financial literacy	54%	[a]	74%	[a]	50%	[a]
High financial literacy	46%	[a]	26%	[a]	50%	[a]
Low health insurance literacy	12%	[a]	26%	[a]	9%	[a]
High health insurance literacy	88%	[a]	74%	[a]	91%	[a]
Risk attitude						
Low risk aversion	52%	[a]	57%	[a]	50%	[a]
High risk aversion	48%	[a]	43%	[a]	50%	[a]
Political affiliation						
Democrat	29%[b]	[a]	29%[b]	[a]	29%	[a]
Republican	21%[b]	[a]	15%[b]	[a]	22%	[a]
Other party/Independent	23%[b]	[a]	22%[b]	[a]	23%	[a]
Missing party affiliation	28%[b]	[a]	35%[b]	[a]	26%	[a]
Gender						
Male	49%	49%	49%	54%	49%	48%
Female	51%	51%	51%	46%	51%	52%
Age						
Younger than 26	15%	18%	26%	22%	12%	17%
26-44	36%	40%	39%	46%	36%	38%
45 and older	49%	42%	35%	33%	52%	45%
Race and ethnicity						
Non-Hispanic White	66%	63%	46%	46%	71%	68%
Non-Hispanic non-White	15%	20%	23%	22%	13%	20%
Hispanic	19%	17%	31%	31%	16%	13%
Education						
No degree	9%	11%	21%	22%	7%	9%
High school or equivalent	28%	28%	37%	36%	26%	26%
Some college	20%	21%	23%	20%	20%	21%
Associate degree	10%	10%	8%	8%	11%	10%
Bachelor's degree and more	32%	30%	12%	14%	37%	34%
Income						
Income lower than 138% of FPL	21%	23%	46%	45%	15%	17%
Income 138-250% of FPL	22%	19%	29%	26%	20%	17%
Income 251-400% of FPL	20%	21%	17%	16%	21%	22%
Income higher than 400% of FPL	37%	38%	9%	13%	44%	44%
Employment						
Employed 2013/2015	64%	[a]	46%	[a]	68%	[a]
Unemployed 2013/2015	24%	[a]	30%	[a]	23%	[a]
Employed 2013/Unemployed 2015	6%	[a]	9%	[a]	5%	[a]
Unemployed 2013/Employed 2015	6%	[a]	15%	[a]	4%	[a]

Continued on next page.

Characteristics	Full sample		Uninsured		Insured	
	ALP	CPS	ALP	CPS	ALP	CPS
	[n=2,742]	[n=122,316]	[n=525]	[n=24,523]	[n=2,217]	[n=97,793]
Family status						
Single	36%	39%	52%	51%	32%	36%
Married	64%	61%	48%	49%	68%	64%
3 and less other household members	84%	83%	80%	80%	85%	84%
More than 3 other household members	16%	17%	20%	20%	15%	16%
Health						
Excellent/very good/good health	87%	89%	80%	88%	89%	89%
Fair/poor health	13%	11%	20%	12%	11%	11%
Health expenditures (last 4 months): $\leq$\$100	67%	[a]	74%	[a]	65%	[a]
Health expenditures (last 4 months): $>$ \$100	33%	[a]	26%	[a]	35%	[a]
State characteristics						
No Medicaid expansion	44%	50%	48%	55%	43%	49%
Medicaid expansion	56%	50%	52%	45%	57%	51%

Data sources: Fall 2013 RAND American Life Panel (ALP) and March 2013 Current Population Survey (CPS), individuals aged 18-64. ***Notes:*** Tables displays weighted averages for individuals younger than 65 with non-missing values in the baseline dimensions. The ALP samples are constructed for those individuals participating prior to October 2013 with valid health insurance information in both fall 2013 and spring 2015. [a] No information available in CPS. [b] Most recent indicated political affiliation in ALP, missing for some participants.

Figure 3.1: Insurance status in spring 2015 for previously uninsured individuals by literacy, risk attitude and political affiliation



Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Figure displays the share of prior uninsured individuals obtaining health insurance between fall 2013 and spring 2015 by literacy, risk attitude and political affiliation (N=525).

3.3 Results

3.3.1 Insurance coverage among the previously uninsured

Figure 3.1 displays the insurance status of respondents uninsured at baseline, by our key variables of interest. Overall, 60.2% of this group had obtained insurance by spring 2015. Those with lower levels of financial and health insurance literacy, the less risk averse, and Republicans were less likely to have obtained coverage. Figure 3.3 in Appendix 3.A shows that in the bivariate analysis, Hispanics, the less educated, larger households, those with lower health expenditures, and those living in states that opted not to expand Medicaid were less likely to have obtained insurance by spring 2015 (all p<0.05).

Table 3.2: Characteristics associated with coverage and selection of health insurance for individuals uninsured in fall 2013

	Any coverage	**Type of coverage selected Multinomial logit [N=525]**			
	Logit [N=525]	**Medicaid [N=71]**	**Marketplace [N=56]**	**ESI [N=100]**	**Other [N=89]**
Literacy					
High financial literacy	1.429*	1.793	1.862*	1.355	1.332
	(0.934,2.185)	(0.794,4.049)	(0.963,3.603)	(0.812,2.262)	(0.648,2.738)
High health insurance literacy	1.464**	1.404	1.543	1.277	1.803***
	(1.025,2.093)	(0.713,2.766)	(0.658,3.617)	(0.651,2.504)	(1.160,2.801)
Risk attitude					
High risk aversion	1.525**	1.272	1.652	1.823***	1.413*
	(1.088,2.138)	(0.617,2.626)	(0.833,3.275)	(1.215,2.735)	(0.960,2.081)
Political affiliation					
Republican	0.357***	0.061***	0.075***	0.640	0.576**
	(0.192,0.666)	(0.012,0.303)	(0.014,0.409)	(0.255,1.610)	(0.348,0.953)
Other party/Independent	0.712	0.308**	0.614	0.854	1.023
	(0.381,1.331)	(0.125,0.761)	(0.231,1.630)	(0.361,2.021)	(0.526,1.993)
Missing party affiliation	0.740	0.585	0.924	0.686	0.787
	(0.477,1.147)	(0.284,1.207)	(0.382,2.234)	(0.338,1.393)	(0.496,1.250)
Pseudo-R2	0.1101	0.1724			

Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Table displays odds ratios (95% CIs) from logit regression (column 1) and relative risk ratios (95% CIs) from multinomial logit regression (columns 2-5) using heteroscedasticity-adjusted standard errors clustered at the state of residence level. Regressions further include controls for gender, age, race and ethnicity, education, income, employment and family status, health and state characteristics. Results in these dimensions are presented in Table 3.5 in Appendix 3.A. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3.2, column 1 presents the results from the logit regression modelling the predicted probability that those uninsured at baseline have insurance at follow-up. For most of our key independent variables, we found significant associations with the probability of obtaining insurance. Those with high financial (OR=1.429; 95% CI, 0.934 to 2.185) or health insurance literacy (OR=1.464; 95% CI, 1.025 to 2.093) were more likely to gain insurance than those with low literacy. Those who were more risk averse were more likely (OR=1.525; 95% CI, 1.088 to 2.138) to obtain insurance than the less risk averse. Republicans were less likely (OR=0.357; 95% CI, 0.192 to 0.666) than Democrats to acquire insurance. We find no significant effects for independents or for those who were not previously asked about their political affiliation.

Columns 2-5 in Table 3.2 report the results of the multinomial logit regression investigating in what types of insurance respondents enrolled in, compared to being uninsured. Those who were more risk averse were not only more likely to obtain insurance, but were also more likely to obtain it via their employer (RRR=1.823; 95% CI, 1.215 to 2.735). Controlling for other characteristics, Republicans were also less likely to enroll in Medicaid and via the Marketplaces (RRR= 0.061; 95% CI, 0.012 to 0.303, and RRR=0.075; 95% CI, 0.014 to 0.409, respectively) than Democrats. Table 3.5 in the Appendix 3.A presents the findings for the demographic controls in the regressions including gender, age, race and ethnicity, education, income, employment and family status, health and state characteristics. The coefficients for these more traditional predictors are in line with earlier studies.

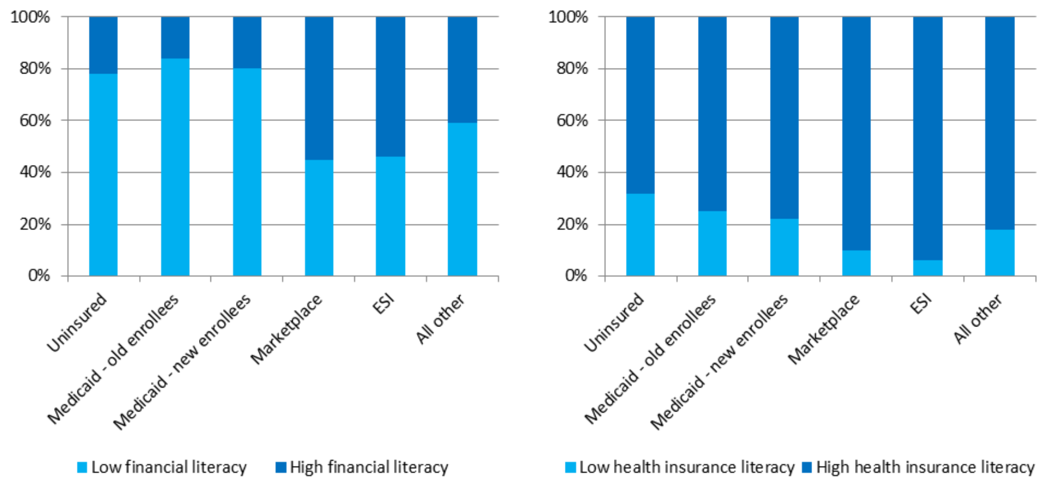
Table 3.3: Characteristics associated with coverage and selection of health insurance for full sample

	Any coverage	**Type of coverage selected Multinomial logit [N=2,742]**			
	Logit [N=2,742]	**Medicaid [N=274]**	**Marketplace [N=148]**	**ESI [N=1,582]**	**Other [N=460]**
Literacy					
High financial literacy	1.598***	1.697**	2.140***	1.382*	1.599*
	(1.142,2.234)	(1.056,2.725)	(1.381,3.316)	(0.940,2.032)	(0.997,2.564)
High health insurance literacy	1.537***	1.853***	1.987**	1.407	1.420**
	(1.221,1.933)	(1.352,2.540)	(1.146,3.447)	(0.760,2.603)	(1.020,1.976)
Risk attitude					
High risk aversion	1.271	1.222	1.258	1.475**	1.214
	(0.952,1.698)	(0.830,1.798)	(0.848,1.868)	(1.073,2.027)	(0.839,1.757)
Political affiliation					
Republican	0.698*	0.370***	0.398**	0.992	0.695*
	(0.474,1.030)	(0.228,0.598)	(0.167,0.949)	(0.627,1.571)	(0.456,1.060)
Other party/Independent	0.836	0.561*	0.771	0.893	1.036
	(0.590,1.184)	(0.302,1.042)	(0.417,1.428)	(0.591,1.350)	(0.724,1.482)
Missing party affiliation	0.936	0.760	1.092	0.962	0.948
	(0.671,1.306)	(0.529,1.092)	(0.602,1.982)	(0.617,1.500)	(0.651,1.381)
Pseudo-R2	0.3488	0.4008			

Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Table displays odds ratios (95% CIs) from logit regression (column 1) and relative risk ratios (95% CIs) from multinomial logit regression (columns 2-5) using heteroscedasticity-adjusted standard errors clustered at the state of residence level. Regressions further include controls for prior insurance coverage type, gender, age, race and ethnicity, education, income, employment and family status, health and state characteristics. Results in these dimensions are presented in Table 3.6 in Appendix 3.A. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure 3.2: Financial and health insurance literacy by type of health insurance coverage in spring 2015

(a) Panel A: Financial literacy **(b)** Panel B: Health insurance literacy



Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Figure displays the share of individuals having low or high levels of literacy by type of health insurance coverage in spring 2015 (N=2,742).

3.3.2 Insurance coverage among the general population

Table 3.3, column 1 shows who among our full sample had insurance coverage at follow-up. Those with high financial and health insurance literacy were each more likely (OR=1.598; 95% CI, 1.142 to 2.234 and OR=1.537; 95% CI, 1.221 to 1.933, respectively) to have insurance in 2015 than those with low literacy. In the multinomial logit (Table 3.3, columns 2-5), those with high levels of financial or health insurance literacy were more likely (RRR=2.140; 95% CI, 1.381 to 3.316 and RRR=1.987, 95% CI, 1.146 to 3.447, respectively) to be covered via the Marketplaces. Those who in 2015 were covered through employers and the Marketplaces had the highest levels of financial and health insurance literacy in 2013 (refer to Figure 3.2).

Highly risk-averse individuals were more likely (RRR=1.475; 95% CI, 1.073 to 2.027) to be covered by employer insurance and somewhat less likely to be uninsured. In terms of political leanings, both Republicans and Independents were less likely (RRR=0.370; 95% CIs, 0.228 to 0.598 and RRR=0.561; 95% CI, 0.302 to 1.042, respectively) to be covered by Medicaid. Table 3.6 in the Appendix 3.A presents the findings for the demographic controls in the regressions, which are consistent with other studies.

3.4 Discussion

When looking at the decisions of Americans about health insurance in the context of the ACA, it is important to consider not just individuals who were uninsured before 2014, when key provisions of the legislation rolled out, but also the general population, as a number of those who chose to enroll in Medicaid and through the Marketplaces already had coverage in 2013 (Carman et al. 2015). In this study, we were able to look at each of these populations and our dataset enabled us to investigate a wider set of characteristics than other research has done to date. We found that several variables that until now have been unrecognized – health insurance literacy, financial literacy, risk aversion, and political affiliation – did, indeed, appear to play a role in the health insurance choices of our sample between 2013 and 2015. Furthermore these effects hold even after controlling for other important determinants of insurance coverage such as income and employment. It is remarkable that the effects of these variables on peoples insurance decisions have the same order of magnitude as standard demographic determinants, such as income and employment status.

The fact that both financial and health insurance literacy were associated with having insurance seems likely to stem from the greater comfort level that more knowledgeable individuals probably have with making health insurance decisions. Those with low health insurance literacy, who answered fewer than 3 of 7 questions correctly, were especially unlikely to have insurance coverage, potentially because they lack experience with the health care system. Because the ACA relies on consumer choice, those who are uninsured and have low health insurance literacy represent a particularly vulnerable population; for them obtaining coverage may be especially difficult. Our findings are robust to alternative specifications using different health measures. It is thus unlikely that our findings for health insurance literacy solely reflect differences in health that are correlated with both the literacy measures and health insurance choices (Table 3.7 in Appendix 3.A). Interestingly, the correlation between financial and health insurance literacy is low, suggesting that they may contribute to insurance coverage through slightly different mechanisms. Our observation that the highly risk averse were significantly more likely than the risk tolerant to obtain health insurance is consistent with predictions from economic theory, where insurance is an important means of avoiding financial risk. We found that risk averse individuals are more likely to obtain coverage through employers, which would occur if risk averse individuals were more likely to select or remain in jobs that offer health insurance. That Republicans were less likely to be insured – and when they were, it was not through Medicaid or the Marketplaces – could be related to strong reservations about the ACA that until now, have run along party lines. However, as party affiliation is missing for a large fraction of our sample – because not all respondents were previously asked to report their affiliation – the results only hold for a possibly selected subset of those with Republican affiliation. Moreover, Republicans were less likely to be uninsured in 2013.

Our study has several limitations. First, our sample size was relatively small. Despite this, we find significant results. Second, our survey had a low cumulative response rate. The rate was on par, however, with other nationally representative surveys conducted by non-governmental organizations that have provided early pictures of the impact of the ACA (Sommers et al. 2015a), and estimates from the ALP closely track administrative data. The trade-off for the small sample size and low response rate was access to data on financial and health insurance literacy, risk aversion, and political affiliation – variables that are usually lacking in data from federal surveys with larger sample sizes and higher cumulative response rates. Third, we relied on survey data, which may be biased due to sample selection or self-reporting. But this does not necessarily distort the accuracy of findings: The American Life Panel provided highly accurate predictions in the 2012 presidential elections, for example (Gutsche et al. 2014).

We focus on two points in time and thus in this work did not distinguish between those who experienced one or multiple health insurance transitions. Health insurance and financial literacy might also contribute to maintaining consistent insurance coverage over this period. This remains scope for future research.

3.5 Conclusion

Understanding which factors affect an individual's decision to obtain health insurance and to enroll via Medicaid versus the Marketplaces versus other channels can help to inform health policy and better target and design outreach and consumer-education programs. Our findings suggest that the barriers to obtaining coverage go beyond the primarily socio-demographic ones which have been established as contributing to gaps in coverage. In that light, policies and programs designed to further reduce the numbers of uninsured – especially through Medicaid and the Marketplaces – should take into account not just ethnicity, employment status, and income, but levels of financial literacy, health insurance literacy, risk aversion, and political affiliation. While we cannot exclude that our results on the literacy measures, risk aversion or political affiliation may be – at least partly – capturing the influence of third factors which we cannot observe, the results still suggest that low financial and health insurance literacy, low risk aversion and being a Republican predict remaining uninsured. Any programs designed to further decrease the number of uninsured adults should take into account that these individuals may have low health insurance and financial literacy. Whether changes to the federal healthcare.gov Marketplace designed to simplify decision making during the 2016 open enrollment period will be enough to help these groups gain coverage remains an open question.

3.A Appendix Figures and Tables

Table 3.4: Descriptive statistics by selected types of health insurance coverage in spring 2015

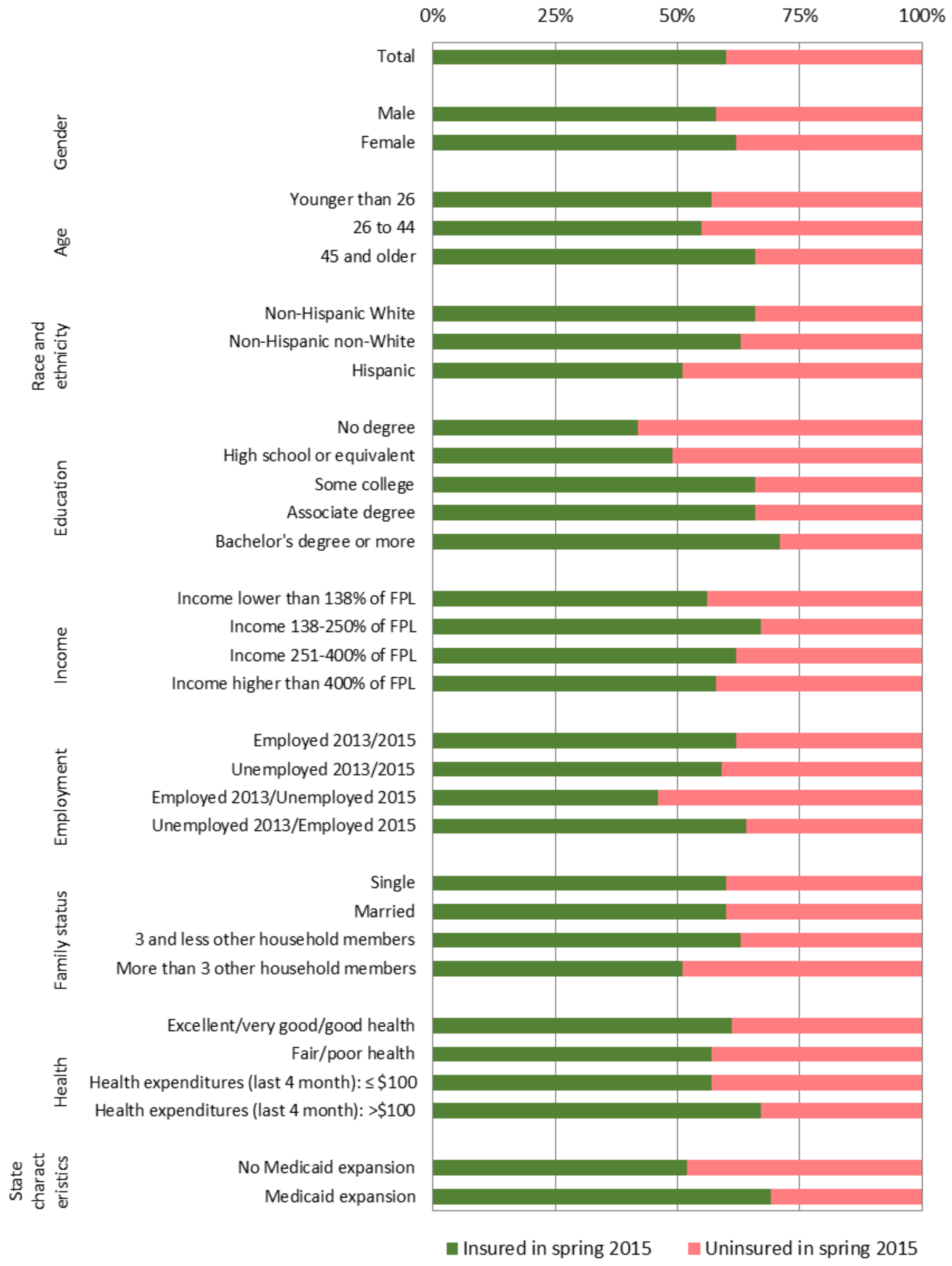
Characteristics	Health insurance coverage in spring 2015 [N=2,742]						Pearson
	Uninsured	Medicaid		Marketplace	ESI	Other	
		Old	New				Chi2
	[N=278]	[N=129]	[N=145]	[N=148]	[N=1,582]	[N=460]	p-value
Literacy							
Low financial literacy	78%	84%	80%	45%	46%	59%	
High financial literacy	22%	16%	20%	55%	54%	41%	0.000
Low health insurance literacy	32%	25%	22%	10%	6%	18%	
High health insurance literacy	68%	75%	78%	90%	94%	82%	0.000
Risk attitude							
Low risk aversion	57%	45%	55%	56%	50%	53%	
High risk aversion	43%	55%	45%	44%	50%	47%	0.282
Political affiliation							
Democrat[a]	28%	37%	33%	32%	28%	28%	
Republican[a]	17%	11%	9%	10%	25%	16%	
Other party/Independent[a]	25%	9%	20%	24%	23%	26%	
Missing party affiliation[a]	29%	43%	38%	34%	25%	30%	0.000
State characteristics							
No Medicaid expansion	56%	34%	24%	38%	45%	41%	
Medicaid expansion	44%	66%	76%	62%	55%	59%	0.000
Gender							
Male	49%	52%	41%	49%	49%	52%	
Female	51%	48%	59%	51%	51%	48%	0.215
Age							
Younger than 26	24%	32%	23%	9%	13%	11%	
26-44	43%	40%	45%	35%	36%	31%	
45 and older	32%	28%	32%	56%	52%	58%	0.000
Race and ethnicity							
Non-Hispanic White	43%	50%	44%	72%	73%	61%	
Non-Hispanic non-White	23%	20%	21%	14%	11%	22%	
Hispanic	34%	31%	35%	15%	15%	17%	0.000
Education							
No degree	23%	35%	20%	5%	4%	15%	
High school or equivalent	39%	28%	36%	27%	24%	32%	
Some college	19%	23%	21%	32%	20%	20%	
Associate degree	7%	8%	10%	10%	11%	8%	
Bachelor's degree and more	11%	6%	13%	26%	41%	25%	0.000
Income							
Income lower than 138% of FPL	46%	74%	60%	24%	7%	34%	
Income 138-250% of FPL	22%	8%	30%	33%	21%	20%	
Income 251-400% of FPL	19%	16%	5%	23%	22%	18%	
Income higher than 400% of FPL	12%	2%	5%	20%	49%	28%	0.000
Employment							
Employed 2013/2015	45%	18%	34%	51%	79%	40%	
Unemployed 2013/2015	28%	70%	46%	35%	13%	46%	
Employed 2013/Unemployed 2015	11%	5%	13%	9%	4%	8%	
Unemployed 2013/Employed 2015	16%	7%	6%	5%	4%	6%	0.000
Family status							
Single	56%	67%	60%	48%	27%	42%	
Married	44%	33%	40%	52%	73%	58%	
$\leq$ 3 other household members	76%	73%	82%	85%	85%	87%	
$>$ 3 other household members	24%	27%	18%	15%	15%	13%	0.000

Continued on next page.

Characteristics	Health insurance coverage in spring 2015						
	Uninsured	Medicaid		Marketplace	ESI	Other	Pearson
		Old	New				Chi2
							p-value
Health							
Excellent/very good/good health	80%	75%	72%	87%	93%	76%	
Fair/poor health	20%	25%	28%	13%	7%	24%	0.000
Health expenditures: $\leq$ \$100	77%	91%	76%	77%	61%	72%	
Health expenditures: >\$100	23%	9%	24%	23%	39%	28%	0.000

Data source: RAND American Life Panel (ALP), individuals aged 18-64. ***Notes:*** Table displays weighted averages for the full ALP sample by type of health insurance coverage in spring 2015. Characteristics measured as of fall 2013. [a] Most recent indicated political affiliation in ALP, missing for some participants.

Figure 3.3: Insurance status in spring 2015 for previously uninsured individuals by standard demographic characteristics



Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Figure displays the share of prior uninsured individuals obtaining health insurance between fall 2013 and spring 2015 by standard demographic characteristics (N=525).

Table 3.5: Further characteristics associated with coverage and selection of health insurance for individuals uninsured in fall 2013

	Any coverage	Type of coverage selected Multinomial logit [N=525]			
	Logit [N=525]	Medicaid [N=71]	Marketplace [N=56]	ESI [N=100]	Other [N=89]
Gender					
Female	1.325	1.382	1.919*	1.255	1.239
	(0.935,1.877)	(0.877,2.175)	(1.000,3.682)	(0.645,2.442)	(0.808,1.902)
Age					
Younger than 26	1.149	1.395	0.460	2.477**	0.689
	(0.566,2.333)	(0.351,5.542)	(0.069,3.078)	(1.120,5.478)	(0.327,1.451)
26-44	0.894	1.211	0.598*	1.325	0.580*
	(0.590,1.355)	(0.625,2.347)	(0.342,1.044)	(0.835,2.103)	(0.315,1.069)
Race and ethnicity					
Non-Hispanic non-White	0.837	0.479*	0.607	1.505	0.769
	(0.569,1.231)	(0.206,1.116)	(0.276,1.332)	(0.923,2.453)	(0.407,1.451)
Hispanic	0.624**	0.449**	0.369**	0.691	0.825
	(0.398,0.979)	(0.208,0.968)	(0.142,0.957)	(0.409,1.166)	(0.521,1.307)
Education					
High school or equivalent	1.061	1.865	0.310	1.661	0.713
	(0.426,2.647)	(0.640,5.432)	(0.036,2.630)	(0.656,4.203)	(0.275,1.846)
Some college	2.316	2.574	3.222	3.909***	1.176
	(0.811,6.618)	(0.407,16.268)	(0.580,17.883)	(1.606,9.512)	(0.285,4.854)
Associate degree	2.345	2.679	3.034	2.468	1.978
	(0.836,6.582)	(0.519,13.822)	(0.412,22.353)	(0.816,7.460)	(0.710,5.506)
Bachelor's degree and more	2.517*	4.507**	3.233	4.062***	1.093
	(0.893,7.094)	(1.154,17.607)	(0.533,19.594)	(1.534,10.753)	(0.320,3.728)
Income					
Income 138-250% of FPL	1.043	0.651	1.238	1.724*	0.842
	(0.696,1.563)	(0.378,1.119)	(0.532,2.883)	(0.941,3.158)	(0.491,1.445)
Income 251-400% of FPL	0.732	0.173***	0.706	1.273	0.798
	(0.386,1.387)	(0.061,0.490)	(0.223,2.234)	(0.596,2.719)	(0.387,1.642)
Income higher than 400% of FPL	0.490*	0.000***	0.622	0.828	0.619
	(0.213,1.124)	(0.000,0.000)	(0.180,2.147)	(0.315,2.176)	(0.235,1.629)
Employment					
Unemployed 2013/2015	0.926	2.312***	0.957	0.186***	1.404
	(0.648,1.322)	(1.518,3.520)	(0.394,2.328)	(0.067,0.514)	(0.886,2.226)
Employed 2013/Unemployed 2015	0.585*	0.832	0.614	0.443	0.669
	(0.322,1.064)	(0.414,1.672)	(0.224,1.680)	(0.131,1.500)	(0.286,1.563)
Unemployed 2013/Employed 2015	1.031	0.903	0.550	1.232	0.898
	(0.674,1.578)	(0.371,2.201)	(0.099,3.054)	(0.804,1.887)	(0.502,1.608)
Family status					
Married	1.052	0.753	0.705	1.477	1.190
	(0.747,1.482)	(0.432,1.313)	(0.420,1.185)	(0.902,2.419)	(0.666,2.126)
More than 3 other household members	0.742	0.585	0.587	0.763	1.030
	(0.488,1.129)	(0.267,1.281)	(0.177,1.943)	(0.363,1.601)	(0.634,1.674)
Health					
Fair/poor health	1.011	1.518	0.527	0.896	1.260
	(0.746,1.372)	(0.866,2.661)	(0.219,1.266)	(0.609,1.317)	(0.788,2.015)
Health expenditures (last 4 months): >$100	1.510**	1.679*	0.722	1.819**	1.684**
	(1.027,2.221)	(0.930,3.031)	(0.331,1.576)	(1.100,3.008)	(1.042,2.721)

Continued on next page.

	Any cove-rage Logit	**Type of coverage selected Multinomial logit**			
		Medicaid	**Marketplace**	**ESI**	**Other**
State characteristics					
Medicaid expansion	2.426***	9.074***	2.062*	1.341	2.108*
	(1.538,3.827)	(5.025,16.384)	(0.928,4.583)	(0.932,1.930)	(0.937,4.740)
Pseudo-R2	0.1101	0.1724			

Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Table displays odds ratios (95% CIs) from logit regression (column 1) and relative risk ratios (95% CIs) from multinomial logit regression (columns 2-5) using heteroscedasticity-adjusted standard errors clustered at the state of residence level. Regressions further include key variables of interest financial and health insurance literacy, risk attitudes and political affiliation. Results in these dimensions are presented in Table 3.2. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3.6: Further characteristics associated with coverage and selection of health insurance for full sample

	Any coverage	**Type of coverage selected Multinomial logit [N=2,742]**			
	Logit [N=2,742]	**Medicaid [N=274]**	**Marketplace [N=148]**	**ESI [N=1,582]**	**Other [N=460]**
Health insurance fall 2013					
Medicaid	9.956***	21.401***	3.372**	5.745***	4.354***
	(6.329,15.660)	(13.384,34.219)	(1.170,9.714)	(2.993,11.029)	(2.288,8.286)
ESI	30.280***	3.205***	4.119***	74.695***	5.534***
	(18.237,50.277)	(1.871,5.492)	(2.083,8.146)	(46.169,120.845)	(3.124,9.803)
All other	7.643***	4.296***	4.672***	4.432***	14.208***
	(5.464,10.689)	(2.836,6.509)	(3.178,6.868)	(2.846,6.901)	(8.883,22.725)
Gender					
Female	1.151	1.196	1.475*	1.209	1.006
	(0.861,1.539)	(0.860,1.663)	(0.998,2.180)	(0.834,1.753)	(0.700,1.444)
Age					
Younger than 26	0.959	1.291	0.435	1.850**	0.474**
	(0.542,1.698)	(0.678,2.458)	(0.095,1.998)	(1.017,3.367)	(0.234,0.960)
26-44	0.765	1.262	0.630***	0.746	0.631*
	(0.551,1.064)	(0.835,1.907)	(0.445,0.892)	(0.505,1.102)	(0.377,1.056)
Race and ethnicity					
Non-Hispanic non-White	0.790	0.641**	0.539*	1.013	0.873
	(0.519,1.204)	(0.420,0.978)	(0.287,1.014)	(0.636,1.613)	(0.517,1.475)
Hispanic	0.557***	0.602**	0.352***	0.629***	0.607**
	(0.417,0.743)	(0.376,0.961)	(0.216,0.572)	(0.453,0.874)	(0.398,0.925)
Education					
High school or equivalent	1.080	0.917	1.175	1.042	1.160
	(0.606,1.923)	(0.526,1.599)	(0.246,5.624)	(0.486,2.234)	(0.659,2.041)
Some college	1.597	1.447	3.252	1.411	1.438
	(0.790,3.230)	(0.664,3.153)	(0.776,13.624)	(0.585,3.402)	(0.712,2.906)
Associate degree	2.057**	1.735	3.375	1.796	1.998**
	(1.109,3.817)	(0.782,3.849)	(0.776,14.677)	(0.814,3.960)	(1.056,3.779)
Bachelor's degree and more	2.280**	1.785	3.519*	2.144*	2.021*
	(1.075,4.835)	(0.837,3.804)	(0.806,15.370)	(0.902,5.096)	(0.932,4.381)
Income					
Income 138-250% of FPL	0.872	0.427***	1.208	1.546*	0.725
	(0.588,1.292)	(0.281,0.649)	(0.656,2.221)	(0.943,2.536)	(0.459,1.145)
Income 251-400% of FPL	0.596	0.139***	0.631	1.171	0.704
	(0.311,1.141)	(0.067,0.287)	(0.312,1.274)	(0.562,2.436)	(0.409,1.213)
Income higher than 400% of FPL	0.701	0.083***	0.365**	1.510	0.823
	(0.402,1.222)	(0.030,0.232)	(0.141,0.942)	(0.858,2.657)	(0.440,1.539)
Employment					
Unemployed 2013/2015	1.188	2.135***	1.029	0.429***	2.178***
	(0.906,1.559)	(1.368,3.332)	(0.611,1.731)	(0.323,0.570)	(1.567,3.028)
Employed 2013/Unemployed 2015	0.416***	1.234	0.683	0.105***	0.754
	(0.228,0.760)	(0.598,2.546)	(0.293,1.595)	(0.045,0.245)	(0.398,1.430)
Unemployed 2013/Employed 2015	0.748	0.598	0.575	0.992	0.746
	(0.465,1.203)	(0.274,1.306)	(0.281,1.177)	(0.610,1.615)	(0.486,1.147)

Continued on next page.

	Any coverage	Type of coverage selected			
		Multinomial logit			
	Logit	**Medicaid**	**Marketplace**	**ESI**	**Other**
Family status					
Married	1.311	0.816	1.069	1.935***	1.271
	(0.913,1.883)	(0.535,1.246)	(0.714,1.601)	(1.239,3.023)	(0.773,2.090)
More than 3 other household members	0.835	0.692	0.695	1.080	0.797
	(0.568,1.226)	(0.414,1.156)	(0.344,1.402)	(0.730,1.598)	(0.440,1.444)
Health					
Fair/poor health	1.085	1.424**	0.587*	0.805	1.318
	(0.846,1.393)	(1.002,2.022)	(0.342,1.010)	(0.576,1.126)	(0.931,1.866)
Health expenditures (last 4 months): >$100	1.260*	1.255	0.853	1.547***	1.160
	(0.979,1.621)	(0.847,1.861)	(0.574,1.266)	(1.166,2.053)	(0.834,1.613)
State characteristics					
Medicaid expansion	2.334***	6.202***	2.146***	1.716***	1.902**
	(1.523,3.576)	(3.725,10.325)	(1.271,3.623)	(1.262,2.334)	(1.099,3.291)
Pseudo-R2	0.3488	0.4008			

Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Table displays odds ratios (95% CIs) from logit regression (column 1) and relative risk ratios (95% CIs) from multinomial logit regression (columns 2-5) using heteroscedasticity-adjusted standard errors clustered at the state of residence level. Regressions further include key variables of interest financial and health insurance literacy, risk attitudes and political affiliation. Results in these dimensions are presented in Table 3.3. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3.7: Robustness checks with respect to definition of health

	Any coverage: Logit (95% CI) [n=525]									
	Baseline (Fair/poor health + expenses >$100)	**Any pre-existing conditions**	**More than 1 doctor visit in past 4 months**	**Over-weight**	**Ever smokes**	**Ever drinks**	**Self-rated health (categories**	**Chances medical bills >$1000: ≥50%**	**Any medical debt**	**All health indicators**
Literacy										
High financial literacy	1.429*	1.448*	1.403	1.414	1.402	1.421*	1.449	1.402	1.428	1.410
	(0.934, 2.185)	(0.942, 2.226)	(0.913, 2.157)	(0.922, 2.168)	(0.914, 2.150)	(0.939, 2.152)	(0.922, 2.277)	(0.903, 2.176)	(0.927, 2.198)	(0.849, 2.340)
High health insurance literacy	1.464**	1.524**	1.458**	1.522**	1.560**	1.492**	1.556***	1.505**	1.487**	1.532**
	(1.025, 2.093)	(1.062, 2.189)	(1.031, 2.062)	(1.059, 2.187)	(1.078, 2.257)	(1.044, 2.134)	(1.119, 2.165)	(1.061, 2.134)	(1.056, 2.093)	(1.062, 2.209)
Risk attitude										
High risk aversion	1.525**	1.503**	1.491**	1.481**	1.493**	1.500**	1.505**	1.479**	1.484**	1.580**
	(1.088, 2.138)	(1.079, 2.094)	(1.062, 2.094)	(1.053, 2.083)	(1.073, 2.076)	(1.072, 2.099)	(1.069, 2.121)	(1.061, 2.061)	(1.065, 2.069)	(1.078, 2.315)
Political affiliation										
Republican	0.357***	0.378***	0.354***	0.367***	0.367***	0.372***	0.343***	0.369***	0.371***	0.328***
	(0.192, 0.666)	(0.211, 0.675)	(0.182, 0.687)	(0.198, 0.682)	(0.197, 0.683)	(0.200, 0.690)	(0.182, 0.646)	(0.198, 0.687)	(0.198, 0.695)	(0.174, 0.620)
Other party/Independent	0.712	0.721	0.688	0.689	0.680	0.686	0.676	0.694	0.705	0.705
	(0.381, 1.331)	(0.395, 1.317)	(0.369, 1.284)	(0.359, 1.322)	(0.359, 1.290)	(0.364, 1.293)	(0.359, 1.272)	(0.365, 1.320)	(0.372, 1.337)	(0.391, 1.274)
Missing party affiliation	0.740	0.774	0.774	0.762	0.715	0.720	0.753	0.762	0.765	0.733
	(0.477, 1.147)	(0.509, 1.179)	(0.507, 1.181)	(0.469, 1.238)	(0.426, 1.200)	(0.429, 1.207)	(0.498, 1.139)	(0.496, 1.171)	(0.494, 1.187)	(0.448, 1.199)
Pseudo-R2	0.110	0.109	0.110	0.107	0.108	0.107	0.110	0.108	0.106	0.133

Data source: RAND American Life Panel, individuals aged 18-64. ***Notes:*** Table displays odds ratios (95% CIs) from logit regressions using heteroscedasticity-adjusted standard errors clustered at the state of residence level. Regressions further include controls for gender, age, race and ethnicity, education, income, employment and family status, health and state characteristics. ***Significance levels:*** *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

4 Early life circumstances predict measures of trust among adults – Evidence from hunger episodes in Post-War Germany

Abstract. *Can a major shock in childhood permanently shape trust? We consider a hunger episode in Germany after WWII and construct a measure of hunger exposure from official data on caloric rations set monthly by the occupying forces providing regional and temporal variation. We correlate hunger exposure with measures of trust using data from a nationally representative sample of the German population. We show that individuals exposed to low caloric rations in childhood have significantly lower levels of trust as adults. This finding highlights that early-life experiences can have long-term effects in domains other than health, where such effects are well-documented.*[1]

4.1 Introduction

Trust has a decisive role in many human interactions. For example, when we consider whether to engage in some form of cooperation with others, it is crucial whether we believe this person is trustworthy (Gambetta 2000). Societies with higher levels of trust tend to have better government institutions and thus higher economic growth (Knack and Keefer 1997; La Porta et al. 1997). Furthermore, measures of trust are important predictors of economic activity in such diverse areas as stock market participation (Guiso et al. 2008) and international trade and investments (Guiso et al. 2009). Unfortunately, not enough is known about the underlying determinants of trust and the reasons for the observed heterogeneity among people.

In this paper, we study in a life-course perspective (Elder 1998) whether a major shock on individual experiences early in life – specifically, the exposure to the experience of hunger – predicts trust later in adult life. We construct a measure of exposure to hunger in Germany after the Second World War (WWII) from data on caloric rations that were set monthly by

[1]This paper is joint work with Iris Kesternich, Department of Economics, University of Leuven, Naamsestraat 69, B-3000 Leuven, Belgium (e-mail: iris.kesternich@kuleuven.be); James P. Smith, The RAND Corporation, 1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138 (email: smith@rand.org); and Joachim K. Winter, Department of Economics, University of Munich, Ludwigstraße 33, 80539 Munich, Germany (e-mail:winter@lmu.de). The data used in this paper were made available to us by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW), Berlin.

the occupying forces in the four occupation zones. We combine these exposure measures with data from a random sample of the contemporary adult German population and document that individuals who were exposed to low caloric rations in their childhood and youth show significantly lower levels of trust many years after the hunger experience.

Our analysis is motivated by a recent wave of research showing that early-life circumstances and shocks, even those experienced *in utero*, do not only predict adult health outcomes (Almond and Currie 2011; Barker 1992, 2004; van den Berg et al. 2016) but also socio-economic outcomes later in life. Adult outcomes affected by such shocks as the experience of war or exposure to prolonged periods of hunger early in life include education and labor market status (Goodman et al. 2011; Jürges 2013; Kesternich et al. 2014), preferences for food consumption (Kesternich et al. 2015), the development of egalitarian motivations (Bauer et al. 2014), and subjective well-being (Bertoni 2015).[2]

This study is also related to an emerging literature in economics arguing that the societal and institutional environment an individual experiences over the course of her life can change preferences and expectations (Black et al. 2015; Fehr and Hoff 2011; Malmendier and Nagel 2011, 2016; Malmendier and Shen 2015). More specifically with respect to social preferences and trust, Putnam et al. (1994) and Bigoni et al. (forthcoming) argue that historical events have shaped preferences for cooperation in Southern and Northern Italy differentially and that the resulting preference heterogeneity can account for the fact that individuals in the South and North react differently to similar institutions and incentives. Also, East and West Germans have been shown to differ in measures of trust and convergence after reunification has been at most very slow, which has been interpreted as the political system shaping trust (Rainer and Siedler 2009).

Despite these recent advances, evidence on the determinants of heterogeneity in trust preferences is still scarce. In particular, it has been difficult to link variation in historical events to variation in trust preferences. In this paper, we exploit the natural experiment induced by variations in caloric supply across regions and over time in the years immediately after the end of WWII in Germany to study this link. Our analysis shows that about one-fifth of the differences in East and West Germany levels of trust go back to the stronger exposure to hunger of East Germans after WWII. The remainder of this paper is structured as follows. In Section 4.2, we describe the data, the measures of trust and exposure to hunger, and the statistical framework. Section 4.3 contains the results. We conclude with a discussion of our findings in Section 4.4.

[2]Hunger has also been shown to have effects in the very short run: In an experimental study by Aarøe and Petersen (2013), respondents with low glucose levels show stronger support for social welfare.

4.2 Methods

4.2.1 Data

We focus in this research on Germany because a high share of the German population was affected by regulated food supply during and after WWII – including especially low levels of caloric rations after WWII that have been shown to be associated with severe hunger. Our study relies on the possibility to match measures of childhood exposure to hunger to adult preference measures of trust.

Measures of adult trust are contained in the German Socio-Economic Panel (SOEP), an annual survey of the German population conducted since 1984.[3] The SOEP has the advantage that each wave is representative of the German population. Furthermore, a rich set of background measures, such as biographic and demographic information, education, and employment histories, are available. However, the panel does not contain direct information on the experience of hunger. Therefore, we enrich the SOEP data with historical data on official daily caloric rations during and after WWII. The historical data provide regional as well as temporal variation in caloric restrictions, and we link them to individual-level data in SOEP using the month of birth and state of residence in childhood.

4.2.2 Measures

Measures of trust. Survey measures of trust preferences are contained in the 2003 wave of the SOEP. Specifically, respondents were asked for their assessments of the following three statements regarding trust: (A) On the whole one can trust people. (B) Nowadays one can't rely on anyone. (C) If one is dealing with strangers, it is better to be careful before one can trust them. Respondents answer each question on a four-point scale from 1 (totally agree), 2 (agree slightly), 3 (disagree slightly) to 4 (totally disagree). These trust questions are similar to those used in the General Social Survey[4] and the World Values Survey[5].

For the SOEP, these earlier trust questions have been refined by Naef and Schupp (2009) to address a critique by Glaeser et al. (2000). Glaeser et al. criticize that these earlier measures are vague, abstract, and hard to interpret. They integrate survey measures of trust with

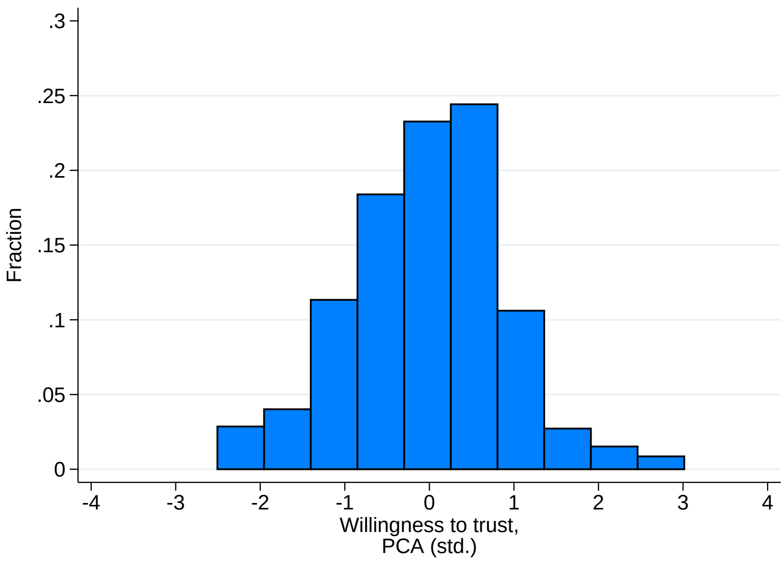
[3]Further information on the SOEP can be obtained online (http://www.diw.de/en/soep).

[4]"Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people? Most people can be trusted / Can't be too careful."

[5]"Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? Most people can be trusted / Need to be very careful."

experimental trust games and find that these standard attitudinal questions do not predict actual behavior in these trust experiments (but rather trustworthiness). Naef and Schupp (2009) develop a new survey measure of trust incorporating the three before mentioned questions. In doing so, they concentrate on James Colemans concept of trust (Coleman 1990) which considers trust rather as a form of behavior than a personal characteristic or trait.

Figure 4.1: Variation in standardized trust measure



Data source: German Socio-Economic Panel. ***Notes:*** Figure displays the trust measure for the hunger sample ($N = 6,503$). The trust measure is constructed as the standardized principal component of agreement with trust questions in the 2003 (general trust, reliance on others, need for caution in dealing with strangers).

They show that the three trust questions used in the SOEP provide a valid and reliable measure of trust in strangers (in contrast to trust in institutions and known others) and correlate with behavior in an incentivized trust game. In this respect, this form of trust appears to be especially relevant for human interactions. The behavioral relevance of this trust measure has been shown by Fehr et al. (2002). They combine the three aforementioned questions by means of factor analysis and show that this combined measure is a significant predictor of the amount that a first-mover sends to the other player in an incentive compatible trust game. Following Fehr et al. (2002) and Dohmen et al. (2012), we combine the answers to the three survey questions into a single trust measure by principal component analysis and standardize the resulting variable. A higher value for this constructed measure is associated with a higher willingness to trust others. We construct these measures for individuals potentially exposed to hunger in childhood. Figure 4.1 illustrates the variation of our trust measure for the hunger sample.

Measures of childhood exposure to hunger. Since the beginning of WWII, food production and distribution in Germany was organized centrally with yearly updates of caloric rations. Food and many other things could only be bought using food stamps. A food stamp could, for instance, indicate that one would be allowed (among others) to buy four pounds of bread and 62 grams of fat per week, as well as one egg per month. After the end of WWII, the administrations of the four occupation zones (British, U.S., French, and Soviet) separately set levels of caloric rations and updated them monthly. Food rationing was not eliminated until 1949 in the former British, U.S., and French zones, which form the Federal Republic of Germany, and 1956 in the German Democratic Republic.

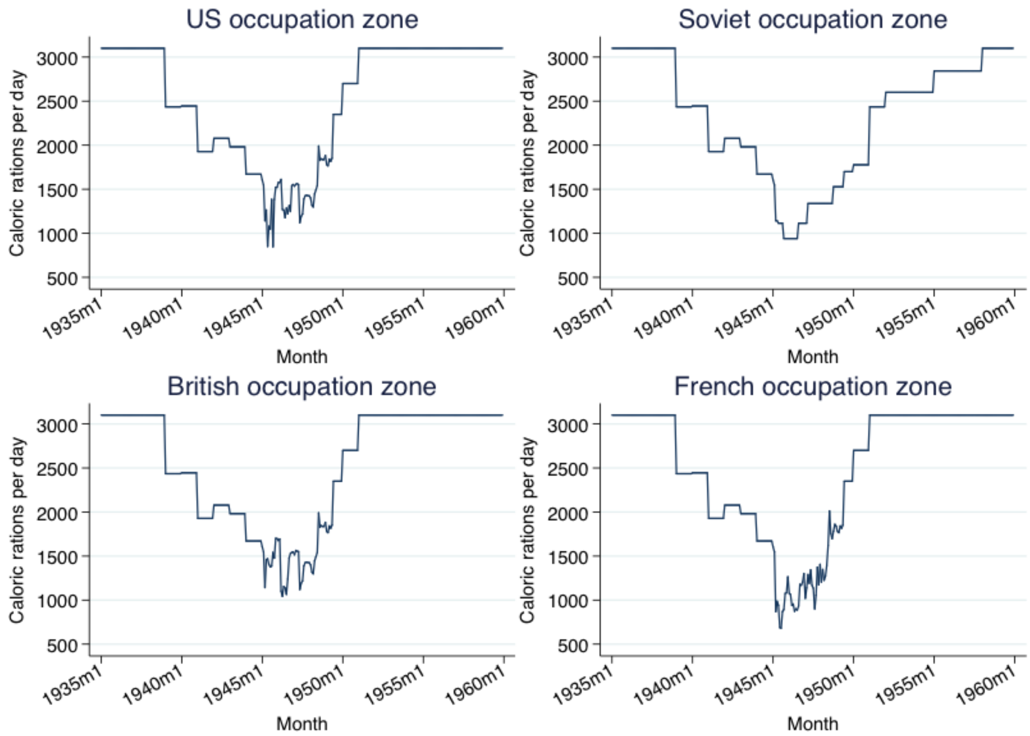
Severe famine conditions in Germany were initially reported in the summer of 1945 (Farquharson 1985), continued through the first years of the Allied occupation, and lasted until 1948. The famine was caused by a combination of supply-side reasons (missing inputs like seeds and fertilizer, missing workforce and machinery as well as losing the pre-war Eastern parts – the so-called food basket for Germany), demand-side reasons (increased population through an influx of refugees from the former Eastern parts of the German Reich), breakdowns in trade and transport within Germany, and the organization and governance of the occupying governments (see Kesternich et al. 2014, for further information).

In this study, we use the monthly area-dependent caloric measures from January 1929 to December 1971 (Kesternich et al. 2015). For the time before 1939 when food supply was not restricted, we use average caloric intake per day from the League of Nations for Germany (Liebe 1947).[6] From 1939 until the end of WWII, daily caloric rations did not differ by area, and new caloric rations were set in September. After the end of the war, caloric rations were updated every four weeks in all but the Soviet zone (Schwarzer, 1995). Rationing was eliminated in July 1948 in the Western zones (Rothenberger 1988) and not before 1956 in the Eastern zone (Schwarzer 1995). Figure 4.2 illustrates this source of regional as well as temporal variation in caloric rations by zone that we exploit in our study. Caloric rations dropped significantly in all zones early in 1945, with a most pronounced drop in the French and U.S. zone. It took much longer for the caloric rations to be back at pre-war levels in the Soviet occupation zone.

Official caloric rations are highly consistent with self-reports of experienced hunger (Kesternich et al. 2015). We compared our collected official caloric rations to self-reports on the existence and duration of experienced hunger that are included in SHARELIFE. SHARELIFE is a retrospective life history survey included in the cross-national European panel

[6]Reported and recommended caloric rations were about 3100 kcal per day (Liebe 1947)

Figure 4.2: Variation in caloric rations over time and occupation zones



Data source: Own collection (refer to the main text). ***Notes:*** Figure displays caloric rations per day for the four occupation zones.

survey SHARE that, however, does not include any measures of trust.[7] We observe a clear relationship between the temporal pattern of caloric rations and self-reported hunger. When caloric rations were not restricted yet in 1938, average daily caloric intake was about 3100 kcal while about 1% of the sample reported hunger. In 1944, caloric rations were set at about 1700 kcal per day and about 10% of the sample reported hunger. When the caloric rations drop to about 1000 kcal in 1945, we observe a spike in self-reported hunger at about 20%. In 1950, caloric rations were mostly back at pre-war levels (except for the Eastern part) and self-reported hunger decreased to about 5% (Kesternich et al. 2015).

As the government and the occupying administrations set caloric rations independently, the sources of restricted food supply were exogenous to the individual. Moreover, we have shown that the exposure to hunger is independent of socio-economic status (Kesternich et al. 2015). To investigate heterogeneous effects of hunger episodes at different stages of childhood, we generate age-specific average daily caloric intakes that depend on the month of birth as well as residence in childhood. We concentrate on being exposed to hunger at ages 0-3 (infant), 4-7 (child), and 8-16 (youth) and further explore the overall exposure at all ages 0-16. For instance, for somebody born in January 1940 and spending childhood in the part of Germany

[7]For a detailed description of SHARE and SHARELIFE see http://www.share-project.org/.

associated with the Soviet zone after WWII, we construct a calorie measure for ages 0-3 taking the average of caloric rations set between January 1940 and December 1943 within this area. In addition, we construct a measure for age-specific variability in daily caloric rations in a similar procedure calculating the standard deviation of the exposed caloric rations at the respective ages depending again on the month of birth and residence in childhood.

4.2.3 Statistical analysis

We estimate ordinary least squares (OLS) regression models in which the dependent variable is the standardized measure of trust obtained from the three survey questions by principal component analysis. To estimate the effect of exposure to hunger in childhood on trust preferences, our key explanatory variables are the average as well as the standard deviation of daily caloric intake at ages 0-16, 0-3, 4-7 and 8-16. In addition to analyzing the association between early life hunger and later life trust for the general population, we investigate the association separately for those growing up in an urban and rural environment. We estimate all models using robust SEs clustered by month of birth $\times$ occupation zone at childhood.

We construct our analytic data set from the 2003 wave of the SOEP, which includes the trust measure. We select 7,941 observations of those who participated in this wave; the selection criteria are possible exposure to WWII – that is being born between 1929 and 1955 and having German nationality or immigrated prior to 1949 (full sample hereafter). Exposure to restricted food supply at different ages in childhood varied by residence in the respective occupation zones as well as by month of birth. We merge the SOEP observations with our hunger exposure measure using the state of residence to identify the relevant occupation zone and the month of birth.

We obtain information on the state of residence in childhood using several mechanisms. First, the SOEP contains a question in 2012 about the town in which the respondent was born.[8] We obtain data on the corresponding state of residence and use this information for those individuals for which the state of residence can be matched to the indicated city with certainty. This information is available for 3,604 individuals (about 45%) of the full sample.

Second, we construct a proxy of the state of residence at childhood using two further sources of information. The SOEP contains a question measuring whether the respondent is still living in the same city or area where he or she spent the majority of childhood up to age 15. If the respondent indicates either "Yes, I still do" or "Yes, I have moved back", the

[8]What city or town were you born in? If there are more than one town with the same name, please also state the name of the nearest city.

respondent's current state of residence provides information on the state of residence in childhood (available for 51% of the full sample).

Furthermore, the SOEP collects information on the federal state of last school attendance (available for 46% of the full sample). We combine the information about the current residency matching the area of childhood as well as the state of school attendance and construct a measure for the state of residence in childhood giving the former information priority (N = 5,868). Comparing this proxy with the reliable state of birth obtained from wave 2012, we find that this information is consistent for about 87% of those individuals for which both information are available (N = 2,759). We use this constructed proxy of the state of residence in childhood for those individuals lacking information on the direct measure based on the town in which the respondent was born. We focus on those for which information on the state of residence in childhood, data on the month of birth, and the trust measure are available (N = 6,503, "hunger sample" hereafter). Note that during the hunger period (from 1945 to 1950), Germans were prohibited to relocate within the country (and abroad) by moving restrictions, so that moving in response to the caloric rations was not an option.

The average calorie measures are divided by 1000, and their signs are reversed so that a higher value is associated with a higher exposure to lower caloric rations that would imply more hunger. For example, a change in average daily caloric rations within a specified period in childhood from 2000 to 1500 calories translates into a change from -2.0 to -1.5 in our transformed variable. To ensure the comparability of the caloric measures based on the average exposure to hunger and the variability in caloric rations, we also divided the standard deviation by 1000. In this respect a higher value is associated with a higher variability of caloric rations.[9]

The data set is further enriched with background characteristics. The background measures contain information on the respondent's age, gender, whether growing up in an urban environment (equal to 1 if growing up in a large/medium/small city, equal to 0 if growing up in the countryside), as well as the education level of both parents (higher degree if parent obtained secondary, intermediate, or upper secondary school degree). Table 4.1 provides the descriptive statistics for the full and the hunger sample. We observe that the samples are very similar regarding their means and standard deviations.

[9]The correlation between the average caloric rations and the variability in caloric rations is 0.870 for age 0-16, 0.579 for age 0-3, 0.650 for age 4-7 and 0.812 for age 8-16.

Table 4.1: Descriptive statistics

	Full sample			Hunger sample		
Variable	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**
Outcome						
Willingness to trust, PCA (std.)	7854	0.00	1.00	6503	0.03	1.00
Calorie measures						
(Negative) average daily caloric intake at age 0-16				6503	-2.60	0.38
(Negative) average daily caloric intake at age 0-3				6503	-2.48	0.61
(Negative) average daily caloric intake at age 4-7				6503	-2.48	0.66
(Negative) average daily caloric intake at age 8-16				6503	-2.71	0.52
Standard deviation daily caloric intake at age 0-16				6503	0.47	0.28
Standard deviation daily caloric intake at age 0-3				6503	0.20	0.17
Standard deviation daily caloric intake at age 4-7				6503	0.18	0.17
Standard deviation daily caloric intake at age 8-16				6503	0.24	0.27
Background measures						
Male	7941	0.50	0.50	6503	0.51	0.50
Age	7941	59.82	7.54	6503	59.26	7.36
Urban area (childhood)	7795	0.60	0.49	6426	0.61	0.49
US zone (childhood)	6713	0.27	0.44	6503	0.27	0.44
UK zone (childhood)	6713	0.33	0.47	6503	0.33	0.47
French zone (childhood)	6713	0.06	0.23	6503	0.06	0.24
Soviet zone (childhood)	6713	0.34	0.47	6503	0.34	0.48
Father: Lower degree	7231	0.80	0.40	5962	0.80	0.40
Father: Higher degree	7231	0.20	0.40	5962	0.20	0.40
Mother: Lower degree	7339	0.86	0.35	6041	0.85	0.36
Mother: Higher degree	7339	0.14	0.35	6041	0.15	0.36

Data sources: German Socio-Economic Panel, own collection of calorie data (refer to the main text). ***Notes:*** The *full sample* contains individuals that participated in wave 2003 of the SOEP, were born between 1929 and 1955 and have German nationality or immigrated prior to 1949. The *hunger sample* contains individuals for which we can construct the trust measure and have data on the month of birth and state of residence in childhood via (1) reliable information on the state of birth contained in wave 2012, or a proxy using (2) the state in which they spent the majority of their childhood up to the age 15 or, if not available, (3) the state in which they last attended school.

4.3 Results

Table 4.2 presents the main coefficients from OLS regressions estimating the relationship between the respondent's willingness to trust and the average daily caloric intake at ages 0-16, 0-3, 4-7 and 8-16, respectively. The first row displays the raw correlations. Subsequent rows sequentially add explanatory variables. In the second row, we control for general background characteristics such as quadratic age, gender, growing up in an urban environment, and different occupation zones. We control for age and occupation zones to make sure that the effect of caloric rations does not merely pick up age and occupation zone effects. We include a dummy for growing up in an urban environment, as historically those living in urban areas where more affected by the hunger episode, and they might also be different regarding trust in strangers.

Table 4.2: OLS regression of adult trust on measures of childhood exposure to hunger

	Main control: Measures of daily caloric intake at			
	All ages	Specific ages (infant, child, youth)		
Regressors	0–16	Age 0–3	Age 4–7	Age 8–16
No other controls				
(Negative) average daily caloric intake	–.184***	–.040*	–.086***	–.103***
	(.034)	(.021)	(.019)	(.026)
+ ***Basic controls***				
(Negative) average daily caloric intake	–.201**	.024	–.054*	–.098
	(.097)	(.035)	(.030)	(.063)
+ ***Parental education***				
(Negative) average daily caloric intake	–.199**	.011	–.053*	–.082
	(.096)	(.035)	(.030)	(.063)
N	6503	6503	6503	6503

Data sources: German Socio-Economic Panel, own collection of calorie data (refer to the main text). ***Notes:*** Values are estimated main coefficients (SE) based on the hunger sample. The standardized principal component of agreement with trust questions in 2003 (general trust, reliance on others, need for caution in dealing with strangers) is used as the dependent variable. The first row displays the OLS regression of the trust measure on the measures of daily caloric intake with no other controls. Subsequent rows sequentially add variables. *Basic controls* contains controls for gender, standardized age and age^2, growing up in an urban environment as well as occupation zone fixed effects; *parental education* contains controls for whether each parent has obtained a higher degree (defined as intermediate, technical, or upper secondary school degree). All estimates are based on OLS regressions with robust standard errors clustered at the month of birth $\times$ occupation zone at childhood level.
Significance levels: *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

In the third row, we additionally include controls for the respondent's socio-economic background proxied by indicators whether each parent has obtained a higher degree of education. Our previous research (Kesternich et al. 2015) shows that this hunger episode affected all socio-economic classes to the same extent. Strictly speaking, we do not need to control for socio-economic background. However, as we do not have individual-level measures of hunger, including these measures helps to control for trends in the socio-economic background that might also be related to trust.

Our main interest lies in childhood exposure to hunger at ages 0-16. We find that experiencing hunger (lower caloric rations) significantly reduces our standardized measure for trust. In particular, a decrease in average daily caloric rations by 1000 kcal decreases our standardized trust measure by about 0.20 standard deviations (statistically significant at the 5 percent level).

We can further characterize the specific ages at which the general hunger experiences had the largest influence on trust. We find the strongest effects for hunger experienced as a child or in one's youth (at ages 4-7 and 8-16, respectively). The point estimates are always largest in the oldest age group. However, only for age group 4-7 the estimates remain significant including our controls (statistically significant at the 10 percent level).

In Table 4.5 in Appendix 4.A, we present robustness checks adding the standard deviation of daily caloric intake. Including both measures of daily caloric intake at the same time, we find that the average measures remain a significant predictor of trust at ages 0-16 and 4-7 (significant at the 5 and 10 percent level, respectively) while the variability of caloric rations does not seem to play a role.

Table 4.3: Distinguishing between effects of childhood exposure to hunger and living in the GDR

Regressors	**With calorie measure**	**Without calorie measure**
(Negative) average daily caloric intake at age 00–16 y	–.199**	
	(.096)	
Basic controls		
Male	.004	.005
	(.024)	(.024)
Age (std.)	.006	–.015**
	(.012)	(.006)
Age^2 (std.)	–.000	.000*
	(.000)	(.000)
Urban area	.028	.028
	(.027)	(.027)
US zone	.018	.018
	(.030)	(.030)
French zone	–.023	–.029
	(.057)	(.057)
Soviet zone	–.156***	–.199***
	(.037)	(.031)
Parental education		
Father: Higher degree	.199***	.199***
	(.037)	(.037)
Mother: Higher degree	.158***	.158***
	(.042)	(.042)
N	6503	6503

Data sources: German Socio-Economic Panel, own collection of calorie data (refer to the main text). ***Notes:*** Values are estimated coefficients (SE) based on the hunger sample. The regression in the first column is identical to the full regression in the first column in Table 4.2. All estimates are based on OLS regressions with robust standard errors clustered at the month of birth × occupation zone at childhood level. The regressions contain dummies for missing observations in the urban area and parents' education that are all insignificant.
Significance levels: *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

In what follows, we focus on the average childhood exposure to hunger at ages 0-16 and its association with adult trust. In Table 4.3, we present results from regressions with the full set of controls. We find that individuals with a higher socio-economic status, proxied by indicators whether each parent of the respondent has obtained a higher school degree, exhibit more trust in strangers. Also, individuals in Eastern Germany exhibit less trust in strangers. Thus, there is an Eastern Germany effect over and above the worse caloric rations. This effect has been documented in the literature and interpreted as an effect of communism on preferences (Rainer and Siedler 2009). When we compare the regression in column (1), including our hunger exposure measure, to the regression in column (2), excluding it, we conclude that about 22 percent of the Eastern Germany effect of lower levels of trust is due to restrictions in caloric rations.

Table 4.4: Exploring the differential effects of childhood exposure to hunger for rural and urban population

	Main control: (Negative) average daily caloric intake at age 0–16					
Regressors	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**
Calorie measure	–.272***	–.132***	–.114	–.295**	–.092	–.305**
	(.055)	(.043)	(.148)	(.131)	(.147)	(.128)
Basic controls						
Male			–.057	.042	–.056	.039
			(.039)	(.032)	(.039)	(.032)
Age 2003 (std.)			–.004	.019	–.011	.017
			(.017)	(.017)	(.017)	(.016)
Age squared 2003 (std.)			–.000	–.000	.000	–.000
			(.001)	(.000)	(.001)	(.000)
US zone (childhood)			–.063	.068	–.054	.051
			(.050)	(.042)	(.050)	(.042)
French zone (childhood)			–.029	–.084	–.028	–.074
			(.080)	(.081)	(.080)	(.080)
Soviet zone (childhood)			–.326***	–.066	–.316***	–.058
			(.061)	(.049)	(.060)	(.048)
Parental education						
Mother: Higher degree					.260***	.142***
					(.099)	(.048)
Father: Higher degree					.172**	.200***
					(.079)	(.043)
N	2504	3922	2504	3922	2504	3922

Data sources: German Socio-Economic Panel, own collection of calorie data (refer to the main text). ***Notes:*** Values are estimated coefficients (SE) based on the hunger sample that is split depending on whether the respondent grew up in an urban or rural environment (excluding those with missing information in this dimension, N=77). The regressions follow the procedure of sequentially adding variables as presented in Table 4.2. All estimates are based on OLS regressions with robust standard errors clustered at the month of birth × occupation zone at childhood level. The regressions contain dummies for missing observations in the parents' education that are all insignificant. ***Significance levels:*** *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

In Table 4.4, we investigate the association between childhood exposure to hunger and adult trust separately for those growing up in an urban and rural environment. We find that both the effect of caloric rations as well as growing up in the Soviet zone shows stark rural-urban differences. First, regarding the caloric restrictions we would expect them to have a stronger impact in urban than in rural regions, because cities were more strongly affected by the hunger episode than rural areas (see Kesternich et al. 2015). As soon as we add our basic controls, we find support for this historical fact: The effect of the caloric rations is smaller and not statistically significant in rural regions, while it shows significant negative association of about 0.30 standard deviations (significant at the 5 percent level) for those growing up in an urban environment. Second, the opposite holds true for the negative effect of living in the GDR on trust preferences: Here, the negative effect of living in the GDR on trust is large and statistically significant for rural regions, while it is small and insignificant for urban respondents. To shed more light on this phenomenon could be scope for further research.

4.4 Discussion

We have shown that exposure to a hunger episode during childhood permanently lowers trust. Investigating different measures of exposure to hunger in childhood, we have shown that rather the general exposure to hunger than the variability of caloric rations was the key factor. The data on trust come from a random sample of the German population in 2003 while the historical exposure data refer to the hunger crisis in Germany after, and mostly immediately after, WWII. Hunger affected all regions and all socio-economic classes, but caloric supplies showed substantial regional and temporal variation. Our results provide a new explanation for the substantial variation in trust preferences that has been observed in several studies. While previous research has shown that trust preferences can be altered in the short-run through the experience of violence, our data suggest that a large shock in a persons childhood can permanently alter trust preferences. While research has demonstrated that trust is a relatively stable adult trait (Caspi and Roberts 2001), our research deals with the early childhood formation of trust. Its subsequent stability in adulthood is very consistent with our perspective. Our results are also in line with other studies showing that childhood experiences can have long-run effects on preferences and expectations.

Do the effects we find for hunger really show the long-term effects of hunger or war per se? For our effects to pick up other war effects, war exposure would need to vary with caloric rations. However, these rations did vary by occupation zone and the zone borders were artificially created. The rations also reflect the disruption of production caused by the creation of the artificial borders, and the loss of the Eastern parts of the German Reich that traditionally used to be Germany's "bread basket". Of course the rations also reflect the destruction of Germany after WWII, but the monthly variation over time was caused by weather conditions, the influx of refugees, and policy choices of the occupying governments. The latter also reflects the food situation in their home countries and thus needs to confiscate food, in particular for the French and Soviet Zones. In an additional robustness check, we included a dummy for those who experienced WWII (born before 1945) in the analysis, finding an insignificant coefficient, while the calorie measure remained significant.

Unlike the experience of bombings that seems to have strengthened a feeling of helping and trusting among the citizens of the affected cities, the experience of extreme food shortages lead to a sense of everybody just fighting for their survival. This interpretation is in line with recent findings on the effects of early childhood health in psychology. Petersen and Aarøe (2015) argue that “low birth weight is utilized as a forecast of a harsh environment, vulnerable condition, or both and, consequently, lowers social trust”. In addition, there was strong mistrust in Germany against the occupying governments who were thought to either not be doing their best to relieve the situation or even actively holding back on or confiscating food supplies. The shock experienced through this long hunger period was strong enough to permanently lower trust.

We discussed that trust plays an important role for human interactions both in the interpersonal and economic context. Considering our results, we would like to stress that the long-run consequences of war might be underestimated when being reduced to financial dimensions due to destruction of buildings, infrastructure, businesses, and so forth. A further dimension to keep in mind should be “soft factors” like trust that affect interactions within society that play a role in the economic area (negotiations, transactions, teamwork at the workplace), but also the private area (interactions with neighbors, finding friends, finding a partner). The large-scale consequences of such shock-induced preference-changes remain scope for future research.

Our findings stand in contrast to the seminal work of Barker (1992) who emphasizes that the largest impact of childhood shocks should be concentrated among the very young. However, Barker studies biological impacts of under-nutrition on health, while we emphasize the behavioral channel, namely, long-term effects on preferences. Our results can either mean that older ages are the sensitive periods for the foundation of trust preferences, or that younger children were protected from the hunger episode. The former interpretation would be in line with recent evidence from Sutter and Kocher (2007) who measure trust preferences at different ages and find that “trust increases almost linearly from early childhood to early adulthood”. One has to note, though, that their study excludes the very young. The latter interpretation is consistent with very young children being better protected as they were still being breastfed.

This finding also adds to the discussion on long-terms effects of war and conflicts. First, experiments conducted around the 2006 Israel-Hezbollah War show that intra-group cooperation was increased during times of conflict. Here, we use a large random sample of a population to show that trust in strangers (in contrast to intra-group preferences as in Gneezy and Fessler 2012) was decreased and that these effects were long-lasting and thus are still present more than 50 years after the experienced shock. Prior research on the long-term effects of WWII show that health, marriage, and labor market outcomes of those surviving are negatively affected (Jürges 2013; Kesternich et al. 2014). If hardship that comes with violent conflicts, in our case hunger, permanently alters preferences, including trust, that are to a certain extent inherited within families (Dohmen et al. 2012; Ljunge 2014), then this could perpetuate negative equilibriums.

4.A Appendix Tables

Table 4.5: Robustness check: General exposure to hunger and variations in exposure to hunger

	Main control: Measures of daily caloric intake at			
	All ages	**Specific ages (infant, child, youth)**		
Regressors	**0–16**	**Age 0–3**	**Age 4–7**	**Age 8–16**
No other controls				
(Negative) average daily caloric intake	–.357***	–.049*	–.093***	–.106**
	(.066)	(.026)	(.025)	(.043)
Standard deviation daily caloric intake	.267***	.053	.039	.008
	(.088)	(.092)	(.099)	(.079)
+ ***Basic controls***				
(Negative) average daily caloric intake	–.208**	.023	–.057*	–.091
	(.097)	(.035)	(.031)	(.070)
Standard deviation daily caloric intake	.076	.106	.042	–.025
	(.117)	(.102)	(.105)	(.090)
+ ***Parental education***				
(Negative) average daily caloric intake	–.203**	.011	–.053*	–.079
	(.096)	(.035)	(.031)	(.070)
Standard deviation daily caloric intake	.038	.118	.011	–.009
	(.117)	(.101)	(.103)	(.089)
N	6503	6503	6503	6503

Data sources: German Socio-Economic Panel, own collection of calorie data (refer to the main text). ***Notes:*** Values are estimated main coefficients (SE) based on the hunger sample. The regressions follow the procedure of sequentially adding variables as presented in Table 4.2. All estimates are based on OLS regressions with robust standard errors clustered at the month of birth × occupation zone at childhood level.

Significance levels: *** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

References

Aarøe, L. and Petersen, M. B. (2013). Hunger games: Fluctuations in blood glucose levels influence social welfare support. *Psychological Science*, 24(12):2550–2556. 101

Adler, G. S. (1994). A profile of the Medicare Current Beneficiary Survey. *Health Care Financing Review*, 15(4):153–163. 18

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500. 1, 8, 11

Almond, D. and Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, 25(3):153–172. 101

Aron-Dine, A., Einav, L., and Finkelstein, A. (2013). The RAND health insurance experiment, three decades later. *The Journal of Economic Perspectives*, 27(1):197–222. 1, 14

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5):941–973. 1, 8, 11

Baicker, K., Congdon, W. J., and Mullainathan, S. (2012). Health insurance coverage and take-up: Lessons from behavioral economics. *The Milbank Quarterly*, 90(1):107–134. 1, 47, 82

Baicker, K. and Goldman, D. (2011). Patient cost-sharing and healthcare spending growth. *Journal of Economic Perspectives*, 25(2):47–68. 9, 14

Barcellos, S. H., Wuppermann, A. C., Carman, K. G., Bauhoff, S., McFadden, D. L., Kapteyn, A., Winter, J. K., and Goldman, D. (2014). Preparedness of Americans for the Affordable Care Act. *Proceedings of the National Academy of Sciences*, 111(15):5497–502. 2, 8, 12, 47, 48, 61, 80, 82

Barker, D. J. P. (1992). *Fetal and infants origins of disease*. BMJ Publishing Group, London. 101, 113

Barker, D. J. P. (2004). The developmental origins of adult disease. *Journal of the American College of Nutrition*, 23(6):588S–595S. 101

Bauer, M., Cassar, A., Chytilova, J., and Henrich, J. (2014). War's enduring effects on the development of egalitarian motivations and in-group biases. *Psychological Science*, 25(1):47–57. 101

Bertoni, M. (2015). Hungry today, unhappy tomorrow? Childhood hunger and subjective wellbeing later in life. *Journal of Health Economics*, 40:40–53. 101

Bhargava, S., Loewenstein, G., and Sydnor, J. (2015). Do individuals make sensible health insurance decisions? Evidence from a menu with dominated options. NBER Working Paper No. 21160. 2, 3, 12, 43, 47, 64, 82

Bigoni, M., Bortolotti, S., Casari, M., and Gambetta, F. (forthcoming). Amoral familism, social capital, or trust? The behavioral foundations of the Italian North-South divide. *The Economic Journal*. 101

Black, S. E., Devereux, P. J., Lundborg, P., and Majlesi, K. (2015). On the origins of risk-taking. NBER Working Paper No. 21332. 101

Blendon, R. J. and Benson, J. M. (2014). Voters and the Affordable Care Act in the 2014 Election. *New England Journal of Medicine*, 371(20):e31(1)–(7). 83

Braun, A., Schmeiser, H., and Schreiber, F. (2016). On consumer preferences and the willingness to pay for term life insurance. *European Journal of Operational Research*, 253(3):761–776. 44, 46

Carman, K. G. and Eibner, C. (2015). Insurance transitions following the first ACA open enrollment period. RAND Working Paper no. RR-948-RC. 81

Carman, K. G., Eibner, C., and Paddoch, S. M. (2015). Trends in health insurance enrollment, 2013-2015. *Health Affairs*, 34(6):1044–1048. 80, 82, 89

Caspi, A. and Roberts, B. W. (2001). Personality development across the life course: The argument for change and continuity. *Psychological Inquiry*, 12(2):49–66. 112

Cassidy, A. (2011). Health Policy Brief: Putting limits on Medigap. *Health Affairs*, September 13, 2011. 33

Cawley, J. and Philipson, T. (1999). An empirical examination of information barriers to trade in insurance. *American Economic Review*, 89(4):827–846. 13

Chandra, A., Gruber, J., and McKnight, R. (2010). Patient cost-sharing and hospitalization offsets in the elderly. *American Economic Review*, 100(1):193–213. 1, 15

Chiappori, P.-A. and Salanié, B. (2000). Testing for asymmetric information in insurance markets. *Journal of Political Economy*, 108(1):56–78. 11, 23

Chiappori, P.-A., Salanié, J. B. B., and Salanié, F. (2006). Asymmetric information in insurance: General testable implications. *The RAND Journal of Economics*, 37(4):783–798. 13

Christensen, S. and Shinogle, J. (1997). Effects of supplemental coverage on use of services by Medicare enrollees. *Health Care Financing Review*, 19(2):5–17. 15

CMS (2015). Medicaid and CHIP: December 2014 monthly applications, eligibility determinations and enrollment report. Centers for Medicare and Medicaid Services, Baltimore, MD. 82

Cohen, A. (2005). Asymmetric information and learning: Evidence from the automobile insurance market. *The Review of Economics and Statistics*, 87(2):197–207. 11, 12, 13, 15, 33

Cohen, A. and Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance*, 77(1):39–84. 9, 11, 12, 13, 14, 23

Coleman, J. S. (1990). *Foundations of Social Theory*. Belknap Press of Harvard University Press, Cambridge, MA. 103

Collins, S. R., Rasmussen, P. W., and Doty, M. M. (2014). Gaining ground: Americans' health insurance coverage and access to care after the Affordable Care Act's first open enrollment period. Commonwealth Fund, New York, NY. 80, 82

Crooker, J. R. and Herriges, J. a. (2004). Parametric and semi-nonparametric estimation of willingness-to-pay in the dichotomous choice contingent valuation framework. *Environmental and Resource Economics*, 27(4):451–480. 45

Cunningham, P. J., Denk, C., and Sinclair, M. (2001). Do consumers know how their health plan works? *Health Affairs*, 20(2):159–166. 2, 8, 13, 47

Cutler, D. and Reber, S. (1998). Paying for health insurance: the trade-off between competition and adverse selection. *The Quarterly Journal of Economics*, 113(2):433–466. 1, 13

Cutler, D. and Zeckhauser, R. (2000). The anatomy of health insurance. In *Handbook of health economics, Volume 1*. 9, 11, 13, 14

Cutler, D. M., Finkelstein, A., and McGarry, K. (2008). Preference heterogeneity and insurance markets: Explaining a puzzle of insurance. *American Economic Review: Papers and Proceedings*, 98(2):157–162. 12, 43, 47, 49, 80

Cutler, D. M. and Lleras-Muney, A. (2006). Education and health: Evaluating theories and evidence. *NBER Working Paper No. 12352*. 2, 12

Davidson, B. N., Sofaer, S., and Gertler, P. (1992). Consumer information and biased selection in the demand for coverage supplementing medicare. *Social Science & Medicine*, 34(9):1023–1034. 15

de Meza, D. and Webb, D. C. (2001). Advantageous selection in insurance markets. *The RAND Journal of Economics*, 32(2):249–262. 2, 9, 12, 47, 49, 83

de Walque, D. (2010). Education, information, and smoking decisions: Evidence from smoking histories in the United States, 1940-2000. *Journal of Human Resources*, 45(3):682–717. 12

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550. 49, 82

Dohmen, T., Falk, A., Huffman, D. B., and Sunde, U. (2012). The intergenerational transmission of risk and trust attitudes. *Review of Economic Studies*, 79(2):645–677. 103, 114

Elder, G. H. (1998). The life course as development theory. *Child Development*, 69(1):1–12. 100

Ericson, K. M. and Starc, A. (2012). Heuristics and heterogeneity in health insurance Exchanges: Evidence from the Massachusetts Connector. *American Economic Review: Papers and Proceedings*, 102(3):493–497. 64

Ettner, S. L. (1997). Adverse selection and the purchase of Medigap insurance by the elderly. *Journal of Health Economics*, 16(5):543–562. 14

Fang, H., Keane, M., and Silverman, D. (2008). Sources of advantageous selection: Evidence from the Medigap insurance market. *Journal of Political Economy*, 116(2):303–350. 10, 14, 18, 19, 20, 24, 25, 35

Farquharson, J. E. (1985). *The Western allies and politics of food. Agrarian management in Postwar Germany.* Berg Publishers, Dover, NH. 104

Fehr, E., Fischbacher, U., von Rosenbladt, B., Schupp, J., and Wagner, G. G. (2002). A nation-wide laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative surveys. *Schmollers Jahrbuch: Journal of Applied Social Science Studies*, 122(4):519–542. 103

Fehr, E. and Hoff, K. (2011). Introduction: Tastes, castes and culture: The influence of society on preferences. *Economic Journal*, 121(556):F396–F412. 101

Finkelstein, A. and McGarry, K. (2006). Multiple dimensions of private information: Evidence from the long-term care insurance market. *American Economic Review*, 96(47):938–956. 2, 9, 12, 13, 44, 47, 49, 64, 83

Finkelstein, A. and Poterba, J. (2004). Adverse selection in insurance markets: Policyholder evidence from the U.K. annuity market. *Journal of Political Economy*, 112(1):183–208. 13

Finkelstein, A. and Poterba, J. (2014). Testing for asymmetric information using "unused observables" in insurance markets: Evidence from the U.K. annuity market. *Journal of Risk and Insurance*, 81(4):709–734. 11

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., and Baicker, K. (2012). The Oregon health insurance experiment: Evidence from the first year. *The Quarterly Journal of Economics*, 127(3):1057–1106. 15, 32

Gallup (2014). Politics affect uninsured Americans' insurance intentions. Gallup, Princeton, NJ. 83

Gambetta, D. (2000). Can we trust trust? In Gambetta, D., editor, *Trust: Making and Breaking Cooperative Relations*, pages 213–227. University of Oxford Press. 100

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L. (2000). Measuring trust. *The Quarterly Journal of Economics*, 115(3):811–846. 102

Gneezy, A. and Fessler, D. M. T. (2012). Conflict, sticks and carrots: War increases prosocial punishments and rewards. *Proceedings of the Royal Society B: Biological Sciences*, 279(1727):219–223. 113

Goodman, A., Joyce, R., and Smith, J. P. (2011). The long shadow cast by childhood physical and mental problems on adult life. *Proceedings of the National Academy of Sciences*, 108(15):6032–6037. 101

Green, D., Jacowitz, K. E., Kahneman, D., and McFadden, D. (1998). Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resource and Energy Economics*, 20(2):85–116. 46

Guiso, L., Sapienza, P., and Zingales, L. (2008). Trusting the stock market. *Journal of Finance*, 63(6):2557–2600. 100

Guiso, L., Sapienza, P., and Zingales, L. (2009). Cultural biases in economic exchange? *The Quarterly Journal of Economics*, 124(3):1095–1131. 100

Gutsche, T. L., Kapteyn, A., Meijer, E., and Weerman, B. (2014). The RAND continuous 2012 presidential election poll. *Public Opinion Quarterly*, 78(S1):233–254. 81, 90

He, D. (2009). The life insurance market: Asymmetric information revisited. *Journal of Public Economics*, 93(9-10):1090–1097. 13

Heiss, F., McFadden, D., and Winter, J. (2006). Who failed to enroll in Medicare Part D, and why? Early results. *Health Affairs*, 25(5):w344–w354. 2, 12, 15, 44, 47, 82

HHS (2015). Health insurance Marketplaces 2015 open enrollment period: March open report. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, Washington, DC. 82

Hogan, C., Lunney, J., Gabel, J., and Lynn, J. (2001). Medicare beneficiaries' cost of care in the last year of life. *Health Affairs*, 20(4):188–195. 19

Hurd, M. D. and McGarry, K. (1997). Medical insurance and the use of health care services by the elderly. *Journal of Health Economics*, 16(2):129–154. 14, 15

Hurd, M. D., Smith, J. P., and Zissimopoulos, J. M. (2004). The effects of subjective survival on retirement and social security claiming. *Journal of Applied Econometrics*, 19(6):761–775. 12

Johnson, E. J., Hassin, R., Baker, T., Bajger, A. T., and Treuer, G. (2013). Can consumers make affordable care affordable? The value of choice architecture. *PLOS ONE*, 8(12):1–6. 65

Johnson, E. J., Hershey, J., Meszaros, J., and Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7(1):35–51. 2, 47

Jürges, H. (2013). Collateral damage: The german food crisis, educational attainment and labor market outcomes of german post-war cohorts. *Journal of Health Economics*, 32(1):286–303. 101, 113

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292. 2, 47

Keane, M. and Stavrunova, O. (2014). Adverse selection, moral hazard and the demand for medigap insurance. Nuffield College Economics Working Paper no. 2012-W10. 15

Kenkel, D. S. (1991). Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 99(2):287–305. 12

Kesternich, I., Heiss, F., McFadden, D., and Winter, J. (2013). Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. *Journal of Health Economics*, 32(6):1313–1324. 44, 45, 46, 64

Kesternich, I., Siflinger, B., Smith, J. P., and Winter, J. K. (2014). The effects of World War II on economic and health outcomes across Europe. *The Review of Economics and Statistics*, 96(1):103–118. 101, 104, 113

Kesternich, I., Siflinger, B., Smith, J. P., and Winter, J. K. (2015). Individual behavior as a pathway between early-life shocks and adult health: Evidence from hunger episodes in Post-War Germany. *The Economic Journal*, 125(588):F372–F393. 101, 104, 105, 109, 111

KFF (2013a). Health insurance coverage of the total population. Kaiser Family Foundation, Washington, DC, http://kff.org/other/state-indicator/total-population/. 2, 16

KFF (2013b). Medigap: Spotlight on enrollment, premiums and recent trends. Kaiser Family Foundation, Washington, DC. 17, 19

KFF (2014a). Employer health benefits survey 2014. Kaiser Family Foundation, Washington, DC. 2, 51

KFF (2014b). Medicare at a glance. Kaiser Family Foundation, Washington, DC. 16, 17

KFF (2014c). Medigap reform: Setting the context for understanding recent proposals. Kaiser Family Foundation, Washington, DC. 16, 17, 33

KFF (2015a). California's previously uninsured after the ACA's second open enrollment period. Kaiser Family Foundation, Washington, DC. 80

KFF (2015b). Data note: Predictors of positive and negative attitudes towards the ACA among non-group insurance enrollees. Kaiser Family Foundation, Washington, DC. 83

KFF (2015c). Kaiser health tracking poll: April 2015. Kaiser Family Foundation, Washington, DC. 83

KFF (2015d). The Medicare Part D prescription drug benefit. Kaiser Family Foundation, Washington, DC. 2

KFF (2015e). Medigap enrollment among new Medicare beneficiaries: How many 65-year olds enroll in plans with first-dollar coverage? Kaiser Family Foundation, Washington, DC. 16

King, M., Ruggles, S., Alexander, J. T., and Warren, R. (2015). Integrated public use microdata series, current population survey: Version 4.0. [machine-readable database]. Minneapolis: University of Minnesota. 49, 81

Knack, S. and Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics*, 112(4):1251–1288. 100

La Porta, R., Lopez-de Silanes, F., Schleifer, A., and Vishny, R. W. (1997). Trust in large organizations. *American Economic Review*, 87(2):333–338. 100

Lemieux, J., Chovan, T., and Heath, K. (2008). Medigap coverage and Medicare spending: A second look. *Health Affairs*, 27(2):469–477. 15

Lewbel, A., McFadden, D., and Linton, O. (2011). Estimating features of a distribution from binomial data. *Journal of Econometrics*, 162(2):170–188. 5, 43, 44, 45, 53, 54, 55, 63, 65, 66, 67, 69, 70

Liebe, H. (1947). Drei Grundprobleme der deutschen Wirtschaft. In Deutsches Institut für Wirtschaftsforschung, editor, *Die deutsche Wirtschaft zwei Jahre nach dem Zusammenbruch. Tatsachen und Probleme.* Nauck, Berlin, Germany. 104

Liebmann, J. and Zeckhauser, R. (2008). Simple humans, complex decisions, subtle subsidies. NBER Working Paper No. 14330. 1, 47, 82

Ljunge, M. (2014). Trust issues: Evidence on the intergenerational trust transmission among children of immigrants. *Journal of Economic Behavior & Organization*, 106:175–196. 114

Loewenstein, G., Friedman, J. Y., McGill, B., Ahmad, S., Linck, S., Sinkula, S., Beshears, J., Choi, J. J., Kolstad, J., Laibson, D., Madrian, B. C., List, J. A., and Volpp, K. G. (2013). Consumers' misunderstanding of health insurance. *Journal of Health Economics*, 32(5):850–862. 2, 8, 12, 13, 47

Long, S. K., Karpman, M., Shartzer, A., Wissoker, D., Kenney, G. M., Zuckerman, S., Anderson, N., and Hempstead, K. (2014a). Taking stock: Health insurance coverage under the ACA as of September 2014. Urban Institute, Washington, DC. 80, 82

Long, S. K., Kenney, G. M., Zuckerman, S., Goin, D. E., Wissoker, D., Blavin, F., Blumberg, L. J., Clemans-Cope, L., Holahan, J., and Hempstead, K. (2014b). The health reform monitoring survey: Addressing data gaps to provide timely insights into the Affordable Care Act. *Health Affairs*, 33(1):161–167. 2, 12, 47

Lusardi, A. and de Bassa Scheresberg, C. (2013). Financial literacy and high-cost borrowing in the United States. NBER Working Paper No. 18969. 48, 80, 82

Lusardi, A. and Mitchell, O. S. (2011a). Financial literacy and retirement planning in the United States. *Journal of Pension Economics and Finance*, 10(4):509–525. 48, 49, 80, 82

Lusardi, A. and Mitchell, O. S. (2011b). Financial literacy around the world: An overview. *Journal of Pension Economics and Finance*, 10(04):497–508. 29

Lusardi, A. and Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, 52(1):5–44. 48

Malmendier, U. and Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk-taking? *The Quarterly Journal of Economics*, 126(1):373–416. 101

Malmendier, U. and Nagel, S. (2016). Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87. 101

Malmendier, U. and Shen, L. S. (2015). Experience effects in consumption. Unpublished working paper. 101

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., Leibowitz, A., and Marquis, M. S. (1987). Health insurance and the demand for medical care: Evidence from a randomized experiment. *American Economic Review*, 77(3):251–77. 1, 14

MCBS (2010). Medicare current beneficiary survey, CY 2010 cost and use, documentation. 18, 19

McCall, N., Rice, T., and Sangl, J. (1986). Consumer knowledge of Medicare and supplemental health insurance benefits. *Health Services Research*, 20(6 Pt 1):633–657. 15

Miyazaki, H. (1977). The rat race and internal labor markets. *The Bell Journal of Economics*, 8(2):394–418. 11

Murphy, J. J., Allen, P. G., Stevens, T. H., and Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics*, 30(3):313–325. 44, 46

Naef, M. and Schupp, J. (2009). Measuring trust: Experiments and surveys in contrast and combination. SOEPpapers on Multidisciplinary Panel Data Research 167, The German Socio-Economic Panel (SOEP), DIW Berlin. 102, 103

Nayga, R. M. (2000). Schooling, health knowledge and obesity. *Applied Economics*, 32(7):815–822. 12

Pauly, M. V. (1968). The economics of moral hazard: Comment. *American Economic Review*, 58(3):531–537. 1, 8, 11

Petersen, M. B. and Aarøe, L. (2015). Birth weight and social trust in adulthood: Evidence for early calibration of social cognition. *Psychological Science*, 26(11):1681–1692. 113

Puelz, R. and Snow, A. (1994). Evidence on adverse selection: Equilibrium signaling and cross-subsidization in the insurance market. *Journal of Political Economy*, 102(2):236–257. 13

Putnam, D., Leonardi, R., and Nanetti, Y. (1994). *Making democracy work: Civic traditions in modern Italy.* Princeton University Press, Princeton. 101

Quincy, L. (2012). Measuring health insurance literacy: A call to action. Health insurance literacy expert roundtable at Kaiser Family Foundation, Washington, DC. 48

Rainer, H. and Siedler, T. (2009). Does democracy foster trust? *Journal of Comparative Economics*, 37(2):251–269. 101, 110

Riley, J. G. (1979). Informational Equilibrium. *Econometrica*, 47(2):331–359. 11

Rothenberger, K. (1988). *Die Hungerjahre nach dem Zweiten Weltkrieg.* Boldt Verlag, Boppard am Rhein, Germany. 104

Rothschild, M. and 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. 1, 8, 11, 47

Schram, A. and Sonnemans, J. (2011). How individuals choose health insurance: An experimental analysis. *European Economic Review*, 55(6):799–819. 65

Schwarzer, O. (1995). Der Lebensstandard in der SBZ/DDR 19945/1989. *Jahrbuch für Wirtschaftsgeschichte*, 1995(2):119–145. 104

Smith, J. D. and Medalia, C. (2015). Health insurance coverage in the United States: 2014 U.S. Census Bureau report P60-253. 80, 81, 82

Sommers, B. D., Gunja, M. Z., Finegold, K., and Musco, T. (2015a). Changes in self-reported insurance coverage, access to care, and health under the Affordable Care Act. *The Journal of the American Medical Association*, 314(4):366–374. 80, 90

Sommers, B. D., Maylone, B., Nguyen, K. H., Blendon, R. J., and Epstein, A. M. (2015b). The impact of state policies on ACA applications and enrollment among low-income adults in Arkansas, Kentucky, and Texas. *Health Affairs*, 34(6):1010–1018. 83

Sommers, B. D., Musco, T., Finegold, K., Gunja, M. Z., Burke, A., and McDowell, A. M. (2014). Health reform and changes in health insurance coverage in 2014. *New England Journal of Medicine*, 371(9):867–874. 80, 82

Spence, M. (1978). Product differentiation and performance in insurance markets. *Journal of Public Economics*, 10(3):427–447. 11

StataCorp (2015). Stata Statistical Software: Release 14. StataCorp LP, College Station, TX. 83

Sutter, M. and Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior*, 59(2):364–382. 113

Sydnor, J. (2010). (Over)insuring modest risks. *American Economic Journal: Applied Economics*, 2(4):177–199. 64

van den Berg, B., Van Dommelen, P., Stam, P., Laske-Aldershof, T., Buchmueller, T., and Schut, F. T. (2008). Preferences and choices for care and health insurance. *Social Science and Medicine*, 66(12):2448–2459. 45

van den Berg, G. J., Pinger, P. R., and Schoch, J. (2016). Instrumental variable estimation of the causal effect of hunger early in life on health later in life. *Economic Journal*, 126(591):465–506. 101

van Rooij, M., Lusardi, A., and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, 101(2):449–472. 82

Vistnes, J. and Cohen, S. B. (2015). Transitions in health insurance coverage over time, 2012-2014 (selected intervals): Estimates for the U.S. civilian noninstitutionalized adult population under the age of 65. Agency for Healthcare Research and Quality, Rockville, MD. 80

Wilson, C. (1977). A model of insurance markets with incomplete information. *Journal of Economic Theory*, 16(2):167–207. 11

Witters, D. (2015). In U.S., uninsured rates continue to drop in most states. Gallup, Washington, DC. 80

Wolfe, J. R. and Goddeeris, J. H. (1991). Adverse selection, moral hazard, and wealth effects in the Medigap insurance market. *Journal of Health Economics*, 10(4):433–459. 14

Zeckhauser, R. (1970). Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *Journal of Economic Theory*, 2(1):10–26. 11