Essays in the Economics of Health and Wellbeing

Hannes Schwandt

TESI DOCTORAL UPF / ANY 2012

DIRECTORS DE LA TESI

Robin Hogarth, Departament d'Economia i Empresa,
Universitat Pompeu Fabra

Jörn-Steffen Pischke, Department of Economics,
London School of Economics and Political Science

Acknowledgments

I am deeply grateful to my supervisors Robin Hogarth at Universitat Pompeu Fabra and Steve Pischke at London School of Economics. Being supervised by two such experienced, accessible and kind scholars made my PhD studies an exciting time full of scientific stimulus and thrilling learning. Robin guided me throughout all stages of my PhD and provided me with great intellectual freedom to develop my research interests. Steve became my co-supervisor during my time as visiting student at LSE and he has greatly supported me since then. It is difficult to overstate the positive (causal) effect that his caring supervision had on my development.

I am thankful to Ada Ferrer-i-Carbonell and Richard Layard who introduced me to wellbeing research at the beginning of my PhD and supported me ever since. A lot of the research reflected in this thesis has been conducted as a visitor at the Center for Economic Performance and I am thankful to Richard Layard for making this arrangement possible.

My research has benefited from comments of many professors and colleagues, including Andrew Clark, Dora Costa, Gabrielle Fack, Guy Mayraz, Joan Monras, Johannes Müller-Trede, Adriana Lleras-Muney, Yona Rubinstein, James Smith, Claudia Steinwender, Thijs van Rens, Till von Wachter, and Amelie Wuppermann. The second thesis chapter has benefited from a research visit at RAND and I am thankful to Arie Kapteyn and Mike Hurd for making this stay possible. Further, Marta Araque and Laura Agusti at the GPEFM have always been of tremendous help.

Finally, I would like to thank Li, my parents and my friends who have accompanied and supported me throughout the past five years.

Abstract

The three chapters of this thesis investigate different aspects of the economics of health and wellbeing. The first chapter tests the rationality of life satisfaction forecasts. Contrary to the rational expectations hypothesis it shows that young people and those who are satisfied strongly overpredict future life satisfaction while the elderly and the unsatisfied strongly underpredict it. The second chapter is about how wealth shocks affect the health of retirees in the US. Results indicate strongly positive effects on physical health, mental health and mortality. The third chapter analyzes the effects of graduating in an unfavorable economic environment on graduates' subsequent income, health insurance and mortality. It finds that recession graduates have significantly lower incomes and worse health insurance coverage. And during the outbreak of the HIV/AIDS epidemic AIDS mortality has been significantly higher among these unlucky cohorts.

Resumen

Los tres capítulos de esta tesis doctoral investigan aspectos de la economía del bienestar y de la salud. El primer capítulo pone a prueba la racionalidad de las predicciones de las personas respecto a la satisfacción global que experimentarán con su vida en el futuro. Se muestra que, en contra de la hipótesis de expectativas racionales, las predicciones de los jóvenes y de quien está satisfecho con su vida son más altas que los niveles de satisfacción realizados posteriormente, mientras las predicciones de las personas mayores y de quien no está satisfecho con su vida son más bajas que los niveles posteriormente experimentados. El segundo capítulo investiga cómo los cambios exógenos de riqueza afectan la salud de una muestra de jubilados en los EEUU. Los resultados indican efectos positivos de la riqueza sobre la salud, tanto física como mental, y un efecto negativo sobre la mortalidad. El tercer capítulo analiza los efectos de graduarse de la universidad en un entorno económico recesivo sobre la salud, la riqueza, y la mortalidad. Graduarse en tiempos de recesión tiene efectos negativos persistentes sobre el salario, la cobertura médica, y - durante la epidemia del SIDA - tambien sobre la mortalidad.

Preface

Health and wellbeing are important aspects of our lives and at the center of many public policies. This doctoral thesis consists of three independent chapters which investigate different aspects of the economics of health and wellbeing. And in the first chapter I analyze errors in people's forecasts about their future life satisfaction. The second chapter is about how wealth shocks affect the health of retirees in the US. In the third chapter I investigate the effects of graduating in an unfavorable economic environment on socio-economic outcomes and mortality.

While there is no direct link between the first and the latter two chapters, all of them are empirical analyses exploiting panel dimensions in the analyzed data. In the first chapter I use the German Socio-Economic Panel to relate people's five-year forecasts of their life satisfaction to the actually realized levels of life satisfaction. In the second chapter I follow retirees in the Health and Retirement Study over time and analyze how their health changes depending on the amount of individual stock holdings and the stock market development. In the third chapter I use repeated cross-sections from the Current Population Survey and the universe of deaths from the US Vital Statistics. The large sample sizes of these data sets allows me to follow cohorts - defined by state and year of birth - over time and analyze how the local unemployment rate around the time of graduation affects these cohorts' income, health insurance coverage and mortality.

Empirical research is typically most useful when it answers causal questions. In the second and third chapter I investigate causal effects of wealth shocks and differences in socio-economic status on health outcomes. How-

ever, correlational studies can also be of great interest if they present fundamental relationships that are at odds with central assumptions made in scientific research. The first chapter claims to be a study of this type.

In the first chapter I explore whether people make systematic forecast errors when being asked about the satisfaction with their lives in 5 years time future life satisfaction. The rational expectations hypothesis assumes that we are on average right and our forecast errors uncorrelated with any information at the time of the forecast. This chapter provides evidence that this might not be the case. I match expected life satisfaction with its later realizations for 15 waves of the German Socio-Economic Panel. The resulting errors are regressed on a broad set of socio-economic variables at the time of the forecast, with a focus on age and current life satisfaction. I find that the young and the very satisfied strongly overpredict their future life satisfaction while the elderly and the very unsatisfied underpredict it. These associations are not driven by economy-wide shocks or cohort effects and therefore contradict the rational expectations hypothesis. Further, estimated effects are highly predictive. Age and life-satisfaction differentials in forecast errors estimated for 1991-1993 predict more than 93% of these differentials in 2000-2002. To the extend that life satisfaction can be interpreted as a proxy for utility these findings might have wide-ranging implications for economic analysis and public policy.

The second and third chapters are about the relationship between wealth and health, and in particular on the direction of causality. The positive correlation between the two is well documented in the economic literature. However, little is known about the causal mechanisms underlying this relationship. I use innovative approaches to identify exogenous variation in wealth and to measure its impact on health outcomes.

In the second chapter, I exploit stock market fluctuation in the wealth of US retirees to analyze the effect of wealth shocks on health. Using data from the Health and Retirement Study I construct exogenous wealth shocks as the interaction of stock holdings with stock market changes. Such constructed wealth shocks have not been used in the literature before. They are highly predictive of actual changes in reported wealth. And they strongly affect health. A 10% wealth shock is associated with an improvement of 2-4% of a standard deviation in physical health, mental health and survival rates. Analyzing individual health conditions I find a strong effect on high blood pressure, smaller effects on heart diseases and strokes and no effect on arthritis, diabetes, lung disease and cancer. The comparison with benchmark regressions indicates that the effect pattern across health outcomes is different from the overall relationship of health and wealth in the data. While wealth shocks affect only particular health conditions, wealth levels have a positive association of similar magnitude with all of them. And for the affected conditions the impact of wealth shocks is 1.5 to 3 times as large as the overall relationship.

In the third chapter I investigates the causal effects of wealth on health from a complementary perspective. While the second chapter is about short-term effects of wealth shocks on elderly health, this chapter is about the long run effects of gradually accumulating differences in income on the mortality of young adults.

Recent studies have found that the unemployment rate graduates face when entering the labor market has a strongly negative and persistant effect on subsequent income. In this chapter I investigate whether this arguably exogenous variation in income and socio-economic status is related to subsequent mortality. Using data from the CPS, I first show that

graduating in a recession is not only associated with lower subsequent income but also with worse health insurance coverage. These effects are quite homogeneous across gender, but much stronger for non-whites than for whites. Next, I show that results carry over when I use the unemployment rate at age 18 instead of the actual graduation age. This allows me as the next step to analyze effects on mortality in the Vital Statistics, that do not report graduation age. In my baseline period, 1979-1991, I find strong positive effects of the unemployment rate at age 18 on mortality at ages 28-33. This effect is mostly driven by AIDS deaths, it is similar across gender but stronger for non-whites. The effect fades out when adding more years to the analysis and is not distinguishable from zero in the overall period from 1979 to 2004. These results suggest that the unemployment rate around graduation affected the mortality of young adults in the US only during the outbreak of the AIDS epidemic. I argue that the negative effects on income and health insurance coverage are the most likely mechanisms underlying this effect.

Contents

List of Figures

List of Tables

Chapter 1

TESTING THE RATIONALITY OF LIFE SATISFACTION FORECASTS: MICRO EVIDENCE FROM THE GERMAN SOCIO-ECONOMIC PANEL.

1.1 Introduction

Rational expectations are a fundamental assumption in many economic models with widespread implications for economic analysis and public policy. It is therefore not surprising that the rational expectations hypothesis (REH) has been discussed and empirically tested since its first

formulation. In recent years the increasing availability of easily accessible panel data has lead to a new wave of economic studies that examine the formation of expectations and the determinants of forecast errors in large samples. Most of these tests focus on expectations about objectively observable choice and market variables like income, consumption, stock market prices or inflation. The REH, however, does not only refer to the unbiased prediction of choices. Rational expectations also imply unbiased prediction of the utility that results from these choices and market outcomes. Since traditionally utility functions were assumed to be stable, economic analysis has so far devoted little attention to the rationality of utility expectations. If tastes do not change, utility is directly inferable from choices, and rational expectations in choice imply rational expectations in utility. But if utility functions are not stable over time, we might not only mispredict choices but also the resulting utility. In this case a test of the REH should allow for both potential sources of biased expectations - biased expectations of what one will get *and* how one will like it.

In psychology there exists a widely scattered literature that analyzes errors in predicting changes in tastes and feelings. These studies on so-called "hedonic forecasting" do the exact counterpart to economic REH testing. They usually use subjective utility data and analyze individual expectations about the utility function, while holding the corresponding choices stable. The collective evidence suggests that tastes change and that these changes are often not correctly anticipated.

For an extensive test of the REH we therefore need to test not only for biased expectations of the choice variables but also allow for biased expectations of the underlying utility function. In the tradition of Rob-

bins' (1932) revealed preferences approach standard economic analysis is mostly focused on objectively observable choices. Utility is derived from these choices but not directly analyzed in an empirical way. In the last decade, however, there has been a growing literature that uses self-reported (subjective) well-being measures as a proxy for utility. A vast number of studies has shown that such measures like self-reported life satisfaction provide insightful information about individual well-being that can and should be used in economic analyses (for a review see Clark, Frijters, and Shields 2008a, or Frey and Stutzer 2002)[1].

In the present study I build on this literature and use self-reported expected and current life satisfaction from the German Socio-Economic Panel (GSOEP) to test the REH. This allows me to include both potential sources of forecast errors in the test, i.e. biased forecasts of choice variables and of the corresponding utility function. Combining individual life satisfaction expectations (for 5 years ahead) with its later realizations over a panel of 17 years, I construct 136,000 forecast errors of 24,000 individuals over 12 observation years (1991 - 2002). These forecast errors I regress on different sets of socio-economic variables, with a focus on age and life satisfaction at the time of the forecast. Under the REH none of these variables should correlate with the forecast error. In different empirical specifications I then examine the role of economy-wide shocks, relax the assumption of cardinality, and test the predictive power of the model. Finally I present related evidence from Cantril (1965) to assess whether

[1] In recent years, subjective happiness measures have been increasingly used in very applied contexts, e.g. to calculate exact income compensation associated with non-monetized factors (such as informal care (van den Berg and Ferrer-i-Carbonell, 2007), unemployment (Kassenboehmer et al., 2009) or air quality (Luechinger, 2009)) and to infer precise policy implications (on the importance for public policy, see Layard (2006)

results generalize to other countries and times.

The remainder of the paper is organized as follows. Section 1.2 reviews the literature, Section 1.3 develops the empirical approach and Section 1.4 describes my data. Section 1.5 and 1.6 present the findings and conclude.

1.2 Literature Overview

1.2.1 The Rational Expectations Hypothesis

First it is important to clarify what is meant by the REH in the literature and how it can be tested. The literature represents the REH commonly as (e.g. Begg, 1982):

$$\begin{aligned} E_t X &= E_t(X \mid I_t) \\ \Leftrightarrow E([E_t X - X] \mid I_t) &= 0 \\ \Leftrightarrow cov([E_t X - X], I_t) &= 0 \end{aligned} \tag{1.1}$$

with:
$E_t X$: expectation at time t of variable X
I_t: information set available at time t according to the relevant economic theory

The REH states, that "expectations ... are essentially the same as the

predictions of the relevant economic theory" (Muth, 1961, p.316). This implies that rational expectations are unbiased and efficient: Forecast errors will have a zero conditional mean and they will be uncorrelated with information available to individual i at time t. Note that this implies that the average forecast error converges to zero as $T \to \infty$ but not necessarily as $N \to \infty$. At a certain period forecast errors of $N \to \infty$ individuals might contain common components due to economy-wide shocks.[2] In the following I will call an individual's expectations "irrational", if they are biased over various periods and/or not all individual information have been used to form them. If forecasts and realizations of X and the information sets are known for various periods, we can test the REH running the following regression:

$$E_t X - X = \beta I_t + \epsilon_t \tag{1.2}$$

with $\beta = 0$ under the REH, given sufficiently high T.

1.2.2 Economic literature on REH testing

The Rational Expectations Hypothesis (REH) has been discussed and tested since its formulation by Muth (1961)[3]. Nevertheless only few pa-

[2]For example, if an unexpected earthquake hits the population, forecasts that were formed before the earthquake will be on average too high even if people have rational expectations. And if the shock is not distributed evenly in the population it will not only affect the regression intercept but also the coefficients on the respective explanatory variables.

[3]Muth (1962) is commonly considered as the "'father of the REH"'. However, it was Tinbergen (1932) who - thirty years earlier - first formulated this hypothesis (see Keuzenkamp, 1991)

pers have tested the REH for households or individuals[4]. The first studies on that issue examined aggregated inflation expectations of consumers, using the Michigan Index of Consumer Sentiment (ICS) (e.g. Maddala et al., 1981 or Batchelor, 1986). Their results are not reliable tests for individual rationality due to microheterogeneity and the resulting aggregation bias. Experimental tests of the REH in the 1990s could not rejected the hypothesis in some contexts, but their most common outcome is that individuals do not hold rational expectations (Dwyer et al., 1993; Swenson, 1997). In recent years a few studies used larger panels to analyze micro-level forecast errors of large numbers of households or individuals. Das and van Soest (1999) analyze income expectations of Dutch households, concluding that "either household expectations are not rational, or macroeconomic shocks occur in a number of consecutive years, or both". These authors do not include age as an explanatory variable. Souleles (2004) uses the Michigan ICS data to analyze household expectations over a fairly long time horizon (20 years). He finds that people repeatedly underestimated disinflation and the severity of business cycles and therefore concludes that household expectations are partly biased and inefficient. Age, even though not in focus of the study, is included and seems to have a strong effect on forecast errors. The older people are, the more they tend to underestimate future financial positions, business conditions, income and inflation. Using the same data, Carroll (2002) finds potentially rational expectation dynamics of households that generate 'stickiness' in aggregate expectations while Branch (2004) develops a rationally heterogeneous expectations hypothesis to explain observed biases in inflation forecasts. Carroll (2002) and Branch (2004) reject the

[4]Early REH tests concentrated on expectations of firms or professional forecasters (for a review see Lovell, 1986).

pure REH but find support of their extended form of the REH. An important exception not rejecting the REH is a recent study Benitez-Silva et al. (2008) who find evidence for the rationality of expectations over education and retirement decisions.

To sum up so far relatively few studies have analyzed forecast errors of households and individuals on a large scale. Further the economic REH literature usually focuses on forecasts over market and choice variables, excluding possible forecast errors due to preferences misprediction. Finally age and the current wellbeing state have not been included or -if so- have not been analyzed in greater detail.

1.2.3 Psychological literature on hedonic forecasting

In psychology a widely scattered literature examines expectation and misprediction of tastes and feelings. Most relevant for this study is the psychological research on so called hedonic (or affective) forecasting, which focuses on the question how people expect that certain life circumstances or choices affect them in future.

Reviewing the relevant literature on hedonic forecasting Loewenstein and Schkade (1999), Gilbert (2006) and Kahneman and Thaler (2006) conclude that people generally underestimate the extent of hedonic adaptation to new life circumstances.[5] Further, utility mispredictions are likely

[5]For example, Kahneman (2000) suggests that the adaptation to winning the lottery as well as to becoming paraplegic after an accident are broadly underestimated. Similarly, Loewenstein and Frederick (1997) find that people expected future changes in life circumstances to affect their overall well-being much more than they believed that matched changes in the past had affected their well-being.

to arise from related biases as the "projection bias" and an (unanticipated) "endowment effect" (Loewenstein et al., 2000).[6]

1.2.4 Happiness research on life satisfaction prediction

First data related to life satisfaction forecasts were collected within a world-wide survey conducted by Cantril around 1960 (see Cantril, 1965), in which respondent had to indicate a step on a present and expected (5 years ahead) individually defined life ladder. Though life ladder rankings represent a different well-being measure, the Cantril data are compared the life satisfaction measure in the GSOEP, to assess the degree to which results from the GSOEP might be generalized to other countries and years (see Section 5). Note, that as a single cross-section in each country Cantril's evidence cannot be used to directly test the REH. One cannot compare a forecast with the same individual's later realization to form individual forecast errors. And even if we have representative data on life satisfaction for the years following Cantril's survey (e.g. it is a well-established fact that average life satisfaction in most western countries remained flat over the whole observation period (Easterlin, 1974) and life cycle happiness usually is found to be flat or u-shaped[7]), it is not possible to rule out common elements forecast errors. Instead a large panel like the GSOEP covering both expected and realized life satisfaction over a long time period is necessary.

[6] However, so far few economic models have tried to incorporate these biases or explain them within the standard economic framework. Only in a recent study, Müller-Trede (2008) shows how random errors in individual utility forecasts can lead to various biases, including the "adaptation bias" and the "projection bias".

[7] A non-exhaustive list includes Clark and Oswald, 1994; Oswald, 1997; Winkelmann and Winkelmann, 1998; Di Tella et al., 2001; Blanchower and Oswald, 2004; Frijters et al., 2004; Senik, 2004; and Shields and Wheatley Price, 2005.

So far only Frijters et al. (2008) have analyzed expected and realized subjective life satisfaction from an earlier version of the GSOEP. These authors focus on forecast errors in East Germany during the time of transition following the German reunification. They find that during the aftermath of the reunification East Germans, the young and the low educated were too optimistic about their future life satisfaction, but forecast errors decreased over the observation period. The authors conclude that during times of transition people are especially bad at predicting their future. However, the observed patterns are likely to be driven by economy-wide shocks and therefore do not necessarily tell much about the rationality of expectations.

Easterlin (2001) reviews the Cantril (1965) survey and concludes that "people typically think that ... they will be better off in the future, although their reports on present happiness remain constant over time" (p. 472). He develops a model in which individuals mispredict utility because they do not anticipate that increasing income leads to increasing aspirations. Even though the findings of the present paper do not entirely match Easterlin's conclusion from the Cantril survey, it offers more detailed evidence on life satisfaction misprediction. This evidence might contribute to test and further develop Easterlin's model. Lacey et al. (2006) surveyed a small subsample (273 individuals) of the US census, in which participants had to rate their own happiness and estimate the happiness of the average 30-year old and the average 60-year old. Their findings suggest that people tend to overestimate the happiness of the average 30-year old and to underestimate happiness of the average 60-year old.

Finally, various studies of Frey and Stutzer (e.g. Frey et al., 2005; Stutzer and Frey, 2007) find evidence of utility misprediction combining life satisfaction data with individual economic decisions like commuting time or watching television. Another branch of papers analyzes the habituation in wellbeing to major life events (e.g. Clark et al., 2008b; or Gardner and Oswald, 2007). However, these studies analyze current wellbeing only, which makes it difficult to evaluate to which extend such habituation is anticipated.

1.3 Empirical Approach

1.3.1 Test specification

In Section 1.2.1 a general test of the REH has been developed. Following this approach we can test the rationality of utility expectations, if we have data on expected utility (e.g. five periods ahead) and its later realizations available:

$$E_t(u_{t+5}(x_{t+5})) - u_{t+5}(x_{t+5}) = \beta I_t + \epsilon_t \tag{1.3}$$

with $\beta = 0$ under the null hypothesis of rational expectations, given sufficiently high T. Forecast errors can arise from misprediction of both future choice variables, x_{t+5}, and the future utility function, u_{t+5}. Economic studies so far have concentrated on testing the rationality of expectations about x_{t+5}, while psychological evidence suggested expectations about u_{t+5} as additional source of biased predictions. Above's test, using expected and realized *utility*, accounts for both possible sources of misprediction, though it cannot disentangle between them. The test examines

who has wrong expectations but not whether they arise from mispredictions of x_{t+5} or u_{t+5} or both.

This study uses expected and realized self-reported life satisfaction (for 5 periods ahead, ie. $E_{i,t}LS_{i,t+5}$ and $LS_{i,t+5}$) as a proxy for expected and realized utility. Further a subset of $I_{i,t}$ ($SOE_{i,t} \subset I_{i,t}$) is observed, consisting of various socio-economic variables. Following the empirical set-up derived above this allows to test:

$$E_{i,t}LS_{i,t+5} - LS_{i,t+5} = \beta SOE_{i,t} + \epsilon_{i,t} \tag{1.4}$$

Note that the dependent variable as a difference cancels out any individual fixed effects that affect both expected and current (and future) life satisfaction. The final shape of the econometric specification is given by the study's focus on the effects of age and current life satisfaction. To allow for a flexible relationship between these variables and the dependent variable, I include 14 age-dummies (referring to age group 17-19, 20-24, 25-29, ... , 80-84) and 11 life satisfaction-dummies (referring to LS = 0, 1, ..., 10).

$$E_{i,t}LS_{i,t+5} - LS_{i,t+5} = \alpha + \beta Age_{i,t} + \gamma LS_{i,t} + \delta X_{i,t} + \epsilon_{i,t} \tag{1.5}$$

with:
Age: a vector of (14-1) dummies representing age groups
LS: a vector of (11-1) dummies representing states of life satisfaction
X: socio-economic controls of 7 categories: (1) **general**: male (female), (low)/middle/high education (2) **job**: in education, unemployed, employed,

pensioner (not employed), apprentice, self-employed, white collar, civil servant (blue collar), temporary (permanent), labor income, tenure (3) **financial**: house income, pension, stocks (4) **family**: household size, married, marriedsep., divorced, widowed, (single), invalid in hh, baby born past year, marriage past year, divorce past year (5) **health**: health status, handicap degree (6) **federal**: 16 federal states, federal gdp, federal u-rate (7) **further**: entry and exit dummies.

Under the null hypothesis of rational expectations and with a sufficient number of observation years we have $\alpha = \beta = \gamma = \delta = 0$. In other words, the REH is rejected if any of these coefficient diverts significantly from zero. Notice, that these coefficients do not have a causal interpretation.[8] Forecast errors might be caused by socio-economic factors, but one could also think of potential reverse causality. For example, higher contemporaneous income might lead to overoptimism when thinking about the future. But overoptimism might also help people getting into better paid jobs. Further, there could be unobserved factors that affect both simultaneously. However, the REH is a hypothesis about correlations and not about causalities and a robust correlation is sufficient to reject the REH.

1.3.2 Dealing with economy-wide shocks

A central challenge of testing the REH using short panels are economy-wide shocks. As pointed out above, even if all individuals have rational expectations, the average of the forecast error over N individuals does

[8]Throughout this paper the term "effect" is used to describe significant coefficients and not a causal relationship.

not need converge to zero if $N \to \infty$. There may be common components in forecast errors in a certain year or period, due to economy-wide shocks. Only the time average of the forecast error must converge (under the REH) to zero as $T \to \infty$ (Chamberlain, 1984, p. 1311). For the present study only 12 consecutive years of observed forecast errors are available. As T=12 might be too small to average out possible economy-wide shocks, observed departures from zero average forecast error (over the whole sample or any subsample) simply reflect aggregate shocks. Further it is very unlikely that aggregate shocks affect the whole population equally. To control for potential interaction effects between time and population groups, I therefore run regressions over subperiods to see whether significant patterns in the forecast error stem only from a short period or whether they are present in all subperiods. If the latter is the case, we can reject the REH with high probability as it is very unlikely that in each period the same shock or a new shock with exactly the same effect structure occurs.

1.3.3 Further issues

The main test specification (equation [1.5]) applies Pooled OLS to analyze the forecast errors. This assumes that expected and realized subjective life satisfaction are cardinal measures. In our case, cardinality cannot be relaxed by simply applying ordinal regression methods like ordered logit. The forecast error (as dependent variable) is the difference of two subjective measures and therefore cardinality is assumed already by generating the dependent variable. However, forecast errors can be redefined into binary variables indicating the sign of the error which implies a great loss of information but allows us to discard cardinality and thereby test

the credibility of the OLS results.

There is no direct mathematical relationship of forecast errors ($E_t(LS_{t+5})$-LS_{t+5}) with current life satisfaction at the time of the forecast (LS_t). Even though the reporting scale is bounded (0-10), somebody very satisfied (LS_t=10) can commit any forecast error from -10 ($E_t(LS_{t+5})$=0 and LS_{t+5}=10) to +10 ($E_t(LS_{t+5})$=10 and LS_{t+5}=0). However, LS_t and $E_t(LS_{t+5})$ might be measured with errors. Due to the bounded scale measurement errors would be negatively correlated with real values, leading to regression to the mean in reported values. Such regression to the mean effects can generate the findings of this study (that the satisfied are too optimistic) if they are larger for reports of LS_t (and LS_{t+5}) than for reports of $E_t(LS_{t+5})$[9]. This does not seem very plausible. If anything, measurement/reporting errors (and thereby regression to the mean effects) should be larger for expected than for current values, as the computation of expectations involves more effort. Larger regression to the mean effects in $E_t(LS_{t+5})$, however, would bias the estimated effect of current life satisfaction downwards, making the positive estimate found in this study more conservative.

The GSOEP is an evolving panel and therefore attrition is present and might lead to biases. Further, some studies have detected an entry bias for the GSOEP, with higher self-reported life satisfaction in the period in which an individual enters the panel (e.g. Frijters, 2005). To control for

[9] In such a case we would observe - even under rational expectations - for true high [low] values of LS^*_t, LS^*_{t+5} and $E^*_t(LS_{t+5})$ on average reports of $E_t(LS_{t+5}) > LS_{t+5}$ [$E_t(LS_{t+5}) < LS_{t+5}$]

such biases, entry and exit dummies are included in all regressions. However, the inclusion of these dummies does not change the estimated effect pattern of age and current life satisfaction.

1.4 Data and Measure

1.4.1 Data

To test the REH on the basis of direct, subjective utility, I use individual data on expected (for five years ahead) and realized life satisfaction from the German Socio-Economic Panel (GSOEP).[10] The GSOEP is a nationally representative panel that has closely followed around 10,500 West Germans each year since 1984 and includes East Germany since 1990. For the present study I use the sub-sample from 1991 to 2007, as only in 1991 expected life satisfaction has been included in the surveys. The number of missing values is fairly small, 98% of all observation report expected life satisfaction and 99.7% report current life satisfaction. To generate forecast errors I match expectations at time t with respective realizations at time $t + 5$. Excluded from the analysis are those who remained in the panel for less than 5 years and those who entered the panel for the first time post-2002 (for whom I do not have data on realizations). Further I exclude individuals of age above 84. Table 1.1 shows the waves for which actual and expected life satisfaction measures are available and the subsample that is used to generated the corresponding forecast errors.

[10]The exact wording of the question is:
"Please answer according to the following scale: 0 means 'completely dissatisfied', 10 means 'completely satisfied'.
-How satisfied are you with your life, all things considered?
-And how do you think you will feel in five years?"

In total I generate 136,618 forecast errors of 23,965 individuals (age 17 - 84) from 12 observation years (1991 - 2002). Federal GDP/c and unemployment rates are taken from Eurostat.

1.4.2 Measure

In recent years there has been a growing literature that uses self-reported (subjective) life satisfaction measures to proxy utility. This literature (the so-called ´´Happiness Research”) documented an increased willingness of economists to use subjective direct utility data (e.g. Layard, 2005; or van Praag et al., 2004). This development has been made possible by the vast number of studies documenting the reliability of such subjective utility data. These studies have collected evidence, that 1) people can generally tell how happy or satisfied they are (usually 99% response rates, Layard; 2003), 2) self-reported happiness corresponds to objectively measurable brain activity (Davidson, 2000) 3) different measures of happiness (e.g. self-reported life satisfaction, well-being, depression) correlate well with each other (Fordyce et al., 1988; Luttmer, 2004), 4) happiness reports are comparable over time and across individuals (e.g. Ng, 1997), 5) subjective happiness data predicts future observed behavior (see Clark, Frijters and Shields (2008a) for an overview) and 6) "assuming ordinality or cardinality of happiness scores makes little difference” Ferrer-i-Carbonell and Frijters, 2004). In the present study I first assume cardinality of self-reported expected and realized life satisfaction (which allows me to use all the information of the forecast error) and then I test whether results hold if only ordinality is assumed.

Regression results in Table 1.2 suggest that expectations about life sat-

isfaction have predictive power about future outcomes besides its correlation with present values of life satisfaction. Current and expected life satisfaction are reported on a 0 to 10 scale. In the presence of measurement errors this boundness can lead to regression to the mean effects. As argued above estimated effects should not be driven by this potential type of bias.

1.5 Findings

1.5.1 Main findings

Before presenting regression results, Figures 1.1-1.4 illustrate average forecast errors over different variables. Figure 1.1 shows the average forecast error over time. At the beginning and at the end of the observation period average forecast errors are positive while in the middle year (1994 to 1997) they are close to zero. The departures from zero average forecast errors are likely to represent macro shocks as they match with the transition period after the German reunification (1991 to 1993) and with Germany's New Economy stock market bubble (1998 to 2001). These two mayor shocks had a different structure which can be seen from the different effect on East and West Germans.[11] Figures 1.2 and 1.4 show the average forecast error over age and over life satisfaction levels, the two central variables of this study. Figures 1.3 and 1.5 illustrate the decomposition in forecasts and later realizations. In the following I test whether the patterns in forecast errors in Figures 1.2 and 1.4 are statistically signif-

[11]The high forecast errors in East-Germany after reunification is in line with Senik (2004) who documents the changing relationship between subjective well-being and income inequality in transition countries.

icant, if they coexist and persist controlling for socio-economic variables and if they are driven by macro shocks or other underlying mechanisms.

Table 1.3 shows the results of three regressions of the forecast error on different sets of explanatory variables. In column 1 forecast errors are regressed on age groups only. In column 2 dummies for current life satisfaction are added, and in column 3 time stable and time varying controls are additionally included.

Coefficients in Column 1 indicate that the effects of age on forecast errors are highly significant. Forecast errors in my sample are constant between age 17 and 34, then strictly decrease until the mid-60s and afterwards stay around that level of the mid-60s. Considering the positive constant, forecast errors in our sample are on average positive for the young (overestimation of future life satisfaction) and negative for the elderly (underestimation of the future). This is in line with negative age effects estimated in the recent literature (Souleles, 2004; Frijters et al., 2008). However, past studies often concluded that people learn over time and older people simply commit less forecast errors. The present results suggest instead that forecast errors of the elderly are negative. Thus they are not necessarily doing a better job in predicting their future than the young, they just seem to be less optimistic about it.

In column 2 of Table 1.3 (11-1) dummies for life satisfaction at the time when the prediction was made are included. The results show that age effects hardly change. Therefore, age effects are not driven by differences in current life satisfaction. Instead, at any state of life satisfaction, the young tend to overestimate their future life satisfaction and the elderly

tend to underestimate it. Secondly, the results describe a highly significant effect of current life satisfaction on the forecast error. The effect is positive, linear and very strong, with an average forecast error difference of three units between the very unsatisfied and the very satisfied. This suggest that the very unsatisfied dramatically underestimate their future life satisfaction (they are too pessimistic), while the very satisfied overestimate it (they are too optimistic). Both groups seem to underestimate the extend to which their life satisfaction gets "back to normal", ie. they underestimate that life circumstances will not stay as extreme as they currently are, or they underestimate how much they will get used to them. This in line with the main conclusion of the psychological literature on hedonic forecasting: people underestimate the extend of hedonic adaptation. These effects are not driven by age.

As column 3 shows, age and life satisfaction effects persist if I include a large set of time stable and time varying controls. This persistence is not surprising for the estimated age effects. Time stable variables hardly change over the life cycle and therefore cannot explain differences between age groups in a representative sample. Time varying variables, other other hand, may heavily depend on age (e.g. health), but their influence on expectations and future life satisfaction is already (partly) captured by current life satisfaction. The estimated effects of the life satisfaction dummies on the other hand become stronger if we include the whole set of variables. This suggests that current life satisfaction is not just capturing the effects on the forecast error of observable socio-economic variables like income, health or marital status, but contains also other dimensions.

Summarizing, Table 1.3 suggests that age and current life satisfaction have significant effects on the forecast errors. These effects occur separately of each other and persist if various socio-economic controls are included.

Table 1.4 reports regression results for three different subperiods in order to test whether the estimated effects are driven by macro shocks in certain periods. As explained above Figure 1.1 describes three time period with clearly differentiated average forecast errors, which are likely to present different macro shocks. Therefore the subperiods in Table 1.4 were chosen to match these three periods. Regression results report that the dummy for East Germany is strongly positive in the transition period after of the German reunification (1991-1993), and vanishing to zero during the time of Germany's New Economy stock market bubble (1998 - 2002).[12] As already discussed above this suggests that different macro shock were at work in these subperiods and that these shocks were not evenly distributed in the population. However, the estimated effects of age and life satisfaction on forecast errors only change slightly between the periods and the overall effect pattern remains the same. Thus the significant results in the previous table were not driven by phenomena in certain periods. Certainly the shocks are not completely evenly distributed among different ages and people of different life satisfaction, but the overall effect patterns remain the same, regardless of the period. That makes it very likely that age and life satisfaction effects do not arise from

[12]This is not surprising. Germany's reunification lead to (ex-post) unjustified optimism in all of Germany, but expectations had been especially unfulfilled in the East. Germany's New Economy stock market bubble on the other hand was driven rather from the rich and "capitalistic" West.

common error components due to economy-wide shocks but represent time-independent effects. As under rational expectations, neither age nor current life satisfaction should be time-independently correlated with the forecast error, this strongly suggests the rejection of the REH.

1.5.2 Additional analysis

Predictive power

In Figures 1.6 and 1.7 the predictive power of a simple model is tested to illustrate the robustness of the estimated effects over different subperiods. A regression specification that only includes age groups and life satisfaction dummies is used to estimate their effects for the subperiod 1991-1993. The resulting coefficients are used to predict individual forecast errors for the subperiod 2000-2002. Finally the average over age and over life satisfaction levels is calculated for predicted and actual 2000-2002 forecast errors. As Figures 1.6 and 1.7 show, the predicted curves fit very well the actual ones. The averages over age groups (over current life satisfaction) of the predicted forecast errors explain 92.3% (95.9%) of the variation in the average of the realized forecast errors. Note that the two subperiods are disconnected by 6 years, which means that age groups in the two subperiods contain completely different individuals.

Relaxing cardinality

Regressions in Table 1.5 test whether the estimated effects depend on the cardinality assumption. To relax cardinality, the discrete forecast errors are transformed into a binary variable that indicates whether forecast errors are i) positive or negative (column 1); ii) positive or zero (column 2);

iii) negative or zero (column 3). As we can see in column 1, the probability to switch from a negative to a positive forecast error decreases over age until age 69, where after it stays around the level of the 60's. In other words young people tend to be too optimistic while the elderly rather see their future too pessimistic. Regression results in column 2 suggest that the probability to commit no error (correct expectations) instead of a positive one follows the same age pattern. Thus the older a person, the more likely she will be realistic instead of too optimistic about her future. That the elderly are not simply better at forecasting, can be seen in column 3. The probability to commit a negative error (being too pessimistic) instead of a zero error is highest for people at age 55 and above. The effect of life satisfaction on the direction of the forecast error is very strong and consistent. People who are rather satisfied with their life, tend to be (i) rather too optimistic (pos. error) than too pessimistic (neg. error) (ii) or realistic (zero error) and (iii) rather realistic (zero error) than too pessimistic (neg. error).[13] These results suggest that the findings in Table 1.3 do not depend on cardinality and and are robust even when the dependent variable is transformed into a binary variable implicating a great loss of information.

Cohort effects

The present study disposes of 12 observation years and contains individuals of age 17 to age 84, which results in a strong correlation of age and year of birth. Therefore it is necessary to ensure that observed age effects

[13]Note, that one group of people forecast particularly well: the highly educated are less likely to commit both positive and negative errors instead of zero errors (column 2 and 3) and they do not tend to neither unjustified optimism nor pessimism (column 1) - they are the most realistic.

do not represent cohort effects. Figure 1.8 illustrates for each birth cohort average forecast errors over age. As we can see the overall age pattern arises within cohorts rather than in between them. Further the inclusion of cohort dummies does not change the overall pattern of the age effects. Finally, related evidence from Cantril (1965) describes similar age effects for birth cohorts born 40 years earlier.

Related evidence: The Cantril surveys

Figure 1.9 illustrates additional evidence from the Cantril surveys (see section 1.2.4) on life ladder rankings in 6 industrialized countries around 1960. Despite considerable differences in levels, these figures tell a common story. In Israel, Japan, Poland, USA, West-Germany and Yugoslavia about 50 years ago young people have been very optimistic about their future life and this optimism strictly decreases with age in the cross-section. The same story is told by the GSOEP data in Figure 1.3 (to compare expectations with *current* values, shift the LSt+5 graph 5 periods to the right) with the only difference that expectations fall below current values around age 59 while life ladder ranking expectations stay above current values for the whole life cycle. This difference might be explained by social changes since 1960 or the small sample size of the Cantril survey, or it represents the conceptual difference between life ladder rankings and life satisfaction. Most important, in combination the Cantril and the GSOEP data suggest that the optimism of the young and its decrease with age is a wide-spread phenomenon in industrialized countries, independent of a particular year or decade. It therefore seems plausible that the strong and 'irrational' age effects on individual forecast errors found in this study are not a unique German experience but can be generalized to other western

countries and other times.

1.6 Conclusion

In this paper I test the rationality of life satisfaction forecasts at different stages (age) and states (current life satisfaction) of life, using 136,000 individual forecast errors from 12 observation years of the GSOEP. There exists a broad economic literature that tests the REH for expectations about choice and market variables, usually assuming preferences to be stable. The psychological literature on "hedonic forecasting" on the other hand analyzes misprediction of preferences given deterministic choices and market outcomes. This paper brings these two branches of literature together by analyzing life satisfaction forecasts without assuming stable preferences nor deterministic choices.

Contrary to the REH I find a strong relationship of age and current life satisfaction with the forecast errors. The young and the satisfied overpredict their future life satisfaction, while the old and the unsatisfied underpredict it. People at age 17-30 *over*estimate their future on average by around 0.6. After age 30 this unjustified optimism decreases until age 60, whereafter people *under*estimate their future by around 0.2. The relationship of forecast error with an individual's current life satisfaction is quite linear, ranging from -2 (cLS=0) to 1 (cLS=10) in Pooled OLS regressions.

The pattern of estimated effects persists if a large set of socio-economic controls is included. Further the pattern is not driven by macro shocks in

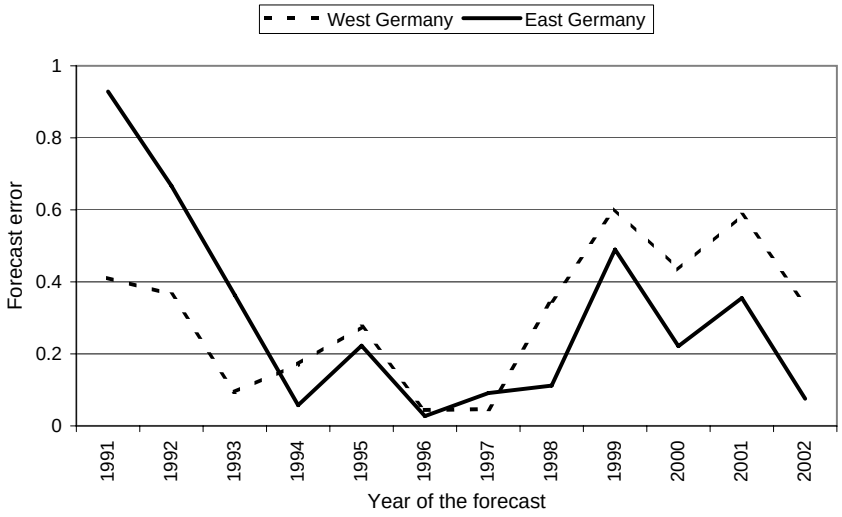
the observation period, even though results suggest strong and unevenly distributed macro shocks during the aftermath of the German reunification (1991-1994) and Germany's New Economy stock market bubble (1998-2002). A test of the model's predictive power shows that estimated effects from one period predict average forecast errors in another (disconnected by 6 years) period very well, i.e. effects are constant over time. Effects are not driven by cohort effects or random measurement errors, nor do they depend on the cardinality assumption.

Additional evidence from the Cantril 1965 survey is presented, indicating that estimated age effects are not a unique German experience of the past decade, but probably were also present in several other industrialized countries around 1960. The effects of current life satisfaction, on the other hand, are in line with the central finding of the psychological literature on "hedonic forecasting". People who are very satisfied or very unsatisfied with their life strongly underestimate the degree of adaptation, i.e. they do not anticipate how quickly life gets 'back to normal'.

To the extent that life satisfaction can be interpreted as a proxy for utility, the systematic forecast errors found in this study might have wide-ranging implications for economic analysis and policy. But before incorporating such biases in economic models and deriving policy implications, we first need to investigate in greater detail the underlying mechanisms and search for possible justifications of such 'irrational' forecasting.

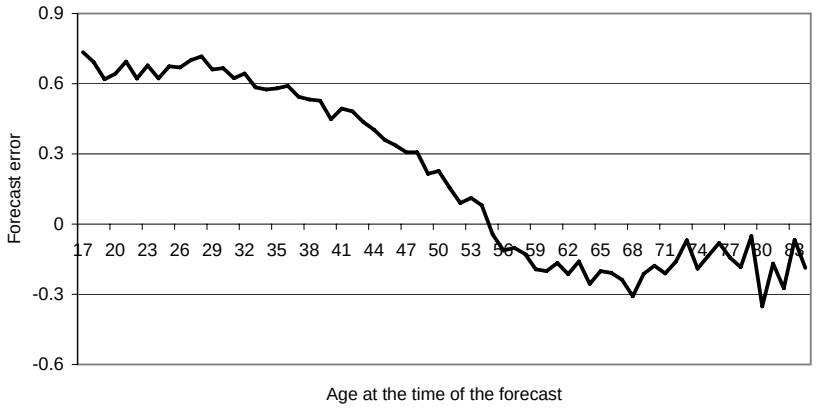
1.7 Tables and Figures

Figure 1.1: Forecast errors in East and West Germany over time



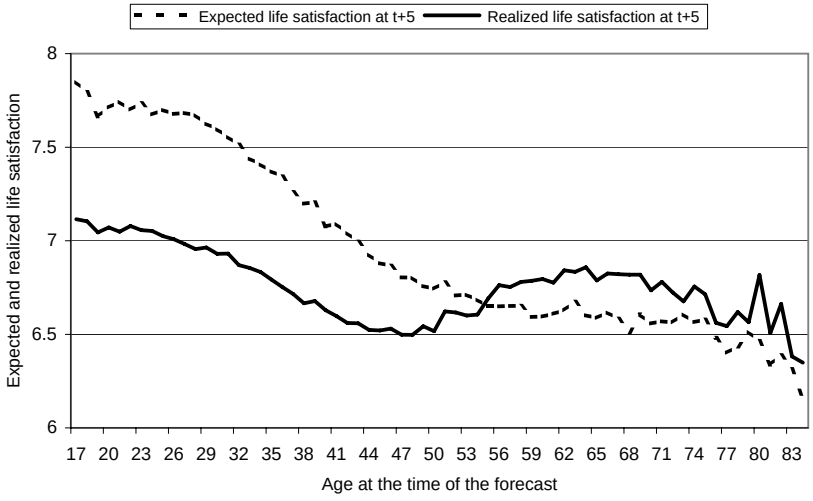
Notes: Forecast errors equal the expected life satisfaction in t+5 minus the realized life satisfaction in t+5. Average forecast errors are plotted separately for East and West Germany over time.

Figure 1.2: Forecast errors over age



Notes: Forecast errors equal expected life satisfaction in t+5 minus realized life satisfaction in t+5. Average forecast errors are plotted over age. Observation period: 1991-2007

Figure 1.3: Decomposed forecast errors over age

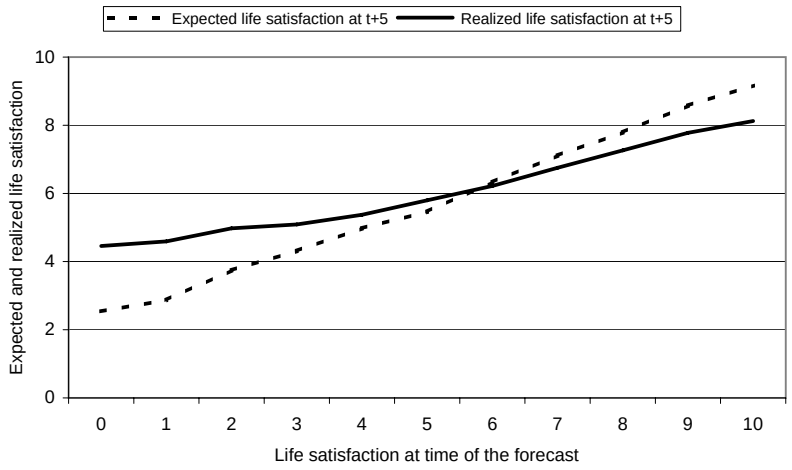


Notes: Average expected and realized life satisfaction over age is plotted. Observation period is 1991-2007

Figure 1.4: Forecast errors over life satisfaction levels

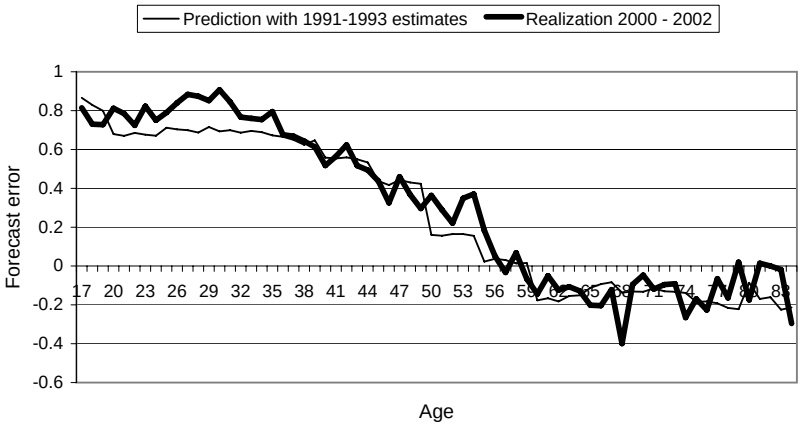
Notes: Forecast errors equal expected life satisfaction in t+5 minus realized life satisfaction in t+5. Average forecast errors are plotted over life satisfaction at the time of the forecast. Period is 1991-2007.

Figure 1.5: Decomposed forecast errors over life satisfaction levels



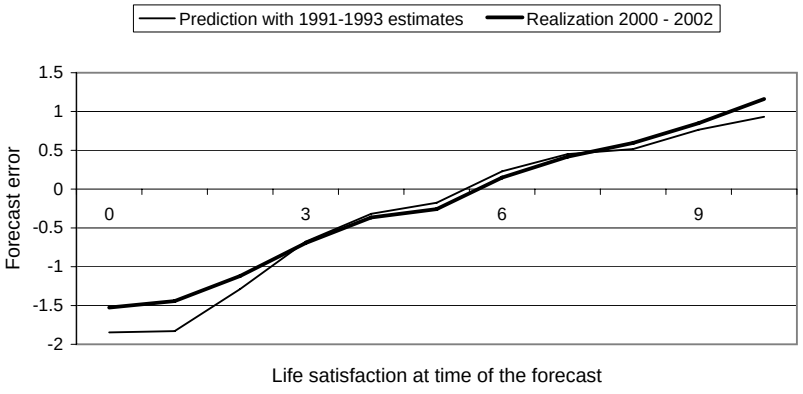
Notes: Average expected and realized life satisfaction over age is plotted. Observation period is 1991-2007

Figure 1.6: Predicting 2000-2002 forecast errors over age using 1991-1993 data



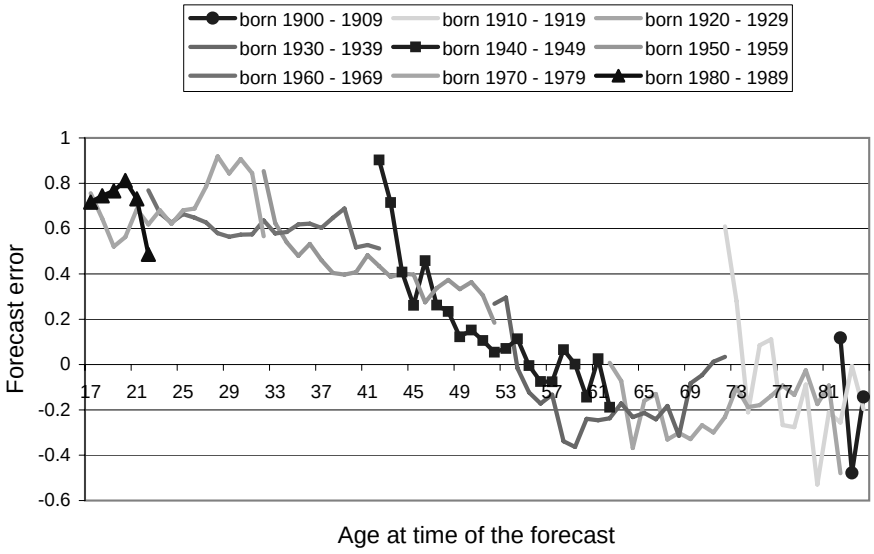
Notes: To obtain the prediction, 1991-1993 forecast errors are regressed on age and current life satisfaction. The resulting estimates are then used in the 2000-2002 data to predict forecast errors. Predicted values are plotted with actual values as averages over age.

Figure 1.7: Predicting of 2000-2002 forecast errors over life satisfaction using 1991-1993 data



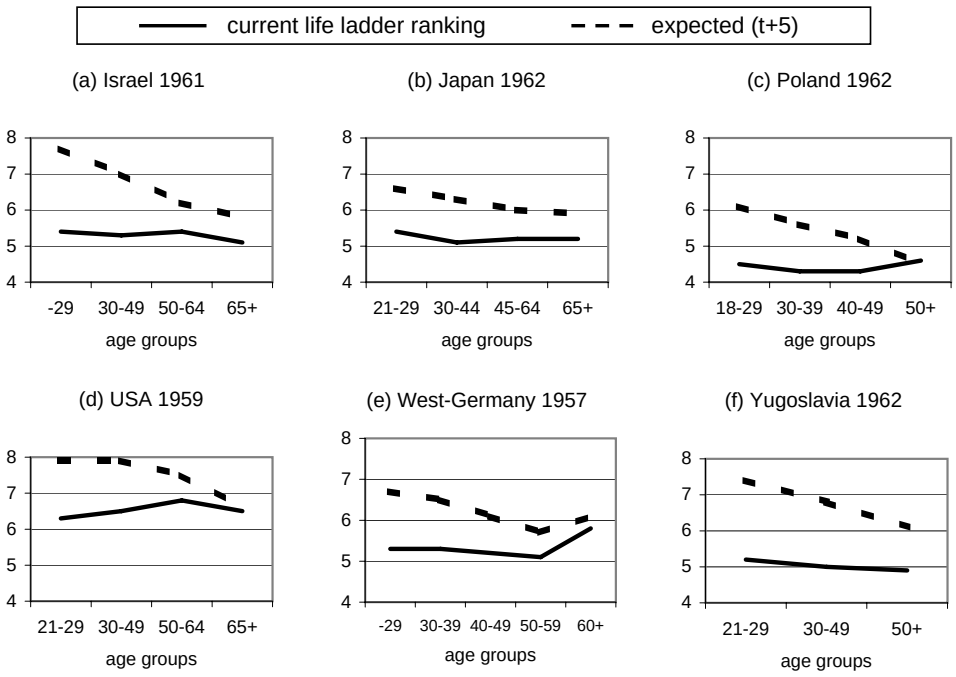
Notes: To obtain the prediction, 1991-1993 forecast errors are regressed on age and current life satisfaction. The resulting estimates are then used in the 2000-2002 data to predict forecast errors. Predicted values are plotted with actual values as averages over life satisfaction.

Figure 1.8: Average forecast errors over age by birth cohorts



Notes: Forecast errors equal the expected life satisfaction in t+5 minus the realized life satisfaction in t+5. Average forecast errors are plotted by birth cohorts over time.

Figure 1.9: Current and expected life ladder ranking over age in six countries around 1960



Notes: Sample sizes are: Israel: 1,170; Japan: 972; Poland: 1,464; USA: 1,549; West-Germany: 480; Yugoslavia: 1,523. Source: Cantril (1965).

Table 1.1: Availability of Data and Generated Forecast Error

	Life Satisfaction	LS 5 periods ahead	Expected LS	Forecast Error
t	**LS_t**	**LS_{t+5}**	**$E[LS_{t+5}]$**	**$E[LS_{t+5}] - LS_{t+5}$**
1984	**x**	**x**	**-**	**-**
1985	**x**	**x**	**-**	**-**
1986	**x**	**x**	**-**	**-**
1987	**x**	**x**	**-**	**-**
1988	**x**	**x**	**-**	**-**
1989	**x**	**x**	**-**	**-**
1990	**x**	**x**	**-**	**-**
1991	**x**	**x**	**x**	**G**
1992	**x**	**x**	**x**	**G**
1993	**x**	**x**	**x**	**G**
1994	**x**	**x**	**x**	**G**
1995	**x**	**x**	**x**	**G**
1996	**x**	**x**	**x**	**G**
1997	**x**	**x**	**x**	**G**
1998	**x**	**x**	**x**	**G**
1999	**x**	**x**	**x**	**G**
2000	**x**	**x**	**x**	**G**
2001	**x**	**x**	**x**	**G**
2002	**x**	**x**	**x**	**G**
2003	**x**	**-**	**x**	**-**
2004	**x**	**-**	**x**	**-**
2005	**x**	**-**	**-**	**-**
2006	**x**	**-**	**-**	**-**
2007	**x**	**-**	**-**	**-**
n	**375,237**	**320,932**	**233,549**	**136,618**

Notes: x = available; - = not available; G = generated.

Table 1.2: Predictive power of expected life satisfaction ($E_t LS_{t+5}$).

	Dependent Variable							
	(1)		(2)		(3)		(4)	
	LS_{t+5}		$houseincome_{t+5}$		$labourincome_t$		$health_{t+5}$	
	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.
LS_t	0.3609***	(107.26)	-0.064	(-0.03)	0.118***	(65.10)	0.026***	(14.04)
$E_t LS_{t+5}$	0.1437***	(45.79)	56.508***	(26.65)	18.7898***	(12.91)	5.2995***	(3.46)
$houseincome_t$			0.667***	(295.38)				
$labourincome_t$					0.6469***	(301.88)		
$health_t$							0.214***	(109.04)
Constant	3.2228***	(172.62)	514.74***	(40.07)	226.55***	(20.80)	0.905***	(88.88)
R-squared	0.2195		0.426		0.5227		0.175	
N	137055		128483		138618		137095	

LS and ELS: reported on a 0 (very unsatisfied) - 10 (very satisfied) scale; income: monthly house income
health: self-rated health status 0 (bad) - 4 (very good); * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.3: Regressions of forecast errors on different control variables

DV: Forecast Errors $E_t[LS_{t+5}] - LS_{t+5}$	(1) age		(2) +LS		(3) +controls	
	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.
age 17-19	-0.040	(-1.05)	-0.105*	(-2.62)	-0.210***	(-4.31)
age 20-24	ref.		ref.		ref.	
age 25-29	0.042	(1.31)	0.044	(1.49)	0.130***	(4.69)
age 30-34	-0.022	(-0.48)	0.011	(0.25)	0.110*	(2.42)
age 35-39	-0.085	(-1.42)	-0.029	(-0.47)	0.084	(1.50)
age 40-44	-0.188***	(-4.90)	-0.108*	(-2.52)	0.031	(0.77)
age 45-49	-0.334***	(-7.75)	-0.237***	(-6.03)	-0.070	(-1.89)
age 50-54	-0.507***	(-11.22)	-0.410***	(-8.67)	-0.241***	(-5.17)
age 55-59	-0.756***	(-13.26)	-0.661***	(-10.99)	-0.517***	(-7.34)
age 60-64	-0.840***	(-16.94)	-0.801***	(-14.06)	-0.682***	(-8.80)
age 65-69	-0.873***	(-14.83)	-0.877***	(-14.42)	-0.735***	(-8.20)
age 70-74	-0.811***	(-18.92)	-0.819***	(-19.35)	-0.666***	(-8.90)
age 75-79	-0.777***	(-12.56)	-0.764***	(-13.23)	-0.590***	(-7.30)
age 80-84	-0.890***	(-8.45)	-0.877***	(-8.37)	-0.740***	(-5.78)
current LS=0			-1.534***	(-11.19)	-1.754***	(-11.75)
current LS=1			-1.370***	(-6.82)	-1.512***	(-8.16)
current LS=2			-0.934***	(-6.89)	-0.989***	(-7.41)
current LS=3			-0.475***	(-7.63)	-0.577***	(-9.44)
current LS=4			-0.099	(-1.61)	-0.132*	(-2.17)
current LS=5			ref.		ref.	
current LS=6			0.410***	(18.23)	0.483***	(21.25)
current LS=7			0.633***	(31.05)	0.782***	(33.06)
current LS=8			0.820***	(32.67)	1.034***	(31.94)
current LS=9			1.056***	(28.27)	1.325***	(41.71)
current LS=10			1.434***	(21.44)	1.703***	(27.19)
Constant	0.605***	(12.48)	-0.024	(-0.51)	0.021	(0.06)
controls[1]	not included		not included		included	
R-squared	0.031		0.084		0.102	
N	136644		136618		126616	

[1] controls: (1) **general**: male (female), (low)/middle/high education (2) **job**: in education, unemployed, employed, pensioner (not employed), apprentice, self-employed, white collar, civil servant (blue collar), temporary (permanent), labor income, tenure (3) **financial**: house income, pension, stocks (4) **family**: household size, married, marriedsep., divorced, widowed, (sinlge), invalid in hh, baby born, marriage or divorce past year (5) **health**: health status, handicap degree (6) **federal**: 16 federal states, federal gdp, federal u-rate (7) **further**: entry and exit dummies. Standard errors are clustered by state; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.4: Regressions of forecast errors over subperiods

DV: Forecast Errors $E_t[LS_{t+5}] - LS_{t+5}$	(1) 1991-1993		(2) 1994-1997		(3) 1998-2002	
	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.
age 17-19	-0.174	(-2.09)	-0.387***	(-5.84)	-0.116	(-1.45)
age 20-24	ref.		ref.		ref.	
age 25-29	0.115*	(2.62)	0.141	(1.95)	0.139*	(2.45)
age 30-34	0.048	(0.69)	0.073	(1.04)	0.125	(1.81)
age 35-39	0.069	(0.75)	0.051	(1.09)	0.074	(0.92)
age 40-44	0.013	(0.15)	0.028	(0.60)	0.012	(0.17)
age 45-49	-0.037	(-0.81)	-0.064	(-0.72)	-0.108	(-1.60)
age 50-54	-0.309***	(-6.11)	-0.283**	(-3.11)	-0.194*	(-2.66)
age 55-59	-0.518***	(-12.13)	-0.552***	(-4.20)	-0.502***	(-5.76)
age 60-64	-0.713***	(-12.37)	-0.701***	(-4.86)	-0.709***	(-8.58)
age 65-69	-0.578***	(-7.82)	-0.722***	(-4.48)	-0.840***	(-7.00)
age 70-74	-0.620***	(-5.50)	-0.545**	(-3.75)	-0.799***	(-8.35)
age 75-79	-0.627**	(-3.31)	-0.344	(-1.73)	-0.765***	(-6.71)
age 80-84	-0.672***	(-5.35)	-0.780*	(-2.90)	-0.835***	(-5.98)
current LS=0	-1.682***	(-4.27)	-1.927***	(-5.78)	-1.736***	(-10.73)
current LS=1	-1.666***	(-7.58)	-1.650***	(-4.46)	-1.367***	(-8.58)
current LS=2	-1.227***	(-5.66)	-1.019***	(-4.26)	-0.865***	(-5.92)
current LS=3	-0.645***	(-4.86)	-0.628***	(-6.11)	-0.530***	(-6.06)
current LS=4	-0.215	(-1.95)	-0.126	(-1.42)	-0.126	(-1.79)
current LS=5	ref.		ref.		ref.	
current LS=6	0.455***	(7.02)	0.505***	(13.30)	0.495***	(16.01)
current LS=7	0.744***	(24.31)	0.797***	(23.22)	0.809***	(19.77)
current LS=8	0.915***	(18.71)	1.103***	(30.12)	1.061***	(19.61)
current LS=9	1.182***	(19.30)	1.443***	(33.08)	1.340***	(22.54)
current LS=10	1.461***	(22.12)	1.780***	(12.23)	1.822***	(22.01)
low education	ref.		ref.		ref.	
med. education	-0.135*	(-2.71)	-0.096**	(-3.38)	-0.064	(-1.71)
high education	-0.149*	(-2.20)	-0.063	(-1.52)	-0.086	(-1.67)
East Germany	0.630***	(7.83)	0.190**	(3.03)	-0.037	(-0.78)
Constant	0.147	(1.46)	-0.034	(-0.37)	0.362**	(3.80)
controls[1]	included		included		included	
R-squared	0.099		0.109		0.103	
N	28182		36746		61688	

[1] same set of controls as in the previous table, except East Germany (instead of federal dummies) and education dummies which are listed.

Table 1.5: Probit regressions of binary forecast error

DV: binary	(1)		(2)		(3)	
Forecast Errors (FE)	1:FE>0; 0:FE<0		1:FE>0; 0:FE=0		1:FE<0; 0:FE=0	
	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.
age 17-19	-0.159***	(-4.72)	-0.128***	(-3.86)	0.038	(0.95)
age 20-24	ref.		ref.		ref.	
age 25-29	0.112***	(4.91)	0.021	(0.94)	-0.105***	(-3.80)
age 30-34	0.102***	(4.33)	-0.013	(-0.59)	-0.129***	(-4.56)
age 35-39	0.044	(1.81)	-0.005	(-0.23)	-0.068*	(-2.36)
age 40-44	-0.012	(-0.49)	-0.063*	(-2.56)	-0.062*	(-2.10)
age 45-49	-0.087***	(-3.40)	-0.102***	(-4.01)	-0.032	(-1.06)
age 50-54	-0.218***	(-8.30)	-0.148***	(-5.59)	0.043	(1.39)
age 55-59	-0.409***	(-15.15)	-0.261***	(-9.49)	0.125***	(3.99)
age 60-64	-0.548***	(-18.20)	-0.337***	(-10.94)	0.188***	(5.52)
age 65-69	-0.585***	(-16.82)	-0.350***	(-9.67)	0.222***	(5.69)
age 70-74	-0.511***	(-13.58)	-0.291***	(-7.40)	0.212***	(5.03)
age 75-79	-0.478***	(-11.15)	-0.210***	(-4.58)	0.256***	(5.30)
age 80-84	-0.586***	(-10.62)	-0.272***	(-4.51)	0.299***	(4.90)
current LS=0	-0.704***	(-9.42)	-0.381***	(-3.98)	0.297***	(4.14)
current LS=1	-0.547***	(-7.53)	-0.047	(-0.46)	0.511***	(6.28)
current LS=2	-0.390***	(-9.34)	0.143*	(2.44)	0.514***	(10.41)
current LS=3	-0.215***	(-7.55)	0.133***	(3.59)	0.346***	(10.53)
current LS=4	-0.074**	(-3.00)	0.162***	(5.24)	0.242***	(8.56)
current LS=5	ref.		ref.		ref.	
current LS=6	0.294***	(17.59)	0.253***	(13.06)	-0.041*	(-2.21)
current LS=7	0.508***	(33.64)	0.209***	(12.37)	-0.310***	(-18.70)
current LS=8	0.787***	(52.35)	0.194***	(11.85)	-0.619***	(-37.68)
current LS=9	1.116***	(56.90)	0.455***	(22.98)	-0.698***	(-30.96)
current LS=10	1.301***	(50.81)	0.512***	(21.30)	-0.868***	(-29.09)
low education	ref.		ref.		ref.	
med. education	-0.045***	(-3.74)	-0.048***	(-3.89)	0.007	(0.54)
high education	-0.013	(-0.77)	-0.100***	(-5.95)	-0.065***	(-3.40)
Constant	-0.013	(-0.17)	0.561***	(7.53)	0.679***	(7.98)
controls	included		included		included	
R-squared						
N	94075		88611		70546	

Comments as in the previous table.

Bibliography

[1] Batchelor, R. (1986) "Quantitative v. Qualitative Measures of Inflation Expectations." *Oxford Bulletin of Economic and Statistics*, 48, 99-120.

[2] Begg, David K. H. (1982) *The Rational Expectations Revolution in Macroeconomics.* Johns Hopkins University Press.

[3] Benitez-Silva, H. and Dwyer, D.S. and Gayle, W.R. and Muench, T.J. (2008) "Expectations in Micro Data: Rationality Revisited." *Empirical Economics*, 34, 381–416.

[4] Blanchflower, D.G. and Oswald, A.J. (2004) "Well-Being over Time in Britain and the USA." *Journal of Public Economics*, 88, 1359-1386.

[5] Branch, W. (2004) "The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations." *Economic Journal*, 114, 592-621.

[6] Cantril, H. (1965) *The Pattern of Human Concerns*. Rutgers University Press, New Brunswick.

[7] Carroll, C. (2003) "Macroeconomic Expectations of Households and Professional Forecasters." *Quarterly Journal of Economics*, 269-298.

[8] Chamberlain, Gary (1984) "Chapter 22: Panel Data." in Zvi Griliches and Michael D. Intriligator (Ed.) *Handbook of Econometrics*, Vol. II, North-Holland.

[9] Clark A.E. and Oswald A.J. (1994) "Unhappiness and Unemployment." *Economic Journal*, 104, 648-659.

[10] Clark, Andrew E., Paul Frijters, and Michael A. Shields. (2008a) "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles." *Journal of Economic Literature*, 46, 1, 95-144.

[11] Clark, A. E., Diener, E., Georgellis, Y. and Richard, E. L. (2008b) "Lags and Leads in Life Satisfaction: a Test of the Baseline Hypothesis." *Economic Journal*, 118, 222-243.

[12] Das, M. and van Soest, A. (1999) "A panel data model for subjective information on household income growth." *Journal of Economic Behaviour and Organization*, Vol. 40, 409-26.

[13] Davidson, R.J. (2000) "Affective style, psychopathology and resilience: Brain mechanisms and plasticity." *American Psychologist*, 55, 1196-1214.

[14] Di Tella R., MacCulloch R.J. and Oswald A.J. (2001) "Preferences over inflation and unemployment: Evidence from surveys of happiness." *American Economic Review*, 91, 335-341.

[15] Dwyer, Gerald P. Jr., Arlington W. Williams, Raymond C. Battalio, and Timothy I. Mason. (1993) "Tests of Rational Expectations in a Stark Setting." *Economic Journal*, 103

[16] Easterlin, R. A. (1974) "Does economic growth improve the human lot?" in P. A. David and R.I. M. W. Reder (Ed.), *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz*, New York: Academic Press Inc.

[17] Easterlin, Richard A. (2001) "Income and Happiness: Towards a Unified Theory." *Economic Journal*, 111, 465-84.

[18] Ferrer-i-Carbonell, A. and Frijters, P. (2004) "How important is methodology for the estimates of the determinants of happiness?" *Economic Journal*, 114, 641-659.

[19] Fordyce, Meredith A. (1988) "A review of research on happiness measures: a sixty second index of happiness and mental health." *Social Indicators Research*, 20, 355-81.

[20] Frey, B.S., Stutzer, A. (2002) "What can economists learn from happiness research?" *Journal of Economic Literature*, 40, 402-435.

[21] Frey, B.S., Benesch C., and Stutzer A. (2005) "Does Watching TV Make Us Happy?" *Journal of Economic Psychology*, 28(3), 2007, pp. 283-313.

[22] Frijters, P., Haisken-DeNew, J. and Shields, M. (2004) "Money does matter! Evidence from increasing real incomes and life satisfaction in East Germany following reunification." *American Economic Review*, 94, 730-740.

[23] Frijters P, Haisken-DeNew JP, Shields MA. (2005) "The causal effect of income on health: Evidence from German reunification." *Journal of Health Economics*, 24; 997-1017.

[24] Frijters, P., Greenwell, H., Shields, M.A., and J.P. Haisken-DeNew (2008) "How rational were expectations in East Germany after the falling of the wall?" *Canadian Journal of Economics*

[25] Gardner, J. and Oswald, A.J. (2007) "Money and mental wellbeing: A longitudinal study of medium-sized lottery wins." *Journal of Health Economics*, vol. 26, pp. 49-60.

[26] Gilbert, Daniel (2006) *Stumbling on Happiness*, New York: Knopf Random House

[27] Kahneman, D. (2000) "Evaluation by Moments: Past and Future" in A. Tversky. D. Kahneman, eds., *Choices, Values, and Frames*, Cambridge: Cambridge University Press.

[28] Kahneman, Daniel, and Thaler, Richard (2006) "Anomalies: Utility Maximisation and Experienced Utility" *Journal of Economic Perspectives*, 20.

[29] Kassenboehmer, S. C. and Haisken-DeNew, J. P. (2009) "You're Fired! The Causal Negative Effect of Entry Unemployment on Life Satisfaction." *The Economic Journal*, 119, 448-462.

[30] Keane, M.P., and Runkle, D.E. (1990) "Testing the Rationality of Price Forecasts: New Evidence from Panel Data." *American Economic Review*, 80, 4 714-735.

[31] Keuzenkamp, H.A. (1991) "A precursor to Muth; Tinbergen's 1932 Model of Rational Expectations." *The Economic Journal*, 101, 1245-53.

[32] Lacey HP, Smith DM, Ubel PA. (2006) "Hope I die before I get old: Mispredicting happiness across the lifespan." *Journal of Happiness Studies*, 7(2):167-182.

[33] Layard, R. (2003) "Happiness: Has Social Science A Clue?" Lionel Robbins Memorial Lectures 2002/03, London School of Economics.

[34] Layard, R. (2005) *Happiness. Lessons from a New Science*, London: Allen Lane.

[35] Layard, R. (2006) "Happiness and public policy: A challenge to the profession." *Economic Journal*, 116(510), C24-C33.

[36] Luechinger, S. (2009) "Valuing Air Quality Using the Life Satisfaction Approach," *Economic Journal*, 119, No. 536 (March), 482-515.

[37] Loewenstein, G. and Frederick, S. (1997) "Predicting Reactions to Environmental Change." in M. Bazerman, D. Messick, A.Tenbrunsel, and K. Wade-Benzoni, eds., *Environment, Ethics, and Behavior*, San Francisco, CA: The New Lexington Press, 52-72.

[38] Loewenstein, G. F., and Schkade, D. (1999) "Wouldn't it be nice? Predicting future feelings." In E. Diener, N. Schwartz, and D. Kahneman (Eds.), *Well-Being: The Foundations of Hedonic Psychology* 85-105. New York, Russell Sage Foundation.

[39] Loewenstein, G., ODonoghue, T. and Rabin, M. (2003) "Projection Bias in Predicting Future Utility." *Quarterly Journal of Economics*, 118(4), 1209-1248.

[40] Lovell, M.C. (1986) "Tests of Rational Expectations Hypothesis." *American Economic Review*, 76-1, 110-124.

[41] Luttmer, Erzo F. (2004) "Neighbors as Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics* 120:3, pp. 963-1002.

[42] Maddala, G., Fisher, R., Lahiri, K. (1981) "A Time Series Analysis of Popular Expectations Data" *Time Series Analysis of Economic Data*, Washington, DC: US Bureau of the Census, 278-289.

[43] Mitchell J. and M. Weale (2007) "The Rationality and Reliability of Expectations Reported by British Households: Micro Evidence from the British Household Panel Survey." *NIESR WP.*

[44] M{uller-Trede, Johannes (2008) "Random Errors - Systematic Effects: Biases in Decision-Making." mimeo. Universitat Pompeu Fabra.

[45] Muth, J. F. (1961) "Rational expectations and the theory of price movements." *Econometrica*, 29, 315-335.

[46] Nickell, S.J. (1981) "Biases in dynamic models with fixed effects." *Econometrica*, 49, 1417-1426.

[47] Ng, Y. K (1997) "A case for happiness, cardinalism, and interpersonal comparability." *Economic Journal*, 107, 1848-58.

[48] Oswald, A. J. (1997) "Happiness and economic performance." *Economic Journal*, 107, 1815-31.

[49] Robbins, L. (1932) *On The Nature and Significance of Economic Science*, Macmillan, London.

[50] Schkade, D. A., and Kahneman, D. (1998) "Does living in California make people happy?: A focusing illusion in judgments of life satisfaction." *Psychological Science*, 9, 340-346.

[51] Senik, C. (2004) "When Information Dominates Comparison: A Panel Data Analysis Using Russian Subjective Data." *Journal of Public Economics*, 88, 2099-2123.

[52] Souleles, N. S. (2004) "Expectations, heterogeneous forecast errors, and consumption: micro evidence from the Michigan Consumer Sentiment Surveys." *Journal of Money, Credit, and Banking*, 36, 39-72.

[53] Stutzer, A. and Frey, B. S. (2007) "What Happiness Research Can Tell Us About Self-Control Problems and Utility Misprediction." in Bruno S. Frey and Alois Stutzer (Ed.), *Economics and Psychology. A Promising New Cross-Disciplinary Field*. Cambridge: MIT Press, pp. 169-195.

[54] Swenson, Charles W. (1997) "Rational Expectations and Tax Policy: Experimental Market Evidence." *Journal of Economic Behavior and Organization*, 32.

[55] Van den Berg, B., and Ada Ferrer-i-Carbonell (2007) "Monetary evaluation of informal care: the well-being evaluation method." *Health Economics*, 16, 1227-44.

[56] Van Praag, B. and Ada Ferrer-i-Carbonell (2004) *Happiness Quantified. A Satisfaction Calculus Approach.* Oxford: Oxford University Press.

[57] Winkelmann L. and Winkelmann R. (1998) "Why are the unemployed so unhappy? Evidence from panel data." *Economica* 65, 1-15.

Chapter 2

WEALTH SHOCKS AND HEALTH OUTCOMES: EVIDENCE FROM STOCK MARKET FLUCTUATIONS.

2.1 Introduction

Richer people are healthier, happier and live longer. Little is known, however, about the causal mechanisms underlying this important correlation of wealth and health. Money might buy health, but health might also reversely affect expenditure and income generation. And third factors such as preferences or life events are likely to affect both simultaneously. The broad existing literature on the wealth-health relationship is skeptical about causal effect of wealth or wealth shocks on adult health in developed countries. So far such effects have been documented only for poor

retirees in poor countries.[1] In this paper I exploit stock market fluctuations in the wealth of US retirees as a source of exogenous wealth shocks. Contrary to the existing literature I find that wealth shocks strongly affect physical health, mental health and survival rates of wealthy retirees in the US.

Over the past two decades every third retiree household in the US held part of its wealth in stocks. And these households invested on average about 20% of their overall remaining life-time wealth in such risky asset. As a consequence the booms and busts in the US stock market over the past two decades generated dramatic unexpected gains and losses in the wealth of stock holding retirees. I analyze this natural experiment using rich micro-data from the Health and Retirement Study (HRS). The HRS is representative of the elderly US population and provides panel data on all wealth components including stock holdings as well as information on physical health, mental health and mortality.

I construct wealth shocks as the interaction of stock holdings with the stock market change. These constructed wealth shocks are highly predictive of changes in reported wealth. And they strongly affect health outcomes. A 10% change in life-time wealth over a two year period is associated with a change of 2-4% of a standard deviation in four different health measures: an index of health conditions, self-reported health, mental health and the probability to survive to the next interview two years ahead. The analysis of individual health conditions reveals a plausible pattern underlying the effect on physical health. Effects are strongest for

[1] For reviews of the literature see Smith (1999), Deaton (2003), Cutler, Deaton and Lleras-Muney (2006), Cutler, Lleras-Muney and Vogl (2011).

hypertension, which we would expect to be most responsive in the short run. Smaller effects I find for heart diseases and strokes which are typically caused by high blood pressure. And there are no effects on arthritis, diabetes, lung disease and cancer which in general take more than two years to be affected by external factors. There is evidence of effect assymetry and heterogeneity across age. Negative shocks tend to have a stronger effect than positive ones. And effects on physical health and survival rates seem to increase with age. Compared to the cross-sectional relationship of wealth and health the estimated effects are large in magnitude.

For a causal interpretation of these estimates constructed wealth shocks must be independent of any unobserved heterogeneity in health changes. Stock market changes are exogenous for the individual retiree but this is not the case for stock holdings. More educated, wealthier and more risk loving individuals typically hold larger fractions of their wealth in stocks. And the observation period covers only a limited number of - on average positive - stock market changes. As a consequence constructed wealth shocks are likely to be correlated with unobserved determinants of stock holdings. For this reason I control separately for the fraction of wealth held in stocks. In other words, I compare health changes for individuals with the same amount of stocks at different points in the stock market cycle. One might still worry that results are driven by a correlation of the stock market with investor types or with the typical investor's health profile. Several robustness checks show that this is unlikely to be the case. This suggests that constructed wealth shocks indeed cause the observed changes in health.

For the interpretation of this relationship as the effects of wealth shocks on health it is further necessary to control for effects of the stock market or the macroeconomic environment that do not run through stock wealth. I argue that retirees without stocks are at least equally strongly affected by potential direct effects as those with stocks. I include time effects to absorb any macroeconomic shocks common to both groups.

Despite a broad existing literature my paper is the first to find positive effects of wealth shocks on elderly health in a developed country. So far such effects have been documented only for poor retirees in Russia (Jensen and Richter 2003) and South Africa (Case 2004). As Cutler, Lleras-Muney and Vogl (2011) summarize in a recent literature review, "... [A] preponderance of evidence suggests that in developed countries today, income does not have a large causal effect on adult health". The most prominent papers providing this evidence can be summarized by three main approaches.

A first set of papers uses approaches related to Granger-causality (Adams et al. 2003; Smith 2005; Michaud and Van Soest 2008). Using the HRS data these papers show that wealth changes and lagged wealth conditional on socio-economic controls do not predict health changes at the micro-level. I replicate Smith’s (2005) findings and discuss Adams et al. (2003) in detail below. Measurement error in self-reported wealth and short-term responses in health seem to be likely explanations of why their results are different from the findings presented in this paper.

Another set of papers analyze aggregate time series of income and health at the state or cohort level (Ruhm 2000; Deaton and Paxson 2001; Deaton

and Paxson 2004; Snyder and Evans 2006; Adda et al. 2009). None of these papers find evidence of a positive relationship of income changes and health changes at the macro-level. Endogeneity and measurement errors are less of an issue in aggregate data. Group averages are independent of individual-specific endogeneity and cancel out random measurement error.[2] However, aggregate income changes might be correlated with macro shocks that also have non-income effects on health. This invalidation of the exclusion restriction, as the authors of these papers note, makes it difficult to infer causal effects from these findings.[3]

A third set of papers exploits lottery winnings as a source of exogenous variation in wealth (Lindahl 2005; Gardner and Oswald 2007; Apouey and Clark 2009). These papers find positive effect on mental health, while results are less conclusive for physical health. A general challenge of lottery studies are small sample sizes and in particular few observations of significant winnings. Further, only positive wealth shocks are observed.

In the present study I combine these different approaches overcoming their individual shortcomings. I merge the rich micro-data from the HRS with aggregate stock market changes to introduce a source of exogenous macro shocks. The interaction of these macro shocks with a micro-level measure of the exposure to these shocks, the amount of stock holdings, allows to better control for potential non-wealth effects of the macroeconomic environment. The resulting setup is in spirit a large-scale lottery

[2] Aggregation at the state or cohort level is equivalent to instrumental variable estimation with a set of state or cohort dummies as instruments (Angrist and Pischke 2009).

[3] For revisions of Ruhm (2000) and of Snyder and Evans (2006) see Miller et al. (2009) and Handwerker (2008), respectively.

framework that allows to analyze the causal effects of both wealth gains and losses on elderly health in the US.

But how plausible are the effects that I find? Should we expect positive physical health effects found for poor retirees in poor countries to carry over to wealthy retirees in the US? Health inputs like medical treatment, medication or mere calorie intake might be affected by wealth shocks for poor retirees in Russia or South Africa. But this is probably less of an issue for stock holding US pensioners, who have enough money left to afford basic pills and food even after a considerable wealth loss. Further, Medicare covers the entire 65+ population in the US so that wealth shocks do not affect basic health insurance coverage unlike for displaced workers. Consumption of healthy food and purchase of a healthy environment could be more responsive determinants of retiree health in the US than basic health inputs. But two years might not be enough time for consumption to affect health outcomes as dramatically as observed.

Other plausible channels are psychological factors such as happiness about pleasant trips that were not affordable before or financial worries and sadness about a lost fortune that had been accumulated as inheritance for the grandchildren. A broad literature in medicine, psychology and biology has documented effects of psychological stress on coronary artery diseases, clinical depression and mortality (Strike and Steptoe 2004). Positive emotions, on the other hand, were found to have positive effects on these health outcomes (for a review see Chida and Steptoe [2008]). This paper finds strong wealth shock effects on high blood pressure and mental health and smaller effects on heart problems and strokes. This is exactly the kind of health response the bio-medical literature would predict if

wealth shocks have an effect on psychological stress.[4] However, this does not rule out that health is also directly affected through health inputs or consumption.

The focus of this study on retirees has several advantages. Compared to younger adults retirees have a lot of wealth and heterogeneity in wealth composition so there is a lot of wealth variation to exploit. Further, as they no longer participate in the labor market effects of stock market shocks running through labor demand are limited. That makes it easier to separate wealth shock effects from other confounding factors. Last, the elderly are closer to the margin of severe health problems than younger adults. This makes it more likely for effects of wealth shocks on latent health to become manifest in observable health outcomes.

However, caution must be exercised when extrapolating from my estimates to other settings. Effects are identified only for stock holding retirees who are on average wealthier, healthier and less risk-averse than those without stocks. Further the estimated effects might not be representative for younger adults who are in better physical shape and flexible in terms of their labor supply to compensate a given wealth shock.[5] Last, my estimates represent the short-term effects of wealth shocks. They might not be representative for the long-run effects of gradually accumulating

[4]The responsiveness of elderly mental health to income related shocks has also been documented by Grip et al. (2009).

[5]Sullivan and von Wachter (2008), however, provide related evidence for younger adults. They show that exogenous job displacements dramatically increase the mortality hazard of male US workers during the years following the job loss. The authors interpret their findings to be consistent with job loss "causing acute stress, which may substantially raise the mortality hazard in the short term."

wealth differences. The comparison with the cross-sectional relationship of wealth and health indeed suggests that the long-run wealth elasticity of health is smaller and more homogeneous across health conditions than the estimated causal effects of wealth shocks.

The remainder of this paper is organized as follows: Section II discusses the identification strategy, Section III describes the data, Section IV the empirical specification. Section V presents the findings and Section VI concludes.

2.2 Identification

This paper seeks to estimate the causal effect of wealth shocks on health. The difficulty of this task is the endogeneity of wealth. Wealth shocks might not only affect health, but health shocks are also likely to reversely affect expenditures and third factors might influence both wealth and health simultaneously. Further, wealth is typically measured with noise leading to attenuation bias. This measurement error problem tends to aggravate in first differences. For these two reasons the simple regression of health changes on wealth changes from observational data might not tell us a lot about the causal effect of wealth shocks on health outcomes.

The ideal experiment to solve the endogeneity problem would be a lottery that randomly assigns wealth losses and gains to people and measures their health before and some time after the assignment. This paper exploits the booms and busts of the US stock market over the past two

decades as a natural experiment that generated considerable wealth gains and losses for retirees owning stocks.[6] This natural experiment comes quite close to the ideal setting. As stock market changes are largely unpredictable for retirees without insider information holding stocks is equivalent to buying lottery tickets.

I construct stock market induced wealth shocks (hereafter *constructed wealth shocks*) as the interaction of the lagged fraction of life-time wealth held in stocks with stock market changes.

$$\frac{s_{i,t-1}}{W_{i,t-1}} \frac{\Delta SP_t}{SP_{t-1}} \tag{2.1}$$

where $s_{i,t-1}$ are past wave's stock holdings, $W_{i,t-1}$ is a measure of past wave's life-time wealth (see below) and $\frac{\Delta SP_t}{SP_{t-1}}$ the percentage change in the S&P500 stock market index between two waves. For example, an individual with 20% life-time wealth held in stocks in the past wave and a 50% stock market increase between the past and the current wave is assigned a 10% positive wealth shock.

To estimate the effects of wealth shocks on health outcomes I regress health changes directly on constructed wealth shocks while controlling for the main effects and demographic covariates:

[6] To my knowledge Coile and Levine (2006) have been the first to exploit this natural experiment. They analyze the impact of stock market movements on retirement decisions, comparing the effects of stock market movements on retirement for groups that are relatively more and less likely to hold stocks. I enhance their approach, using the exact fraction of wealth held in stocks instead of a binary indicator of stock market exposure which increases the power of the analysis.

$$\Delta H_{i,t} = \alpha + \beta \frac{s_{i,t-1}}{W_{i,t-1}} \frac{\Delta SP_t}{SP_{t-1}} + \gamma \frac{s_{i,t-1}}{W_{i,t-1}} + \vartheta_t + \delta X_{i,t} + \epsilon_{i,t} \qquad (2.2)$$

where $H_{i,t}$ are different health measures, $\frac{s_{i,t-1}}{W_{i,t-1}} \frac{\Delta SP_t}{SP_{t-1}}$ are constructed wealth shocks, ϑ_t are year fixed effects and $X_{i,t}$ predetermined demographic controls. Health measures are regressed in first differences because wealth shocks can only explain changes but not past levels in health. Taking first differences therefore cleans the dependent variable of unexplainable variation while it does not reduced the number of observations since the construction of wealth shocks already requires a lag.

For the interpretation of β as the causal effect of wealth shocks on health two conditions must be satisfied. Constructed wealth shocks are causal only if they are independent of any unobserved heterogeneity in health changes. Further, their effect on health captured by β must run exclusively through changes in stock wealth.

2.2.1 Are constructed wealth shocks causal?

Stock market changes are largely unpredictable (for a review of the finance literature on market efficiency see Malkiel [2003]) and therefore random for the individual retiree. However, the observation period covers only a limited number of on average positive stock market changes. As a consequence constructed wealth shocks are on average higher (more

positive) for those with more stocks. At the same time stock holdings are not random. The richer, the more educated and the more risk loving typically hold larger fractions of their wealth in stocks. This results in a correlation of constructed wealth shocks with unobservable determinants of stock holdings. Regressing health measures in first differences cancels out unobserved heterogeneity that is constant over time. But determinants of stock holdings might not only correlate with health levels but also with health profiles over time so that first differences alone do not rule out potential endogeneity.[7] Therefore it is important to control separately for the lagged fraction of wealth held in stocks ($\frac{s_{i,t-1}}{W_{i,t-1}}$).

This means I compare health changes for individuals with the same amount of stocks at different points in the stock market cycle. Or in terms of the lottery analogy, I measure the health response to lottery winnings and losses conditional on the amount of lottery tickets bought.

One concern might be that retirees with the same fractions of wealth held in stocks at different points in the stock market cycle are not comparable. A retiree with 20% wealth in stocks at the beginning of a boom might be different from a retiree with 20% in stocks right before a crash. The observation period covers only a limited number of stock market changes so that there could be a spurious correlation of stock market changes with the type of investor. Also, individuals do not rebalance portfolios con-

[7]For example, individuals who anticipate a health risk might want to reduce financial risks and redistribute their portfolio from stocks to safer assets. Or people with less education have more declining health profiles due to worst health behavior and at the same time hold less stocks due to less financial literacy. Given a limited number of on average positive stock market changes these examples of reverse causality or simultaneity would imply a positive correlation of constructed wealth shocks and health changes.

tinuously. So a retiree with 20% in stocks who does not rebalance her portfolio will end up with 33% in stocks when the stock market doubles. To rule out that results are driven by such potential correlation of the stock market cycle with the type of investor I present two stage least squares regressions in which I instrument actual stock holdings with the individual's average stock holdings over the entire observation period or initial stock holdings in the first period. Both average and initial stock holdings are constant over time for a given individual. Hence they are uncorrelated with were we are in the stock market cycle. Although all results carry over with the instrumental variable estimate, the more efficient ordinary least squares estimates are chosen as baseline specification.

Another way to check whether estimated effects are driven by changes in investor types is to include predetermined demographic controls ($X_{i,t}$). If the relationship of health changes and constructed wealth shock is driven by changes in the type of investors then the inclusion of controls like gender, age, education, region, etc. should change the coefficient on wealth shocks. However, adding a wide range of demographic controls to the baseline specification hardly changes any of the estimates. Following a similar logic, it is insightful to look at simple OLS benchmark regressions of health on wealth and other socio-economic variables. Assume that constructed wealth shocks merely pick up a relationship of health with the investors' socio-economic background. Then the OLS benchmark regressions should reveal a similar pattern across different health measures as found for constructed wealth shocks. However, the pattern emerging from OLS regressions is clearly different.

Still one might worry that the stock market correlates coincidentally with

health profiles of those retirees who tend to hold a lot of stocks. A brief look at the stock market development during the observation period in Figure 2.1 suggests that this is unlikely to be the case. Even though there is an overall positive trend between 1992 and 2009, negative and positive changes follow each other towards the end of the sample. It is hard to imagine that health profiles of stock holders just happen to follow these ups and downs by chance. However, health profiles might be slightly correlated with the overall trend in the stock market leading to a spurious relationship of health changes with constructed wealth shocks. If this were the case we should find a similar pattern in OLS benchmark regressions of health changes on socio-economic background. But as for health levels, the pattern emerging from benchmark regressions of health changes on lagged wealth is very different from the effects of constructed wealth shocks. Finally, I present regressions where I include placebo shocks in addition to the real wealth shocks. I interact the lagged stock fraction with changes in the unemployment rate instead of the stock market. And I interact the stock market change with the lagged wealth fraction held in bonds instead of stocks. Both of these two measures are strongly correlated with constructed wealth shocks. Still, their inclusion does not significantly change my estimates (results reported in the Web Appendix[8]).

To sum up, it seems unlikely that a correlation of the stock market cycle with investor types or with investors' health profiles is driving the results. This suggests that constructed wealth shocks are indeed causing the observed changes in health. To interpret this causal effect as the effect of wealth shocks on health it is necessary to control for stock market effects

[8]The Web Appendix is available at http://www.econ.upf.edu/jobmarket/schwandt.html

on health that might not run through stock wealth.

2.2.2 Are effects running exclusively through stock wealth?

Stock market changes might not only determine the valuation of stock holdings but also correlate with prices of other non-stock wealth holdings such as bonds or real estate. A way to test for such correlation is to look at the comovement of the stock market with the wealth of households that do not own stocks. Figure 2.1 compares the S&P500 with the coefficients from regressions of wealth changes on wave dummies for retirees with stocks and without stocks in the previous period. For retirees with stocks they follow the up's and down's in the S&P500, especially after wave 4 when lagged stock holdings are measured more precisely (see Data section). But for retirees without stocks wealth changes are positive in all waves and seem uncorrelated with the stock market. More detailed regressions taking into account the precise month of interviews are presented in the Findings section. Again, the stock market is highly predictive for wealth changes of stock holders, while the effect on wealth for those without stocks is essentially zero. This suggests that there is not much an effect of the stock market on non-stock wealth.

But the stock market or more broadly the macroeconomic environment might also affect health through non-wealth channels. For example, a macroeconomic environment in which stock markets collapse might have negative effects on the individual's employment which would probably not only affect her wealth but also directly her health. As the sample is restricted to retiree households effects running through the individual's

employment status are limited. But retirees might be troubled about their children becoming unemployed or their grand children not finding a job after graduating from high school. Further we could think of the provision of public goods that might depend on the macroeconomic environment and have a direct effect on pensioners' health. And retirees could be stressed and fearing social instability when hearing apocalyptic news about the economy in the media. However, it seems reasonable to assume that these direct effects are at least as strong for retirees who do not hold stocks as for those with stocks. Retirees without stocks tend to be poorer, less educated and more risk averse. If anything, they depend more on public goods, suffer more from bad news and their children are the first to get fired when it comes to mass lay-offs in a recession. To control for potential direct effects in a conservative way I therefore include time fixed effects (ϑ_t). Before describing the data and the final empirical specification in detail a few issues remain to be discussed.

2.2.3 Measurement and scaling issues

Constructed wealth shocks under- or overestimate actual wealth shocks if retirees' expectations of stock market returns systematically differ from zero. Luckily the last four HRS waves include a question about the likelihood that the stock market increases within the following year. Figure 2.3 in the Appendix plots monthly averages for this question together with the S&P500. Expectations are strikingly low: even those with stocks expect on average only a 45-60% chance that the stock market will increase. Furthermore, expectations seem to be slightly correlated with the stock market. Following Dominitz and Manski (2007) I transform

expected probabilities about stock market increases into expected stock market returns and adjust for them when constructing wealth shocks. As expectations are only marginal compared to actual stock market changes their inclusion decreases estimates only slightly. For better comparability of my results with other studies I therefore do not include expectations in the baseline regressions.

Changes in reported wealth are not only endogenous but also notorious for attenuation bias due to measurement error. Constructed wealth shocks help to minimize this kind of bias because they rely on levels instead of changes in self-reported wealth. Notice that the other component of constructed wealth shocks, changes in the S&P500, represent average stock market returns. Average returns do not account for individual portfolio compositions which are not observed in the data. However, the resulting measurement error in constructed wealth shocks is negatively correlated with actual returns but uncorrelated with constructed wealth shocks, i.e. the regressor of interest. This kind of measurement error implies less precise estimates but no attenuation towards zero.

Constructed changes in stock wealth ($s_{i,t-1}\frac{\Delta SP_t}{SP_{t-1}}$) are divided, or rescaled, by a measure of life-time wealth ($W_{i,t-1}$), i.e. is the discounted sum of current wealth holdings and expected future pension income (see Data section for details). The rationale behind this rescaling is that the effect of a given wealth shock is likely to depend on the initial wealth level. A \$50,000 loss might not be noteworthy for the very rich but is painful for the poorer. And what matters is not just what an individual possesses at the time of the shock but also what she expects to earn in the future. If she has high annual income and still many years to live a given wealth

loss can be easily compensated by dissaving. Taking into account not just current wealth but also future income makes sense especially for retirees. They typically have constant pension income and a limited time horizon of remaining years to live. An additional advantage of rescaling by life-time wealth instead of current wealth is that life-time wealth has fewer zeros or negative values which have to be excluded from the analysis. Results, however, are not driven by the inclusion of life-time wealth. The overall effect pattern remains the same when rescaling wealth shocks by current wealth instead of life-time wealth (see Web Appendix).

In the baseline specification (equation 2.2) I regress changes in health directly on constructed wealth shocks. An alternative specification would be the two-stage least squares regression with constructed wealth shocks as an instrument for changes in reported wealth. Such a specification would provide us with estimates that are scaled in terms of the average change in reported wealth associated with a given constructed wealth shock. But reported wealth is net of consumption. And as people tend to adapt their consumption to wealth shocks, changes in reported wealth tend to be systematically smaller than the original wealth shock. Regressions of changes in reported wealth on constructed wealth shocks indeed provide evidence of such consumption smoothing (see Findings section). From a policy perspective, however, we are interested in estimates in terms of the actual wealth shock and not in terms of the wealth change that remains after people have adapted their consumption. This is why the direct regression of health changes on constructed wealth shocks is chosen as the baseline regression. Results from two-stage least squares regressions are reported in the Appendix (Table 2.19).

2.3 Data

The data used in this study come from the first 9 waves of the Health and Retirement Survey (HRS), covering the years 1992 to 2009.[9] The HRS is a biannual panel starting in 1992 with 12,654 individuals representing US adults aged 51 - 61 in 1992 (born during the years 1931 - 1941). In 1998 and 2004 new cohorts were added to keep the sample representative for those of age 51 and older. Per household one so-called financial respondent is interviewed about her and the other family members' income and wealth holdings. Other questionnaire items such as health measures are reported by all household members. The sample of this study is restricted to financial respondents, who report wealth and stock holdings and non-zero retirement income in the previous wave, and their spouses if existent. Further I restrict the sample to singles and couples who were retired in the previous wave, i.e. either (i) both financial respondent and spouse were neither working for pay (i.e. neither working, nor part-time working, nor partly retired) nor unemployed or (ii) both considered themselves completely retired. The final regression sample consists of about 39,500 person-year observations, of which 19,000 refer to singles. The interview month is known, so that the HRS data can be matched to monthly stock market data from the Standard & Poor's 500 stock market index (S&P500).[10] Constructed wealth shocks are generated for financial re-

[9]The data is drawn from the RAND HRS file. Variables that are not included in the RAND file were added from the HRS raw data. The AHEAD waves are not included as they have been found to suffer from systematic underreporting of stock wealth (Rohwedder et al. 2006)

[10]The S&P500 is the weighted average of 500 of the biggest actively traded companies in the US and therefore represent a broad indicator of the US stock market. However, using the Dow Jones Industrial Average, which represents only 30 companies delivers similar results.

spondents and matched to spouses. Interviews which start in one month and end in a later month are dropped as well as spouse interviews that are conducted in a different month from the financial respondent.

2.3.1 Wealth data

The HRS contains detailed information on income and wealth holdings. Financial information is reported in exact amounts and unfolding response brackets are offered if exact amounts are unknown. This study uses cleaned and partly imputed wealth data from the RAND HRS file. Current household wealth ($A_{i,t}$) consists of net housing wealth, real estate wealth, vehicles, business wealth, individual retirement accounts (IRAs), stocks and mutual funds, checking and savings accounts, CDs, savings bonds and treasury bills, bonds, other savings, and debts.

I construct a measure of life-time wealth ($W_{i,t}$) as the sum of current wealth and discounted expected future income.

$$W_{i,t} = A_{i,t} + E(\sum_{\tau=0}^{T-t} \frac{Y_{t+\tau}}{(1+r)^{t+\tau}}) \tag{2.3}$$

with $Y_{i,t}$ income and r the real annual interest rate. Current wealth and *past* earnings are well documented in the HRS. Fortunately, retiree income - consisting of pensions and annuities ($PIA_{i,t}$), old age social security ($SS_{i,t}$) and veteran benefits ($VetBen_{i,t}$) - can be expected to stay constant (in real terms) after the first receipt until the individual's end of

life. Hence we can take past year's annual income from pensions, annuities, old age social security and veteran benefits as the expectation for future income.[11] Interest rate expectations (set to 3%) are assumed to stay constant as well. Further, the survival probability is needed. I calculate (τ)-year survival rates by age (t), gender (g) and 10-year birth cohort (c) using the SSA life tables.

$$W_{i,t} = A_{i,t} + (SS_{i,t} + PAI_{i,t} + VetBen_{i,t}) \sum_{\tau=1}^{T-t} \frac{E(S_{t+\tau}|t_i, g_i, c_i)}{(1+r)^{t+\tau}} \quad (2.4)$$

Social security benefits pose a potential problem as there are financial incentives to delay take-up to age 65 (Coile et al. 2002). For retirees below age 65 who do not report receiving social security it is not clear whether they are postponing or whether they are not entitled to social security payments. I present robustness checks excluding all households with one or both spouses below age 65.

Different life expectancies within households, i.e. within couples, are a further complication. Typically wives can expect to survive their husbands, but it would be demanding to calculate all different survival constellations and the corresponding exact survivor benefit amounts. For simplicity a couple's life-time wealth is calculated by applying the couple's mean life expectancy to the sum of the couple's total annual income. Restricting the sample to singles in order to avoid this simplified life-time wealth formula for couples does not affect the pattern of the estimated ef-

[11] The HRS reports monthly (past month's) income which is multiplied by 12 to obtain future annual income.

fects (see robustness checks). The same holds true if I use current wealth ($A_{i,t}$) instead of life-time wealth to rescale wealth shocks (see Web Appendix).

A central ingredient for constructing wealth shocks is the amount of stock holdings. Direct stock holdings are well documented in each wave, but they do not include stocks held in IRAs. Retirees often hold considerable fractions of their wealth in (often various) IRAs. To calculate the total amount of stock holdings it is therefore important to know the percentage of each IRA invested in stocks.

In 2006 and 2008 for each IRA the exact percentage invested in 'stocks and mutual funds' is reported. In the 1998 to 2004 waves three categories indicate whether IRAs are invested 'mostly in stocks', 'mostly in interest-earning assets', or 'about evenly split'. I translate these categories into 100%, 0%, and 50% invested in stocks, which results in roughly the same investment distribution in 2004 as for the exact information in 2006 and 2008. The assumption of a stable investment distribution between 2004 and 2006/2008 for US IRAs is checked with data from the Survey of Consumer Finances (SCF), a US representative triennial survey with about 22,000 households per wave. The SCF reports exact information on the IRA fraction invested in stock for 2004 and 2007. The cumulative distribution function does not change significantly between SCF 2004 and SCF 2007, indicating that IRA investment distributions in the US were indeed stable over that period.

For the three initial HRS waves, 1992 to 1996, no information is available on IRAs invested in stocks. In order not to lose these entire waves,

IRAs in these years are assigned the average IRA stock investment rate of the year 1998 (52%). This adds a considerable amount of noise and results are tested against the exclusion of these waves.

Table 2.7 summarizes sample characteristics and main wealth measures per HRS wave (for further wealth summary statistics see Table 2.11 in the Appendix). In 1998 older than average cohorts are added and younger cohorts in 2004, leading to discontinuous jumps in these measures. Retiree rates increase with age, but even at age 70 for 30% of the households at least one spouse is still in the labor force. The fourth and fifth row show the information available on the fraction of IRAs invested in stocks and the respective imputed values. The regression sample includes all households who were retired in the previous wave and reported wealth, non-zero retiree income and stock holdings. In the regression sample on average about half the life-time wealth is held in current wealth and about 1/3 of all households hold at least some stocks. Since wealth shocks are constructed for households with stocks, these are the 'treated'. They are on average twice as wealthy as retirees without stocks and hold about 20% of their life-time wealth in stocks. Due to the assumption that any IRA is invested 52% in stocks the fraction of households with stocks is inflated in the first three waves. For the same reason stock holdings in these waves are artificially low because many poor households with small IRA accounts that in reality do not own any stocks are included in the group of stock holders.

The final two rows of Table 2.7 display average stock market changes between interviews and the resulting constructed wealth shocks. The booms and busts around the New Economy stock market bubble and the finan-

cial crisis, which are covered by the observation period, can be clearly seen. Averages of constructed wealth shocks per wave roughly resemble the average stock market change multiplied by the average fraction held in stocks in the previous period and range from -6% to +8%.

Figure 2.2 plots constructed wealth shocks and the S&P500 over time. Each circle represent one household and is placed at the month of the interview. Wealth shocks roughly range from -30% to +40%. These are dramatic changes. For a retiree who has about 10 years remaining to live a 10% loss in life-time wealth equals the amount of planned expenditures for a whole year. If she is smoothing consumption, she will have to spend 10% less than planned every month until the end of her life. If a fixed part of her wealth is planned for inheritance or emergencies, consumption has to decrease by even more. Notice that these dramatic wealth shocks are constructed and might not correspond to real changes in wealth. Their predictive power is assessed in the Findings section.

2.3.2 Health data

I use different health measures from the HRS as dependent variables: An index of health conditions, individual health conditions, self-reported health, self-reported change in health, a mental health index as well as survival to the next interview. For better comparability of these measures which are reported on different scales and represent health circumstances of different severity, all measures are transformed the following way. First, measures of bad health are inverted such that higher values of a measure always refer to better health. This means that a positive

coefficient on wealth shocks always refers to an improvement in the respective health measure. To make effect sizes comparable across measures, I follow an approach of van Praag and Ferrer-i-Carbonell (2008) and assign to the categories of each measure the expected value of a standard normal variable conditional on being between the category's lower and upper cut-off points implied by an ordered probit fitted on the raw sample fraction. Changes in these transformed health measures are then regressed via OLS on constructed wealth shocks and controls (van Praag and Ferrer-i-Carbonell [2008] refer to this as 'probit-adapted OLS'). Summary statistics of original and transformed health measures are reported in the Appendix, Tables 2.13 and 2.14.

The index of health conditions equals the sum of conditions which have *ever* been diagnosed by a doctor according to the respondent. The HRS questionnaire includes eight conditions: high blood pressure, heart disease, stroke, arthritis, cancer, diabetes, lung disease and psychiatric problems. These health conditions are also analyzed in separate regressions. In theory the wording of the question only allows for new ever-diagnosed conditions to appear but never to disappear. In the data, however, a significant number of people report a condition in one wave but neglects the same condition in a future wave. Including these cases tends to increase the significance of the results. It is therefore likely that such 'wrong' answers are not mere noisy but contain information about actual or perceived changes in the respondent's health. Individuals might understand the question wrongly (overlooking the 'ever') or repress the memory of a cured disease. One should therefore be aware that at least for a fraction of respondents these questions only indicate the current prevalence of a condition.

For self-reported health respondents are asked to rate their current health as poor, fair, good, very good or excellent. An additional question, self-reported changes in health, asks whether compared to the previous interview health is worse, the same, or better.[12] Self-reported changes in health are regressed directly in levels and not in first differences as the question already implies a health change.

The mental health index sums a subset of eight questions from the 20 question CES-D depression score, which has been developed to diagnose clinical depression. Six questions indicate whether the respondent experienced the following emotions all or most of the time during the past week: felt depressed, everything is an effort, sleep is restless, felt alone, felt sad, and could not get going. Two questions, that are subtracted from the index, indicate whether the respondent felt happy and enjoyed life, all or most of the time during the past week. Like the health conditions index, the mental health index is inverted for regressions so that higher values indicate better mental health. Due to coding differences I do not include the mental health index in the first two waves.

Deaths of survey participants are documented since the third wave (1996). In so-called exit surveys a proxy respondent (usually a surviving family member) is interviewed about time and circumstances of the death. Thus deaths are well documented and not just one possible reason for an observed panel attrition. 'Survival', used as the dependent variable in the

[12]Wave 1-6 offers 5 categories to rate the health changes: much worse, somewhat worse, same, somewhat better, much better. For comparability with wave 7-9 the first two and the last two categories are recoded as worse and better, respectively.

baseline regressions, indicates whether the respondent survives until the next interview. This means that survival from t to t+1 is regressed on wealth shocks from t-1 to t. Therefore only individuals up to wave 8 can be included in the survival regressions.

2.4 Empirical Specification

The identification strategy outlined above leads to the following empirical specification:

$$\Delta H_{i,t} = \alpha + \beta \frac{s_{h(i),t-1}}{W_{h(i),t-1}} \frac{\Delta SP_{m(i,t)}}{SP_{m(i,t-1)}} + \gamma \frac{s_{h(i),t-1}}{W_{h(i),t-1}} + \vartheta_t + \delta X_{i,t} + \epsilon_{i,t} \quad (2.5)$$

with indices:
i: Individual
$h(i)$: Household of (i)
t: HRS wave (biannual)
$m(i,t)$: Month of the interview of individual (i) in wave (t)

and variables:
$\Delta H_{i,t}$: Health outcomes
SP: Standard & Poor's 500 stock market index
s_{t-1}: Lagged stock holdings
W_{t-1}: Lagged life-time wealth
ϑ_t: Year dummies
$X_{i,t}$: Demographic controls: age, age^2, age^3, years of education and 1

dummy for sex, 2 for race, 4 for region of residence, 4 for degree, and 7 for lagged marital status.

Changes in different health measures are regressed via OLS on the interaction of stock market changes with the lagged fraction of life-time wealth held in stocks (constructed wealth shocks) while controlling separately for the 'main effects', i.e. the lagged stock fraction and year dummies. Including a full set of year x month dummies leads to very similar results (see Web Appendix). Health outcomes and demographics vary at the individual level, wealth at the household level and the stock market at the monthly level. As explained in the Data section all health measures are transformed such that changes are interpreted in terms of standard deviations and positive changes always refer to a health improvement. Alternative transformations such as OLS with standardized measures or ordered probit regressions with original measures lead to very similar results.[13] Standard errors are clustered by households. Clustering at the level of individuals, interview dates or stock market changes result in very similar standard errors.

Predetermined demographic controls such as age, gender, race or lagged martial status may be included to decrease the variance of the regression residual and thereby increase the precision of the estimates. The inclusion of demographic controls should not change the point estimate of constructed wealth shocks if the latter are (conditionally) independent. Summary statistics of demographic controls are reported in the Appendix, Table 2.12.

[13]The advantage of probit-adapted OLS over standardizing is that it takes into account a possibly unbalanced distribution of the sample over the measure's different categories.

2.5 Findings

2.5.1 Predictive power of constructed wealth shocks

Constructed wealth shocks are highly predictive of changes in reported wealth. As reported in column (1) of Table 2.2 the regression of changes in reported wealth on constructed wealth shocks and controls yields a highly significant coefficient of about 0.74. Excluding the first waves for which the information on stocks in IRAs is noisy results in a slight increase to 0.83. This means that a constructed wealth shock of 10% corresponds to a change in reported wealth by about 8%. As argued above, retirees are likely to adapt their consumption to wealth shocks. The estimated coefficient suggests that out of a 10% wealth shock 2% goes into consumption.[14] In column (3) and (4) of Table 2.2 the exact stock fraction is substituted by a dummy for stock holdings. Again stock market changes are highly predictive of wealth changes for those with stocks. A 10% change in the stock market leads to a 1.6% change in the wealth of stock holders.

Notice that the stock market effect on those without stocks (i.e. the coef-

[14]This implies a propensity to consume out of stock wealth of 20%. Compared to the literature that has found estimates ranging from 1-5% this seem very large (Poterba 2000). A possible explanation could be the old age of the sample. Consumption smoothing implies that the propensity to consume out of a given wealth shock increases with age. If you have less years to live a given shock has to be smoothed over fewer years. But the coefficient on wealth shocks might also be attenuated due to measurement error in the lagged stock fraction. The 20% estimate should probably not be overinterpreted.

ficient on 'stock market change') is slightly negative in the overall sample and not significantly different from zero for the years with exact data on stock holdings. This gives further support to the conclusion of Figure 2.1, that there is not much of an effect of the stock market cycle on the wealth of retirees without stocks. Further, the R^2 is extremely low despite the inclusion of a broad set of demographic controls. This indicates that reported wealth in first differences is a noisy measure. Despite this noise constructed wealth shocks do a good job in picking up actual changes in reported wealth. Let us now turn to the effects of these wealth shocks on health outcomes.

2.5.2 Effects of wealth shocks on health outcomes

Table 2.3 reports the baseline regressions of five health measures (rows) on constructed wealth shocks. Regressions in column (1) include as controls only the main effects, i.e. the lagged fraction of wealth held in stocks, the stock market change and year fixed effects. In column (2) a broad set of demographics is added and in column (3) the first four waves with noisy information on stocks in IRAs are excluded. All estimates displayed in this and the following tables refer to the coefficient on constructed wealth shocks. A positive coefficient refers to a health improvement in terms of standard deviations.

The regressions in the first column indicate a positive effect of constructed wealth shocks on the index of health conditions, self-reported change in health, the mental health index and survival. Effects are of similar size, ranging from 0.2 to 0.35. This means that a 10% wealth shock is asso-

ciated with a change of 2-3.5% of a standard deviation in the different health measures. Only for changes in self-reported health is the coefficient small and not significantly different from zero. Including a broad set of demographic controls hardly changes any of the coefficients. This provides confidence that constructed wealth shocks are independent of unobserved heterogeneity. If the estimates were strongly affected by the inclusion of predetermined controls we should be worried about the exogeneity of constructed wealth shocks. Excluding the first four waves slightly increases most of the estimates but the overall pattern does not change. The increase in the coefficients could be driven by the greater precision in stock holding data but also by the changing age composition of the sample over time or by qualitative changes in the nature of stock market shocks.

In Table 2.4 I repeat these regressions separately for the eight health conditions from the health conditions index. As in the previous regressions, all health conditions are transformed such that positive coefficients indicate a health improvement (i.e. a lower chance to get the respective health condition) in terms of standard deviations. A problem of the analysis of various health conditions is that the chance of wrongly rejecting the null increases with every additional regression.[15] In the present setup, however, significant estimates would be in line with causal effects of constructed wealth shocks for some health conditions while they would not for others. Health changes are regressed on wealth shocks over a period of on average two years. If estimated effects on health measures are causal

[15]In general one can correct for this problem by either reducing the number of tests (as done above by summarizing conditions into one index) or by adjusting p-values (Anderson 2008).

they must be driven by diseases that are responsive to environmental factors and that do not take a lot of time to develop. The regressions in Table 2.4 reveal a strongly positive effect of wealth shocks on high blood pressure, smaller effects on heart disease, strokes and psychiatric problems and no significant effect on arthritis, cancer, diabetes and lung disease. As in the regressions for health measures the inclusion of demographic controls hardly changes estimates. When the first waves are excluded in column (3) the effect on high blood pressure remains large and significant. The coefficients for heart disease, strokes and psychiatric problems, however, decrease slightly while standard errors increase rendering the effects insignificant. However, any combination of the wealth shock coefficients from the heart disease, stroke or psychiatric problems regressions is still jointly significant. For arthritis, cancer, diabetes and lung disease, on the other hand, neither pairs nor groups of three or four yield joint significance.

These heterogeneous effects across different physical health conditions are plausible (for a medical text book describing these conditions see Fauci et al. 1998). High blood pressure is the most responsive health problem in the short run and arises from both psychological stress as well as unhealthy nutrition and behavior. Moreover, high blood pressure is a cause for heart problems and strokes. Therefore a positive effect on heart problems and strokes is what one should expect given the strong effect on high blood pressure.Effects on arthritis, diabetes, lung diseases or cancer would be less plausible. Arthritis is determined by genetic disposition. Diabetes is driven by genes as well as by obesity. One could think of a response in body weight to stress, but such an indirect effect might take more than 1-2 years. And I do not find an effect of wealth shocks on

body weight. Lung diseases are typically driven by smoking or unhealthy environments at work and take a long time to develop. Regarding cancer there is a psycho-medical literature discussing stress as a potential cause, but such effects remain highly controversial (Chida et al. 2008).

The effect on psychiatric problems in Table 2.4 seems rather small compared to the large effect on the mental health index in Table 2.3. However, it is likely that not every depression that might reveal itself in the mental health index has been diagnosed by a doctor. And people might be more likely to underreport the diagnosis of psychiatric problems than individual symptoms of depression, such as restless sleep or feeling alone, as the latter carry less stigma. Looking at individual depression symptoms from the mental health index does not reveal a single driver such as hypertension for the health conditions index (results reported in the Appendix, Table 2.15). This is what we should expect. The mental health index does not represent a list of different diseases but a collection of symptoms associated with clinical depression. Any single symptom is not necessarily a sign of depression but what makes it a mental health problem is having many of the symptoms at the same time.

Note that the effect on the two-year survival rate in Table 2.3 is exactly what we should expect given the effects on mental health and in particular on high blood pressure. High blood pressure related health problems are the number one killer in the Western world (Cutler, Deaton and Lleras-Muney 2006). And the sample of analyzed elderly is already at the margin of death. On average 10% do not survive the next two years. So it does not take a massive effect on latent health for them to be pushed over this threshold.

2.5.3 Interaction with age, gender and the sign of shocks

Table 2.5 investigates the heterogeneity of effects across age and gender and the effect symmetry between negative and positive wealth shocks. The coefficients of wealth shocks interacted with the respective categories are displayed as well as the significance level of their difference. Overall, interaction terms are not estimated with much precision which is not surprising given that estimates in the overall sample are already quite noisy. But the effect heterogeneities that are strong enough to be detected are plausible.

The age interactions are strongly different in the survival regression. Wealth shocks affect survival rates for the elderly six times as much as for the younger group. The effect on the index of health conditions shows up with a similar age pattern. Effects are twice as large for the elderly. These differences are not significant, but the joint hypothesis of equality in both the survival and the health conditions regression can be rejected at the 5% level. For self-reported health and mental health no clear age differential arises. This pattern across health measures makes sense. Both mortality and health conditions show up in the data only if an individual is pushed over a certain health threshold. As the health distribution shifts with age towards worse health the density around this threshold increases with age. This means that we should observe a larger effect on mortality and health conditions for the elderly even if the effect on latent health is the same across age groups. Mental and self-reported health, on the other hand, are more continuous so that health deterioration over age does not automati-

cally imply stronger effects on these measures.

There are no significant gender differences. Mental and self-reported health seem to be stronger affected for women which would be in line with the literature on gender differences in mental health. However, the estimated differentials are imprecise and not even significantly different from zero in joint tests. These results do not imply that effects are the same for males and females but it seems that estimates are not driven by gender.

The effect asymmetry one typically has in mind when thinking of wealth shocks is that negative shocks outweigh positive ones. For survival this seems to be the case. A 10% wealth loss decreases the likelihood to survive almost six times as much as a 10% wealth gain would increase it. For the other health measures (except for the mental health index) the effect asymmetry goes in the same direction but differences are rather small and not significantly different from zero. Notice, however, that in the survival regression the last wave of mostly negative shocks is not included (since the survival coding requires knowledge of the vital status in the following period, see Data section). This means that the effect of negative shocks is estimated only for shocks in 2002/3, while regressions for the other health measures include negative shocks in 2002/3 and in 2008/9. This could explain why the asymmetry is stronger for survival if negative shocks had stronger effects in 2002/3 than in 2008/9.

Repeating regressions for the other health measures excluding the last wave indeed strongly increases the effects of negative wealth shocks (reported in the Appendix, Table 2.16). This suggests that negative shocks

in 2002/3 had a significantly stronger effect than positive shocks, while this cannot be said for negative shocks in 2008/9. There are two plausible reasons for why the negative shocks in 2008/9 might appear to have a relatively weak impact on health. First, in 2008/9 the stock market had just started to collapse in the months when the interviews were conducted while in 2002/3 the market had been going down already for about two years. This means that in the last wave there was less time for the effects of negative wealth shocks to become manifest in health outcomes. Second, in contrast to the burst of the New Economy stock market bubble in 2002/3, the 2008/9 stock market crisis came along with a collapse of the housing market and the overall economy. As argued in the Identification section the direct health effects of an economic crisis might be stronger for retirees without stocks. In this case the inclusion of time fixed effects attenuates the effect of negative shocks in 2008/9 but not in 2002/3.

2.5.4 Comparison with estimates from the literature

Smith (2005) and Adams et al. (2003) have analyzed the HRS data and do not find evidence of causal effects of wealth changes on health. It is important to clarify why my estimates are different from their findings.

Using a sample of employed individuals from the HRS Smith (2005) shows that changes in stock wealth conditional on socio-demographic controls do not correlate with changes in health. Since my sample consists of retirees, these findings for employed individuals do not necessarily contradict my results. But it is insightful to replicate this specification in my sample. In column (2) of Table 2.6 I substitute constructed wealth

shocks by reported changes in stock wealth, replicating the measure of wealth shocks in Smith (2005). The resulting coefficients are essentially zero. In column (3) changes in reported wealth are divided by life-time wealth. Estimates remain small and largely insignificant. These findings of zero effects are remarkable because we would expect potential endogeneity left in stock wealth changes to bias the estimate up and not towards zero. However, a more severe problem than potential endogeneity might be measurement error in reported stock wealth. Regressions in Table 2.2 have shown that changes in overall wealth are quite noisy and this is likely to be the case as well for changes in stock wealth. This suggest that changes in reported stock wealth might be too noisy to uncover an existing relationship of wealth changes and health outcomes.

Adams, Hurd, McFadden, Merrill and Ribeiro (2003) develop an innovative approach related to Granger causality and find that lagged wealth conditional on a broad set of socio-economic variables is not Granger-causing changes in health for almost all health measures in the HRS. However, in a recent study Stowasser, Heiss, McFadden and Winter (2011) repeat the analysis of Adams et al. (2003) using the full range of data available in the HRS. In these extended data they reject Granger causality only for three out of 40 health conditions: for cancer, female lung disease and male hypertension. The rejection for hypertension, the condition for which I find strongest effects, could be explained by contemporaneous wealth shock effects. The approach of Adams et al. (2003) tests for a causal effect of lagged wealth on health changes. If it does not take long for hypertension to respond to a wealth shock then a lagged wealth shock might already affect lagged hypertension and no effect would be left in the first difference. If effects are not permanent, this could even imply an inverted effect on

the first difference.

2.5.5 Comparison with the cross-section

How large are the estimated effects? A good way to assess the effect size is the comparison with the cross-sectional relationship of wealth and health. Regressing health on wealth in levels does not allow for a causal interpretation as the coefficient on wealth also reflects reverse causality and omitted third factors. But one would expect such endogeneity to bias the coefficient upwards. Such benchmark regressions provide an upper bound for the (average) causal effect of wealth on health in the sample, in particular if few additional controls are included.

As reported in Table 2.7 the coefficient on wealth is highly significant in all benchmark regressions. However, the estimated effects of wealth shocks are about 25% above this benchmark for the index of health conditions and the mental health index. In other words, a 10% negative wealth shock leads to a slightly larger health decline than the health gap that is associated with a 10% wealth difference in the data. Lose 10% of your life-time wealth in the stock market and you end up with slightly worse health than your neighbor who has been 10% poorer before. Benchmark regressions for individual health conditions in Table 2.12 indicate that this is still not the whole story. While wealth shocks affect only particular conditions the cross-sectional wealth gradient is strongly significant and of similar size for all health conditions, except for cancer.[16] And for

[16]For cancer the gradient is inverted meaning that richer people are more likely to have cancer. This reversal has been documented in other data sets but is so far largely

hypertension, heart disease and strokes the wealth shock effect is about twice the size of the benchmark gradient. This means that after a stock market induced wealth loss you will suffer more from hypertension and related diseases than your ex-ante poorer neighbor. But your neighbor is still more likely to have arthritis, diabetes and lung disease.

The differences between the baseline and cross-sectional estimates suggest that the effects of wealth shocks are different from the average causal effects of wealth on health in the sample. This seems plausible. Someone owning $500k can afford better health care and healthier consumption than somebody owning $300k which over time accumulates to a better health stock. This however is a different effect from losing $200k in a stock market crash, which involves high blood pressure and psychological factors such as stress and depression rather than just a slight change in health inputs.

The comparison with the cross-section also provides confidence that my estimates are not driven by a coincidental correlation of the stock market with the socio-economic status of stock market investors. If this were the case, we should observe a similar pattern of effects across health conditions as in the benchmark regressions. But the pattern is clearly different. Still one might worry that effects are driven by a correlation with the typical health *profiles* of investors. Possibly at older ages richer people tend to get more hypertension and related diseases simply because they have done well at younger ages. Tables 2.17 and 2.18 in the Appendix report benchmark regressions of health changes on wealth levels. Again

unexplained.

the pattern in these regressions is very different from the causal estimates.

2.5.6 IV strategies and alternative sample specifications

The results in Table 2.9 provide further evidence that results are not driven by a coincidental correlation of the stock market with investor types. In 2SLS regressions the interaction of the stock market changes with the individual's average stock fraction or the individual's initial stock fraction are used as instruments for constructed wealth shocks. Both the average and the initial stock fraction are time constant and therefore uncorrelated with the up's and down's of the stock market. The first two columns of Table 2.9 repeat the baseline results. In column (3) constructed wealth shocks are instrumented by the interaction of the stock market with the individual's average stock fraction over the whole observation period, while in column (4) the average for the post-1997 waves is taken. In column (5) the initial 1998 stock fraction is taken instead. Despite the significant loss of information that is implied by this strategy, most estimates in columns 3-5 remain significant. More importantly they do not go to zero but -if anything- tend to increase.

Regressions in Table 2.10 show that results are robust against various changes in the sample specification. In column (2) all financial respondents and their spouses regardless of their employment status are included as long as some kind of retirement income is reported for the household. This increases the sample size by about sixty percent, but coefficients remain largely the same. In column (3) only households are included in which both spouses are above age 64. This rules out the possibility that

results are driven by the group of pre-retirement age pensioners who are typically selected into the sample through bad health. In column (4) only single households are included. In column (5) all households without stocks in the previous period are excluded so that the sample is restricted to the 'treated'. Despite the decrease in the sample size by two-thirds estimates do not change much and the overall effect pattern remains the same. In the final column the bottom quartile from the life-time wealth distribution is excluded, which again changes estimates only slightly.

2.6 Conclusion

This paper provides evidence that wealth shocks have strongly positive effects on health outcomes of stock holding retirees in the US. A 10% wealth shock is associated with an improvement of 2-4% of a standard deviation in physical health, self-reported health, mental health and survival rates. Analyzing individual health conditions I find a strong effect on high blood pressure, smaller effects on heart diseases and strokes and no effect on arthritis, diabetes, lung disease and cancer. The analysis of interaction terms reveals that effects on physical health and mortality are significantly stronger for the elderly. Further, negative shocks tend to have stronger effects than positive shocks. The comparison with the cross-sectional relationship of wealth and health indicates that the estimated causal effects of wealth shocks are larger than the long-run wealth elasticity of health.

These findings are an important contribution to the broad economic lit-

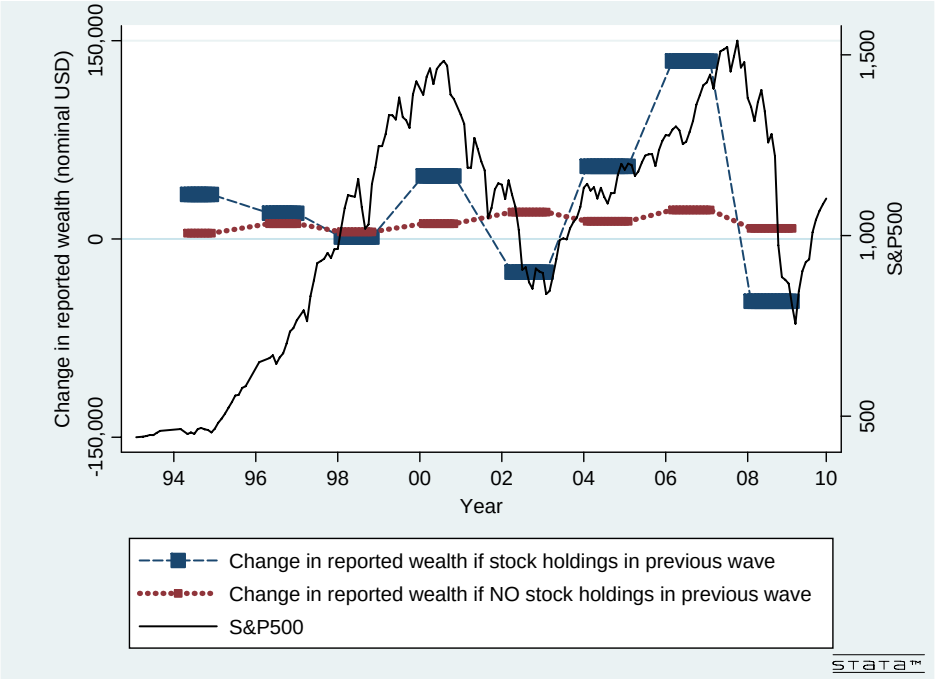
erature on the wealth-health relationship. So far positive effects of wealth changes on elderly health have been found only for poor retirees in Russia and South Africa. This paper is the first to document such effects for wealthy retirees in a wealthy country. I uncover these effects with a new measure to identify stock market fluctuations in the wealth of US retirees. This measure, the interaction of stock holdings with stock market changes, is of interest beyond the context of health economics. It could also be used to study, for example, the effects of unearned income on labor supply, savings and in particular on consumption.[17]

The pattern of affected health conditions found in this study point to a story in which psychological factors play an important role. Psychological factors as central mechanism linking economic shocks and health outcomes are in line with the results of Sullivan and von Wachter (2008). They find strong mortality effects of lay-offs for displaced workers in the US and argue that psychological reactions are the most likely mechanism underlying these effects. These could be psychological reactions to the arrival of news about future consumption as well as reactions to actual changes in consumption. Applying the empirical strategy developed in this paper to data sets that allow to study consumption behavior in detail would be a promising path for future research. Of particular use would be consumption data in combination with information on individual stock portfolio compositions. Precise information on individual stock holdings allows to construct high-frequency individual-specific wealth shocks which would greatly increase the power of such analysis without the need of extended time series of stock market changes.

[17] See Coile and Levine (2006) for a study that uses a similar approach to analyze the effects of stock market movements on retirement (as discussed above).

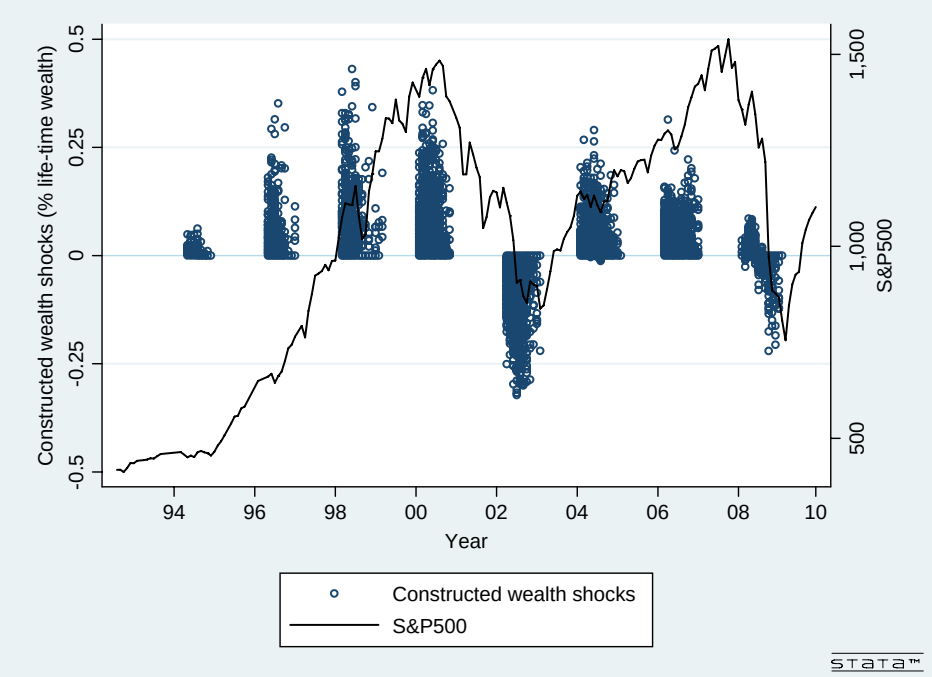
2.7 Tables and Figures

Figure 2.1: Changes in Reported Wealth and the S&P500



Average changes in reported wealth for retiree households with and without stocks in the previous period are plotted per HRS wave. The length of the bars indicates the time period in each wave over which interviews were conducted. Information on stock ownership is noisy in the first three waves. For further details on wealth measures and sample restrictions see the Data section.

Figure 2.2: Constructed Wealth Shocks and the S&P500



Constructed wealth shocks, i.e. the interaction of the previous fraction of life-time wealth held in time with the stock market change between interviews ($\frac{s_{i,t-1}}{W_{i,t-1}} \frac{\Delta S\&P_t}{S\&P_{t-1}}$), are plotted over time with the S&P500. Each circle represents the constructed wealth shock of one household and is placed in the figure at the exact month of the household's interview in t.

Table 2.1: HRS Sample Characteristics and Summary Statistics (Means).

HRS wave	1	2	3	4	5	6	7	8	9
Year	1992-1993	1994-1995	1996-1997	1998-1999	2000-2001	2002-2003	2004-2005	2006-2007	2008-2009
Full HRS sample									
N	12,583	11,354	11,142	22,420	20,775	19,542	21,353	19,687	18,446
Age	55	57	59	67	68	69	67	69	70
% retiree households	0.22	0.30	0.38	0.63	0.66	0.70	0.64	0.68	0.70
Information '% of IRA in stocks'	none	none	none	3 categories	3 categories	3 categories	3 categories	exact %	exact %
Imputed % of IRA in stocks	52%	52%	52%	0, 50, 100%	0, 50, 100%	0, 50, 100%	0, 50, 100%	exact %	exact %
Regression sample									
N	749	1,739	2,422	7,719	9,026	9,152	9,270	8,990	6,745
Current wealth (nominal USD)	154,207	215,799	236,489	244,332	289,050	318,799	355,095	442,771	431,243
Life-time wealth (nominal USD)	330,419	472,135	481,758	406,933	442,261	480,235	529,501	610,719	747,729
Fraction owning stocks (excl IRAs)	0.27	0.32	0.30	0.29	0.30	0.29	0.29	0.26	0.25
Fraction owning stocks (incl IRAs)	0.33	0.42	0.43	0.32	0.32	0.31	0.31	0.29	0.27
...those owning stocks (incl IRAs)									
N	166	487	699	1,821	2,069	2,057	2,049	1,892	1,328
Life-time wealth (nominal USD)	607,805	771,061	776,806	763,855	825,181	899,127	1,036,519	1,226,753	1,240,089
% life-time wealth held in stocks	0.11	0.11	0.11	0.19	0.20	0.19	0.20	0.21	0.23
S&P500 change since past interview	-	0.09	0.46	0.71	0.32	-0.32	0.14	0.16	0.02
Constructed wealth shock	-	0.01	0.05	0.08	0.06	-0.06	0.03	0.03	0.003

New cohorts were added to the HRS sample in 1998 and 2004. Retiree households refer to singles or couples with neither working for pay nor being unemployed. The regression sample includes all households that were retired and reported their wealth, retiree income and stock holdings in the previous wave. Life-time wealth is the sum of current wealth and expected future discounted retiree income (see Data section). In the first three waves: due to the uniform imputation of the fraction of IRA invested in stocks the fraction of stock owners (incl. IRAs) is inflated and the fraction of life-time wealth held in stocks is deflated. Average constructed wealth shocks (last row) roughly resemble the product of the previous wave's fraction of life-time wealth in stocks and the S&P500 change since past interview. Further wealth summary statistics are reported in the Appendix.

Table 2.2: Regressions of Changes in Reported Wealth on Constructed Wealth Shocks.

Dependent Variable: Wealth Change	Full sample (1)	Year>1999 (2)	Full sample (3)	Year>1999 (4)
Constructed wealth shock	0.739***	0.827***		
	(0.154)	(0.194)		
Lagged stock fraction	-0.305***	-0.307***		
	(0.066)	(0.069)		
Lagged stock ownership x Stock market change			0.145***	0.222***
			(0.039)	(0.053)
Lagged stock ownership			-0.057**	-0.055**
			(0.024)	(0.025)
Stock market change	-0.056**	-0.032	-0.058*	-0.048
	(0.028)	(0.033)	(0.032)	(0.038)
Demographics	✓	✓	✓	✓
n	29,904	26,971	29,904	26,971
R^2	0.003	0.003	0.002	0.002

The dependent variable is the change in reported household wealth divided by lagged life-time wealth ($\Delta A_{i,t}/W_{i,t-1}$). Constructed wealth shocks are the interaction of the lagged stock fraction and the stock market change ($\frac{s_{i,t-1}}{W_{i,t-1}} \frac{\Delta SP_t}{SP_{t-1}}$). 'Lagged stock ownership' is a dummy indicating stock ownership in the previous period. Regressions include only one observation per sample household and year. For details on wealth measures and sample restrictions see the Data section. Standard errors, in parenthesis, are clustered by household.

Table 2.3: Baseline Regressions of Health Measures on Wealth Shocks

Dependent Variable	(1)	(2)	(3)
Δ Index of Health Conditions	0.220***	0.225***	0.249***
	(0.063)	(0.063)	(0.070)
n	32,079	32,079	30,048
Δ Self-reported Health	0.067	0.096	0.212
	(0.117)	(0.117)	(0.136)
n	39,292	39,292	35,045
Self-reported Change in Health	0.340***	0.332***	0.235*
	(0.116)	(0.114)	(0.128)
n	39,315	39,315	35,066
Δ Mental Health Index	0.291**	0.313**	0.408**
	(0.139)	(0.139)	(0.160)
n	32,880	32,880	31,043
Survival	0.161*	0.188**	0.233**
	(0.098)	(0.091)	(0.107)
n	31,333	31,333	27,301
Controls			
Main effects	✓	✓	✓
Demographics		✓	✓
Restricted to year>1999			✓

The coefficient on constructed wealth shocks ($\frac{s_{i,t-1}}{W_{i,t-1}} \frac{\Delta SP_t}{SP_{t-1}}$) is displayed. A positive coefficient refers to a health *improvement* in the respective dependent variable in terms of standard deviations (see Data section). 'Main effects' are the lagged fraction of wealth held in stocks ($\frac{s_{i,t-1}}{W_{i,t-1}}$) and year dummies. 'Demographics' are age, age^2, age^3, dummies for gender and race, eight dummies for lagged marital status, five region dummies, five education dummies, and years of education. In column (3) the first three waves with no data on stocks in IRAs are excluded. The estimation method used is OLS. Standard errors, in parenthesis, are clustered by households.

Table 2.4: Baseline Regressions of Health Conditions on Wealth Shocks

Dependent Variable	(1)	(2)	(3)
Δ High blood pressure	0.166***	0.167***	0.172***
	(0.057)	(0.057)	(0.064)
n	37,064	37,064	33,990
Δ Heart disease	0.118*	0.123*	0.113
	(0.067)	(0.067)	(0.079)
n	38,479	38,479	34,327
Δ Stroke	0.092*	0.089*	0.082
	(0.050)	(0.050)	(0.060)
n	39,010	39,010	34,896
Δ Diabetes	0.019	0.010	-0.007
	(0.036)	(0.036)	(0.039)
n	38,493	38,493	34,638
Δ Cancer	0.050	0.058	0.043
	(0.047)	(0.048)	(0.055)
n	38,870	38,870	34,792
Δ Arthritis	0.000	0.005	0.021
	(0.066)	(0.066)	(0.072)
n	36,843	36,843	33,829
Δ Lung disease	0.047	0.049	0.043
	(0.034)	(0.034)	(0.038)
n	38,301	38,301	34,428
Δ Psychiatric problems	0.072*	0.073*	0.059
	(0.040)	(0.040)	(0.046)
n	37,819	37,819	34,037
Controls			
Main effects	✓	✓	✓
Demographics		✓	✓
Restricted to year>1999			✓

Positive coefficients refer to health *improvements* in the respective dependent variable in terms of standard deviations. Further comments as in Table 2.3.

Table 2.5: Regressions of Health Measures on Wealth Shocks Interacted with Age, Gender and Sign of Shock

	Interaction category								
	Age			Gender			Sign of shocks		
Dependent Variable	<=75 (1)	>75 (2)	Δ (p-value) (3)	Male (4)	Female (5)	Δ(p-value) (6)	Shocks≤ 0 (7)	Shocks≥ 0 (8)	Δ (p-value) (9)
Δ Index of Health Conditions	0.140*	0.299***	0.214	0.186*	0.241***	0.671	0.354*	0.235**	0.629
	(0.080)	(0.099)		(0.103)	(0.080)		(0.212)	(0.116)	
Δ Self-reported Health	0.093	0.099	0.982	-0.032	0.168	0.371	0.140	-0.116	0.549
	(0.145)	(0.186)		(0.175)	(0.150)		(0.382)	(0.206)	
Self-reported Change in Health	0.316**	0.166	0.509	0.138	0.470***	0.145	0.204	0.136	0.889
	(0.152)	(0.170)		(0.178)	(0.145)		(0.425)	(0.224)	
Δ Mental Health Index	0.334*	0.305	0.918	0.148	0.436**	0.283	0.267	0.350	0.871
	(0.187)	(0.208)		(0.204)	(0.183)		(0.474)	(0.223)	
Survival	0.066	0.369**	0.105	0.221	0.147	0.696	1.189***	0.180	0.029
	(0.081)	(0.169)		(0.156)	(0.111)		(0.428)	(0.171)	
Controls (interacted)									
Main effects	✓				✓			✓	
Demographics	✓				✓			✓	

The coefficients on constructed wealth shocks ($\frac{s_{i,t-1}}{W_{i,t-1}} \frac{\Delta SP_t}{SP_{t-1}}$) interacted with the two respective subgroups are displayed. 'Δ (p-value)' indicates the significance level of the difference between the two interacted coefficients. Positive coefficients refer to a health *improvement* in the respective dependent variable in terms of standard deviations (see Data section). All controls are interacted (except time effects in the 'sign of shocks' regressions) and all waves are included. 'Main effects' are the lagged fraction of wealth held in stocks ($\frac{s_{i,t-1}}{W_{i,t-1}}$) and year dummies. 'Demographics' are age, age^2, age^3, dummies for gender and race, eight dummies for lagged marital status, five region dummies, five education dummies, and years of education. Numbers of observations are the same as in the baseline regressions, Table 2.3, column (2). The estimation method used is OLS. Standard errors, in parenthesis, are clustered by households.

Table 2.6: Regressions of Health Measures on Changes in Reported Stock Wealth

	Specification of wealth shocks		
	Baseline	Using changes in *reported* stock wealth	
	$\frac{\Delta SP_t}{SP_{t-1}} s_{i,t-1}/W_{i,t-1}$	$(s_{i,t} - s_{i,t-1})/10,000$	$(s_{i,t} - s_{i,t-1})/W_{i,t-1}$
Dependent Variable	(1)	(2)	(3)
Δ Index of Health Conditions	0.225***	0.00007	0.002
	(0.063)	(0.00005)	(0.009)
n	32,079	32,079	32,079
Δ Self-reported Health	0.096	0.00007	0.032*
	(0.117)	(0.00008)	(0.017)
n	39,292	39,292	39,292
Self-reported Change in Health	0.332***	0.00006	0.011
	(0.114)	(0.00006)	(0.017)
n	39,315	39,315	39,315
Δ Mental Health Index	0.313**	0.00004	-0.010
	(0.139)	(0.00011)	(0.020)
n	32,880	32,880	32,880
Survival	0.188**	-0.00004	-0.012
	(0.091)	(0.00008)	(0.019)
n	31,333	31,333	31,333
Controls			
Main effects	✓	✓	✓
Demographics	✓	✓	✓

The coefficient on wealth shocks as defined at the top of each column is displayed. $\frac{\Delta SP_t}{SP_{t-1}}$ = percentage change in the S&P500; $s_{i,t}$ = stock wealth; $W_{i,t}$ = life-time wealth (see Data section). A positive coefficient refers to a health *improvement* in the respective dependent variable in terms of standard deviations. 'Main effects' are the lagged fraction of wealth held in stocks ($\frac{s_{i,t-1}}{W_{i,t-1}}$) and year dummies. 'Demographics' are age, age^2, age^3, dummies for gender and race, eight dummies for lagged marital status, five region dummies, five education dummies, and years of education. The estimation method used is OLS. Standard errors, in parenthesis, are clustered by households.

Table 2.7: Benchmark Regressions of Health Measures on Ln of Life-time Wealth

	Regression specification	
	Baseline	Benchmark
	$\Delta H_{i,t}$ on $\frac{\Delta S\&P_t}{S\&P_{t-1}} \frac{s_{i,t-1}}{W_{i,t-1}}$	$H_{i,t}$ on $lnW_{i,t}$
Dependent Variable ($H_{i,t}$)	(1)	(2)
Health Conditions Index	0.225***	0.175***
	(0.063)	(0.009)
Self-reported Health	0.096	0.291***
	(0.117)	(0.007)
Mental Health Index	0.313**	0.232***
	(0.139)	(0.008)
Controls		
Main effects	✓	
Demographics	✓	
Age, age^2, male		✓

In column (1) the coefficient on constructed wealth shocks is displayed; for further comments on the baseline regressions see Table . In column (2) the coefficient from OLS regressions of health levels on ln life-time wealth is displayed. Only age, age^2 and male are included as controls such that life-time wealth proxies for the overall socio-economic status. The inclusion of further controls decreases the coefficients on ln life-time wealth. There are no level equivalents for 'self-reported change in health' and for 'survival'. Benchmark regressions of health changes on ln life-time wealth are reported in the Appendix.

Table 2.8: Benchmark Regressions of Health Conditions on Ln of Lifetime Wealth

	Regression specification	
	Baseline	Benchmark
	$\Delta H_{i,t}$ on $\frac{\Delta S\&P_t}{S\&P_{t-1}} W_{i,t-1}$	$H_{i,t}$ on lnW_t
Dependent Variable ($H_{i,t}$)	(1)	(2)
High blood pressure	0.167***	0.077***
	(0.057)	(0.007)
Heart disease	0.123*	0.053***
	(0.067)	(0.007)
Stroke	0.089*	0.060***
	(0.050)	(0.005
Arthritis	0.005	0.054***
	(0.066)	(0.006)
Cancer	0.058	-0.045***
	(0.048)	(0.006)
Diabetes	0.010	0.093***
	(0.036)	(0.006)
Lung disease	0.049	0.063***
	(0.034)	(0.005)
Psychiatric problems	0.073*	0.093***
	(0.040)	(0.006)
Controls		
Main effects	✓	
Demographics	✓	
Age, age^2, male		✓

For comments see previous table.

Table 2.9: 2SLS Regressions with Average or Initial Stock Holdings as Instrument for Actual Stock Holdings

	Regression specification				
	Baseline		2SLS		
			IV for constructed wealth shocks		
			$\frac{\Delta S\&P_t}{S\&P_{t-1}}[\frac{s_i}{W_i}]^{average}$		$\frac{\Delta S\&P_t}{S\&P_{t-1}}[\frac{s_i}{W_i}]^{1998}$
Dependent Variable	(1)	(2)	(3)	(4)	(5)
Δ Index of Health Conditions	0.225***	0.249***	0.293***	0.318***	0.316***
	(0.063)	(0.070)	(0.074)	(0.082)	(0.102)
n	32,079	30,048	32,079	30,048	19,441
Δ Self-reported Health	0.096	0.212	0.016	0.129	0.023
	(0.117)	(0.136)	(0.134)	(0.157)	(0.207)
n	39,292	35,045	39,292	35,045	22,847
Self-reported Change in Health	0.332***	0.235*	0.357***	0.223	0.170
	(0.114)	(0.128)	(0.130)	(0.144)	(0.173)
n	39,315	35,066	39,315	35,066	22,863
Δ Mental Health Index	0.313**	0.408**	0.390**	0.502***	0.513**
	(0.139)	(0.160)	(0.163)	(0.188)	(0.233)
n	32,880	31,043	32,880	31,043	20,047
Survival	0.188**	0.233**	0.190*	0.252**	0.076
	(0.091)	(0.107)	(0.108)	(0.126)	(0.146)
n	31,333	27,301	31,333	27,301	19,237
Controls					
Year dummies	✓	✓	✓	✓	✓
Demographics	✓	✓	✓	✓	✓
Lagged stock fraction	✓	✓			
Average stock fraction 92-06			✓		
Average stock fraction 98-06				✓	
1998 stock fraction					✓
Restricted to year>1999		✓		✓	✓

The coefficient on constructed wealth shocks ($\frac{s_{i,t-1}}{W_{i,t-1}}\frac{\Delta SP_t}{SP_{t-1}}$) is displayed. The estimation method used is OLS in columns (1) and (2), and 2SLS in columns (3)-(5). The instrument for constructed wealth shocks in (3) is $\frac{\Delta S\&P_t}{S\&P_{t-1}}[\frac{s_i}{W_i}]^{average}_{92-06}$, in (4) is $\frac{\Delta S\&P_t}{S\&P_{t-1}}[\frac{s_i}{W_i}]^{average}_{98-06}$, and in (5) is $\frac{\Delta S\&P_t}{S\&P_{t-1}}[\frac{s}{W}]^{1998}$. First stage t-statistics for these instruments are about 134, 132 and 63, respectively. Further comments as in Table 2.3.

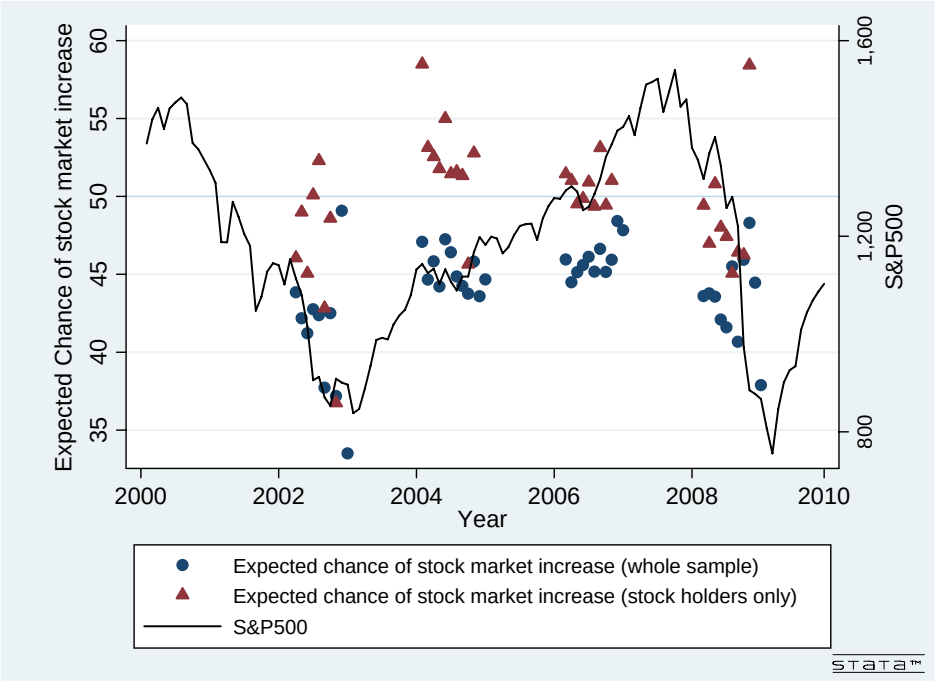
Table 2.10: Alternative sample specifications

	Sample specification					
	Baseline	Incl. non-retirees	Age>64	Singles only	$s_{i,t-1} > 0$	Excl. poorest quartile
Dependent Variable	(1)	(2)	(3)	(4)	(5)	
Δ Index of Health Conditions	0.225***	0.190***	0.293***	0.415***	0.142*	0.219***
	(0.063)	(0.050)	(0.079)	(0.108)	(0.075)	(0.065)
n	32,079	50,477	23,826	15,185	12,051	23,718
Δ Self-reported Health	0.096	0.137	0.233	0.252	0.144	0.066
	(0.117)	(0.085)	(0.153)	(0.201)	(0.141)	(0.124)
n	39,292	62,617	27,889	18,637	14,505	28,475
Self-reported Change in Health	0.332***	0.327***	0.182	0.353*	0.253*	0.184
	(0.114)	(0.090)	(0.141)	(0.184)	(0.140)	(0.120)
n	39,315	62,648	27,913	18,658	14,507	28,486
Δ Mental Health Index	0.313**	0.211**	0.390**	0.469*	0.260	0.332**
	(0.139)	(0.106)	(0.175)	(0.241)	(0.169)	(0.145)
n	32,880	51,104	24,443	15,695	12,658	24,842
Survival	0.188**	0.154**	0.195	0.238	0.206*	0.204**
	(0.091)	(0.063)	(0.125)	(0.175)	(0.111)	(0.095)
n	31,333	50,503	21,423	14,613	11,827	22,483
Controls						
Main effects	✓	✓	✓	✓	✓	✓
Demographics	✓	✓	✓	✓	✓	✓

Column (2): non-retired individuals are included (as long as some kind of retirement income is reported for HH). (3): HH are excluded if financial respondent or spouse is below age 65. (4): Only single retiree households are included. (5): HH without stocks in (t-1) are excluded. (6): HH from quartile of the life-time wealth distribution are excluded. Further comments as in Table 2.3

2.8 Appendix

Figure 2.3: HRS Expectations of an Increase in the Stock Market and the S&P500



Monthly averages of the following question in the HRS are plotted: ’By next year at this time, what is the percent chance that mutual fund shares invested in blue chip stocks like those in the Dow Jones Industrial Average will be worth more than they are today?’ Averages for months with less than 25 responses are not displayed.

Table 2.11: Summary Statistics Wealth Measures

Wealth measure	Symbol (1)	Mean (2)	Std. dev. (3)
Reported household wealth (nominal USD)	$A_{i,t}$	334,285	1,002,583
Change in reported household wealth (nominal USD)	$\Delta A_{i,t}$	24,519	1,039,176
Household life-time wealth (nominal USD)	$W_{i,t}$	523,615	3,403,015
Relative change in reported household wealth	$\frac{\Delta A_{i,t}}{W_{i,t-1}}$	0.122	1.419
Fraction of life-time wealth held in stocks	$\frac{s_{i,t}}{W_t}$	0.061	0.141
Percentage change in the S&P500	$\frac{S\&P_t}{S\&P_{t-1}}$	0.160	0.302
Constructed wealth shocks	$\frac{s_{i,t-1}}{W_{i,t-1}} \frac{\Delta S\&P_t}{S\&P_{t-1}}$	0.006	0.039

Comments as in Table 2.1.

Table 2.12: Summary Statistics Demographic Controls

Variable	Mean	Std. dev.	Variable	Mean	Std. dev.
Sex			*Education*		
Female	0.621		Years of education	11.59	3.41
			Less than high school	0.316	
Age			GED diploma	0.045	
Age	72.78	9.63	High-school graduate	0.324	
Age^2	5,389	1,418	Some college	0.177	
Age>75	0.388		College and above	0.138	
Race			*Marital status (lagged)*		
White	0.820		Married	0.518	
Black	0.150		Married, spouse absent	0.009	
Other	0.030		Partnered	0.016	
			Separated	0.013	
Region			Divorced	0.079	
Northeast	0.164		Separated/divorced	0.005	
Midwest	0.246		Widowed	0.329	
South	0.411		Never married	0.032	
West	0.178				
Other	0.001				

Standard deviations are omitted for binary variables.

Table 2.13: Summary Statistics of Health Measures.

	Transformation							
	Original					Standardized		
	Levels		First difference			Levels		First difference
	Range	Mean (Std. dev.)	Range	Mean (Std. dev.)	Inversed	Mean (Std. dev.)		Mean (Std. dev.)
Health measure	(1)	(2)	(3)	(4)	(5)	(6)		(7)
Index of health conditions	[0;...;8]	2.350 (1.445)	[-5;...;5]	0.257 (0.615)	yes	0.000 (0.975)		-0.175 (0.417)
Self-reported health	[0;...;4]	1.904 (1.145)	[-4;...;4]	-0.081 (0.938)	no	0.000 (0.959)		-0.069 (0.790)
Self-reported change in health	[-1;0;1]	-0.225 (0.598)	-	-	no	0.000 (0.871)		-
Mental health index	[0;...;8]	1.694 (2.020)	[-8;...;8]	0.038 (1.831)	yes	0.000 (0.934)		-0.018 (0.861)
Survival	[0;1]	0.930 (0.255)	-	-	no	0.000 (0.527)		-

For comments see the Data section.

Table 2.14: Summary Statistics of Health Conditions.

	Transformation							
	Original					Standardized		
	Levels		First difference			Levels	First difference	
Health condition	Range (1)	Mean (Std. dev.) (2)	Range (3)	Mean (Std. dev.) (4)	Inversed (5)	Mean (Std. dev.) (6)	Mean (Std. dev.) (7)	
High blood pressure	[0;1]	0.601	[-1;0;1]	0.049	yes	0.000	-0.079	
		(0.490)		(0.240)		(0.788)	(0.386)	
Heart disease	[0;1]	0.307	[-1;0;1]	0.044	yes	0.000	-0.073	
		(0.461)		(0.256)		(0.761)	(0.422)	
Stroke	[0;1]	0.103	[-1;0;1]	0.020	yes	0.000	-0.038	
		(0.304)		(0.184)		(0.590)	(0.358)	
Arthritis	[0;1]	0.201	[-1;0;1]	0.030	yes	0.000	-0.052	
		(0.401)		(0.177)		(0.701)	(0.309)	
Cancer	[0;1]	0.157	[-1;0;1]	0.026	yes	0.000	-0.048	
		(0.364)		(0.172)		(0.660)	(0.313)	
Diabetes	[0;1]	0.121	[-1;0;1]	0.022	yes	0.000	-0.041	
		(0.326)		(0.160)		(0.617)	(0.303)	
Lung disease	[0;1]	0.674	[-1;0;1]	0.045	yes	0.000	-0.073	
		(0.469)		(0.241)		(0.769)	(0.395)	
Psychiatric problems	[0;1]	0.175	[-1;0;1]	0.025	yes	0.000	-0.045	
		(0.380)		(0.185)		(0.679)	(0.330)	

For comments see the Data section.

Table 2.15: Regressions of Mental Health Index Items on Wealth Shocks

Dependent Variable	(1)	(2)	(3)
Δ Felt depressed	0.258**	0.251**	0.260*
	(0.119)	(0.119)	(0.138)
Δ Everything is an effort	0.129	0.143	0.128
	(0.124)	(0.124)	(0.143)
Δ Sleep is restless	0.146	0.155	0.223
	(0.138)	(0.139)	(0.161)
Δ Felt alone	0.207*	0.212*	0.251*
	(0.121)	(0.122)	(0.141)
Δ Felt sad	0.220*	0.223*	0.258*
	(0.128)	(0.128)	(0.146)
Δ Could not get going	-0.002	0.019	0.113
	(0.134)	(0.134)	(0.152)
Δ Felt happy	0.057	0.055	0.086
	(0.112)	(0.112)	(0.128)
Δ Enjoyed life	0.074	0.079	0.123
	(0.099)	(0.100)	(0.113)
Controls			
Main effects	✓	✓	✓
Demographics		✓	✓
Restricted to year>1999			✓

The coefficient on constructed wealth shocks is displayed. A positive coefficient refers to an *improvement* in the respective dependent variable, e.g. feeling *less depressed* or feeling *more happy*, in terms of standard deviations. Further comments as in Table 2.3.

Table 2.16: Baseline and Interacted Regressions of Health Measures, Excluding Wave 9

	Full sample	Wave 9 excluded				
	Baseline	Baseline	Interacted (sign of shock)			
Dependent Variable	(1)	(2)	Shocks$\leq$ 0 (3)	Shocks$\geq$ 0 (4)	Δ (p-value) (5)
Δ Index of Health Conditions	0.225***	0.212***	0.547*	0.249*	0.360
	(0.063)	(0.064)	(0.293)	(0.133)	
Δ Self-reported Health	0.096	0.079	0.064	-0.194	0.676
	(0.117)	(0.120)	(0.583)	(0.217)	
Self-reported Change in Health	0.332***	0.263**	1.378**	0.059	0.054
	(0.114)	(0.116)	(0.616)	(0.252)	
Δ Mental Health Index	0.313**	0.328**	0.436	0.276	0.825
	(0.139)	(0.142)	(0.704)	(0.253)	
Survival	0.188**	0.188**	1.189***	0.180	0.029
	(0.091)	(0.091)	(0.428)	(0.171)	
Controls					
Main effects	✓	✓		✓	
Demographics	✓	✓		✓	

The coefficient on constructed wealth shocks, in columns (3) and (4) interacted with the sign of shocks, is displayed. Column (5) reports the significance level of the difference between the interacted coefficients. The coefficients for survival in columns (1) and (2) are identical as survival is not observed in wave 9. Further comments as in Table 2.5.

Table 2.17: Benchmark Regressions of Changes in Health Measures on ln of Life-time Wealth

	Regression specification	
	Baseline	Benchmark
	$\Delta H_{i,t}$ on $\frac{\Delta S\&P500}{S\&P500}\frac{s_{i,t-1}}{W_{i,t-1}}$	$\Delta H_{i,t}$ on lnW_t
Dependent Variable	(1)	(2)
Δ Health Conditions Index	0.225***	0.000
	(0.063)	(0.002)
Δ Self-reported Health	0.096	-0.011***
	(0.117)	(0.003)
Self-reported Change in Health	0.332***	0.069***
	(0.114)	(0.006)
Δ Mental Health Index	0.313**	-0.014***
	(0.139)	(0.004)
Survival	0.188**	0.056***
	(0.091)	(0.004)
Controls		
Main effects	✓	
Demographics	✓	
Age, age^2, male		✓

In column (1) the coefficient on constructed wealth shocks is displayed. In column (2) the coefficient on the ln life-time wealth from OLS regressions of health changes on ln life-time wealth is displayed. Only age, age^2 and male are included as controls such that life-time wealth proxies for the overall socio-economic status. For further comments on the baseline and benchmark regressions see Tables 2.7 and 2.8, respectively.

Table 2.18: Benchmark Regressions of Changes in Health Conditions on ln of Life-time Wealth

	Regression specification	
	Baseline	Benchmark
	$\Delta H_{i,t}$ on $\frac{\Delta S\&P500}{S\&P500}\frac{s_{i,t-1}}{W_{i,t-1}}$	$\Delta H_{i,t}$ on lnW_t
Dependent Variable	(1)	(2)
Δ High blood pressure	0.167***	-0.003
	(0.057)	(0.002)
Δ Heart disease	0.123*	-0.001
	(0.067)	(0.002)
Δ Stroke	0.089*	0.009***
	(0.050)	(0.002
Δ Arthritis	0.005	-0.010***
	(0.066)	(0.002)
Δ Cancer	0.058	-0.003*
	(0.048)	(0.001)
Δ Diabetes	0.010	0.008***
	(0.036)	(0.002)
Δ Lung disease	0.049	0.004***
	(0.034)	(0.002)
Δ Psychiatric problems	0.073*	0.005***
	(0.040)	(0.002)
Controls		
Main effects	✓	
Demographics	✓	
Age, age^2, male		✓

For comments see previous table.

Table 2.19: Constructed Wealth Shocks as Instrument for Changes in Reported Wealth

	Regression specification			
	Baseline		2SLS	
	$\frac{\Delta S\&P500}{S\&P500}\frac{s_{i,t-1}}{W_{i,t-1}}$ as regressor		$\frac{\Delta S\&P500}{S\&P500}\frac{s_{i,t-1}}{W_{i,t-1}}$ as IV for reported wealth changes	
Dependent Variable	(1)	(2)	(3)	(4)
Δ Index of Health Conditions	0.225***	0.249***	0.303***	0.304***
	(0.063)	(0.070)	(0.107)	(0.111)
First stage F-statistic	-	-	22.94	19.27
Δ Self-reported Health	0.096	0.212	0.129	0.253
	(0.117)	(0.136)	(0.158)	(0.168)
First stage F-statistic	-	-	29.03	22.85
Self-reported Change in Health	0.332***	0.235*	0.444**	0.299*
	(0.114)	(0.128)	(0.174)	(0.163)
First stage F-statistic	-	-	29.22	22.90
Δ Mental Health Index	0.313**	0.408**	0.400**	0.471**
	(0.139)	(0.160)	(0.194)	(0.210)
First stage F-statistic	-	-	26.30	21.71
Survival	0.188**	0.233**	0.255*	0.282**
	(0.091)	(0.107)	(0.131)	(0.140)
First stage F-statistic	-	-	29.32	22.01
Controls				
Main effects	✓	✓	✓	✓
Demographics	✓	✓	✓	✓
Restricted to year>1999		✓		✓

Columns (1) and (2) display the coefficients on constructed wealth shocks in the baseline regressions. Columns (3) and (4) display the coefficient on changes in reported wealth ($\Delta A_{i,t}/W_{i,t-1}$) in 2SLS regressions with constructed wealth shocks as instrument. First-stage regressions are reported in Table . First-stage F-statistics vary across health measures due to differences in the number of observations. For further comments see Table 2.3.

Bibliography

[1] Adams, Peter, Michael D. Hurd, Daniel McFadden, Angela Merrill and Tiago Ribeiro (2003) "Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status." *Journal of Econometrics*, 112:1, 3-56.

[2] Adda, Jerome, James Banks, and Hans-Martin von Gaudecker (2009) "The Impact of Income Shocks on Health: Evidence from Cohort Data," *Journal of the European Economic Association*, 7 (6), 1361-1399.

[3] Anderson, Michael L. (2008) "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedaian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103 (484):1481-1495.

[4] Angrist, Joshua D. and J-Steffen Pischke (2009) *Mostly Harmless Econometrics: An Empiricists Companion*, Princeton, NJ: Princeton University Press.

[5] Apouey, Benedicte and Andrew Clark (2010) "Winning Big but Feeling No Better? The Effect of Lottery Prizes on Physical and Mental Health" *IZA Discussion Paper* Nr. 4730.

[6] Case, Anne (2004) "Does Money Protect Health Status? Evidence from South African Pensions" in David Wise (Ed.), *Perspectives on the Economics of Aging*, Chicago: University of Chicago Press.

[7] Chida, Yoichi, and Andrew Steptoe (2008) "Positive psychological wellbeing and mortality: a quantitative review of prospective observational studies." *Psychosomatic Medicine*, 70, 741-756.

[8] Chida, Yoichi, Mark Hamer, Jane Wardle, and Andrew Steptoe (2008) "Do stress-related psychological factors contribute to cancer incidence and survival: a systematic review and meta-analysis." *Nature Clinical Practice Oncology*, 5, 466-475.

[9] Coile, Courtney C. and Phillip B. Levine (2006) "Bulls, Bears, and Retirement Behavior." *Industrial and Labor Relations Review*, 59 (3), 408-429.

[10] Coile, Courtney, Peter Diamond, Jonathan Gruber, and Alain Jousten (2002) "Delays in Claiming Social Security Benefits," *Journal of Public Economics*, 84 (3), 357-85.

[11] Cutler, David M., Deaton, Angus, Lleras-Muney, Adriana (2006) "The Determinants of Mortality." *Journal of Economic Perspectives*, 20 (3).

[12] Cutler, David M., Lleras-Muney, Adriana, and Vogl, Tom (2011) "Socioeconomic Status and Health: Dimensions and Mechanisms." in S. Glied and P. Smith (Ed.), *Oxford Handbook of Health Economics*, Oxford University Press.

[13] Deaton, Angus (2003) "Health, inequality, and economic development" *Journal of Economic Literature*, 41/1: 113-158.

[14] Deaton, Angus and Christina Paxson (2001) "Mortality, Education, Income, and Inequality among American Cohorts" in David Wise (Ed.), *Themes in the Economics of Aging*, Chicago: University of Chicago Press.

[15] Deaton, Angus and Christina Paxson (2004) "Mortality, income, and income inequality over time in Britain and the United States," in David Wise (Ed.), *Perspectives on the Economics of Aging* Chicago: University of Chicago Press.

[16] Dominitz Jeff, Manski Charles F. (2007) "Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study." *Journal of the European Economic Association*, 5(2-3):369-79.

[17] Fauci, A. S., E. Braunwald, K. J. Isselbacher, J. D. Wilson, J. B. Martin, D. L. Kasper, S. L. Hauser and Longo, D. L. (1998) *Harrisons Principles of Internal Medicine*. McGraw-Hill, New York, NY.

[18] Gardner, Jonathan and Andrew J. Oswald (2007) "Money and mental wellbeing: A longitudinal study of medium-sized lottery wins" *Journal of Health Economics*, 26, 1, 49-60.

[19] Grip, Andries de, Maarten Lindeboom, and Raymond Montizaan (2009) "Shattered Dreams: The Effects of Changing the Pension System Late in the Game," *IZA Discussion Paper*, Nr. 4034.

[20] Handwerker, Elizabeth Weber (2011) "What can the Social Security Notch tell us about the impact of additional income in retirement?" *Journal of Economic and Social Measurement*, 36, No. 1-2, 71-92

[21] Jensen, Robert and Kaspar Richter (2003) "The health implications of social security failure: Evidence from the Russian pension crisis." *Journal of Public Economics*, 88.

[22] Lindahl, Mikael (2005) "Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income." *Journal of Human Resources*, 40; 144-168.

[23] Malkiel, Burton G. (2003) "The efficient market hypothesis and its critics." *Journal of Economic Perspectives*, 17, 59-82.

[24] Michaud, Pierre-Carl, and Arthur Van Soest (2008) "Health and Wealth of Elderly Couples: Causality Tests using Dynamic Panel Data Models." *Journal of Health Economics*, 27 (5): 1312-25.

[25] Miller, Douglas L., Marianne E. Page, Ann Huff Stevens and Mateusz Filipski (2009) "Why are recessions good for your health?" *American Economic Review: Papers & Proceedings*, 99:122-127.

[26] Poterba, James M. (2000) "Stock market wealth and consumption." *Journal of Economic Perspectives*, 14 (2000), pp. 99-118.

[27] Rohwedder, Susann and Steven Haider. (2006) "Increases in Wealth Among the Elderly in the Early 1990s: How Much is Due to Survey Design?" *Review of Income and Wealth*, 52, 4.

[28] Ruhm, Christopher J. (2000) "Are recessions good for your health?" *Quarterly Journal of Economics*, 115/2: 617-650.

[29] Smith, James P. (1999) "Healthy bodies and thick wallets: The dual relation between health and economic status." *Journal of Economic Perspectives*, 13/2: 145-166.

[30] Smith, James P. (2005) "Consequences and Predictors of New Health Events" in David Wise (Ed.) *Advances in the Economics of Aging*, Chicago: University of Chicago Press.

[31] Snyder, Stephen S., and William N. Evans. (2006) "The impact of income on mortality: evidence from the Social Security notch." *Review of Economics and Statistics*, 88/3: 482-495.

[32] Stowasser, Till, Florian Heiss, Daniel L. McFadden and Joachim K. Winter (2011) "Healthy, Wealthy and Wise? Revisited: An Analysis of the Causal Pathways from Socio-Economic Status to Health" in David Wise (Ed.) *Investigations in the Economics of Aging*, Chicago: University of Chicago Press.

[33] Strike, Philip C. and Andrew Steptoe (2004) "Psychosocial Factors in the Development of Coronary Artery Disease" *Progress in Cardiovascular Diseases*, 46, 4, 337-347.

[34] Sullivan, Daniel and Till von Wachter (2009) "Job Displacement and Mortality: An Analysis using Administrative Data" *Quarterly Journal of Economics*, 124 (3), 1265-1306.

[35] van Praag, Bernard M.S. and Ada Ferrer-i-Carbonell (2008) *Happiness Quantified: A Satisfaction Calculus Approach* revised edition, Oxford: Oxford University Press.

Chapter 3

UNLUCKY COHORTS: INCOME, HEALTH INSURANCE AND AIDS MORTALITY OF RECESSION GRADUATES.

3.1 Introduction

A number of recent studies have shown that graduates who face a bad economic environment when entering the labor market suffer strong and persistent income losses.[1] In this paper I investigate for a sample of young adults the effects of graduating in unfavorable economic conditions on

[1] Oreopoulos, von Wachter and Heisz 2012; Genda, Kondo and Ohta 2010; Kahn 2009; Beaudry and DiNardo 1991

further socio-economic outcomes and on mortality. I find that unlucky recession cohorts do not only earn less. They also have worse health insurance coverage. And during the outbreak of the HIV/AIDS epidemic mortality rates have been significantly higher for these cohorts. The negative effects on income and health insurance seem the most plausible mechanisms underlying the effects on mortality.

The strong relationship of socio-economic status and health outcomes such as mortality has been documented by a broad literature in economics as well as in medicine and biology.[2] However, so far little is known about the causal mechanisms underlying this relationship. A number of studies have exploited income or wealth shocks to investigate short term effects on health outcomes.[3] Exogenous and persistent income differences that would allow to investigate gradually accumulating long-term effects of income and socioeconomic status on health are typically hard to find. Income differences due to economic conditions at graduation offer the opportunity for such an analysis.

In this paper I use data from the CPS to replicate the effects of the graduation unemployment rate on income found in the previous literature. I find that a one percentage point increase in the unemployment rate at graduation leads to an annual income loss of initially 3 % which decreases to 1.5 % over the subsequent 15 years. I document similar effects on alternative income measures. Further, I find strongly negative effects on the

[2]For reviews of the economic literature see Deaton (2003) or Cutler, Deaton and Lleras-Muney (2006), Cutler, Lleras-Muney and Vogl (2011).

[3]For example Jensen and Richter (2003), Case (2004), and Gardner and Oswald (2007)

probability to be covered by employer-provided health insurance and a positive effect on Medicaid coverage. These effects on income and health insurance are stronger for non-whites than for whites but similar across gender. Next, I show that results largely carry over to a specification in which I use the unemployment rate at age 18 instead of the actual graduation age. This allows me to analyse effects on mortality in the Vital Statistics which do not report the year of graduation. For the baseline period, 1979 to 1991, I find strong positive effects of the unemployment rate at age 18 on mortality at ages 28 to 33. This effect is driven by AIDS deaths, it is homogeneous across gender but stronger for non-whites. The effect fades out when adding more years to the analysis and is not distinguishable from zero in the overall period from 1979 to 2004. These results indicate that during the outbreak of a deadly epidemic recession graduates face much higher mortality rates. The timing of the most significant periods as well as the pattern across gender and race suggest that the negative effects on income and health insurance coverage are important mechanisms underlying these findings.

The income effects that I find are similar to those reported in the literature. Using the CPS and the PSID, Beaudry and DiNardo (1991) show that an one percentage point increase in the unemployment rate at job start leads to a subsequent income loss of 2-3 %. [4] Oreopoulos, von Wachter and Heisz (2012), hereafter OWH, use Canadian administrative data on college graduates and find an effect of -2 % on income that fades to zero

[4]Further, they document a dominant effect of the minimum unemployment rate since job start. They interpret this finding as supportive evidence for a labor market model in which workers are protected from the contemporaneous unemployment rate through their employment contracts but at the same time mobile to move to other jobs or to renegotiate contracts in order to benefit from improvements in economic conditions.

within 10 years. Stronger effects they document for the bottom tercile in the skill distribution. This group suffers initial income losses of 3 % and even after ten years annual income is depressed by 2 %. These effects for the lower skilled tercile are very similar to those presented in Beaudry and DiNardo (1991). The main mechanism OWH identify is that recession graduates take jobs at poorer quality firms when entering the labor market. In subsequent years high skilled graduates move to higher quality employers. Mobility across employers and industries is increased for this group and the average quality of the employer improves over time. The lower skilled graduates, on the other hand, are stuck at low quality firms.[5] This mechanism is in line with my findings. There is not much of an effect on employment status in the CPS. And the negative effects on income as well as employer-provided health insurance indicates that recession graduates take lower quality jobs.

My findings relate to the economic literature that analyses effects of health insurance on health outcomes. The RAND Health Insurance Experiment investigated the effects of randomized insurance co-payments for insured individuals, finding little evidence of effects on this intensive margin (Newhouse et al. 1993). The currently ongoing Oregon health care experiment randomizes at the extensive margin instead. First results show strong effects on physical and mental health (Finkelstein et al. 2012). Card et al. (2009) exploit age discontinuities in the eligibility for Medicare while Currie and Gruber (1996a, 1996b) analyze changes in eligibility rules for Medicaid. They find positive effects on the (health and) survival of elderly, newborns and children, respectively. My findings sug-

[5]This mechanism of catchup through increased mobility is in line with the model developed in Beaudry and DiNardo (1991).

gest that, during the outbreak of a deadly epidemic, health insurance also matters for the survival of young adults.

The effects of the business cycle on mortality has been investigated in existing papers. Ruhm (2000) shows in the US Vital Statistics that mortality is procyclical, i.e. negatively related to the contemporaneous unemployment rate. Miller at al. (2010) show that the procyclicality of mortality is driven by age groups that are not participating in the labor market: the young and in particular the very old. They argue that personal health care is cheaper in times of high unemployment, implying that the findings of Ruhm (2000) are in fact driven by a positive income effect. Van den Berg, Lindeboom and Portrait (2006) analyze the effect of the business cycle at the time of birth on mortality later in life. They find that those born in a recession live a few years less than those born in a boom time. While there might be potential selection into giving births, ie. that recession-mothers differ from boom-time-mothers, their results provide further evidence of the dramatic effects of the business cycle on people's lives.

3.2 The HIV/AIDS epidemic

For the interpretation of the effects on AIDS deaths found in this study, a brief overview of the HIV/AIDS epidemic is necessary. HIV is a virus that causes AIDS, a condition which progressively destroys the immune system making the body vulnerable to opportunistic infections and eventually leading to death. The main modes of exposure are male to male sex (46 % among persons reported with AIDS), followed by injection drug

use (25 %) and heterosexual contact (11 %) (CDC 2001). Without treatment, AIDS is developed about 10 years after infection with HIV and leads to death in about 9 months. HIV/AIDS treatments increase survival time by about 4 to 12 years but to date there is no known cure (Tassie 2002, King et al. 2003).

In Figures 3.11 to 3.14 I illustrate overall deaths and AIDS deaths for the analyzed cohorts by cohorts and years. These numbers come from the Vital Statistics. They are not population estimates but represent the universe of deaths in the population that I am analyzing.

The first AIDS deaths occurred in 1981 but only in 1986 the epidemic started to take off (Figure 3.11 [b]). In the late 1980s and early 1990s the number of AIDS deaths increased steadily until 1995 whereupon it dropped abruptly and stayed at a rather constant level after 1997. 1995 to 1997 was the time that more effective antiretroviral therapies became widely available (Chiasson et al. 1999). The peak around 1995 can also be seen in overall death (panel [a]). For the most affected cohorts every 4th to 5th death in the early 1990s was caused by AIDS (panel [c]).

A major legislative development that does not directly show up in the mortality data is the Ryan White Act that has been signed in August 1990. This program provides essential medications to those infected persons who are uninsured or are unable to afford the cost of HIV related medicines. It started with a budget of $257m in 1991 increasing up to $2.3b in 2011. As of 2011, the program still serves more than half a million low-income people who are infected with HIV/AIDS. Of those 33 % are uninsured and 56 % are underinsured (Johnson 2011).

Figures 3.12 (a) to (f) show these numbers separately for men and women. For the included birth cohorts (i.e. in the respective age range) men die in greater numbers in all of the years (panel [a] and [b]). Still the hump around 1995 is more pronounced for men. Panels (c) and (d) show impressively that AIDS has been - at least until the late 1990s - mainly a male epidemic. During the peak years the number of male AIDS deaths is about four times higher among males for the most affected cohorts. Relative to overall deaths, AIDS death rates for these cohorts are about twice as large for men than for women. After 1997, however, male and female AIDS death rates strongly converge and equalize for the youngest cohorts in the early 2000s.

Figures 3.13 (a) to (f) present these graphs split up by race, showing deaths for whites and non-whites. Overall deaths (panels [a] and [b]) are higher for whites due to the larger population share of this group. About 84 % of the sample is white (see Table 3.1). However, during the outbreak of the epidemic AIDS deaths of whites are just about one third higher for the most affected cohorts (panels [c] and [d]) and similarly high for younger cohorts. After 1997 there are more AIDS deaths among non-whites than among whites across all cohorts. Panels (e) and (f) show AIDS deaths relative to race-specific overall deaths. During the peak years 20 % of overall deaths are due to AIDS for whites in the most affected cohorts, while for non-whites in these cohorts almost every third death is caused by AIDS. After 1997 AIDS death fraction among white cohorts converge towards 2-3 % while non-white cohorts seem to converge to a higher level around 10 % .

There is an important distinction between the observed gender and race AIDS gap. While the gender gap is mostly driven by biological sex-specific modes of exposure the race gap reflects socio-economic differences (Chaisson 1995). In other words, the race gap in AIDS deaths implies that HIV/AIDS epidemic has dramatically increased socio-economic inequalities in mortality among young adults in the US.

In Figure 3.14 I illustrate this increase in the race mortality gap more directly. In this figure I plot the fraction of non-white deaths over time together with CPS estimates of the fraction of the non-white population.[6] In the absence of a race mortality gap these two graphs should lie on top of each other. Over the sample period the fraction of the non-white population increases smoothly from 14 % to 20 % due to younger cohorts with higher non-white shares that enter the sample over time. In 1979 the fraction of non-white deaths is 4 percentage points higher than the fraction of the non-white population. Thus, already before the outbreak of the AIDS epidemic non-whites died at a higher rate. In Figure 3.15 I plot the ratio of the non-white to white death rate that is implied by the numbers underlying Figure 3.14. The 4 percentage point gap in Figure 3.14 implies that the non-white death rate is about 30 % above the death rate of whites. In the following years up to 1985 the gap increases only slightly, from 30 % to 40 %. However, with the outbreak of the AIDS epidemic the gap increases dramatically, peaking in 1991. In this year about 16 % of the population is non-white but almost every third death comes from

[6]Notice that in the Vital Statistics data I only include decedents born in the US while the CPS is a sample of the current population. As the non-white population is more likely to be born outside the US, the CPS estimates are likely to overstate the fraction of native non-whites. Therefore Figure 3.14 is likely to understate the race gap.

this minority group (Figure 3.14). The non-white death rate is about 110 % above the death rate of whites (Figure 3.15). After 1991 the race gap decreases linearly, reaching the 1985 level in 2004.

This uncovers an interesting pattern that is not evident in the previous figures. Unlike overall AIDS deaths the race gap does not peak in 1995. And it does not discontinuously drop after 1995 when antiretroviral therapies became widely available. This suggests that the pattern around 1995 observed in the previous figures represent time effects that are homogeneous across race (and thus probably across socio-economic status). Instead the race gap peaks in 1991. This is the year when the White Ryan program started which was particularly directed to low-income groups and people with no or little health insurance. Figure 3.15 suggests that this program might have been more effective than overall AIDS deaths indicate.

The source of the observed dramatic differences in AIDS deaths across race remains an open question. Higher infection rates are likely to play an important role but even in samples of HIV infected subjects strong differences in survival rates by race and socio-economic status remain (Rothenberg et al. 1987, Chaisson 1991). Chaisson et al. (1995) find that these differences are not driven by biological factors. They show that among patients with HIV infections who received medical care from a single urban center disease progression or survival rates were not associated with race or socio-economic status (nor with sex or injection drug use). This points to a story in which access and use of medical care might play an important role.

3.3 Data

The data used in this paper come from two main sources: The US Vital Statistics (Vital Stats) which provides information on every single death in the US starting in 1968; and the March Supplement of the Current Population Survey (CPS), which provides a vast set of socio-economic characteristics for a repeated cross-section since 1962. About 55,000 individuals are surveyed per year in the 1960s, about 150,000 in the 1970s to 1990s and about 200,000 after 2000. In order to relate deaths and socio-economic outcomes from these two data sets to the graduation unemployment rate information on the respondent's (decedents) year and the state of graduation are required.

The death files from the Vital Stats include the date and the cause of death along with core demographic characteristics such as age, sex and race. Information on the decedent's education, which would allow to calculate the year of graduation, is included only after 1989. I use the year when the decedent was 18 as a proxy for the graduation year. From 1979 to 2004 the death certificates include the state of the decedent's birth which I use as proxy for the state of graduation. This restricts the sample to decedents born in the US.

The CPS reports the number of years of education until 1991. This allows to calculate the respondent's year of graduation as calendar year minus age plus 6 plus years of education. The state of current residence is included in all waves, which I use as a proxy for the state of graduation.

State-level unemployment rates are available from the Bureau of Labor

Statistics only since 1976. Therefore I have to exclude individuals who graduated before 1976 when using the actual graduation year and individuals who were born before 1958 (i.e. of age 18 before 1976) when using age 18 as proxy for graduation age. Further I restrict my main analysis to the years 1979 to 1991 when both state of birth is available in the Vital Stats and years of education is reported in the CPS. In this sample respondents or decedents have at most 15 years of potential experience (when using the actual graduation unemployment rate) or are at most of age 33 (when using the age 18 unemployment rate). The extended sample includes the years 1979 to 2004. Table 3.1 presents summary statistics for these different samples.

I supplement the Vital Stats with intercensus population estimates from the US Census to calculate death rates. Note that in my mortality regressions I exploit variation in deaths at the level of year of birth x state of birth x calendar year. The appropriate denominator to calculate death rates would therefore be population estimates at the same level of aggregation. Unfortunately, to my knowledge population estimates at the level of state of birth do not exist.

The US Census only reports state of current residence. Using the size of a cohort currently residing in a given state as a proxy for the size of the same cohort born in this state is likely to bias my estimates upwards. It seems plausible to expect that graduates facing a recession migrate to other states with more favorable labor market conditions. In this case the graduate cohort remaining in the recession state shrinks while the cohort in the more favorable state increases. Using cohort sizes by state of residence is therefore likely to lead to higher death rates for recession cohorts

and lower rates for boom cohorts. In the CPS analysis I am using state of residence and I find that migration does not seem to introduce a selection bias as migration is neither gender nor race specific. However, even non-selective migration from bust to boom states would unbalance population estimates and lead to the described bias in death rates.

I therefore use cohort sizes at age zero as a proxy for the initial population base. In the year of birth only a small fraction of a cohort is subject to migration and this migration is not a response to a recession that is 18 years ahead - as long as there are no cycles of 18 years. Obviously, using the initial cohort size to calculate death rates does not take into account the shrinking due to these deaths.[7] But death rates are still small in the analyzed age range and regression include age fixed effects that control for the natural shrinking rate. And any bias due to faster shrinking of more affected cohorts would attenuate effects in subsequent years.[8]

3.4 Empirical Strategy

I seek to analyze the relationship of economic conditions at the year of graduation with socio-economic outcomes and mortality. For a causal in-

[7] It seems straight-forward to adjust for cohort-specific deaths using the Vital Stats. But as noted above prior to 1979 death certificates do not include the state of birth so that any cohort born before 1979 would be subject to 'blind' years for which we do not know how many deaths occurred.

[8] If affected cohorts shrink faster then the population base that 'generates' deaths is smaller than for these cohorts than for less affected ones and as a consequence the same death rate implies a smaller number of absolute deaths. Dividing by the initial rather than an updated cohort size leads to artificially low death rates for affected cohorts.

terpretation we need the economic conditions to be randomly assigned.

The economic conditions at different stages in one's life are unpredictable and whether you graduate in a recession or a boom time can be considered to be basically luck. One might be worried though that graduating in one year or the other might be endogenous - we could think of a story those with higher potential earnings manage to continue schooling in a recession until the job market improved. But the existing studies shows that endogenous graduation timing is not a big issue (OWH, Genda et al. 2010, Kahn 2009). Also, in the main specifications I am using the year when a cohort is of age 18 as predicted graduation year which is independent of business cycle fluctuations. Further, some states may permanently have worse economic conditions than others and people in these states might be different a priori. And a recession might come along with other macro shocks that directly affect our outcomes of interest. But I control for state and cohort fixed effects to take care of such potential confounding factors. The key variables to identify the effect of graduating in a recession on subsequent outcomes are the year of graduation and the state of graduation.

In the CPS I proxy state of graduation by the state of current residence. This could be a problem as graduates that face high local unemployment might seek to migrate to states with more favorable labor market conditions. If those with higher potential income are better in escaping from local recessions by migrating to better states this would introduce an upward bias. We would only observe those members of a recession cohort in their state of graduation who did not make it to a better state and who would have had lower incomes in any case. OWH show that regional mo-

bility is small and short lived for Canadian graduates. I provide evidence that such selection seems not to be relevant in the CPS data either.

In the Vital Statistics I use the year when the decedent was of age 18 as a proxy for the graduation year. The unemployment rate at age 18 is obviously only a rough proxy for the economic conditions that a cohort faces over the years it is entering the labor market.[9] Alternatively I could use a weighted average of the unemployment rates over all graduation years of a cohort, with the cohort fractions graduating in each year as weights. Results tend to be similar for such a specification. But using the economic conditions over a time window of several graduation years limits the number of cohorts and calendar years that can be included.[10] An advantage of this simple proxy over the actual year of graduation is that it rules out potential selection due to endogenous graduation timing.

I use the state of birth (instead of the state of residence, which is also available) as a proxy for the state of graduation. This is a conservative choice. Endogenous migration of graduates from recession to boom states

[9]In the CPS data I show that the majority of the typical cohort in my sample graduates after 12 years of education. Given that children typically start school at age 6 the modal graduation year in my sample is at age 18. Also OWH find that those with lower (hypothetical) education suffer more from graduating in a recession. In order to capture as much of the recession effects as possible it is therefore useful to choose an age at which high school graduates rather than college graduates enter the job market.

[10]State-level unemployment rates are available only since 1976. Let's say we want to use the average unemployment rates for the years in which a cohort is of age 16-24 and only include a cohort once it turned 25. The oldest birth cohort for which state unemployment rates are available from age 16 onwards is born in 1960. And the first calendar year in which we could include this cohort is 1985. Using instead age 18 unemployment rates, the oldest included birth cohort is born 1958 and the first calendar year this cohort can be included is 1977.

would in this case attenuate estimates towards zero. The state of birth unemployment rate that I assign to a decedent would be above the unemployment rate that she actually faced when graduating. In other words, I would overstate the regressor of interest and therefore underestimate the corresponding coefficient.

My empirical strategy consists of three steps. First, I replicate OWH's specification in the CPS data, using the (reconstructed) actual year of graduation and state of residence as a proxy for state of graduation. Second, I show in the same sample that results largely carry over to a specification in which I use the unemployment rate at age 18 instead of the actual graduation year. Third, I turn to the Vital Statistics data to estimate the effect of the age 18 unemployment rate on mortality using state of birth as proxy for state of graduation

The key variable of interest, the graduation unemployment rate, varies only across states and cohorts. I follow OWH and collapse the individual-level data at the level of state, cohort and calendar year. In the CPS specification using the actual graduation year the data is collapsed by state-of-residence x graduation year x calendar year. Using instead the age 18 unemployment rate I collapse the CPS data by state-of-residence x year-of-birth x calendar year. The Vital Statistics data are collapsed by state-of-birth x year-of-birth x calendar year. All regressions are weighted by the corresponding cell sizes. Standard errors are clustered at the cohort x state level to account for cohort specific serial correlation.

As OWH I use the following specification with the actual graduation year:

$$\bar{y}_{g,s,t} = \alpha + \beta_e u^G_{g,s} + \gamma_e + \delta_g + \lambda_s + \theta_t + \rho u_{g,s,t} + \epsilon_{g,s,t} \tag{3.1}$$

The indices g, s, t and e refer to the graduation cohort, state, calendar year and years of potential experience (years since graduation). $\bar{y}$ are different socio-economic outcomes collapsed at the level of graduation year, state and calendar year. γ, δ, λ and θ are the coefficients on unrestricted experience, graduation cohort, state and calendar year fixed effects (i.e. dummy variables for each value of the respective variable). The coefficient vector of interest, β_e, contains the coefficients on the interaction of the unemployment rate at the year of graduation ($u^G_{g,s}$) with dummies for the individual years since graduation. This means the effects of the graduation unemployment rate is allowed to vary for every year following graduation.

The different fixed effects control for the typical experience profile, for nation-wide cohort effects, for state-specific time-constant effects as well as nation-wide contemporaneous shocks. Therefore, the coefficient vector β_e captures deviations from the typical experience profiles in the different outcomes that are uncorrelated contemporaneous nation-wide shocks and related to cohort-state specific variations in graduation unemployment rates. I include the current state unemployment rate, $u_{g,s,t}$, in all regressions to isolate the effects of the graduation unemployment rate from the subsequent evolution of the local labor market conditions (year fixed effects control for the current business cycle at the national level).

Using the unemployment rate at age 18 as a proxy for the graduation unemployment rate changes the regression equation slightly:

$$\bar{y}_{c,s,t} = \alpha + \beta_a u_{c,s}^{A18} + \gamma_a + \delta_c + \lambda_s + \theta_t + \rho u_{c,s,t} + \epsilon_{c,s,t} \tag{3.2}$$

With c indicating the year of birth (instead of the year of graduation) and a years of age. $u_{c,s}^{A18}$ refers to the unemployment rate in state s in the year when birth cohort c was of age 18. I include age dummies instead of dummies for potential experience ('e' in the previous equation). In the mortality regressions $\bar{y}_{c,s,t}$ refers to the sum of deaths or to death rates at the level of year of birth x state of birth x calendar year level.

3.5 Findings

3.5.1 The effect of the graduation unemployment rate on socio-economic outcomes

Given the large number of estimated coefficients of interest in each regression it is most convenient to present the results in a graphical way. Tables with the corresponding regression results are reported below for the core regressions.

Figures 3.1 - 3.9 illustrate the estimated effects of the unemployment rate around graduation on various socio-economic outcomes. For each outcome, three sets of specifications are estimated. The first set (a) uses the unemployment rate at the actual graduation year, which is calculated as

the year of birth plus 6 plus years of education (equation 1). The second set (b) takes the year when the individual was of age 18 as proxy for the graduation year (equation 2). The third set (c) uses the same specification as (b) but also includes the CPS waves 1992 to 2004 for which years of education are not reported. The effects of the graduation unemployment rate are interacted with dummies that indicate the number of years passed since the graduation date (in [a]) or the years of age (in [b] and [c]). This allows the effect of graduating in a recession to vary over time.

Figure 3.1 presents the effects of graduating in a recession on the natural logarithm of income. For the regression results underlying this figure see Table 3.2. Figure 3.1 (a) illustrates a strongly negative and persistent effect of the unemployment rate at the actual year of graduation on annual wages. A 1 percentage point higher unemployment rate decreases wages in the first year after graduation by about 3 %. This is a huge effect. A typical recession with an increase in unemployment by 5 percentage points would be associated with an income loss of about 15 %. This effect if strongly persistent over time. Even 15 years after leaving high school or college a recession graduate faces an annual income loss of about 10 %.

These effects are about twice as large as the estimates reported in OWH for Canadian college graduates. Importantly, however, I do not only include college graduates but also graduates from high school and even high school drop-outs. OWH show that those with lower income degrees suffer more from graduating in a recession. Their estimated effects for the bottom tercile are remarkably similar to my estimates, starting at -3 % the year after graduation and increasing to about -1.7 % after ten years.

Figure 3.1 (b) illustrates the estimated effects of the unemployment rate at the year when the individual was 18 on wages at age 19 to 34. Until age 24 the estimates are very similar to those in Figure 3.1 (a) when using the actual year of graduation. A cohort that faces a 1 percentage point higher unemployment rate at age 18 suffers an income decrease of about 4 % around age 20. After age 24, however, this negative effect of the age 18 unemployment rate fades out reaching zero at age 29, while this is not the case when using the actual graduation year.

This difference in estimated effects could point to a negative selection due to endogenous graduation timing that might explain part of the negative effect at higher years of potential experience in Figure 3.1 (a). But as shown below selection seems not to be a driving force. A more plausible explanation seems to be that at higher ages the fraction of "treated" people who actually graduated at age 18 decreases and so does the average effect of the age 18 unemployment rate.

Figure 3.1 (c) repeats this specification including the years 1992 to 2004 for which no information on the exact years of education is available. Confidence intervals are smaller due to larger sample sizes. Further the effect at age 20 is particularly strong while it is only slightly negative thereafter. It is difficult to explain why results are different from Figure 3.1 (b), in particular as no comparison with a specification using actual graduation year is possible. The main take away for the (c) specification is that results do not change much when including more waves. And since the mortality regressions are mostly focused on the period 1979 to 1991, differences of individual estimates in the extended sample do not need to be overinterpreted at this point.

Figures 3.2 (a)-(c) repeat these regressions for the natural logarithm of household income for the subsample of household heads and their spouses. Results are very similar to the estimates in Figures 3.1.

Figure 3.3 illustrates that there does not seem to be much of an effect of the graduation unemployment rate on subsequent employment status. This might seem surprising, but it should be kept in mind that I control for the current unemployment rate in order to isolate the effect of the unemployment rate at entry from the subsequent evolution of labor market conditions. Without this control there are - as one would expect - strongly negative effects on employment status.

Figure 3.4 presents effects on the probability to live in a household with household income below the official poverty threshold. Not surprisingly, these figures largely mirror the effects for household income illustrated in Figures 3.2. Notably, the effects in specifications (a) and (b) are even more similar than for household income.

Figure 3.5 (a) reveals that recession graduates are also less likely to be married. Facing a 5 % higher unemployment rate at graduation reduces the likelihood that you are married in the subsequent years by about 2 percentage point. When using the age 18 unemployment rate (Figure 3.5 [b]) the effect of marriage becomes slightly weaker and less precise, but the pattern is similar until age 25 where after the point-estimates are essentially zero.

Figures 3.6 and 3.7 illustrate the effects on health insurance coverage (for

corresponding regression results see Tables 3.3 and 3.4). Graduating in a recession with a 5 % higher unemployment rate decreases the probability to be covered by employer health insurance by 5 percentage point while the chance to have Medicaid increases by 2 percentage point. Given average coverage rates in the sample of 38 % for employer health insurance and 7 % for Medicaid, these effects on health insurance are large (13 % and 28 %, respectively). For employer health insurance the estimated effect shifts towards zero after 5 years or after age 23 but point estimates stay below zero for 14 years or until age 32. For Medicaid coverage point estimates go to zero after 8 years or age 27. In all cases, however, these estimates indicate that recession graduates suffer an extended period of worse health insurance coverage.

Finally the results in Figure 3.8 demonstrate that the race composition is similar in recession and non-recession cohorts. This suggests that selection due to migration or - in specification (a) - due to endogenous graduation timing is not an important issue. If those with better socio-economic background and therefore higher potential earnings would delay graduation or would migrate to states with lower unemployment we should see a negative relationship between the graduation unemployment rate and the probability of being white - since race strongly reflect socio-economic status. But the race composition is unrelated to the economic conditions both in the actual year of graduation as well as at age 18. When including CPS waves up to year 2004 there seems to be a positive effect after age 25. While this pattern seems difficult to explain it poses some doubt on whether one should rely on this identification strategy in the extended sample without being able to check results with the actual year of graduation.

There are three main results from these CPS figures. First, I replicate the findings from OWH in the CPS data. Effects on income tend to be stronger but the overall pattern is similar. Second, I show that there are also effects on marital status and strong effects on health insurance coverage. The latter finding is particularly relevant when looking at health outcomes such as mortality. Third, using the unemployment rate at age 18 instead of the actual year of graduation results in very similar effects. This allows to implement this identification strategy in data sets that contain age but no information on the years of education, as the Vital Statistics.

Before turning to the mortality data, however, it is useful to look whether different socio-economic groups are affected differently by graduating in a recession. Figures 3.9 to 3.16 present results by gender and by race (white vs. non-white). I only show results for the age 18 unemployment rate and the CPS waves 1979 to 1991 (specification [b]), as this is the main specification that I am using the mortality regressions. In some cases I omit confidence intervals for the non-white regression estimates to keep reasonable ranges for the commonly scaled figures. For regression tables corresponding to Figures 3.9, 3.14 and 3.15 see Tables 3.5 to 3.7.

Across gender estimated effects are quite homogeneous for most outcomes, with two exceptions. The negative effect of graduating in a recession on the probability of being married seems to be driven by males (Figures 3.12 [a] and [b]). This is plausible. For men the decision to get married is mostly about the cost of paying for a family and being less flexible to take jobs in another region. Both of these factors imply that

having a worse job due to graduating in a recession results in a lower probability to get married. For women, who are typically not the main earner in a family, marriages are likely to relax the constraints implied by having a worse job. The other exception is Medicaid coverage. As illustrated in Figures 3.15 (a) and (b), the positive effect on the probability to be covered by Medicaid is more prolonged for women. This also makes sense because Medicaid is particularly directed to women. This is surprising given that we might expect stronger effects on job quality (i.e. less income and worse health insurance coverage) to lead also to lower marriage rates.

Effects are less homogeneous across race. Confidence intervals are larger for the non-white regressions due to the smaller sample size, but estimates clearly suggest a stronger effect of the age 18 unemployment rate for this subgroup. The effects on individual income, household income, poverty as well as insurance coverage are about twice to three times as large as for whites. Only for the probability to be married effects do not seem to differ between these two groups.

This heterogeneity across race is in line with OWH’s finding that those with lower potential income are affected more strongly by graduating in a recession. A further explanation might be that among whites there are relatively more individuals with more than 12 years of schooling, who are therefore not "treated" by the age 18 unemployment rate. Finding that effects on income and health insurance coverage are similar across gender but stronger for non-whites than whites is useful when turning to health outcomes such as mortality. If effects on health run through these channels we should observe similar heterogeneities in health outcomes.

Figures 3.16 (a) - (d) test for an effect of the age 18 unemployment rate on cohorts' race composition in the gender regressions and on the gender composition in the race regressions. In neither of the four regressions significant effects are detectable providing further evidence that selection into graduation years and/or states of residence is not an important issue.

To sum up, the CPS data revealed strongly negative effects of graduating in a recession on income and health insurance which are similar across gender but stronger for non-whites. Marital status is negatively affected for men but not for women and this effect does not seem to differ between whites and non-whites.

3.5.2 The effect of the age 18 unemployment rate on mortality

In this section I explore to which extent graduating in a recession has an effect on the graduates' subsequent mortality. As the years of education are not reported in the death certificates in many years, I use throughout the unemployment rate at age 18 as a proxy for economic conditions at graduation for all mortality regressions (equation 2).

All regressions control for state of birth, cohort, calendar year and age fixed effects as well as for the current state unemployment rate. I restrict the main analysis to the years 1979 to 1991, the time period for which I could compare my identification strategy in the CPS to the specification using the actual year of graduation.

Figure 3.17 illustrates the effect of the age 18 unemployment rate on the number of deaths per cohort (defined by state of birth and year of birth). Until age 27 effects are virtually zero. Starting at age 28, 10 years after the age for which the unemployment rate has been taken, estimates become positive and increase linearly with every additional year of age. At age 33 the point estimate reaches 22 deaths. This implies that a 1 percentage point increase in the unemployment rate at age 18 leads to 22 additional deaths of 33 year-olds per state and birth cohort, or 1,122 additional deaths of 33 year-olds in the overall US in a given year.

How big is this effect in terms of the living population or in terms of the average death rate? And might the effect be driven by changes in cohort sizes? Notice that I am controlling for state and for year of birth fixed effects. This takes care of potential changes in cohort sizes at the national level and time-constant differences in state sizes. However, cohort sizes might not only change at the national but also at the state level. And a particularly large cohort in a given state might lead to a higher unemployment rate in this state when this cohort turns 18 and enters the labor market. At the same time a larger cohort implies a higher number of deaths and the increasing effect after age 27 might just reflect the natural increase in the death rate as cohorts become older.

Figure 3.18 shows on the left y-axis the effect on the death rate in terms of 10,000 cohort members.[11] This means that every observation for a given

[11] I take cohort sizes from the annual intercensus estimates. As these estimates are only available per state of residence I use cohort sizes by state of residence at age zero as a proxy for initial cohort sizes by state of birth. See the data section for a more detailed discussion.

state-of-birth x year-of-birth cohort is divided by the same constant.[12] In this specification a small hump shows up already at age 25, but the overall pattern is very similar to that in Figure 3.17. While effects are close to zero for the early 20s, estimates linearly increase after age 26. At age 33 the effect reaches 0.5, meaning that the effect of 22 deaths displayed in Figure 3.17 refers to 5 deaths per 100,000 cohort members.

On the right y-axis of Figure 3.18 the effect is displayed in terms of the natural logarithm of the death rate, approximating the percentage change in the death rate caused by a 1 percentage point increase in the unemployment rate at age 18. The effect at age 33 reaches 0.04, meaning that the 5 additional deaths per 100,000 cohort members refer to a 4 % increase in the cohort's death rate. In other words, a typical recession at age 18 - an increase of the unemployment rate by 5 percentage points - leads to a 20 % increase in death rates at age 33. This seems to be a strong effect.

How plausible is such a large mortality effect for an age range in which causes of death are not strongly driven by income or health insurance (among young adults accidents, suicides and violence are the most common causes of death in non-epidemic times). Could this relationship be particularly strong in the analyzed time period? To answer this question I extend the analyzed time period in the next two figures. Indeed, when including the years 1992 to 1995 the effects decrease strongly (Figure 3.19). The estimate at age 33 decreases from 22 to 15 additional deaths.

[12] Alternatively I could take the logarithm of the number of births and include state-of-birth x year-of-birth fixed effects. This does not allow to estimate the level of the effects but the effect changes over age resulting from such a specification are very similar to those presented in Figure 3.18.

When including the years 1992 to 2004 in Figure 3.20, the effect diminishes to 2.3 and is no longer distinguishable from zero. Furthermore, effects seem to be slightly negative until age 27. It is difficult to interpret these slightly negative effects for the extended time period for which I cannot analyze my identification strategy in the CPS. However, Figures 3.19 and 3.20 suggest that the strong positive age effect displayed in Figures 3.17 and 3.18 is driven by the years prior to 1992.

The rise of the AIDS epidemic falls into this time period and AIDS deaths would be a plausible driver underlying the estimated effects, killing mostly young adults in the age range that I am analyzing. Fortunately, the cause of deaths are reported in the Vital Statistics which allows to run separate regressions for AIDS and non-AIDS deaths.

Figures 3.21 (a) and (b) illustrate estimates from such separate regressions. As it can be seen the positive and increasing effect on mortality after age 27 is mostly driven by AIDS deaths. Estimates for non-AIDS deaths in Figure 3.21 (b) do also increase after age 28 but this increase is small compared to the effects in Figure 3.21 (a) and for most years not significant. A potential explanation for the similar, though much weaker effect pattern in Figure 3.21 (b) could be that not all AIDS deaths are diagnosed as such in the death certificate. Overall, the results in Figure 3.21 clearly suggest that the effects on mortality are driven by AIDS deaths. But why do estimated effects decrease when including the years 1992 to 1995 (Figure 3.19), a period during which overall AIDS deaths were still on the rise (see Figure 3.12)? A possible explanation could be the Ryan White Act HIV/AIDS program that started in 1991 and was directed in

particular to infected people with little or no health insurance coverage. As explained above this program might not have reverted overall trends in AIDS deaths as much as the introduction of more effective antiretroviral therapies in 1995. But it seems to have softened socio-economic differences in mortality rates already since 1992 (Figure 3.15).

In Figures 3.22 (a) to (d) I analyze the effects separately for men and women as well as for whites and non-whites. For such subgroup analysis effects in terms of absolute numbers of deaths would not be informative because these groups vary considerably in size. One reasonable way to rescale effects would be to construct group specific death rates. However, as mentioned above population estimates by state of birth x year of birth are difficult to obtain. Constructing even smaller cells by gender and race would make this task yet more difficult. A useful shortcut is to take non-AIDS deaths as a proxy for cohort size and use the ratio of AIDS deaths to non-AIDS deaths as dependent variable. To the extend that the age 18 unemployment rate also positively affects non-AIDS deaths (e.g. if not all AIDS deaths are categorized as such in the death certificates) estimated effects on these ratios are attenuated towards zero.

The comparison of the mortality effects across gender and race resembles astonishingly the heterogeneity found for income and health insurance effects in the CPS. The effects of the age 18 unemployment rate on AIDS death ratios is very similar for males and females (Figures 3.22 [a] and [b]). The race regressions, however, reveal much stronger effects for non-whites than for whites (Figures 3.22 [c] and [d]). Effects for non-whites are about twice as large as those for whites.

3.6 Interpretation

I find a strong positive effects of the age 18 unemployment rate on AIDS deaths at age 28 to 33 during the outbreak of the AIDS epidemic. Including years after 1991 in my analysis the estimated effects decrease. Further the effects are homogeneous across gender but stronger for non-whites.

How plausible are such effects? Given the rather short time period which seems to be driving these effects - the first decade of the AIDS epidemic - one might wonder whether my specification merely picks up a spurious correlation of the business cycle with the dramatic increases in AIDS deaths. However, notice that all regressions include cohort as well as calendar year fixed effects. This means that any variation at the national level - both at the time when the unemployment rate is measured as well as when deaths occur - is taken out.

Still, could there be a spurious or confounded relationship at the level of individual states? Possibly those states that experienced a local recession in the late 1970s just happened to be hit particularly strong by the AIDS epidemic a few years later? In this case we should expect my estimates to reflect the general gender and race gaps in the AIDS epidemic as described above. For non-whites who were in general hit much more strongly by the epidemic (Figure 3.13) I do find larger effects. However, estimated effects of the age 18 unemployment rate are very similar for male and females even though the gender gap in overall AIDS deaths is

even more pronounced than the race gap (Figures 3.12 and 3.13). Moreover, I find the same effect pattern across gender and race for income and health insurance in the CPS. In other words, the pattern of mortality effects is precisely what we should expect if these effects are driven by the same causes as the effects in the CPS but not if they are due to a spurious relationship of local recessions with the outbreak of the AIDS epidemic.

Hence, it seems plausible that the differences in AIDS deaths are actually caused by differences in the economic conditions when these cohorts entered the labor market. These effects could be driven by two mechanisms. Income losses and worse health insurance coverage among recession graduates may lead to a faster disease progression and hence a shortened survival time given an HIV infection. Alternatively, infection rates might be higher among recession graduates.

Recession graduates might be more exposed to HIV infections if they tend to be more engaged in unsafe sex and injection drug use (the main modes of exposure, see above). We could think of a story in which recession graduates have more sex partners or opportunity costs of drug use are lower due to unemployment. However, the CPS results for marital status show that - if anything - we should expect to see a difference in mortality effects across gender, as males seem to be less likely to get married when graduating in worse economic conditions (Figure 3.18). Across race, on the other hand, there is no differential effect on marital status but the mortality effects are stronger. To the extent that married individuals have less sex partners this evidence suggests that higher infection rates through increased sexual contacts is not the main driver behind the mortality effects.

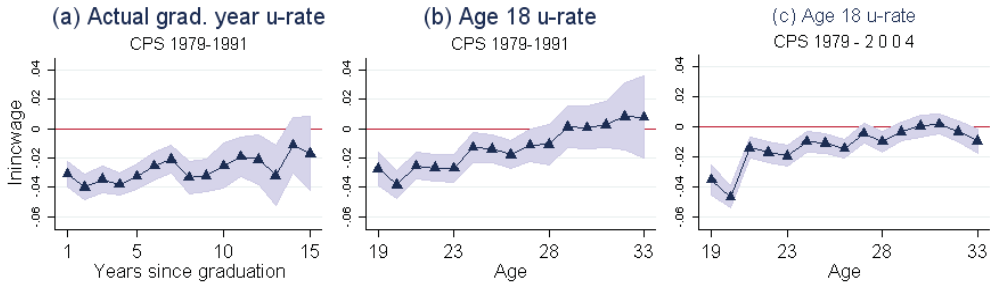
The opportunity cost story, on the other hand, seems not appealing given that there is not much of an effect on the probability of being unemployed. As mentioned in the introduction, the weak effects on employment status are in line with the literature. The mechanism through which worse economic conditions lower the socio-economic status of graduates is that they take lower quality jobs and not that they do not find a job at all. Further I find that recession cohorts have worse health insurance. This also implies that the direct costs of unhealthy behavior are larger for these cohorts.

Finally, the weakening of the mortality effects after 1991 coincides with the introduction of the Ryan White Act which provided medical treatment to those with insufficient or no health insurance coverage. If the estimated mortality effects were solely due to higher infection rates and not due to worse health insurance coverage, the Ryan White Act should have decreased AIDS mortality rates uniformly across cohorts and not have weakened my estimates during a period in which AIDS deaths were still on the rise.

However, these arguments are rather suggestive and not strong evidence against infection rates driving my results. Testing for this mechanism directly with data on HIV/AIDS prevalence rates will be the next step to gain further insights.

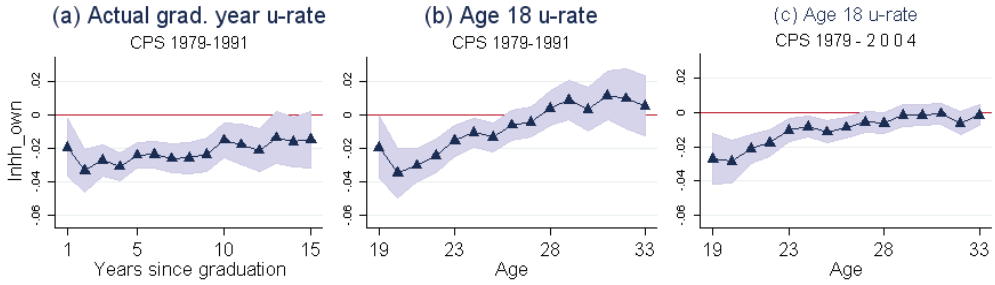
3.7 Tables and Figures

Figure 3.1: Effect of Unemployment Rate around Graduation on: Ln Wage



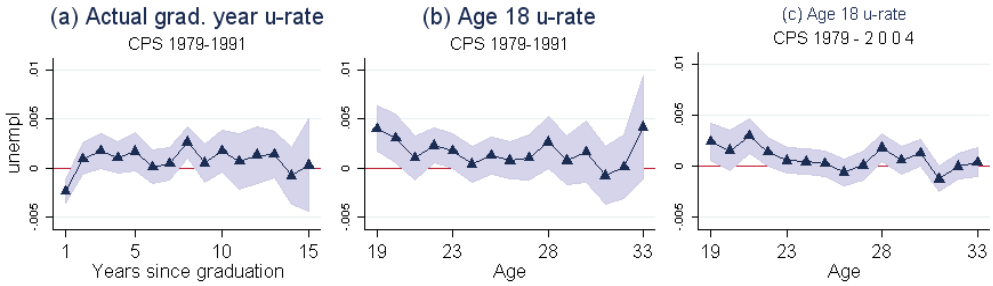
Notes: The dependent variable in all panels is the ln of individual wage. Panel (a) displays the coefficients of the unemployment rate at the year of graduation interacted with dummies for the years since graduation. Panels (b) and (c) displays the coefficients of the unemployment rate at age 18 interacted with dummies for age 19 to 33. Shaded areas indicate 95 % confidence intervals. Regressions control for cohort, state, calendar year, and years since graduation/age fixed effects and for the current unemployment rate. For corresponding regression results see Table 3.2.

Figure 3.2: Effect of Unemployment Rate around Graduation on: Ln Household Income



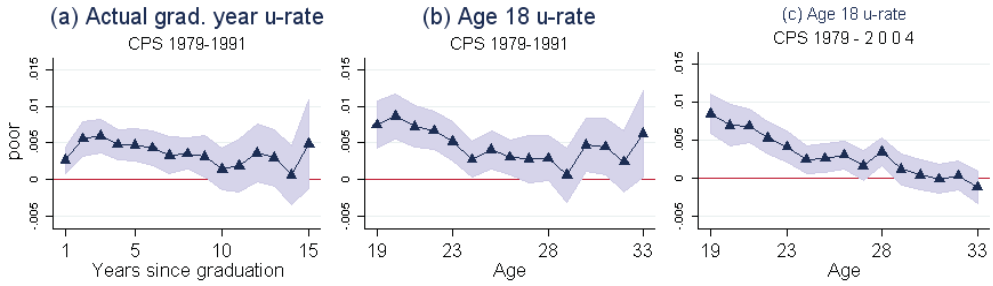
Notes: The dependent variable in all panels is the ln of household income. For further explanations see the notes below Figure 3.1.

Figure 3.3: Effect of Unemployment Rate around Graduation on: Unemployed



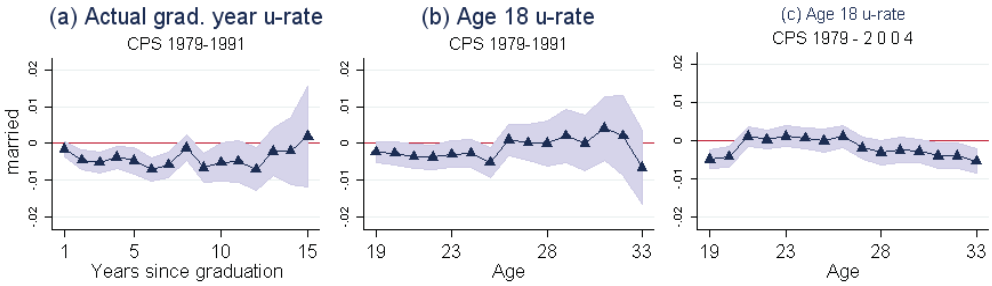
Notes: The dependent variable in all panels is a dummy variable indicating whether the respondent is unemployed. For further explanations see the notes below Figure 3.1.

Figure 3.4: Effect of Unemployment Rate around Graduation on: Below Poverty Threshold



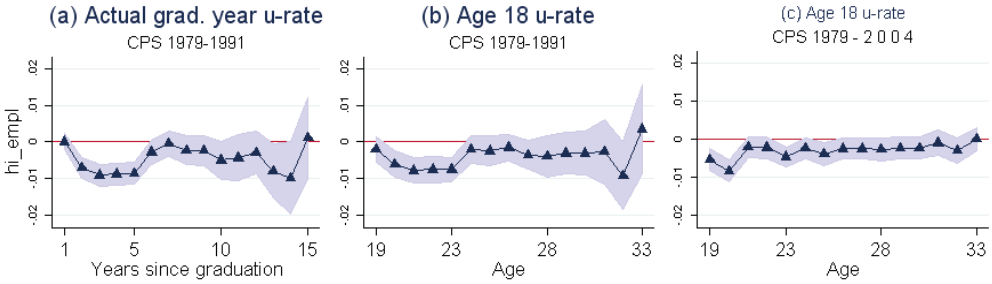
Notes: The dependent variable in all panels is a dummy variable indicating whether the respondent's family income is below the official poverty threshold. For further explanations see the notes below Figure 3.1.

Figure 3.5: Effect of Unemployment Rate around Graduation on: Married



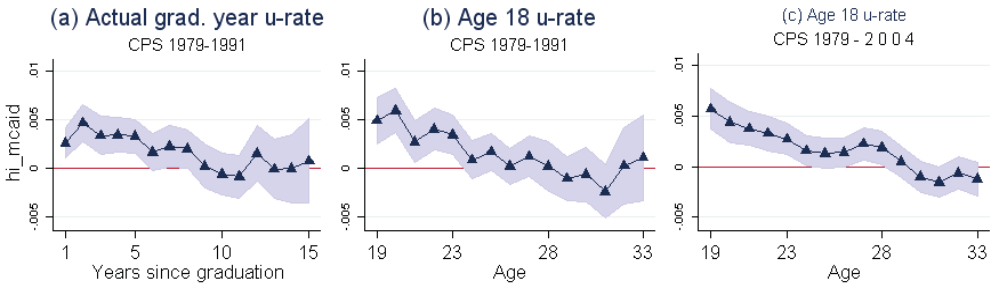
Notes: The dependent variable in all panels is a dummy variable indicating whether the respondent is married. For further explanations see the notes below Figure 3.1.

Figure 3.6: Effect of Unemployment Rate around Graduation on: Employer Health Insurance



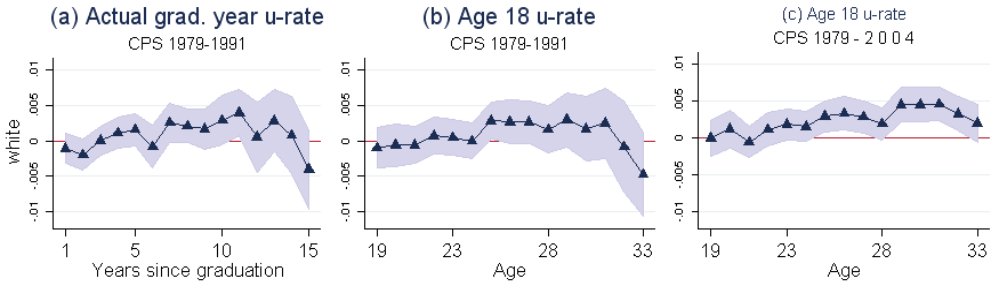
Notes: The dependent variable in all panels is a dummy variable indicating whether the respondent is covered by employer health insurance. For further explanations see the notes below Figure 3.1.For corresponding regression results see Table 3.3.

Figure 3.7: Effect of Unemployment Rate around Graduation on: Medicaid



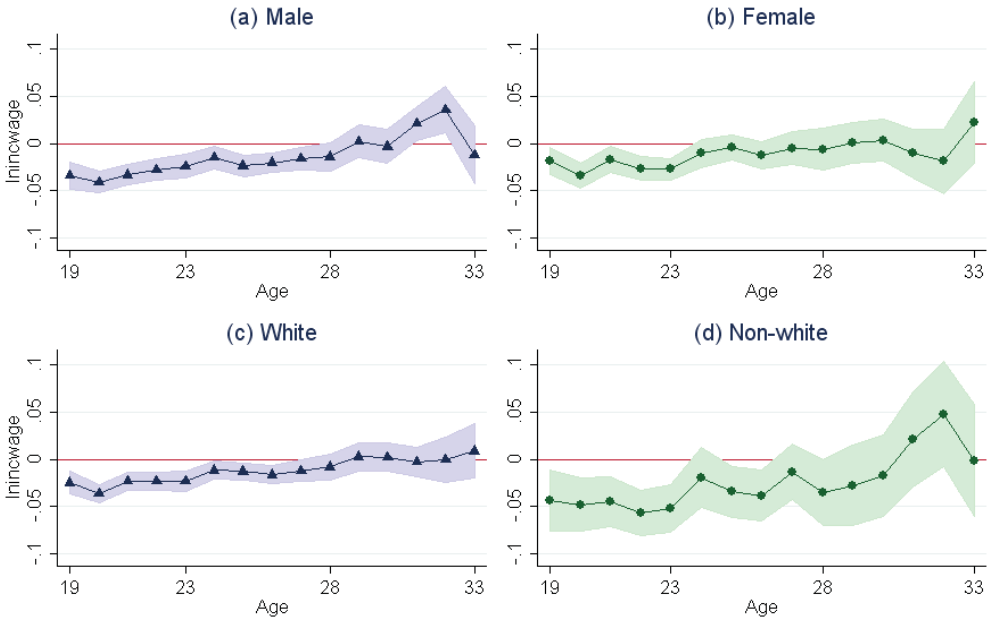
Notes: The dependent variable in all panels is a dummy variable indicating whether the respondent is covered by Medicaid. For further explanations see the notes below Figure 3.1. For corresponding regression results see Table 3.4.

Figure 3.8: Effect of Unemployment Rate around Graduation on: White



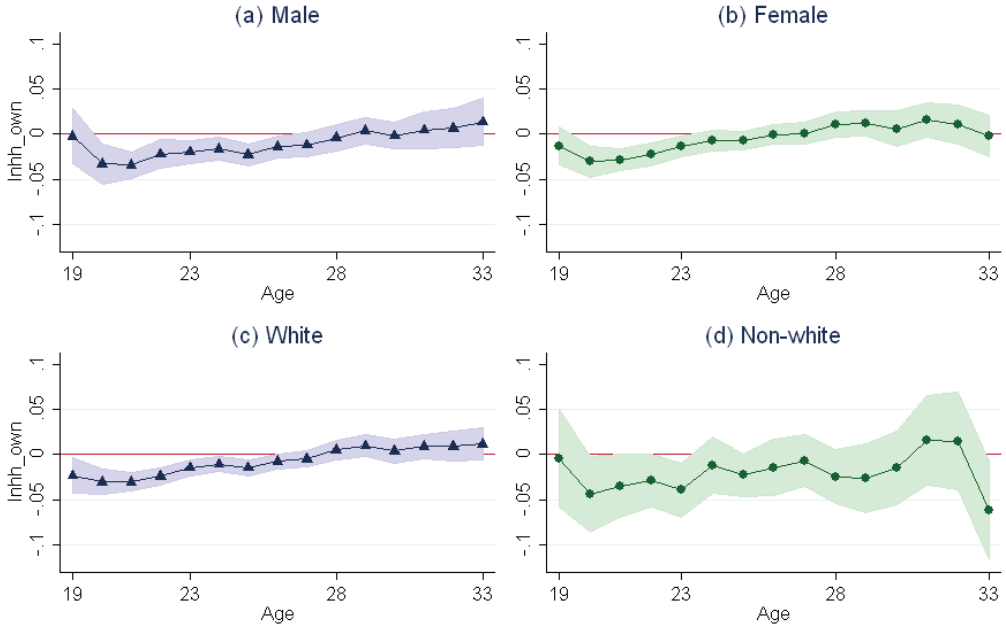
Notes: The dependent variable in all panels is a dummy variable indicating whether the respondent is white. For further explanations see the notes below Figure 3.1.

Figure 3.9: Effect of Age 18 Unemployment Rate on: Ln Wage



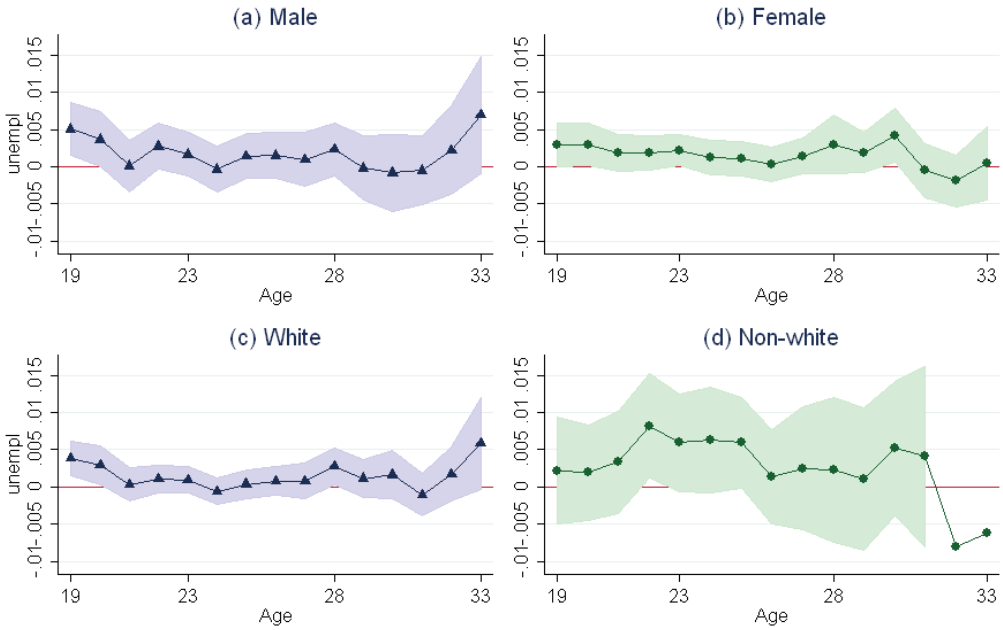
Notes: The dependent variable is the ln of individual wages. The coefficients of the unemployment rate at age 18 interacted with dummies for age 19 to 33 are displayed from separate regressions for male, female, white and non-white. Shaded areas indicate 95 % confidence intervals. Regressions control for cohort, state, calendar year, and years since graduation/age fixed effects and for the current unemployment rate. For corresponding regression results see Table 3.5.

Figure 3.10: Effect of Age 18 Unemployment Rate on: Ln Household Income



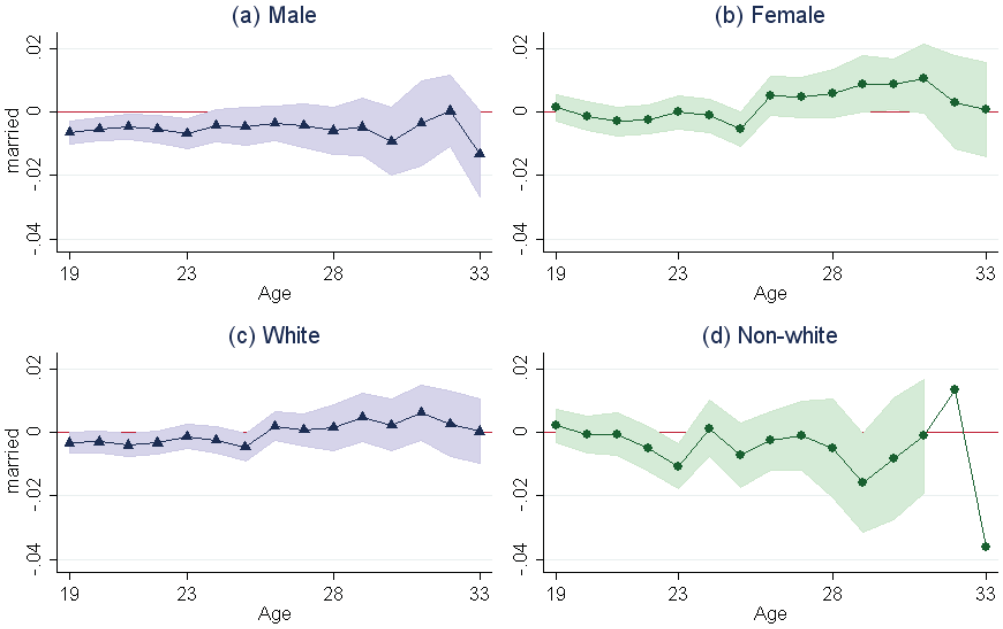
Notes: The dependent variable in all panels is the ln of household income. For further explanations see the notes below Figure 3.9.

Figure 3.11: Effect of Age 18 Unemployment Rate on: Unemployed



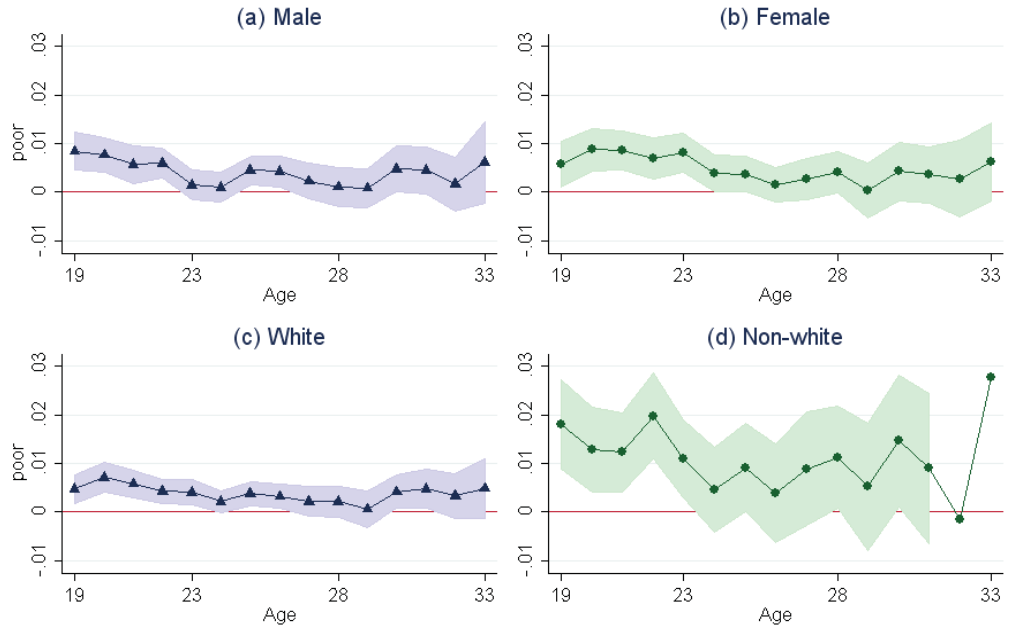
Notes: The dependent variable in all panels is a dummy variable indicating whether the respondent is unemployed. Confidence intervals for the last two coefficients in panel (d) are suppressed to keep scales in reasonable ranges. For further explanations see the notes below Figure 3.9.

Figure 3.12: Effect of Age 18 Unemployment Rate on: Below Poverty Threshold



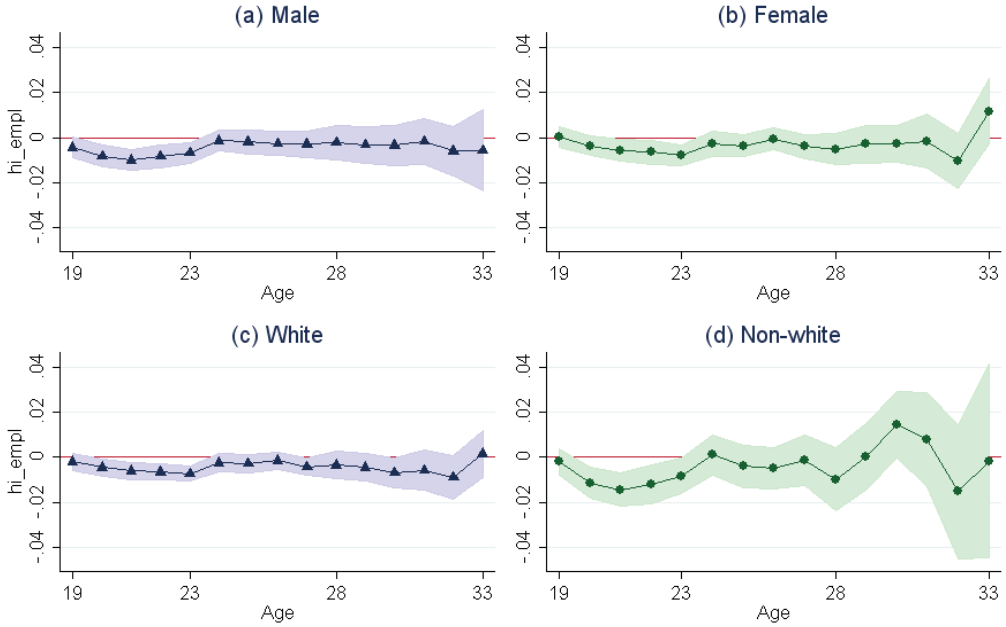
Notes: The dependent variable is a dummy variable indicating whether the respondent is married. Confidence intervals for the last two coefficients in panel (d) are suppressed to keep scales in reasonable ranges. For further explanations see the notes below Figure 3.9.

Figure 3.13: Effect of Age 18 Unemployment Rate on: Married



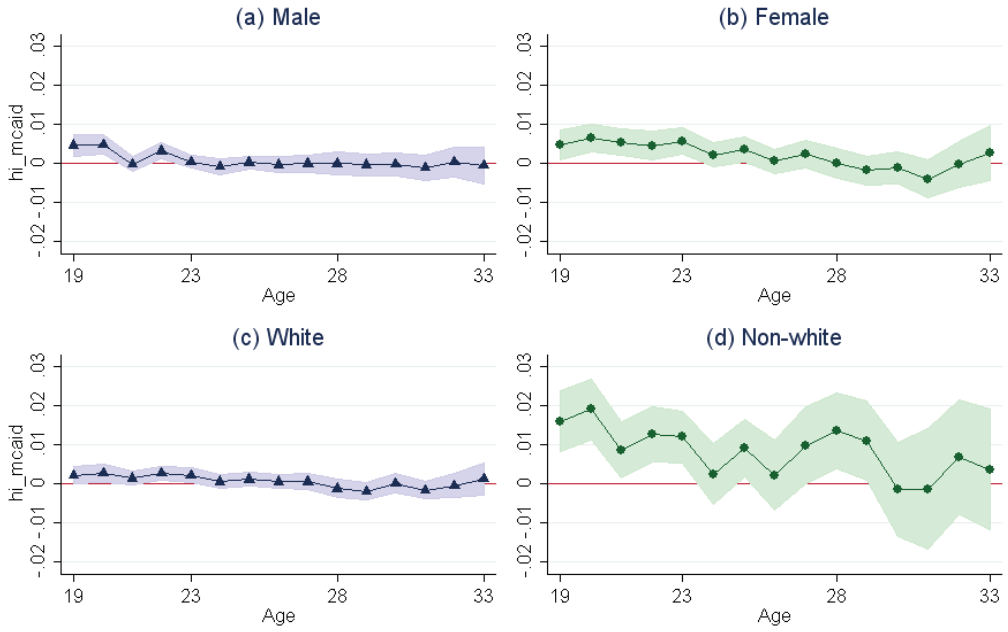
Notes: The dependent variable in all panels is a dummy variable indicating whether the respondent's family income is below the official poverty threshold. Confidence intervals for the last two coefficients in panel (d) are suppressed to keep scales in reasonable ranges. For further explanations see the notes below Figure 3.9.

Figure 3.14: Effect of Age 18 Unemployment Rate on: Employer Health Insurance



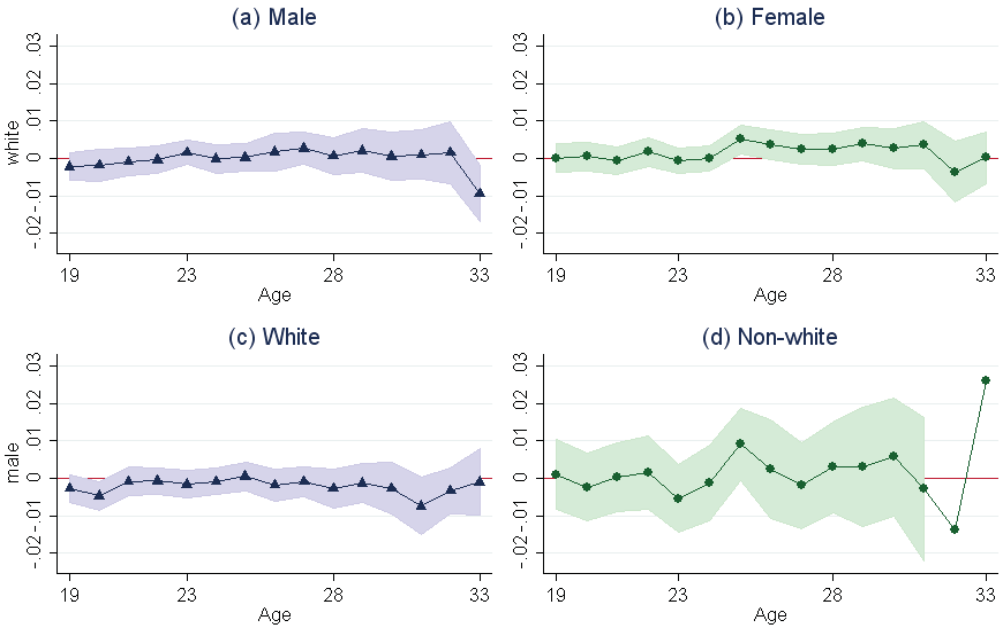
Notes: The dependent variable in all panels is a dummy variable indicating whether the respondent is covered by employer health insurance. For further explanations see the notes below Figure 3.9.For corresponding regression results see Table 3.6.

Figure 3.15: Effect of Age 18 Unemployment Rate on: Medicaid



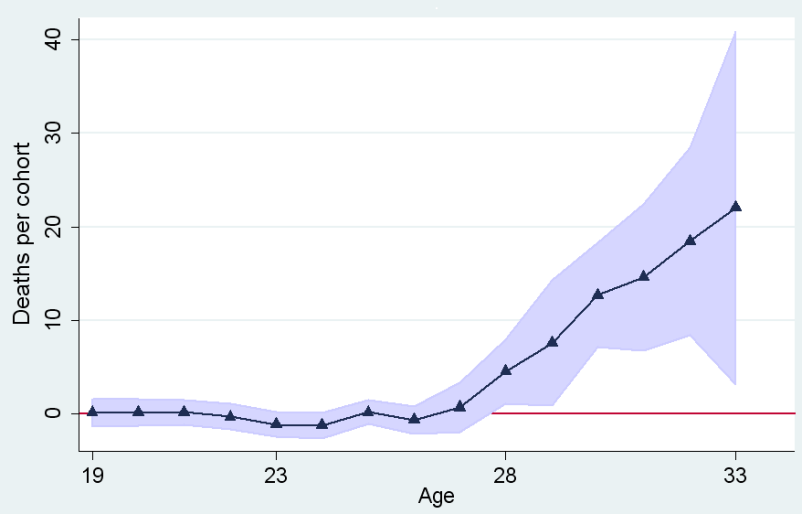
Notes: The Dependent variable in all panels is a dummy variable indicating whether the respondent is covered by Medicaid. For further explanations see the notes below Figure 3.9. For corresponding regression results see Table 3.7.

Figure 3.16: Effect of Age 18 Unemployment Rate on: White



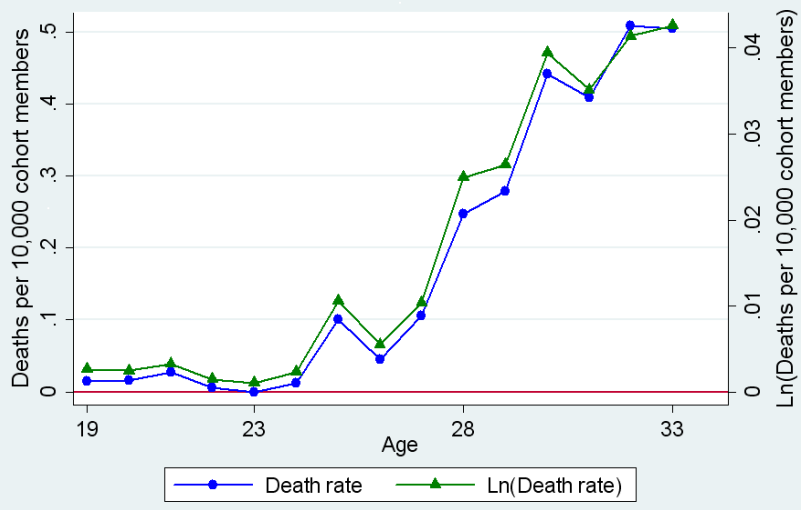
Notes: The dependent variable is a dummy variable indicating whether the respondent is male in panels (a) and (b) and white in (c) and (d). Confidence intervals for the last two coefficients in panel (d) are suppressed to keep scales in reasonable ranges. For further explanations see the notes below Figure 3.9.

Figure 3.17: Effect of Age 18 Unemployment Rate on Number of Deaths per Cohort, US Vital Statstics 1979 - 1991



Notes: Coefficients from the regression of the number of deaths per cohort on the unemployment rate at age 18 interacted with dummies for age 19 to 33 are displayed. Regressions include the current unemployment rate as well as fixed effects for year-of-birth, state, calendar year and age. Shaded areas indicate 95 % confidence intervals. For corresponding regression results see Table 3.8 column (1).

Figure 3.18: Effect of Age 18 Unemployment Rate on Death Rates, US Vital Statstics 1979 - 1991



Notes: The dependent variable is the number of deaths per 10,000 cohort members on the left and the ln of deaths per 10,000 cohort members on the right-hand side. For further comments see Figure 3.17.

Figure 3.19: Effect of Age 18 Unemployment Rate on Number of Deaths per Cohort, US Vital Statstics 1979 - 1995

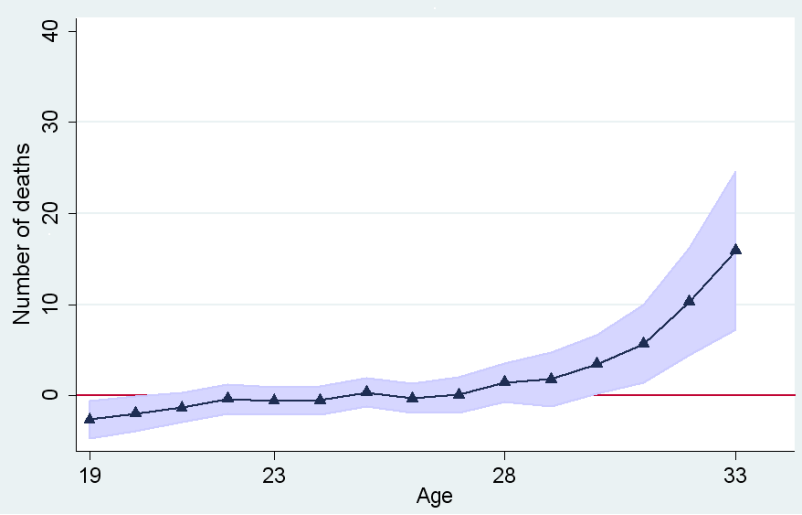
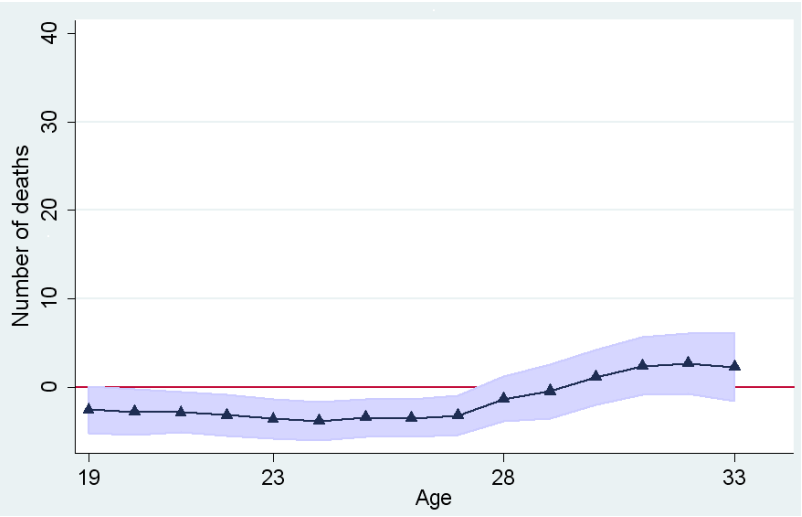
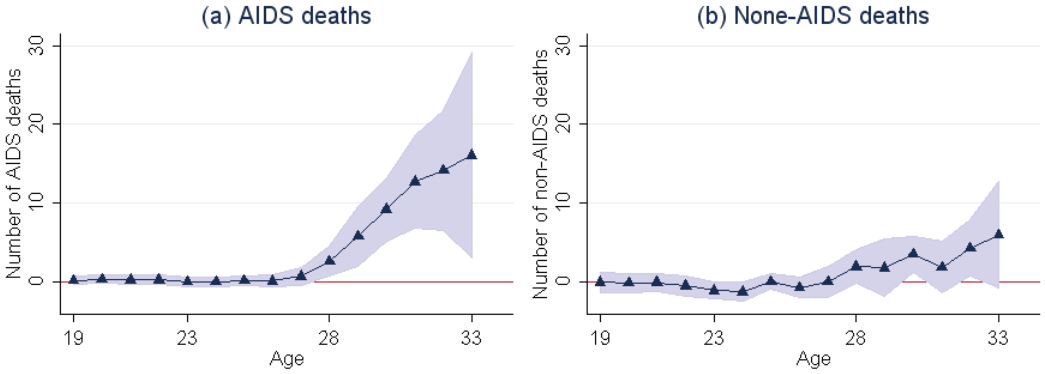


Figure 3.20: Effect of Age 18 Unemployment Rate on Number of Deaths per Cohort, US Vital Statstics 1979 - 2004



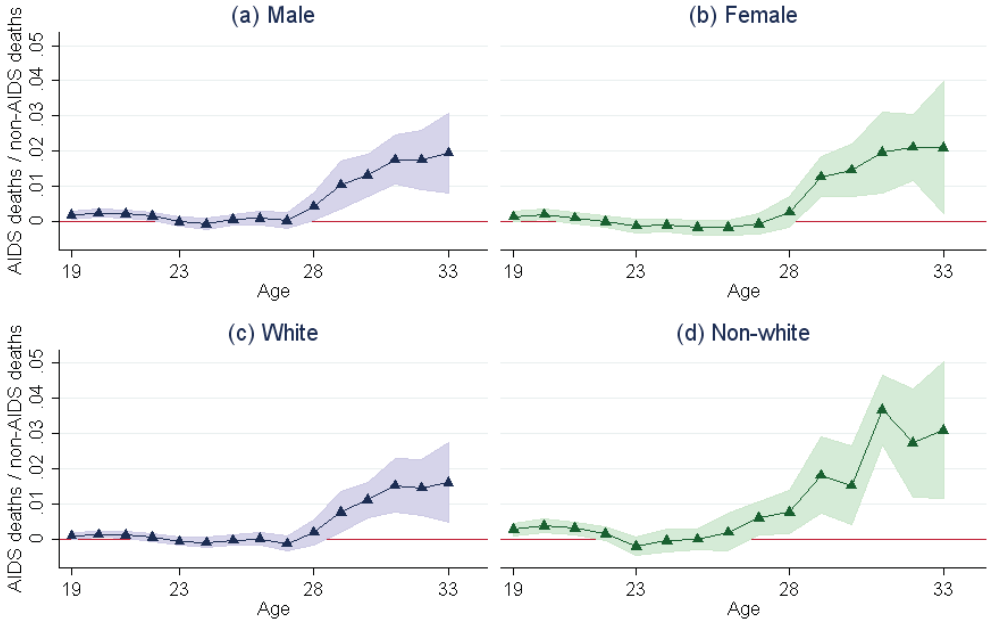
Notes: For further comments see Figure 3.19. For corresponding regression results see Table 3.8 col (3).

Figure 3.21: Effect of Age 18 Unemployment Rate on AIDS and non-AIDS Deaths, US Vital Statstics 1979 - 1991



Notes: Coefficients from separate regressions with AIDS deaths and Non-Aids Deaths as dependent variable are displayed. For further comments see Figure 3.19. For corresponding regression results see Table 3.9.

Figure 3.22: Effect of Age 18 Unemployment Rate on AIDS Death Ratios (AIDS / non-AIDS), US Vital Statstics 1979 - 1991



Notes: Coefficients from separate regressions for males, females, whites and non-whites with AIDS death ratios as dependent variable are displayed. For further comments see Figure 3.19. For corresponding regression results see Table 3.10.

Table 3.1: Descriptive Statistics

	Included years		
	1979-1991		1979-2004
	Model 1: Actual grad. year	Model 2: Age 18 grad year	Model 2: Age 18 grad year
Cohort restrictions	Graduation year>1975	Year when age 18>1975	Year when age 18>1975; age<34
	(1)	(2)	(3)
Current Population Survey			
Age (mean)	23.55	24.09	25.56
% male	50.07%	49.65%	50.04%
% white	83.98%	84.06%	82.16%
Years of education (mean)	12.56	12.98	13.08
Ln income (mean)	9.16	9.35	9.48
% unemployed	7.32%	7.85%	6.92%
% married	35.68%	39.03%	40.26%
% below poverty threshold	13.96%	13.85%	13.92%
% covered by employer health insurance	38.30%	42.36%	41.02%
% covered by Medicaid	7.31%	6.91%	8.14%
N	385,870	265,264	615,650
Vital Statistics - Universe of death			
Number of deaths	-	523,471	1,247,521
Male deaths	-	389,161	910,956
White deaths	-	395,413	903,284
AIDS deaths	-	26,614	85,873
Male AIDS deaths	-	22,381	67,886
White AIDS deaths	-	17,189	48,112
Annual intercensus estimates			
Initial population base	-	57,600,000	102,000,000
(size of analyzed cohorts at year of birth)			

Notes: Model (1) and (2) refer to regression equations (1) and (2), respectively. There are no death statistics for model (1) as the Vital Statistics do not contain the actual graduation year. Samples are restricted to graduation cohorts that graduated after 1975 / birth cohorts that were of age 18 after 1975, because state-level unemployment rates are available only since 1976. For further details see the Data section.

Table 3.2: Regressions of Ln Income on Graduation / Age 18 Unemployment Rate

	baseline: 1979-1991		extended: 1979-2004
Effect of Unemployment Rate Around	Grad. Year u-rate	Age 18 u-rate	Age 18 u-rate
Graduation on Ln Income	(1)	(2)	(3)
Effect at Year 1	-0.031	-0.027	-0.035
	(0.004)	(0.006)	(0.005)
Effect at Year 2	-0.04	-0.038	-0.047
	(0.005)	(0.005)	(0.004)
Effect at Year 3	-0.035	-0.025	-0.014
	(0.005)	(0.005)	(0.004)
Effect at Year 4	-0.038	-0.027	-0.017
	(0.004)	(0.005)	(0.004)
Effect at Year 5	-0.032	-0.027	-0.019
	(0.004)	(0.005)	(0.004)
Effect at Year 6	-0.025	-0.013	-0.01
	(0.005)	(0.005)	(0.004)
Effect at Year 7	-0.021	-0.014	-0.011
	(0.005)	(0.005)	(0.003)
Effect at Year 8	-0.033	-0.018	-0.014
	(0.006)	(0.005)	(0.003)
Effect at Year 9	-0.032	-0.011	-0.004
	(0.006)	(0.006)	(0.003)
Effect at Year 10	-0.025	-0.011	-0.01
	(0.008)	(0.007)	(0.004)
Effect at Year 11	-0.019	0.001	-0.003
	(0.007)	(0.007)	(0.003)
Effect at Year 12	-0.021	0.001	0.001
	(0.009)	(0.007)	(0.004)
Effect at Year 13	-0.032	0.003	0.002
	(0.010)	(0.008)	(0.003)
Effect at Year 14	-0.011	0.008	-0.003
	(0.010)	(0.012)	(0.003)
Effect at Year 15	-0.017	0.008	-0.009
	(0.013)	(0.014)	(0.004)
State FE, year FE	✓	✓	✓
Graduation cohort FE	✓		
Years since graduation FE	✓		
Birth cohort FE		✓	✓
Age FE		✓	✓
Number of observations (collapsed)	5,967	5,966	7,955
Underlying observations (individual level)	385,870	265,263	615,645
R^2	0.81	0.85	0.84

Notes: Column (1) displays the coefficients from the regression of ln income on the graduation year unemployment rate interacted with dummies for each year since graduation (regression equation [1]). Columns (2) and (3) display the coefficients on the age 18 unemployment rate interacted with dummies for age 19-23 (equation [2]). 'FE' refers to 'fixed effects', i.e. a set of dummies for the different values of the respective variable. For a graphical presentation see Figure 3.1.

Table 3.3: Regressions of Employer Health Insurance Coverage on Graduation / Age 18 Unemployment Rate

	baseline: 1979-1991		extended: 1979-2004
Effect of Unemployment Rate Around	Grad. Year u-rate	Age 18 u-rate	Age 18 u-rate
Graduation on Employer Health Insurance	(1)	(2)	(3)
Effect at Year 1	0.000	-0.002	-0.005
	(0.001)	(0.002)	(0.001)
Effect at Year 2	-0.007	-0.006	-0.008
	(0.001)	(0.002)	(0.001)
Effect at Year 3	-0.009	-0.008	-0.002
	(0.002)	(0.002)	(0.001)
Effect at Year 4	-0.009	-0.008	-0.002
	(0.001)	(0.002)	(0.001)
Effect at Year 5	-0.009	-0.008	-0.005
	(0.002)	(0.002)	(0.001)
Effect at Year 6	-0.003	-0.002	-0.002
	(0.002)	(0.002)	(0.001)
Effect at Year 7	-0.001	-0.003	-0.004
	(0.002)	(0.002)	(0.001)
Effect at Year 8	-0.002	-0.002	-0.002
	(0.002)	(0.002)	(0.001)
Effect at Year 9	-0.002	-0.004	-0.002
	(0.002)	(0.002)	(0.001)
Effect at Year 10	-0.005	-0.004	-0.003
	(0.003)	(0.003)	(0.002)
Effect at Year 11	-0.004	-0.003	-0.002
	(0.003)	(0.003)	(0.002)
Effect at Year 12	-0.003	-0.003	-0.002
	(0.003)	(0.003)	(0.001)
Effect at Year 13	-0.008	-0.003	-0.001
	(0.004)	(0.005)	(0.002)
Effect at Year 14	-0.010	-0.009	-0.003
	(0.005)	(0.005)	(0.002)
Effect at Year 15	0.001	0.003	0.000
	(0.006)	(0.006)	(0.002)
State FE, year FE	✓	✓	✓
Graduation cohort FE	✓		
Years since graduation FE	✓		
Birth cohort FE		✓	✓
Age FE		✓	✓
Number of observations (collapsed)	5,814	5,814	7,880
Underlying observations (individual level)	378,611	262,328	612,714
R^2	0.70	0.75	0.71

Notes: The dependent variable is the probability to be covered by employer-provided health care. Further comments as in Table 3.2. For a graphical presentation see Figure 3.6.

Table 3.4: Regressions of Medicaid Coverage on Graduation / Age 18 Unemployment Rate

	baseline: 1979-1991		extended: 1979-2004
Effect of Unemployment Rate Around	Grad. Year u-rate	Age 18 u-rate	Age 18 u-rate
Graduation on Medicaid Coverage	(1)	(2)	(3)
Effect at Year 1	0.003	0.005	0.006
	(0.001)	(0.001)	(0.001)
Effect at Year 2	0.005	0.006	0.004
	(0.001)	(0.001)	(0.001)
Effect at Year 3	0.003	0.003	0.004
	(0.001)	(0.001)	(0.001)
Effect at Year 4	0.003	0.004	0.003
	(0.001)	(0.001)	(0.001)
Effect at Year 5	0.003	0.003	0.003
	(0.001)	(0.001)	(0.001)
Effect at Year 6	0.002	0.001	0.002
	(0.001)	(0.001)	(0.001)
Effect at Year 7	0.002	0.002	0.001
	(0.001)	(0.001)	(0.001)
Effect at Year 8	0.002	0.000	0.001
	(0.001)	(0.001)	(0.001)
Effect at Year 9	0.000	0.001	0.002
	(0.001)	(0.001)	(0.001)
Effect at Year 10	-0.001	0.000	0.002
	(0.001)	(0.001)	(0.001)
Effect at Year 11	-0.001	-0.001	0.000
	(0.001)	(0.001)	(0.001)
Effect at Year 12	0.002	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
Effect at Year 13	0.000	-0.002	-0.002
	(0.002)	(0.001)	(0.001)
Effect at Year 14	0.000	0.000	-0.001
	(0.002)	(0.002)	(0.001)
Effect at Year 15	0.001	0.001	-0.001
	(0.002)	(0.002)	(0.001)
State FE, year FE	✓	✓	✓
Graduation cohort FE	✓		
Years since graduation FE	✓		
Birth cohort FE		✓	✓
Age FE		✓	✓
Number of observations (collapsed)	5,814	5,814	7,880
Underlying observations (individual level)	378,611	262,328	612,714
R^2	0.19	0.34	0.32

Notes: The dependent variable is the probability to be covered by Medicaid. Further comments as in Table 3.2. For a graphical presentation see Figure 3.7.

Table 3.5: Regressions of Ln Income on Age 18 Unemployment Rate by Gender and Race

	Sample			
Effect of Age 18 Unemployment Rate on Ln Income, 1979-1991	Male (1)	Female (2)	White (3)	Non-white (3)
Effect at Age 19	-0.034	-0.018	-0.025	-0.044
	(0.007)	(0.007)	(0.006)	(0.016)
Effect at Age 20	-0.041	-0.034	-0.036	-0.048
	(0.006)	(0.007)	(0.005)	(0.015)
Effect at Age 21	-0.033	-0.017	-0.023	-0.045
	(0.006)	(0.007)	(0.005)	(0.014)
Effect at Age 22	-0.028	-0.026	-0.023	-0.057
	(0.006)	(0.006)	(0.005)	(0.012)
Effect at Age 23	-0.024	-0.027	-0.023	-0.052
	(0.006)	(0.006)	(0.006)	(0.013)
Effect at Age 24	-0.015	-0.010	-0.011	-0.019
	(0.006)	(0.008)	(0.005)	(0.016)
Effect at Age 25	-0.024	-0.004	-0.013	-0.035
	(0.006)	(0.007)	(0.005)	(0.014)
Effect at Age 26	-0.020	-0.012	-0.016	-0.038
	(0.005)	(0.007)	(0.005)	(0.014)
Effect at Age 27	-0.016	-0.005	-0.012	-0.013
	(0.006)	(0.009)	(0.006)	(0.015)
Effect at Age 28	-0.014	-0.006	-0.008	-0.036
	(0.008)	(0.011)	(0.007)	(0.018)
Effect at Age 29	0.002	0.001	0.003	-0.028
	(0.009)	(0.011)	(0.008)	(0.022)
Effect at Age 30	-0.003	0.003	0.002	-0.017
	(0.009)	(0.011)	(0.007)	(0.022)
Effect at Age 31	0.022	-0.011	-0.003	0.021
	(0.009)	(0.013)	(0.008)	(0.026)
Effect at Age 32	0.036	-0.019	0.000	0.048
	(0.012)	(0.018)	(0.012)	(0.029)
Effect at Age 33	-0.012	0.022	0.009	-0.001
	(0.016)	(0.022)	(0.015)	(0.030)
State FE, year FE	✓	✓	✓	✓
Birth cohort FE	✓	✓	✓	✓
Age FE	✓	✓	✓	✓
Number of observations (collapsed)	5,962	5,962	5,961	4,814
Underlying observations (individual level)	128,733	136,521	227,089	37,767
R^2	0.80	0.69	0.84	0.55

Notes:The coefficients from separate regressions for males, females, whites and non-whites of ln income on the age 18 unemployment rate interacted with dummies for age 19-23 are displayed. 'FE' refers to 'fixed effects', i.e. a set of dummies for the different values of the respective variable. For a graphical presentation see Figure 3.9.

Table 3.6: Regressions of Employer Health Insurance Coverage on Age 18 Unemployment Rate by Gender and Race

	Sample			
Effect of Age 18 Unemployment Rate on	Male	Female	White	Non-white
Employer Health Insurance, 1979-1991	(1)	(2)	(3)	(3)
Effect at Age 19	-0.004	0.000	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.003)
Effect at Age 20	-0.008	-0.004	-0.005	-0.011
	(0.002)	(0.002)	(0.002)	(0.003)
Effect at Age 21	-0.010	-0.006	-0.006	-0.014
	(0.002)	(0.002)	(0.002)	(0.004)
Effect at Age 22	-0.008	-0.007	-0.006	-0.012
	(0.003)	(0.003)	(0.002)	(0.004)
Effect at Age 23	-0.007	-0.008	-0.007	-0.008
	(0.002)	(0.002)	(0.002)	(0.004)
Effect at Age 24	-0.001	-0.003	-0.002	0.001
	(0.002)	(0.003)	(0.002)	(0.005)
Effect at Age 25	-0.002	-0.004	-0.003	-0.004
	(0.003)	(0.003)	(0.002)	(0.005)
Effect at Age 26	-0.003	0.000	-0.001	-0.005
	(0.003)	(0.002)	(0.002)	(0.005)
Effect at Age 27	-0.003	-0.004	-0.004	-0.001
	(0.003)	(0.003)	(0.002)	(0.006)
Effect at Age 28	-0.002	-0.005	-0.004	-0.010
	(0.004)	(0.004)	(0.003)	(0.007)
Effect at Age 29	-0.003	-0.003	-0.004	0.000
	(0.004)	(0.004)	(0.003)	(0.008)
Effect at Age 30	-0.003	-0.003	-0.007	0.015
	(0.005)	(0.004)	(0.003)	(0.008)
Effect at Age 31	-0.002	-0.002	-0.006	0.008
	(0.005)	(0.006)	(0.005)	(0.011)
Effect at Age 32	-0.006	-0.011	-0.009	-0.015
	(0.006)	(0.006)	(0.005)	(0.015)
Effect at Age 33	-0.006	0.012	0.002	-0.002
	(0.009)	(0.008)	(0.005)	(0.022)
State FE, year FE	✓	✓	✓	✓
Birth cohort FE	✓	✓	✓	✓
Age FE	✓	✓	✓	✓
Number of observations (collapsed)	5,813	5,813	5,812	5,006
Underlying observations (individual level)	127,367	134,961	224,561	37,767
R^2	0.68	0.54	0.72	0.45

Notes: The dependent variable is the probability to be covered by employer-provided health care. Further comments as in Table 5. For a graphical presentation see Figure 3.14.

Table 3.7: Regressions of Medicaid Coverage on Age 18 Unemployment Rate by Gender and Race

	Sample			
Effect of Age 18 Unemployment Rate on Medicaid Coverage, 1979-1991	Male (1)	Female (2)	White (3)	Non-white (3)
Effect at Age 19	0.005	0.005	0.002	0.016
	(0.001)	(0.002)	(0.001)	(0.004)
Effect at Age 20	0.005	0.006	0.003	0.019
	(0.001)	(0.002)	(0.001)	(0.004)
Effect at Age 21	0.000	0.005	0.001	0.009
	(0.001)	(0.002)	(0.001)	(0.004)
Effect at Age 22	0.003	0.005	0.003	0.013
	(0.001)	(0.002)	(0.001)	(0.004)
Effect at Age 23	0.000	0.006	0.002	0.012
	(0.001)	(0.002)	(0.001)	(0.003)
Effect at Age 24	-0.001	0.002	0.001	0.002
	(0.001)	(0.002)	(0.001)	(0.004)
Effect at Age 25	0.000	0.004	0.001	0.009
	(0.001)	(0.002)	(0.001)	(0.004)
Effect at Age 26	0.000	0.000	0.001	0.002
	(0.001)	(0.002)	(0.001)	(0.005)
Effect at Age 27	0.000	0.002	0.001	0.010
	(0.001)	(0.002)	(0.001)	(0.005)
Effect at Age 28	0.000	0.000	-0.001	0.014
	(0.001)	(0.002)	(0.001)	(0.005)
Effect at Age 29	0.000	-0.002	-0.002	0.011
	(0.001)	(0.002)	(0.001)	(0.005)
Effect at Age 30	0.000	-0.001	0.000	-0.001
	(0.001)	(0.002)	(0.001)	(0.006)
Effect at Age 31	-0.001	-0.004	-0.002	-0.001
	(0.002)	(0.002)	(0.001)	(0.008)
Effect at Age 32	0.000	0.000	0.000	0.007
	(0.002)	(0.003)	(0.002)	(0.008)
Effect at Age 33	-0.001	0.003	0.001	0.004
	(0.002)	(0.004)	(0.002)	(0.008)
State FE, year FE	✓	✓	✓	✓
Birth cohort FE	✓	✓	✓	✓
Age FE	✓	✓	✓	✓
Number of observations (collapsed)	5,813	5,813	5,812	5,006
Underlying observations (individual level)	127,367	134,961	224,561	37,767
R^2	0.26	0.25	0.31	0.22

Notes: The dependent variable is the probability to be covered by Medicaid. Further comments as in Table 3.5. For a graphical presentation see Figure 3.15.

Table 3.8: Regressions of Deaths on Age 18 Unemployment Rate

	Included Years		
Effect of Age 18 Unemployment Rate on	1979-1991	1979-1995	1979-2004
Number of Deaths per Cohort	(1)	(2)	(3)
Effect at Age 19	0.137	-2.596	-2.507
	(0.757)	(1.059)	(1.322)
Effect at Age 20	0.176	-1.966	-2.744
	(0.710)	(0.993)	(1.322)
Effect at Age 21	0.193	-1.304	-2.810
	(0.663)	(0.816)	(1.165)
Effect at Age 22	-0.275	-0.337	-3.115
	(0.689)	(0.817)	(1.201)
Effect at Age 23	-1.084	-0.558	-3.536
	(0.657)	(0.777)	(1.146)
Effect at Age 24	-1.222	-0.493	-3.784
	(0.698)	(0.790)	(1.111)
Effect at Age 25	0.211	0.367	-3.397
	(0.653)	(0.793)	(1.083)
Effect at Age 26	-0.632	-0.289	-3.458
	(0.757)	(0.823)	(1.077)
Effect at Age 27	0.707	0.111	-3.131
	(1.340)	(0.987)	(1.141)
Effect at Age 28	4.567	1.478	-1.320
	(1.737)	(1.080)	(1.291)
Effect at Age 29	7.606	1.817	-0.409
	(3.427)	(1.520)	(1.554)
Effect at Age 30	12.714	3.465	1.179
	(2.840)	(1.626)	(1.582)
Effect at Age 31	14.603	5.687	2.413
	(3.980)	(2.169)	(1.648)
Effect at Age 32	18.458	10.316	2.703
	(5.094)	(2.991)	(1.723)
Effect at Age 33	22.040	15.923	2.303
	(9.624)	(4.429)	(1.939)
State FE, year FE	✓	✓	✓
Birth cohort FE	✓	✓	✓
Age FE	✓	✓	✓
Number of observations (collapsed)	5,928	8,987	15,872
Underlying population	468,440,679	687,489,342	1,146,700,000
R^2	0.97	0.95	0.92

Notes: Coefficients from regressions of the number of deaths per cohorton on the age 18 unemployment rate interacted with dummies for age 19-23 are displayed. Columns differ by the range of included years. Underlying population refers to the size of the included cohorts multiplied by the year in which they are analyzed. 'FE' refers to 'fixed effects', i.e. a set of dummies for the different values of the respective variable. For a graphical presentation see Figures 17, 19 and 20.

Table 3.9: Regressions of AIDS and Non-AIDS Deaths on Age 18 Unemployment Rate

Effect of Age 18 Unemployment Rate on AIDS and non-AIDS Deaths, 1979-1991	AIDS Deaths (1)	Non-AIDS Deaths (2)
Effect at Age 19	0.208	-0.071
	(0.276)	(0.659)
Effect at Age 20	0.323	-0.147
	(0.281)	(0.638)
Effect at Age 21	0.283	-0.090
	(0.295)	(0.606)
Effect at Age 22	0.238	-0.513
	(0.301)	(0.652)
Effect at Age 23	0.014	-1.098
	(0.309)	(0.550)
Effect at Age 24	0.016	-1.238
	(0.317)	(0.642)
Effect at Age 25	0.159	0.052
	(0.335)	(0.519)
Effect at Age 26	0.074	-0.706
	(0.388)	(0.668)
Effect at Age 27	0.687	0.020
	(0.557)	(1.000)
Effect at Age 28	2.617	1.951
	(0.989)	(1.081)
Effect at Age 29	5.800	1.806
	(1.932)	(1.893)
Effect at Age 30	9.217	3.497
	(2.046)	(1.194)
Effect at Age 31	12.745	1.858
	(3.040)	(1.650)
Effect at Age 32	14.164	4.293
	(3.868)	(1.841)
Effect at Age 33	16.094	5.946
	(6.677)	(3.455)
State FE, year FE	✓	✓
Birth cohort FE	✓	✓
Age FE	✓	✓
Number of observations (collapsed)	5,928	5,928
Underlying population	468,440,679	468,440,679
R^2	0.66	0.98

Notes: Coefficients from separate regressions for AIDS deaths and Non-AIDS deaths are displayed. Further comments as in Table 3.8. For a graphical presentation see Figures 21.

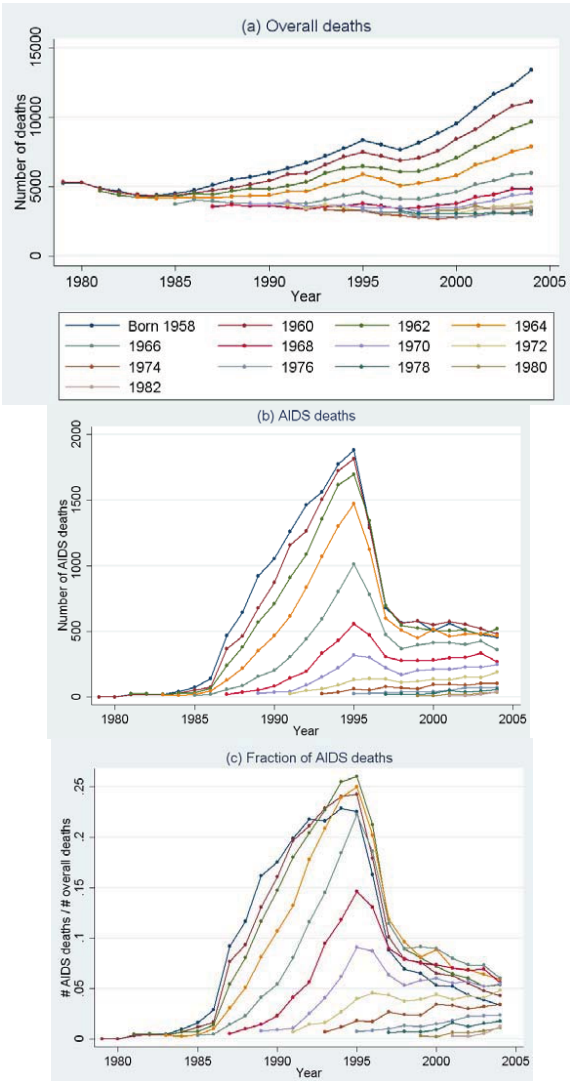
Table 3.10: Regressions of AIDS/Non-AIDS Death Ratios on Age 18 Unemployment Rate by Gender and Race

	Sample			
Effect of Age 18 Unemployment Rate on	Male	Female	White	Non-white
AIDS Deaths Ratio, 1979-1991	(1)	(2)	(3)	(3)
Effect at Age 19	0.002	0.001	0.001	0.003
	(0.001)	(0.001)	(0.001)	(0.001)
Effect at Age 20	0.002	0.002	0.001	0.004
	(0.001)	(0.001)	(0.000)	(0.001)
Effect at Age 21	0.002	0.001	0.001	0.003
	(0.001)	(0.001)	(0.000)	(0.001)
Effect at Age 22	0.001	0.000	0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Effect at Age 23	0.000	-0.001	-0.001	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Effect at Age 24	-0.001	-0.001	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.002)
Effect at Age 25	0.001	-0.002	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Effect at Age 26	0.001	-0.002	0.000	0.002
	(0.001)	(0.001)	(0.001)	(0.003)
Effect at Age 27	0.000	-0.001	-0.001	0.006
	(0.001)	(0.002)	(0.001)	(0.002)
Effect at Age 28	0.004	0.003	0.002	0.008
	(0.002)	(0.002)	(0.002)	(0.003)
Effect at Age 29	0.010	0.013	0.008	0.018
	(0.003)	(0.003)	(0.003)	(0.006)
Effect at Age 30	0.013	0.015	0.011	0.015
	(0.003)	(0.004)	(0.003)	(0.006)
Effect at Age 31	0.018	0.020	0.015	0.037
	(0.004)	(0.006)	(0.004)	(0.005)
Effect at Age 32	0.018	0.021	0.015	0.027
	(0.004)	(0.005)	(0.004)	(0.008)
Effect at Age 33	0.019	0.021	0.016	0.031
	(0.006)	(0.010)	(0.006)	(0.010)
State FE, year FE	✓	✓	✓	✓
Birth cohort FE	✓	✓	✓	✓
Age FE	✓	✓	✓	✓
Number of observations (collapsed)	5,927	5,841	5,927	5,300
R^2	0.83	0.44	0.78	0.59

Notes: Coefficients from separate regressions for males, females, whites and non-whites are displayed. Further comments as in Table 3.8. For a graphical presentation see Figures 22.

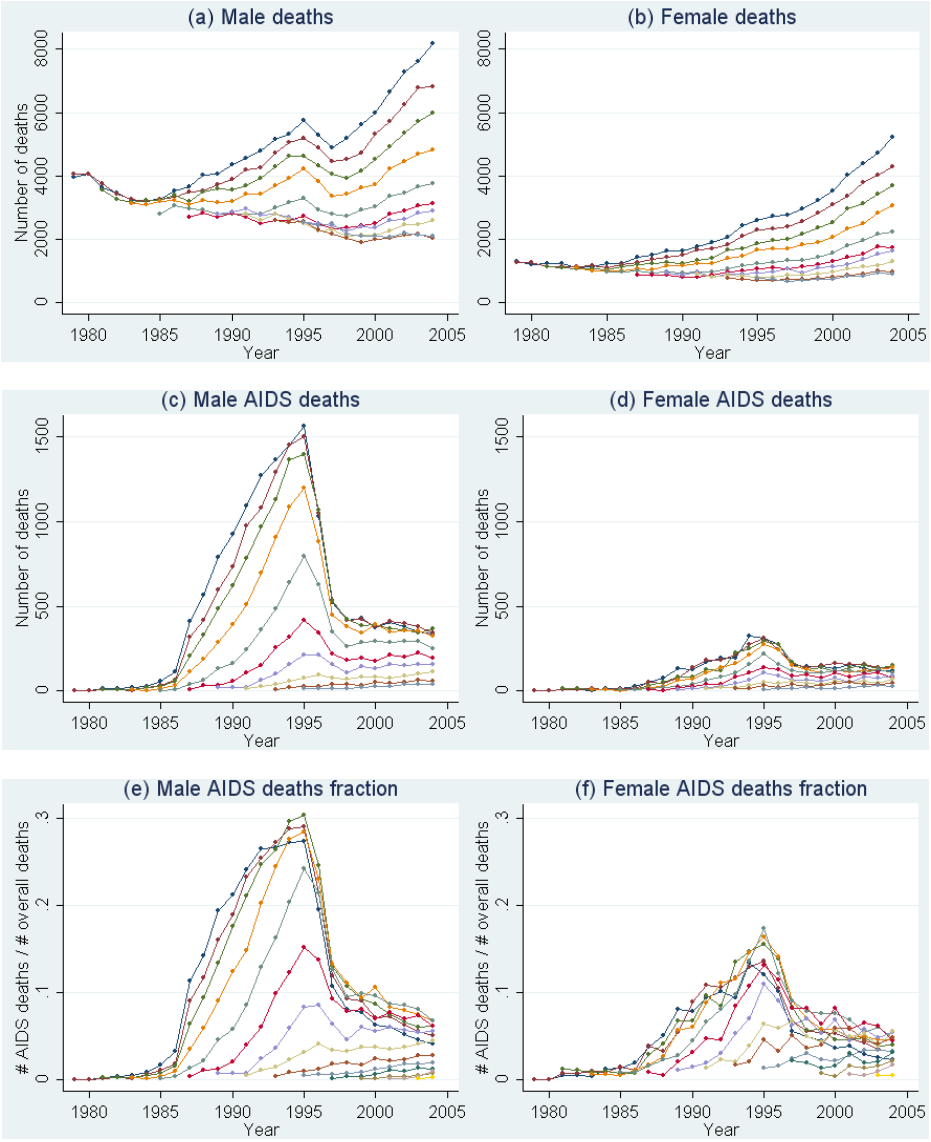
3.8 Appendix

Table 3.11: Deaths, AIDS Deaths, and AIDS Death Fractions by Cohorts



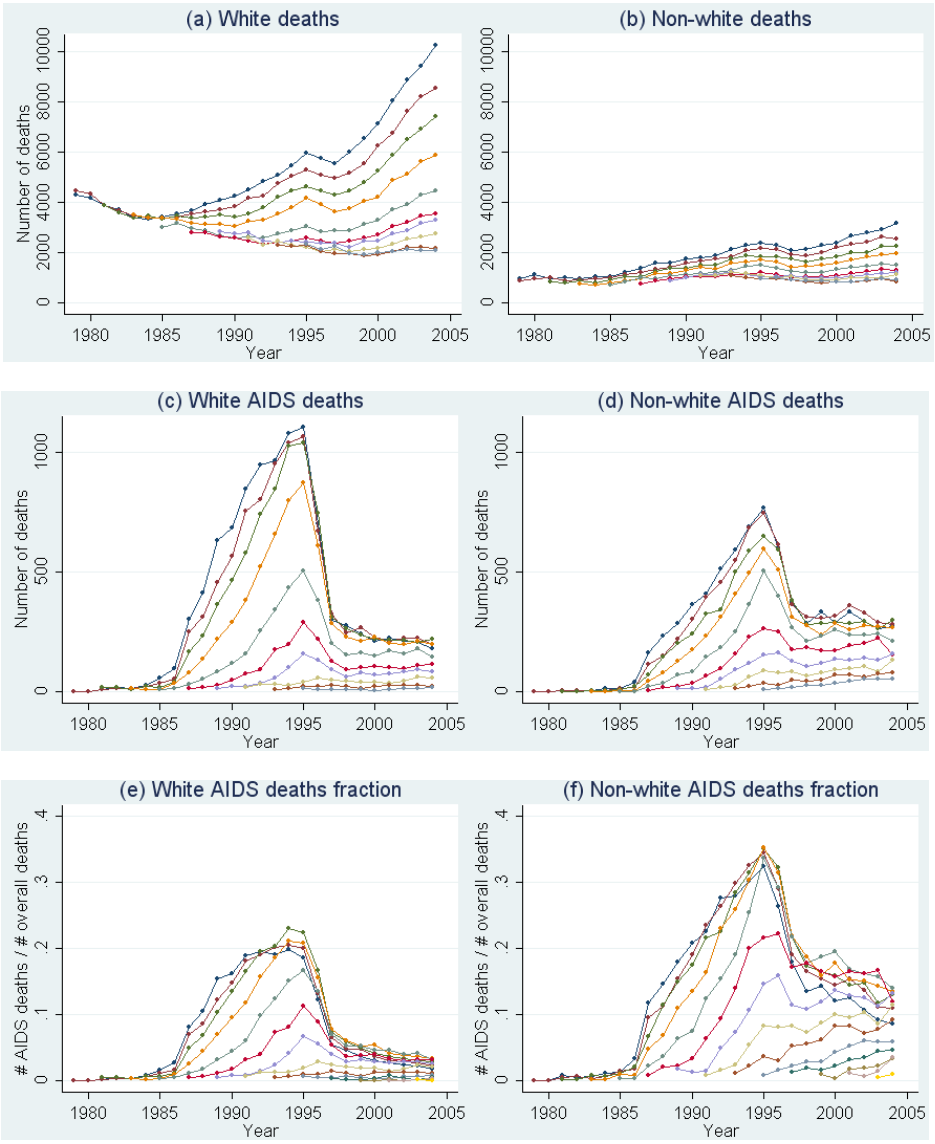
Notes: Figures present the universe of overall deaths and of AIDS deaths, and AIDS death fractions for the sample cohorts, taken from the Vital Statistics. For panels (b) and (c) the same legend applies as in panel (a).

Table 3.12: Deaths, AIDS Deaths, and AIDS Death Fractions by Cohorts and Gender



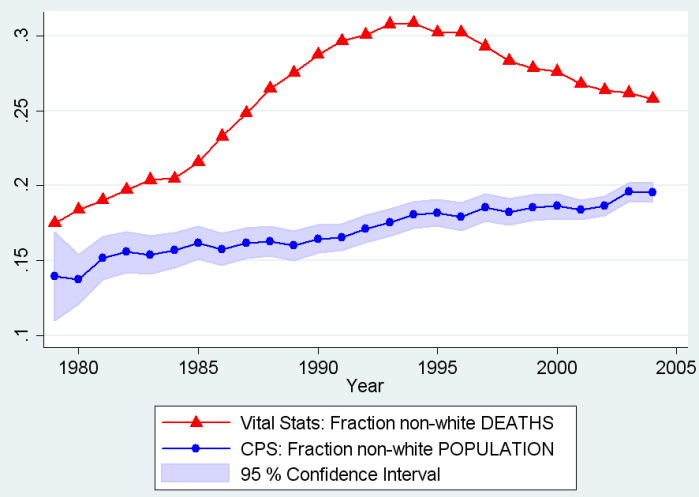
Notes: Figures present - separately for males and females - the universe of overall deaths and of AIDS deaths and AIDS death fractions for the sample cohorts, taken from the Vital Statistics. The same legend as in Figure A1 (a) applies.

Table 3.13: Deaths, AIDS Deaths, and AIDS Death Fractions by Cohorts and Race



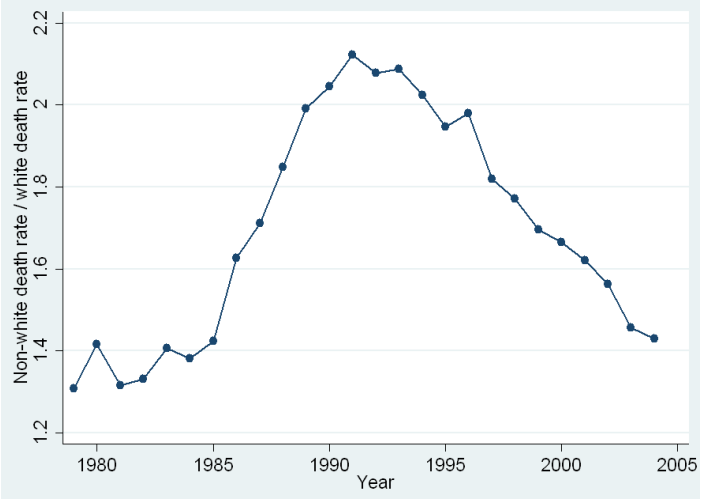
Notes: Figures present - separately for whites and non-whites - the universe of overall deaths and of AIDS deaths and AIDS death fractions for the sample cohorts, taken from the Vital Statistics. The same legend as in Figure A1 (a) applies.

Table 3.14: Fractions of Non-White Deaths and of the Non-White Population over Time



Notes: The fraction of non-white deaths refers to the ratio of non-white deaths divided by overall deaths, calculated from the Vital Statistics. The fraction of non-white population is estimated from the CPS. For a discussion see the section on the HIV/AIDS epidemic.

Table 3.15: Ratio of Non-White Death Rate / White Death Rate over Time



Notes: The ratio of the non-white death rate divided by the white death rate is plotted. This ratio combines the information from the Vital Stats and the CPS used in Figure A4. The ratio is calculated in each year by: [number of non-white deaths / fraction non-white population] / [number of white deaths / fraction white population]

Bibliography

[1] Beaudry, Paul and John DiNardo (1991) "The Effect of Implicit Contracts on the Movements of Wages over the Business Cycle" *Journal of Political Economy*, 99(4): 665-688.

[2] Card, David, Carlos Dobkin and Nicole Maestas (2009) "Does Medicare Save Lives?" *Quarterly Journal of Economics*, Vol 124, No 2, pp. 597-636.

[3] Case, Anne (2004) "Does Money Protect Health Status? Evidence from South African Pensions in David Wise (Ed.)," *Perspectives on the Economics of Aging*, Chicago: University of Chicago Press.

[4] CDC - U.S. Centers for Disease Control and Prevention. (2001) "HIV and AIDS-United States, 1981-2000." *MMWR Morbid Mortal Weekly Report*, 50:430-434.

[5] Chaisson RE, Fuchs E, Stanton D, et al. (1991) "Racial heterogeneity of HIV antigenemia in people with HIV infection." *AIDS*, 5:177-80.

[6] Chaisson RE, Keruly JC, Moore RD. (1995) "Race, sex, drug use, and progression of human immunodeficiency virus disease." *The New England Journal of Medicine*, 333:751-6.

[7] Chiasson MA, Berenson L, Li W, et al. (1999) "Declining HIV/AIDS mortality in New York City." *Journal of Acquired Immune Deficit Syndrome*, 21:59-64.

[8] Currie, Janet and Gruber Jonathan (1996a) "Health Insurance Eligibility, Utilization of Medical Care, and Child Health." *Quarterly Journal of Economics*, 111:431466.

[9] Currie, Janet and Gruber Jonathan (1996b) "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women." *Journal of Political Economy*, 104:12631296.

[10] Cutler, David M., Lleras-Muney, Adriana, and Vogl, Tom (2011) "Socioeconomic Status and Health: Dimensions and Mechanisms" in In S. Glied and P. Smith, eds., *Oxford Handbook of Health Economics*, Oxford University Press, 2011.

[11] Deaton, Angus (2003) "Health, inequality, and economic development" *Journal of Economic Literature*, 41/1: 113-158.

[12] Finkelstein A, Taubman S, Wright B, Bernstein M, Gruber J, et al. (forthcoming) "The Oregon Health Insurance Experiment: Evidence from the First Year." *Quarterly Journal of Economics*

[13] Gardner, Jonathan and Andrew J. Oswald (2007) "Money and mental wellbeing: A longitudinal study of medium-sized lottery wins" *Journal of Health Economics*, 26, No. 1 (January), 49-60.

[14] Genda, Yuji., Ayako Kondo, and Souichi Ohta. (2010) "Long-term Effects of a Recession at Labor Market Entry in Japan and the United States," *Journal of Human Resources*, Vol. 45, pp. 157-196.

[15] Jensen, Robert and Kaspar Richter (2003) "The health implications of social security failure: Evidence from the Russian pension crisis." *Journal of Public Economics* 88.

[16] Johnson, Judith A. (2011) "The Ryan White HIV/AIDS Program" *Congressional Research Service*, 2011; 7-5700.

[17] Kahn, Lisa (2010) "The Long-Term Labor Market Consequences of Graduating College in a Bad Economy." *Labour Economics*, 17(2): 303-316.

[18] King JT, Justice AC, Roberts MS, Chang CH, Fusco JS et al. (2003) "Long-Term HIV/AIDS Survival Estimation in the Highly Active Antiretroviral Therapy Era." *Medical Decision Making*, 23 (1): 9-20.

[19] Miller, Douglas L., Marianne E. Page, Ann Huff Stevens and Mateusz Filipski (2009) "Why are recessions good for your health?" *American Economic Review: Papers & Proceedings*, 99:122-127.

[20] Newhouse, Joseph P, and the Insurance Experiment Group (1993) "Free for All?: Lessons from the RAND Health Insurance Experiment." *Harvard University Press, Cambridge, MA.*

[21] Oreopoulos, Philip, Till von Wachter, and Andrew Heisz (2009) "The Short- and Long-Term Career Effects of Graduating in a Re-

cession" *American Economic Journal: Applied Economics*, 4(1): 1-29.

[22] Rothenberg R, Woelfel M, Stoneburner R, Milberg J, Parker R, Truman B. (1987) "Survival with the acquired immunodeficiency syndrome: experience with 5833 cases in New York City." *The New England Journal of Medicine*, 317:1297-302.

[23] Ruhm, Christopher J. (2000) "Are recessions good for your health?" *Quarterly Journal of Economics*, 115/2: 617-650.

[24] van den Berg, Gerard J., Maarten Lindeboom, and France Portrai (2006) "Economic Conditions Early in Life and Individual Mortality." *American Economic Review*, 96 (March): 290-302.