Towson University
Department of Economics
Working Paper Series

Working Paper No. 2010-08

Subjective Health Expectations

by

By Juergen Jung and Kim P. Huynh

May, 2014

Subjective Health Expectations*

Kim P. Huynh† Bank of Canada

Juergen Jung‡ Towson University

May 9, 2014

Abstract

Subjective health expectations are derived using data from the Health and Retirement Study (HRS). We first use a Bayesian updating mechanism to correct for focal point responses and reporting errors of the original health expectations variable. We then test the quality of the health expectations measure and describe its correlation with various health indicators and other individual characteristics. Our results indicate that subjective health expectations do contain additional information that is not incorporated in subjective mortality expectations and that the rational expectations assumption cannot be rejected for subjective health expectations. Finally, the data suggest that individuals younger than 70 years of age seem to be more pessimistic about their health than individuals in their 70's.

JEL Classification: I10, D84, C11, C23

Keywords: Focal Points, Bayesian Updating of Expectations, Rational Health Expectations, Work Limiting Health Problems.

*Formerly circulated as "Correcting Focal Point Biases in Subjective Health Expectations: An Application to the RAND-HRS Data." We would like to thank Chetan Dave, Gerhard Glomm, Michael Kaganovich, Wayne R. Gayle, Rusty Tchernis and Pravin Trivedi for many helpful comments. We are also grateful to Li Gan for making Matlab code available to us. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada.

†Bank of Canada, 234 Wellington Street, Ottawa, Ontario K1A 0G9, Canada. E-mail: kim@huynh.tv

‡Corresponding author: Juergen Jung, Department of Economics, Towson University, 8000 York Road, MD 21252-0001, Phone: (812) 345-9182, E-mail: jjung@towson.edu

1 Introduction

Standard macroeconomic models impose rational expectations as a modeling device. However, Manski (2004) makes the case for the more frequent use of direct measures of subjective expectations in economics and foresees the need for research on how such expectations are formed. From the empirical literature such as Benitez-Silva and Ni (2008), Hurd (2009), and Khwaja, Sloan and Chung (2007), we know that self reported subjective probability measures in household surveys have strong predictive power. This is especially true when individuals have considerable private information which is common when dealing with health and mortality issues. In addition, subjective mortality expectations have been used to explain retirement, cf. O'Donnell, Tappa and Doorslaer (2008), and savings decisions cf. Bloom et al. (2006). However, results obtained from using subjective mortality expectations data are somewhat different from results obtained using health outcome variables. The latter are shown to have a much stronger influence on retirement decisions than the expectations variables, cf. Siddiqui (1997) and Dwyer and Mitchell (1998). Most of the literature, such as Hurd (1989), Hurd and McGarry (1995), Hurd and McGarry (2002), Gan et al. (2004), Van Solinge and Henkens (2010), that connects subjective expectations with labor supply and retirement decisions concentrates on the role of mortality expectations and ignores subjective expectations about future health issues.

In this paper, we argue that expectations about future health events are significantly different from expectations about future mortality. Auld (2002) finds that changes in health do not have the same effects as changes in life-expectancy, although both are closely linked. Fortunately the Health Retirement Study (HRS), a panel data survey that covers the elderly in the U.S. from 1992 onwards, reports measures about subjective health expectations. More specifically, these data provide information about individuals' expectations about future work limiting health problems and the actual occurrence of such health problems which allows us to systematically analyze subjective expectations about future health.

Our results can be summarized as follows. Subjective health expectations about future work limiting health problems seem to consistently predict health outcomes. The rational expectations hypothesis cannot be rejected for subjective health expectations. Since a high percentage of the answers to the health expectations question are focal points, we correct for focal points using a Bayesian algorithm. After filtering the subjective expectations from reporting errors they become even more similar to rationally formed expectations. Younger cohorts are more pessimistic about their future health than older cohorts. Finally, we are able to show that subjective expectations about future health events carry additional information that is not contained in subjective survival expectations. We therefore argue that it is important to understand how individuals form beliefs

about their future health in order to improve economic models that simulate household decision making over the life-cycle.

To justify the importance of the results of the health expectations variable we first analyze the health outcome variable and connect it to results from the literature on disability. We then test how well subjective expectations about such work limiting health problems predict health outcomes and whether such expectations are formed rationally. The health variable that we use for this analysis is *work limiting health problems* from the HRS data. We refer to this variable as *Health-Problem* from now on. This binary variable indicates whether an impairment or health problem exists and limits the kind or amount of paid work the respondent is able to perform. The HRS does not define what a work limiting health problem is. We therefore quantify *Health-Problem* by investigating its correlation with various health indicators as well as demographic and income variables. The HRS data set is very rich in detailed questions about the health status of its respondents, so that regressions of this form will give an indication of which health problems are more likely to constitute work limiting health problems.

An obstacle with using subjective health expectations are focal point responses. In wave 1, 18.2 percent of respondents indicate a zero probability of acquiring a work limiting health problem within the next 10 years, whereas 4.7 percent think that they will have a work limiting health problem with probability one. There is a third focal point at 50 percent. Roughly 30 percent of respondents expect a work limiting health problem with probability one-half. We find similar results for the remaining waves. Expectations of zero and one are not very sensible while 50 percent is coarse at best. Perry (2005) finds that individuals answering with focal responses of zero and one on an expected mortality question are on average less educated, hold fewer assets and have lower income than the rest of the sample whereas respondents reporting a 50 percent chance of surviving up to a target age look essentially the same as the rest of the sample. He therefore suggests that answers of zero and one may be more a sign of poor understanding of the question than of optimism or pessimism. In order to correct for focal point responses a Bayesian updating mechanism is used following a procedure developed by Gan, Hurd and McFadden (2005). We then test whether health expectations are consistent with health outcomes, formed rationally, and provide information that is not contained in other expectations variables such as mortality expectations.

In this paper we make several contributions to the literature. First, we contribute to the literature that links disability to work limiting health problems. Kapteyn, Smith and Soest (2008) report that disability from various health impairments is highly correlated with work limiting health problems. It has also been reported that disability and other work impairments do have strong effects on labor market decisions (e.g. Gannon (2005) or Kerkhofs and Lindeboom (1995)).

An important research objective will therefore be to understand how individuals form expectations about possible future health impairments and how such expectations influence current economic decisions. We do make progress in this direction by finding that the rationality of health expectations cannot be rejected. In a similar setting Ludwig and Zimper (2007) present various models of learning of subjective mortality expectations.[1]

We also contribute to the literature that tests whether subjective expectations are formed rationally. Rationality assumptions have been tested using survey data by Bernheim (1990), Das and Soest (1997), Das and Soest (1999), Benitez-Silva et al. (2008), and more recently by Dave (2011). Pesaran and Weale (2006) provide a summary of the literature on expectations in household surveys. There is some evidence that individuals form expectations in a rational manner (i.e. Bernheim (1990), Benitez-Silva et al. (2008)) but this depends on the particular context that a decision maker is in. Dave (2011), for instance, does not find prove for either rational or adaptive expectations about capital expenditures on machinery and equipment in Canadian manufacturing plants that instead exhibit a regressive expectations process. Our approach most closely follows Bernheim (1990) and Benitez-Silva et al. (2008) and similar to their results we also cannot reject the rational expectations hypothesis.

The paper is structured as follows. The next section describes the data. Section 3 analyzes the subjective health expectations measure. We correct for focal point responses, test for consistency, informational content and rationality of these self-reported health expectations. Section 4 adds a discussion about the importance of work limiting health problems in economic modeling. Section 5 concludes the paper. A detailed appendix is available on the authors website.[2]

2 Data

The Health and Retirement Study (HRS) is a longitudinal household survey data set for the study of retirement and health among the elderly in the United States. With the goal of making the data more accessible to researchers, the RAND Center for the Study of Aging, with funding and support from the National Institute on Aging (NIA) and the Social Security Administration (SSA), created several products, including the RAND HRS Data File, which is a user-friendly file derived from all HRS waves. It is a composite data set that combines four cohort studies to construct a nationally representative data base of the older population in the U.S. The cohorts are the AHEAD cohorts born before 1924, the CODA cohorts born between $1924-1930$, the HRS cohorts born between $1931-1941$ and the War Baby cohorts born between $1942-1947$. The

[1] We leave the analysis of learning health expectations and the impacts of health expectations on labor market and savings decisions for future research.

[2] http://pages.towson.edu/jjung/papers/TechnicalAppendixSubHealthExp.pdf

largest of these surveys is the Health and Retirement Study (HRS) conducted by the Institute for Social Research at the University of Michigan. It covers a broad range of topics, including health, income, assets, employment, retirement, insurance, and family structure. In this paper we use six waves of data from surveys conducted every two years between $1992 - 2002$.

The majority of respondents in wave 1 of the HRS were 51 to 61 years old when the survey was first conducted in 1992. The baseline survey included $12,652$ persons, or $7,600$ households, with oversampling of Mexican-Americans, African-Americans and residents of Florida. Juster and Suzman (1995) present a general overview of the HRS, Wallace and Herzog (1995) review the health measures in particular and Hurd and McGarry (1995) evaluate the subjective probabilities of survival. In the following we will concentrate on the population aged between 40 and 80 years in wave 1 who will turn 52 and 92 years respectively in wave 7. Wave 7 does not contain the health expectations variable anymore so that we can only use waves $1 - 6$ (years 1992 - 2002). We will restrict the descriptive statistics to these waves so that we are left with $39,442$ observations. Figure 1 contains histograms of the age distributions in the various waves and table 6.1 presents summary statistics of the pooled data of wave $1 - 6$.

3 Health Expectations

We call the variable concerning individual expectations about future work limiting health problems *Health-Expectation*. The exact wording of the survey question is:

"What about the chances that your health will limit your work activity during the next 10 years?"

Respondents can answer with a number from 0 to 100, where 0 indicates an individuals subjective believe that there is absolutely no chance of developing a work limiting health problem and 100 means that the person is absolutely certain that a health problem will develop. Histograms of *Health-Expectation* by gender for waves 2 to 6 are reported in figure 2.

It is not entirely clear what individuals form expectations about when asked about their expectation concerning future work limiting health problem. The data do, however, contain a variable called "work limiting health problems" which is defined as follows:

"*Now we want to ask how your health affects paid work activities. Do you have any impairment or health problem that limits the kind or amount of paid work you can do?*"

In order to quantify this variable we use a discrete-choice model to estimate what causes such work limiting health problems.[3] We find that measures for earnings are negatively correlated with health problems whereas asset holdings turn out to be not significant in explaining this type

[3]Results of detailed regressions of this kind are available in a technical appendix from the authors' website.

of health problem. Men are more likely to develop work limiting health problems and, perhaps surprisingly, age is negatively correlated with work limiting health problems. Finally, healthy lifestyle choices like regular exercise are negatively correlated with health problems. However, after controlling for endogeneity using a more sophisticated instrumental variables estimator using lagged values as regressors to ensure exogeneity, lifestyle choices are not significant anymore, with the exception of smoking.[4]

The standard criticism concerning the use of self reported data in this context is that individuals tend to answer that they do have work limiting health problems to justify that they are out of work. Estimates therefore tend to overstate the health effects on hours worked. See French (2003) for a discussion of this issue. Other issues with self-reported mortality and health data include perception differences by age and socio-economic status (e.g. Sen (2006), Crossley and Kennedy (2002)) as well as nationality (e.g. Jürges (2006)).

3.1 Summary Statistics

We next present non-parametric statistics on $40 - 55$ year old individuals of wave 1 and divide the sample into subgroups by educational attainment and wealth and income quantiles.

Table 2 presents health expectations across all waves according to educational attainment. Comparing the mean expectations we see that in wave 1, college educated individuals have lower expectations about having work limiting health problems in the future than their less educated counterparts. College and above report a 34.2 percent probability versus 43.9 percent, 42.1 percent and 38.6 percent for no high school, GED and high school graduates respectively. This pattern repeats itself across all waves, although in later waves as the population gets older the expectations of higher educated individuals moves closer to expectations of lower educated groups.

Table 3 summarizes health expectations according to wealth quantiles. We find a similar convergence pattern as in the classification by educational attainment. Individuals in high wealth and income quantiles (the income specific table is available in a technical appendix) have lower average subjective expectations about having a work limiting health problem within the next ten years. As the population gets older the expectations converge somewhat for both wealth and income quantiles.

We next compare the subjective health expectations with the actual occurrence of work limiting health problems a decade later. We report mean values of health expectations in wave 1 (*Health-Expectation)* and compare them to the realizations of health limiting problems in wave 6 (*Health-Problem).* Results are reported in table 4.

[4]We follow Wooldridge (2005) and Hernandez-Quevedo, Jones and Rice (2008) in constructing such a dynamic regression model. The technical appendix, available upon request, contains the details of this model.

In table 4 it seems that health expectations are fairly inconsistent when compared with realized health problems approximately 10 years later. To see this compare the average *Health-Expectation* in wave 1 to the average *Health-Problem* in wave 6. However, if one accounts for individuals who either left the survey or died from wave 1 to wave 6 (we unfortunately cannot distinguish between the two cases) then health expectations become more consistent for the age group 45 and above (compare mean of variable *Health-Problem-A* in table 4). Smith, Taylor and Sloan (2001) also report that attrition between waves is approximately 20 percent that is not due to death. Adjusting for this they find that the death rates in the HRS data corresponds fairly well to the decennial life table measures. The same holds true in for a sub-sample where we only include individuals without work limiting health problems in wave 1. It also appears that males slightly underpredict future health problems, whereas females slightly overpredict health problems. From these summary statistics we conclude that expectations about future work limiting health problems are formed reasonably, that is, consistent with later realizations of such health problems.

3.2 Adjusted Subjective Health Expectations

In figure 2 we saw earlier that self-reported expectations show focal point responses, especially high at 0 percent, 50 percent and 100 percent. Respondents who report a 100 percent chance of developing work limiting health problems have on average lower income, asset holdings and education. All other focal respondents (0 percent and 50 percent probability of developing work limiting health problems) are similar to the rest of the sample as can be seen from the non-parametric estimates in table 5. If a respondent thinks that there is absolutely no chance, a zero probability, of having a work limiting health problem within the next 10 years, the question arises why one could not just take this value and postulate that the respondent will use exactly this expectation in her decision process on other economic choice variables (i.e. consumption, savings, etc.). Since economists are ultimately interested in modeling these decision processes, why not work with this probability?

We argue that focal point responses cannot reflect true probabilities because it is unreasonable to assume that health expectations that cover a decade can be made with absolute certainty. We also assume that individuals know this when they actually make their optimizing decision and simply misreported their subjective probabilities in the survey, so that it makes sense to correct this reporting error. We correct for this reporting error by creating a new variable for health expectations called *Adjusted-Health-Expectation*. There are multiple ways to accomplish this. First, we could either replace focal point answers of zero with 0.01 and focal point answers of 100 with 99.9 respectively. This method is suggested in Picone et al. (2004). Another method, suggested by Gan, Hurd and McFadden (2005), uses a Bayesian updating mechanism to smooth

the focal points. This methods is briefly described as follows.

Using observed outcome probabilities of the variable *Health-Problem* we can construct a non-parametric estimate of the population health hazard rate $\lambda_{0a}(a+t)$ per age group a, where subscript 0 indicates that this is the hazard rate of the population, and $a+t$ indicates the hazard rate t years from age a. We present the population hazard rate in the top panel of figure 3 for each age group. The subjective health hazard rate of an individual i of age a can then be expressed as scaled hazard function:

$$\lambda_{ia}(a+t) = \gamma_i \lambda_{0a}(a+t),$$

where $\lambda_i(t)$ is the subjective hazard rate, γ_i is an idiosyncratic adjustment factor. A value $\gamma_i > 1$ indicates a "pessimistic"individual and a value $\gamma_i < 1$ indicates an "optimistic" individual. To construct a measure of γ_i the Bayesian updating model developed in Gan, Hurd and McFadden (2005) is used.

This method assumes that the prior survival probability distribution (probability of surviving without a health problem) at a future point in time is a truncated normal distribution between zero and one. The conditional density of the observed survival probability is assumed to be a censored normal distribution between zero and one which allows for the focal points. We estimate the parameters of the truncated and censored normal distributions using a log likelihood procedure to acquire estimates of $(\sigma_1, \sigma_2, \Psi)$, where σ_1 and σ_2 are standard deviations of the respective normal distributions and Ψ measures the population's average degree of optimism. We then use the posterior density mean as the individual's estimated subjective survival probability. This mean will never be at the boundary of the interval from zero to one so that the adjusted subjective survival probabilities do not contain any more focal points. The details of this procedure as well as a detailed description of the algorithm is available in a technical appendix on the authors' website or in the exposition of Gan, Hurd and McFadden (2005).

The procedure only corrects responses of individuals who did not have a work limiting health problem at the time of the survey, so that some of the focal responses still remain in the sample. Figure 2 shows the variable *Adjusted-Health-Expectation* next to the original one. We see that focal point responses are greatly reduced.

3.3 Are Health Expectations Formed Rationally?

We next employ the framework developed in Bernheim (1990) and Benitez-Silva et al. (2008) to test whether expectations about work limiting health problems are formed rationally. We use adjusted health expectations data unless indicated otherwise. An individual is trying to predict a variable X and has access to certain information during period t. We denote this information

set by Ω_t. In period $t+1$ the information set is augmented by newly available information ω_{t+1}, so that the new information set is $\Omega_{t+1} = (\Omega_t, \omega_{t+1})$. In our model we impose that individuals form expectations according to

$$X_t^e = E\left(X|\Omega_t\right),$$

where E is the expectations operator. This guarantees that errors in expectations will be uncorrelated with the set of variables known at time t. It then follows that

$$E\left(X_{t+1}^e|\Omega_t\right) = E\left[E\left(X|\Omega_t, \omega_{t+1}\right)|\Omega_t\right] = E\left[X|\Omega_t\right] = X_t^e.$$

The evolution of expectations is

$$X_{t+1}^e = X_t^e + \eta_{t+1}, \tag{1}$$

where the expectations error is $\eta_{t+1} = X_{t+1}^e - E\left[X_{t+1}^e|\Omega_t\right]$ and

$$E\left[\eta_{t+1}|\Omega_t\right] = 0. \tag{2}$$

From expression (1) and (2) we can derive a regression framework to test for the rational expectations hypothesis, that is

$$X_{t+1,i}^e = \alpha + \beta X_{t,i}^e + \gamma \Omega_{t,i} + \epsilon_{t,i}, \tag{3}$$

where i indexes the individual, α is a constant, and γ is a parameter vector that estimates the effect of information in period t on period $t+1$ expectations. The rational expectations (RE) hypothesis then implies that $\alpha = \gamma = 0$ and $\beta = 1$ (strong RE). Weak rationality, according to Bernheim (1990), assumes $\gamma = 0$ and tests for $\alpha = 0$ and $\beta = 1$. In both cases expectations follow a random walk.

Estimating expression (3) with a simple OLS procedure could be misleading due to measurement errors in the dependent variable. We already mentioned that there are focal point responses in the subjective expectations variables. These lead to trimodal error distributions instead of normal error distributions. Also, noisy self-reports and omitted variables can make estimation more complex. Individuals may exaggerate or underestimate their expectations or have other motives to misrepresent them. We partly corrected for these problems by reducing focal point responses using the procedure described above.

A test for weak rationality would assume that $\gamma = 0$ in equation (3) so that the actual estimation is for

$$X_{t+1,i}^e = \alpha + \beta X_{t,i}^e + \epsilon_{t,i}. \tag{4}$$

Expectations are not formed rationally whenever the estimates for β are not close to one and

the estimate for the intercept α is significantly different from zero (see table 6, second column). From this we would conclude that health expectations are not formed according to our theory, so that we would have to reject the weak rationality hypothesis. The same holds true for strong rationality as can be seen in the first column of table 6. Similar results have been found in Dutch household survey data on income expectations by Das and Soest (1997) and Das and Soest (1999).

We next follow Bernheim (1990) who claims that one should instrument the variable *Health-Expectation* with other subjective expectations variables. The use of these variables as instruments is based on the assumption that individuals' expectations are internally consistent, in the sense that all expectations are based on the same information. We therefore use the subjective mortality expectations *Expectation-to-Live-to-75* and *Expectation-to-Live-to-85* as instruments for work limiting health expectations, *Health-Expectation* and estimate an instrumental variables estimator.

Column 3 to 6 in table 6 report the regression results under strong rationality and weak rationality assumptions where we also report results using the raw health expectations data that still contains large numbers of focal point responses (in columns 4 and 6 respectively). In table 6 we only use data from wave 1 and wave 2, however, we find similar results when using the entire panel. When using the instrumental variables procedure, the coefficients on *Health-Expectation* are indistinguishable from one and the intercepts are not significant. When testing the strong rationality assumption we find that most regressors that stand for information matrix Ω_t are insignificant as well. We therefore are not able to reject the rational expectations hypothesis anymore so that is appears that the expectations variable *Health-Expectation* follows a random walk. Using the adjusted health expectations variable that minimizes the impact of focal point responses results in slightly stronger results in favor of rational expectations. In general, we need to be careful with our interpretation because tests of this kind have low power.

As a robustness check we also ran these tests for different age groups separately (e.g. $40-50$, $55-60$, and $60-65$) to see whether agents become more or less "rational" as they get older. We find that, indeed, the rational expectations hypothesis can be rejected for younger cohorts but cannot be rejected as the cohorts get older (the particular results are available upon request from the authors).

Finally, we observe that agents younger than 70 seem to be more pessimistic about their health than agents older than 70 (see bottom panel in figure 3). The self reported probabilities of contracting a health problem are consistently lower than the actual outcome probabilities of health problems of these age cohorts. We find similar results when estimating a relatively large hazard scaling parameter using the Bayesian updating procedure of Gan, Hurd and McFadden (2005). The technical appendix contains details of these parameter estimates. Ludwig and Zimper (2007) in a similar study highlight the relative pessimism of the younger generations for subjective

mortality expectations.

4 The Informational Content of Subjective Health Expectations

We next analyze whether there is additional information in the subjective expectations about work limiting health problems *Health-Expectation*, that is not contained in subjective expectations about mortality, that is the expectation to live to age 75 (*Expectation-to-Live-to-75*) and the expectation to live to age 85 (*Expectation-to-Live-to-85*). This is of particular interest as sometimes mortality expectations are used in lieu of health expectations in economic models of household behavior (e.g. Hurd (1989) uses mortality expectations as health proxies).

Table 7 reports a Probit regression of $Health\text{-}Problem_{2002}$ in wave 6 on $Health\text{-}Expectation_{1992}$ formed in wave 1, as well as mortality expectations formed in wave 1 (i.e. $Expectation\text{-}to\text{-}Live\text{-}to\text{-}75_{1992}$, and $Expectation\text{-}to\text{-}Live\text{-}to\text{-}85_{1992}$). We find that even after including mortality expectations into our regression model, the health expectations variable stays significant (column 5 in table 7). This indicates that there is additional information in subjective health expectations that is not covered by subjective mortality expectations. Approximating health expectations with mortality expectations as it is often done in the literature is therefore a very strong assumption as it neglects "significant" information. We therefore consider it an improvement to use subjective health expectations information directly when modeling health impairments in household decision models.

To further test the informational content of subjective health expectations we estimate probit models of future *Health-Problems* in wave 2, 3, 4, 5, and 6 based on expectations about these health problems formed in wave 1. Table 8 contains the results. We find an interesting change of the predictive power of subjective health expectations formed in wave 1. It appears that subjective health expectations are stronger predictors for health problems within 2 years and within 8 and 10 years. However, health problems that start within the next 4 to 6 years are less well predicted.

To test the extent of the additional information carried in subjective health expectations precisely one would have to incorporate subjective health expectations into a consumption-savings model and compare the predictions based on this model to predictions based on models using objective realizations of health states. Only then can one safely quantify the additional effect that subjective health expectations carry. Modeling a life-cycle model and calibrating or estimating it would go beyond the scope of this paper and is left for future research. Gan et al. (2004), however, do find significant improvements in using subjective survival expectations in such a modeling environment. This should give an indication that a similar result is possible using subjective health expectations.

5 Conclusion

In this paper, we investigate work limiting health problems and the expectations about work limiting health problems. Since subjective health expectations suffer from strong focal point bias, we use a Bayesian correction mechanism and correct for focal point responses. We show that work limiting health expectations are strongly correlated with other variables measuring health impairments and as such work limiting health expectations can be linked to the literature on disabilities. We also find that individuals with higher education and higher income report lower subjective probabilities to contract work impairing health problems and that self reported expectations about future work limiting health problems match up with the actual realization rates of such health problems. Therefore, we conclude that expectations are formed consistently.

Finally, we find that subjective health expectations do follow somewhat different patterns than mortality expectations and that they do contain additional information that is not contained in subjective mortality expectations. We test the assumption of rational expectations and do not reject it for subjective health expectations. In addition, agents younger than 70 seem to be more pessimistic about their health than agents older than 70 which is consistent with similar findings about subjective mortality expectations. The results of our empirical work illustrates how macroeconomic models can reconcile rational expectations as a modeling device with the empirical data. To our knowledge this paper presents the first systematic analysis of subjective health expectations. We leave the interesting questions about learning health expectations, the theoretical details about the formation of such expectations, and the application of subjective health expectations in dynamic household modeling for future research.

References

Auld, Christopher M. 2002. "Disentangling the Effects of Morbidity and Life Expectancy on Labor Market Outcomes." *Health Economics* 11(6):471–483.

Benitez-Silva, Hugo, Debra S. Dwyer, Wayne-Roy Gayle and Thomas J. Muench. 2008. "Expectations in Micro Data: Rationality Revisited." *Empirical Economics* 34(2):381–416.

Benitez-Silva, Hugo and Huan Ni. 2008. "Health Status an Health Dynamics in an Empirical Model of Expected Longevity." *Journal of Health Economics* 27:564–584.

Bernheim, Douglas B. 1990. How Do the Elderly Form Expectations? An Analysis of Responses to New Information. In *Issues in the Economics of Aging*, ed. David Wise. University of Chicago Press pp. 259–285.

Bloom, David E., David Canning, Michael Moore and Younghwan Song. 2006. "The Effect of Subjective Survival Probabilities on Retirement and Wealth in the United States." NBER Working Paper 12688.

Crossley, Thomas F. and Steven Kennedy. 2002. "The Reliability of Self-Assessed Health Status." *Journal of Health Economics* 21:643–658.

Das, Marcel and Arthur Van Soest. 1997. "Expected and Realized Income Changes: Evidence from the Dutch Socio-Economic Panel." *Journal of Economic Behavior and Organization* 32(1):137–154.

Das, Marcel and Arthur Van Soest. 1999. "A Panel Data Model for Subjective Information on Household Income Growth," Journal of Economic Behavior & Organization." *Journal of Economic Behavior & Organization* 40(4):409–426.

Dave, Chetan. 2011. "Are Investment Expectations Rational, Adaptive or Regressive?" *Economic Inquiry* 49(1):212–225.

Dwyer, Debra S. and Olivia S. Mitchell. 1998. "Health Problems As Determinants of Retirement: Are Self-Rated Measures Endogenous?" *Journal of Health Economics* 18:173–193.

French, Eric. 2003. "The Effects of Health, Wealth, and Wages on Labor Supply and Retirement-Behavior." Federal Reserve Chicago, Working Papers Series WB 2000-02.

Gan, Li, Guan Gong, Michael Hurd and Daniel McFadden. 2004. "Subjective Mortality Risk and Bequests." NBER Working Paper 10789.

Gan, Li, Michael Hurd and Daniel McFadden. 2005. Individual Subjective Survival Curves. In *Analyses in Economics of Aging*, ed. David Wise. The University of Chicago Press pp. 377–411.

Gannon, Brenda. 2005. "A Dynamic Analysis of Disability and Labour Force Participation in Ireland 1995-2000." *Health Economics* 14(9):925–938.

Hernandez-Quevedo, Cristina, Andrew M. Jones and Nigel Rice. 2008. "Persistence in Health Limitations: A European Comparative Analysis." *Journal of Health Economics* 6:1472–1488.

Hurd, Michael D. 1989. "Mortality Risk and Bequests." *Econometrica* 57(4):779–813.

Hurd, Michael D. 2009. "Subjective Probabilities in Household Surveys." *Annual Review of Economics* 1:543–564.

Hurd, Michael D. and Kathleen McGarry. 1995. "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study." *The Journal of Human Resources* 30 (Supplement)(Supplement):S268–S292.

Hurd, Michael D. and Kathleen McGarry. 2002. "The Predicitive Validity of Subjective Probabilities of Survival." *The Economic Journal* 112:966–985.

Jürges, Hendrik. 2006. "True Health vs. Response Styles: Exploring Cross-Country Differences in Self-ReportedHealth." MEA-University of Mannheim.

Juster, F. Thomas and Richard Suzman. 1995. "An Overview of the Health and Retirement Study." *Journal of Human Resources* 30(Supplement):S7–S56.

Kapteyn, Arie, James P. Smith and Arthur Van Soest. 2008. "Dynamics of Work Disability and Pain." *Journal of Health Economics* 27:496–509.

Kerkhofs, M. and M. Lindeboom. 1995. "Subjective Health Measures and State-Dependent Reporting Errors." *Health Economics* 4(3):221–235.

Khwaja, Ahmed, Frank Sloan and Sukyung Chung. 2007. "The Relationship Between Individual Expectations and Behaviors: Mortality Expectations and Smoking Decisions." *Journal of Risk and Uncertainty* 35:179–201.

Ludwig, Alexander and Alexander Zimper. 2007. "A Parsimonious Model of Subjective Survival Expectancy." Mannheim Research Institute for the Economics of Aging (MEA).

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica* 72(5):1329–1376.

O'Donnell, Owen, Federica Tappa and Eddy Van Doorslaer. 2008. "Can Subjective Survival Expectations Explain Retirement Behaviour?" DNB Working Paper, Nr. 188.

Perry, Michael. 2005. "Estimating Life Cycle Effects of Survival Probabilites in the Health and Retirement Study." University of Michigan Retirement Research Center, Working Paper 2005-103.

Pesaran, M. Hashem and Martin R. Weale. 2006. Survey Expectations. In *Handbook of Economic Forecasting*, ed. C.W. J. Granger G. Elliott and A. Timmermann. North-Holland pp. 715–770.

Picone, Gabriel, A. Frank, A. Sloan and Donald H. Taylor Jr. 2004. "Effects of Risk and Time Preference and Expected Longevity on Demand for Medical Tests." *Journal of Risk and Uncertainty* 28(1):39–53.

Sen, Amartya. 2006. "Health: Perception Versus Observation." *British Medical Journal* (324):860–861.

Siddiqui, S. 1997. "The Impact of Health on Retirement Behaviour: Empirical Evidence from West Germany." *Health Economics* 6(4):425–438.

Smith, Kerry V., Donald H. Taylor and Frank A. Sloan. 2001. "Longevity Expectations and Death: Can People Predict Their Own Demise?" *American Economic Review* 91(4):1126–1134.

Van Solinge, Hanna and Kene Henkens. 2010. "Living Longer, Working Longer? The Impact of Subjective Life Expectancy on Retirement Intentions and Behaviour." *European Journal of Public Health* 20(1):47–51.

Wallace, Robert B. and Regula A. Herzog. 1995. "Overview of the Health Measures in the Health and Retirement Study." *Journal of Human Resources* 30(Supplement):S84–S107.

Wooldridge, Jeffrey M. 2005. "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity." *Journal of Applied Econometrics* 20:39–54.

6 Appendix: Tables and Figures

6.1 Tables

Variable	Mean	Std. Dev.	N
Health Expectation	39.69	28.27	39442
Adjusted Health Expectation	43	16.08	35200
Health Problem	0.11	0.31	39415
Expectation to Live to 75	67.71	27.12	34085
Expectation to Live to 85	45	31.14	25934
Age	57.73	6.39	39442
Male	0.46	0.5	39442
$>$ 12 Years Education	0.46	0.5	39372
Living with Partner	0.78	0.41	39404
Full Time Employed	0.83	0.38	39442
Part Time Employed	0.03	0.17	39442
Asset Holdings in $1,000	0.31	0.97	39442
Earnings in $1,000	26.67	40.03	39442
Log of Out-of-Pocket Expenses	5.60	2.47	39442
Log of Health Expenses	7.13	2.13	39442
Smoker	0.21	0.41	39285
Physical Effort at Work	0.36	0.48	37658
Body Mass Index	27.32	4.94	39046
Mother Still Alive	0.39	0.49	38789
Father Still Alive	0.16	0.37	38705
Very Good Health	0.35	0.48	39439
Good Health	0.3	0.46	39439
Fair Health	0.11	0.31	39439
Poor Health	0.02	0.14	39439

Table 1: Summary statistics. Source HRS 1992-2002.

	No High School	GED	High School	Some College	College
			Wave 1: 1992		
Mean	42.56	43.42	37.14	36.37	32.23
Std.Dev.	29.84	26.89	27.87	27.64	23.63
N	766	190	1452	963	887
			Wave 2: 1994		
Mean	39.84	39.66	35.28	32.53	30.00
Std.Dev.	30.45	29.12	26.31	28.52	24.56
N	821	211	1556	1045	905
			Wave 3: 1996		
Mean	41.33	39.94	37.91	35.71	34.04
Std.Dev.	30.90	28.67	28.59	27.59	24.69
N	513	154	1153	793	735
			Wave 4: 1998		
Mean	42.35	41.76	39.49	36.71	35.56
Std.Dev.	29.39	28.30	27.66	26.63	25.39
N	470	136	1027	694	661
			Wave 5: 2000		
Mean	43.52	46.33	42.98	40.48	38.48
Std.Dev.	28.51	28.83	27.05	27.04	24.24
N	377	120	821	599	595
			Wave 6: 2002		
Mean	44.20	46.91	44.70	40.98	39.57
Std.Dev.	29.79	28.24	27.35	27.05	25.70
N	304	97	690	508	504

Table 2: Health Expectations by Educational Attainment.
We follow the wave 1 age group of 40-55 year old individuals up to wave 6.

	1st Quantile	2nd Quantile	3rd Quantile	4th Quantile
		Wave 1: 1992		
Mean	42.95	38.83	36.04	32.54
Std.Dev.	29.44	26.63	27.06	26.54
N	834	1127	1162	1135
		Wave 2: 1994		
Mean	39.32	36.62	33.14	30.52
Std.Dev.	30.31	27.85	26.44	25.68
N	921	1179	1243	1199
		Wave 3: 1996		
Mean	41.05	39.15	35.36	34.15
Std.Dev.	28.95	28.48	27.23	27.23
N	639	892	937	885
		Wave 4: 1998		
Mean	40.74	38.87	38.38	36.47
Std.Dev.	30.05	27.04	26.19	26.59
N	590	804	826	774
		Wave 5: 2000		
Mean	46.78	41.38	41.95	37.41
Std.Dev.	28.16	26.06	27.13	25.39
N	473	673	732	638
		Wave 6: 2002		
Mean	45.24	42.08	42.68	41.23
Std.Dev.	29.09	27.55	26.60	26.63
N	396	578	599	534

Table 3: Health Expectations per Wealth Quantiles.
We follow the wave 1 age group of 40-55 year old individuals up to wave 6.

	MALE					
	Full Sample			**Sub-sample**		
	Age 40-45	**45-50**	**50-55**	**Age 40-45**	**45-50**	**50-55**
Wave 1: 1992						
Mean(Health Expectation)	0.33	0.36	0.39	0.35	0.35	0.37
N	32	206	1757	30	187	1607
Wave 6: 2002						
Mean(Health Problem)	0.48	0.41	0.43	0.21	0.15	0.18
N	29	196	1587	24	166	1340
Wave 6-A: 2002						
Mean(Health Problem-A)	0.56	0.38	0.46	0.51	0.46	0.49
N	43	251	2197	35	204	1790

	FEMALE					
	Full Sample			**Sub-sample**		
	Age 40-45	**45-50**	**50-55**	**Age 40-45**	**45-50**	**50-55**
Wave 1: 1992						
Mean(Health Expectation)	0.31	0.36	0.39	0.30	0.34	0.37
N	262	753	1551	242	683	1443
Wave 6: 2002						
Mean(Health Problem)	0.17	0.24	0.30	0.11	0.17	0.21
N	287	844	1913	250	708	1559
Wave 6-A: 2002						
Mean(Health Problem-A)	0.33	0.40	0.47	0.44	0.47	0.49
N	354	1067	2483	299	890	1999

Table 4: Consistency of the Health Expectations Variable.
The table reports the average of self-reported health expectations in 1992 and the realizations of actual health problems 10 years later in 2002. Health Problems-A counts all individuals having left the survey (due to death or attrition) as having a health problem, so that the variable Health Problem is set to one for such individuals.
The sub-sample reports the average of self-reported health expectations of individuals without work limiting health problems in 1992 and thus comprises a healthier subset of individuals.

Variable	Mean	Std. Dev.	N
Total sample			
Age	51.70	3.23	25348
Income	221866.05	542110.74	25205
Earnings	21688.21	33724.88	25205
Education years	12.65	3.05	25272
Male (1/0)	0.65	0.48	25348
No focal points			
Age	51.68	3.23	18143
Income	228428.72	582477.98	18000
Earnings	20449.79	35667.47	18000
Education years	12.63	3.16	18080
Male (1/0)	0.66	0.47	18143
Health-Expectation = 50			
Age	51.95	3.04	4109
Income	187931.40	389059.87	4109
Earnings	24876.26	25136.10	4109
Education years	12.75	2.69	4102
Male (1/0)	0.63	0.48	4109
Health-Expectation = 0			
Age	51.41	3.48	2602
Income	242115.61	474436.84	2602
Earnings	25966.31	32981.16	2602
Education years	12.78	2.69	2596
Male (1/0)	0.66	0.47	2602
Health-Expectation = 100			
Age	52.00	3.16	494
Income	158344.27	413996.77	494
Earnings	17761.48	21078.79	494
Education years	11.77	3.20	494
Male (1/0)	0.61	0.49	494

Table 5: Summary by Expected Work Limiting Health Problem: Age 40-60

	(1)	(2)	(3)	(4)	(5)	(6)
	Strong R.E.	Weak R.E.	IV-Strong R.E.	IV-Strong R.E.-raw	IV-Weak R.E.	IV-Weak R.E.-raw
Health Expectation				10.04***		10.05***
				(0.14)		(0.10)
Adjusted Health Expectation	0.26***	0.40***	10.03***		10.03***	
	(0.02)	(0.02)	(0.14)		(0.09)	
Expectation to Live to 75	-0.02					
	(0.01)					
Expectation to Live to 85	-0.05***					
	(0.01)					
Age	0.69***		-0.06	0.20		
	(0.07)		(0.16)	(0.15)		
Male	-10.46**		-0.81	0.10		
	(0.62)		(0.76)	(10.52)		
> 12 Years Education	-0.62		-0.84	-10.62		
	(0.61)		(0.74)	(10.42)		
Living with Partner	-10.08*		0.10	0.14		
	(0.65)		(0.84)	(10.55)		
Employed Full Time	10.49		-0.07	20.81		
	(10.09)		(10.41)	(20.57)		
Employed Part Time	20.01		20.93	40.46		
	(10.98)		(20.35)	(40.29)		
Asset Holdings	-0.16		-0.50	-0.73		
	(0.32)		(0.40)	(0.78)		
Earnings in $1,000	0.00		0.00	0.01		
	(0.01)		(0.01)	(0.01)		
Log of Out-of-Pocket Expenses	-0.10		-0.01	-0.01		
	(0.11)		(0.13)	(0.24)		
Log of Health Expenses	0.10		0.11	0.24		
	(0.13)		(0.15)	(0.29)		
Smoker	-0.57					
	(0.66)					
Daily Physical Activity	-0.84					
	(0.58)					
Physical Effort Work	-0.35		-10.25*	-20.60**		
	(0.58)		(0.70)	(10.33)		
Body Mass Index	-0.06		-0.08	-0.13		
	(0.06)		(0.07)	(0.13)		
Mother Alive	10.15**					
	(0.51)					
Father Alive	-10.39**					
	(0.63)					
Very Good Health	0.46		-10.14	-10.86		
	(0.61)		(0.81)	(10.56)		
Good Health	20.40***		0.41	-0.39		
	(0.75)		(10.03)	(20.06)		
Fair Health	20.78**		-0.99	-20.86		
	(10.29)		(10.76)	(30.43)		
Poor Health	-0.97		-110.21	-200.85		
	(50.79)		(60.95)	(130.80)		
N	3295	3295	3295	3295	3295	3295
R^2	0.21	0.13	.	.	.	.

Table 6: Tests for Strong and Weak Rational Expectations.
We use a linear probability model and regress adjusted as well as non-adjusted health expectations measures. We only use data from year 1994 and the variable $Health - Expectation$ from 1996 which leaves us with 3295 observations. The models estimated in the table are: Strong Rational Expectations, Weak Rational Expectations, Strong Rational Expectations using an instrumental variables estimator, Strong Rational Expectations with IV using on-adjusted raw data, Weak Rational Expectations with IV, and Weak Rational Expectations with IV using raw, non-adjusted, data.
Regressors not reported include industry and region dummy variables as well as health indicator variables from the principle components analysis reported earlier. Significance levels are denoted *, **, and *** for 0.10, 0.05, and 0.01, respectively.

	(1) Probit-1-raw	(2) Probit-2	(3) Probit-3	(4) Probit-4	(5) Probit-5
Adjusted Health Expectation		.013***	.013***	.013***	.013***
		(.003)	(.003)	(.003)	(.003)
Health Expectation	.007***				
	(.001)				
Expectation to Live to 75			-.0004		-.0007
			(.002)		(.002)
Expectation to Live to 85				-.0001	.0003
				(.001)	(.002)
N	2270	2270	2270	2270	2270

Table 7: Information content of *Health-Expectation*. Dependent variable is *Health-Problem* (work limiting health problems) in year 2002. Independent variables are from year 1992.
We also control for an identical set of regressors as in Table 6 - Model 1 as well for industry and regional effects (not reported due to space constraints). Significance levels are denoted *, **, and *** for 0.10, 0.05, and 0.01, respectively while standard errors are in parentheses.

	(1) Probit-1-1994	(2) Probit-1996	(3) Probit-1998	(4) Probit-2000	(5) Probit-2002
Adjusted Health Expectation	.014***	.007**	.011***	.013***	.013***
	(.004)	(.003)	(.004)	(.003)	(.003)
Expectation to Live to 75	.002	.002	-.004	.001	-.0007
	(.003)	(.003)	(.003)	(.002)	(.002)
Expectation to Live to 85	.0006	-.002	-.0007	-.002	.0003
	(.002)	(.002)	(.002)	(.002)	(.002)
N:	2114	1908	1928	1879	2270

Table 8: Information content of *Health-Expectation*. Dependent variable is *Health-Problem* (work limiting health problems) in year 1994, 1996, 1998, 2000, and 2002. Independent variables are from year 1992.
We also control for an identical set of regressors as in Table 6 - Model 1 as well for industry and regional effects (not reported due to space constraints). Significance levels are denoted *, **, and *** for 0.10, 0.05, and 0.01, respectively while standard errors are in parentheses.

6.2 Figures

Figure 1: Age distributions, waves 1 to 6

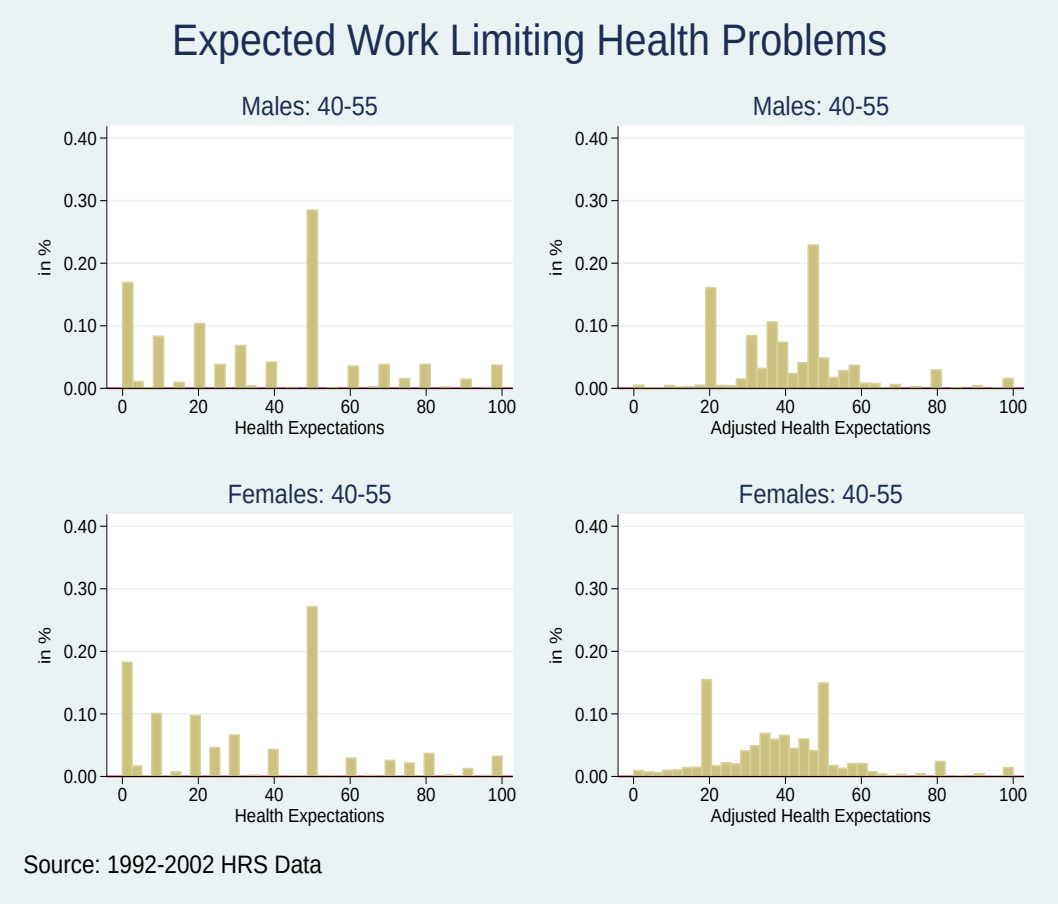

Figure 2: Histogram of Health Expectations (expected work limiting health problems) and Adjusted Health Expectations of Males and Females Between Age 40-60

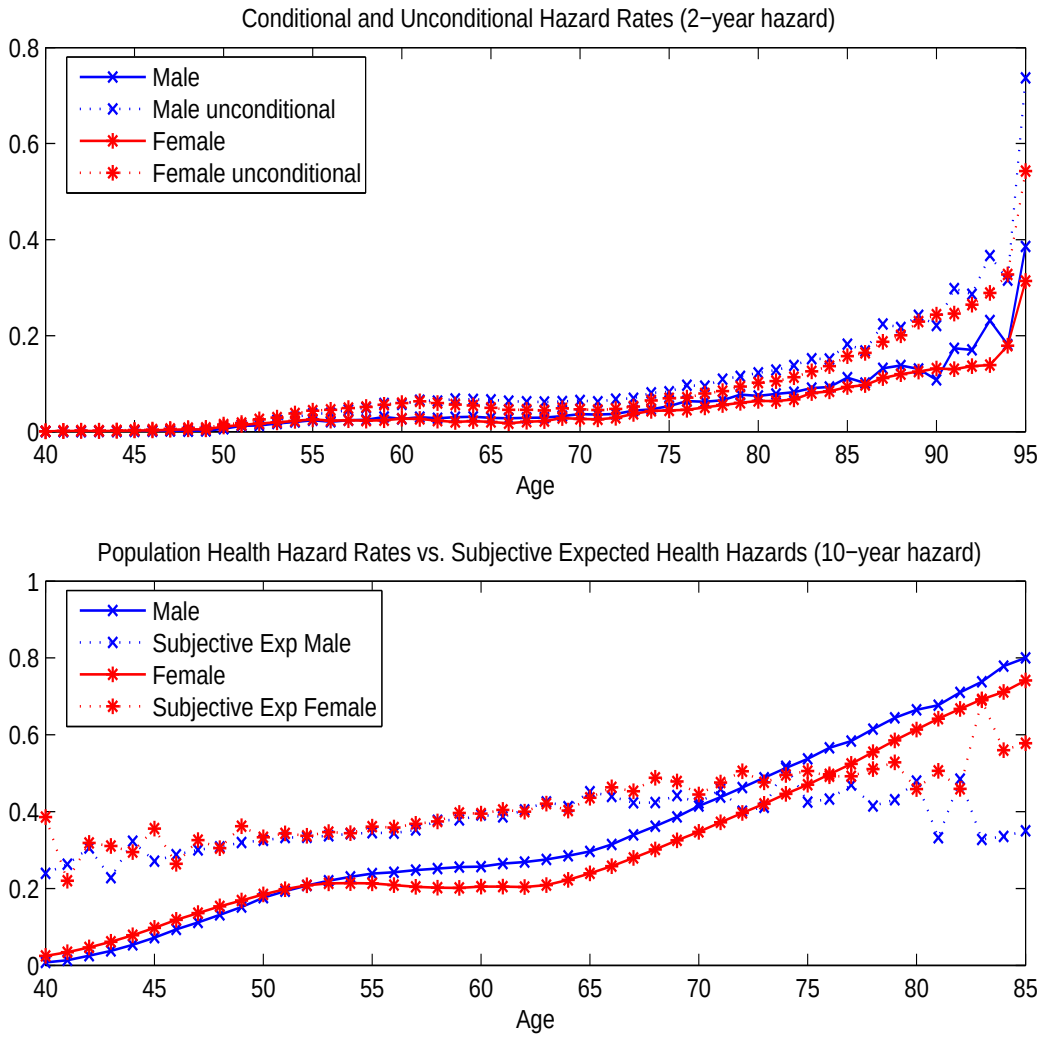

Figure 3: The top panel depicts conditional and unconditional hazard rates of developing a Health Problem (i.e., a work limiting health problem).
The conditional hazard rate only counts individuals that transition from a state without having a health problem to a state of having a health problem. The bottom panel compares cumulative population hazard rates over a 10 year horizon with subjective health hazard rates over the same time frame. Naturally the hazard rates in the bottom panel are higher than the 2 year hazard rates in the top panel. The data is from RAND-HRS,wave 1-6.
Note: We use the conditional hazard rate (conditional on being without a health problem) for our calculations as this counts only the newly sick as failures in the hazard calculation. The unconditional hazard rate also counts individuals that transition from a state of having a health problem to the state of (still) having a health problem two years later. It is for this "double counting" that the unconditional hazard rate is higher.