Health Stocks and Health Flows in an Empirical Model of Expected Longevity

Hugo Benítez-Silva*
SUNY-Stony Brook

Huan Ni †
SUNY-Stony Brook

December 27, 2005

Abstract

Expected longevity is an important factor influencing older individual's decisions such as consumption, savings, claiming of Social Security benefits, and labor supply. It has also been shown to be a good predictor of actual longevity, which in turn is highly correlated with health status. We define health both as a stock and a flow, and analyze the effects of these two types of measures on subjective probabilities of living to a certain age. Health is a continuous latent variable, and researchers have used a variety of self-reported health measures as indicators of health status. However, health flows have been used much less by empirical researchers, and how to measure them appropriately is still an open question. Using longitudinal data from the Health and Retirement Study, we directly incorporate health flows in explaining the variation in expected longevities, and compare two alternative health flow measures: the self-reported health change, and the computed health change based on self-reports of health status. Up to 33% of the sample reports changes in their health status inconsistently with a direct question on health changes over time, and the reports of another 20% of the sample can lead to information losses if computed changes in self-reported status are used to assess changes in actual health. These potentially serious problems raise doubts regarding the use and interpretation of the computed health changes and lagged measures of self-reported health as controls for health dynamics in a variety of empirical settings. Our empirical results show that after controlling for both subjective and objective measures of health status and unobserved heterogeneity in reporting, self-reported health changes are a more appropriate measure of health dynamics.

JEL codes: I12, J14

Keywords: Expected Longevity, Health Status, Health Flows, Panel Data, Health and Retirement Study

*hugo.benitez-silva@sunysb.edu. Benítez-Silva is grateful for the financial support from the Spanish Ministry of Science and Technology through project number BEC2002-04294-C02-01 and to the Fundación BBVA. Additional financial support from the National Institute on Aging under grant 1 P01 AG022481-01A1 on a related project is also gratefully acknowledged. Any remaining errors are our own.

†huni@ic.sunysb.edu ; tel: 631-902-5158; fax: 631-632-7516; Economics Department, SUNY-Stony Brook, Stony Brook, NY 11794-4384.

1 Introduction and Motivation

Expected longevity is an important factor influencing older individual's decisions such as consumption, savings, claiming of Social Security benefits, and labor supply. It has also been shown to be a good predictor of actual longevity (see Hurd et al. (1999), Smith et al. (2001), Hurd and McGarry (2002), and Siegel et al. (2003)), which in turn is highly correlated with health. We define health both as a stock and a flow, and analyze the effects of these two types of measures on subjective probabilities of living to a certain age. Health is a latent variable, and researchers have used a variety of self-reported indicators as proxies for health status. However, health flows have been used much less by empirical researchers, and how to measure them appropriately is still an open question.

An additional motivation for studying the determinants of expected longevity has to do with the fact that it might be a more appropriate variable to use in a variety of models compared to using aggregate indicators of life expectancy. For example, life tables may not be correct for the average individual, since due to observed and unobserved heterogeneity individuals would surely have expected longevities that differ from the population averages provided by the life tables. Based on individual subjective survival probabilities, Gan et al. (2003) have developed an index to compare these self-reported probabilities with life tables. The comparison shows there is significant individual heterogeneity in the Asset and Health Dynamics of the Oldest Old data, and ignoring it may cause biases in many behavioral models.[1] Furthermore, individuals make decisions based on their expected longevity instead of the actual one, which they are unlikely to know ahead of time.

It is important, therefore, to understand how individuals form their longevity expectations, and how these evolve over time, before we use them in any forward looking micro models that account for uncertainty, or in any empirical model that tries to control for heterogeneity regarding life expectancies. In this paper we provide one of the first analysis of how longevity expectations evolve dynamically as a function of both health stocks and health flows. Notice that health stocks give

[1] In an additional research piece using the same data set, Gan et al. (2004) have compared how expected longevities and life tables help the fit of a dynamic consumption and saving model. The authors find that using subjective survival probabilities instead of those from the life tables gives better predictions on saving behaviors of this aging population.

us information on an individual's health level at a certain point in time, while health flows provide information about the direction of change in health status since the period before that same point in time, capturing health dynamics.

The dynamic component of health can come from a variety of sources. First, individuals invest in their health through, for example, preventive visits to health professionals, changes in their diets, exercise, and changes in habits such as smoking or drinking. The results of these investments on a person's health status are uncertain, and in many cases can be the result of previous diagnoses of particular conditions. Second, health depreciates in a fairly continuous fashion through the natural aging process and through the worsening of chronic conditions. Third, health can deteriorate in a significant and rather discontinuous way as negative health shocks occur. The effects on a person's health of the last two types of events is also uncertain and it can depend on the interaction with the first type of investments, as well as on attitudes towards treatment and recovery, and towards life in general. No measure is likely to capture all these components, therefore we will use a variety of indicators hoping to account for most of these effects.

In empirical research, health stocks have been measured by both subjective and objective health status, such as self-rated health, incidence of chronic diseases, Activities of Daily Living (ADLs), Instrumental Activities of Daily Living (IADLs), and measures of health limitation. Self-rated health has been found to play a significant role in explaining individuals' mortality by many researchers. Mossey and Shapiro (1982) are credited with one of the first convincing pieces on this issue. Since then, it has been a robust finding in studies among working-age adults (Miilunpalo et al. (1997)), older adults (Lee (2000) and Mete (2005)), and among all groups of the population by socioeconomic status (Burström and Fredlund (2001)). Interpreting the evidence up to that point, Idler and Benyamini (1997) present a review on 27 studies, and conclude that self-rated health is an independent predictor of mortality, even controlling for objective measures of health. They propose a variety of interpretations to these results, one of which states that "self-rated health is a dynamic evaluation, judging trajectory and not only current level of health." Ferraro and Kelley-Moore (2001) using panel data also find support for the link between self-reported health and mortality. Moreover, they argue that self-rated health can be considered a dynamic measure of health.

But does self-rated health—while controlling for other objective health measures—capture all the information regarding health flows? And if not, which other variables should be used to capture the flow component of health on expectations and decisions? A number of recent papers present evidence of the effects of health changes on retirement expectations. Benítez-Silva and Dwyer (2005) show that (computed) health changes affect older American workers' retirement expectations innovations. McGarry (2004) has shown that self-reported health declines have a negative effect on older workers' retirement expectations, and hinted at the point that using computed differences of health status might lead to information loss. In an earlier study, Deeg et al. (1989) found that the effects of self-rated health diminishes once we control for self-reported health declines in a study of mortality.

We hope to achieve three main goals in this paper. First, we re-open the debate regarding the validity of self-reported health status, by discussing its longitudinal validity linked to the fact that it can fail to capture meaningful health changes among respondents, and the fact that it is unlikely to be the best measure to capture health dynamics. We also emphasize that its use in panel studies can exacerbate the biases intrinsic to measuring a continuous latent variable, due to cut-point shifts (the cut-points between health categories might not be fixed over time for a given individual, or instead they might be areas dividing health categories) and reference groups effects (individuals might report their health as compared with people their age). Second, we investigate different measures of health flows and identify the most appropriate ones. We conclude that self-reported health changes are more effective measures of health flows than computed measures using health stock differences, or lagged measures of health status, since the direct health change question is less exposed to biases, and avoids loss of information. Third, we will show that incorporating information on health flows could help our understanding of how individuals form their expectations of longevity. Our overall results suggest that self-reported health flows should be used in empirical and behavioral models as an important complement to standard measures of health, in order to capture the dynamics of health, and to consistently estimate the effects of self-reported health status. We also find that computed health changes do play a significant role in explaining changes in expected longevity, suggesting that this measure has some information that is not perfectly anticipated by individuals but affects how longevity expectations evolve over time.

The rest of the paper is organized as follows: Section 2 describes and critically analyzes measures of health stocks and health flows both from a conceptual as well as an empirical perspective. Section 3 presents the data used in our econometric estimations. In Section 4, we discuss the empirical methods used in this study. Section 5 presents the results of our panel analysis of longevity expectations formation, and Section 6 concludes.

2 Measuring Health: Self-reports, Stocks, and Flows

In this section we first provide a discussion regarding the validity of self-reported health measures, both in the cross-sectional and longitudinal sense. Second, we analyze a variety of potential problems related to the use of self-reported health status to proxy for health flows.

The validity of self-reported measures of health variables has been the topic of many studies in recent years. In general, the concept of validity is taken to mean whether this measure accurately reflects health status, and therefore can be considered as a useful indicator in empirical and behavioral economic models. The discussion in the economic literature regarding validity of self-reported health measures is mainly focused on the issue of cross-sectional validity, even if authors were using panel and time-series data on the variables of interest. Researchers have been concerned with the use of these variables as possible regressors or indicators of the behavior, decisions, and choices under study, because of the possible econometric problems that using these measures could introduce. Most research has concentrated on understanding how well self-reported health could capture the particular health of an individual at a point in time. The dynamic considerations have been limited to either using lagged health status in the models, or using health transitions as a reflection of actual changes in health.[2] To our knowledge, no study in the economic literature has tried to discuss the advantages and disadvantages of different types of health flow measures in empirical work, or has focused on whether self-reported health could capture actual health changes (this is

[2]Bound et al. (1999) emphasize the dynamic nature of health, and find that lagged health indicators can play a role in retirement behavior. They do not compare different possible measures of health dynamics, especially they do not use self-reported health changes.

what we will call longitudinal validity).

2.1 Cross-Sectional Validity of Self-Reported Health Measures

The traditional literature on the (cross-sectional) validity of self-reported health measures stems, in part, from the fact that there is an increasingly large number of surveys that ask many questions about individuals' self-assessment of, for example, their health, or labor market opportunities. While in general these questions are regarded as very useful, they still raise a host of potential problems. To date, there seems to be little agreement as to the (cross-sectional) validity of such measures, even if they are used pervasively in empirical work.

As shown in Table 1 below, using all available waves of the Health and Retirement Survey, self-reported health is valid in the sense of its correlation with a host of variables that are measuring health status in both subjective and objectives ways. There is a strong monotonic relationship between the categorical health indicator and the different measures we present, and the correlations are quite high. People with better self-reported health status expect to live longer, have less chronic diseases (defined here as an index between 0 and 1, with increments of 1/7 for each of 7 illnesses), have less difficulties in performing activities of daily living (defined here as an index between 0 and 1, with increments of 1/23 for each of 23 indicators of activities and instrumental activities of daily living), and visited the doctor much less in the year before the interview. Some of the measures presented in the columns can be problematic. First, all of them are self-reported. Second, doctor visits can be considered a measure of health care utilization and therefore its correlation with any contemporaneous health measure should not necessarily be interpreted as evidence of the validity of the health indicator. Third, it could be argued that the index of chronic diseases, even if in most questions individuals are asked to refer to a previous diagnose by a professional, is also proxying for health care seeking behavior. Fortunately, even if the product of self-reports, the index of ADLs and IADLs is probably the cleanest measure, and the one with the highest correlation with the self-assessed health measure. In any case, all the measures point in the same direction, leaving little doubt about the information content of the self-reported health measure.

Table 1: Self-Reported Health Status and Related Measures

	Pliv75		Pliv85		Index of Chronic Disease		Index of ADL		Doctor Visits	
Health Status	Mean	Sd.	Mean	Sd.	Mean	Sd.	Mean	Sd.	Mean	Sd.
Excellent Health	0.779	0.228	0.608	0.293	0.102	0.123	0.050	0.079	4.28	6.17
Very good Health	0.713	0.242	0.522	0.295	0.169	0.142	0.083	0.103	5.90	7.97
Good Health	0.636	0.276	0.443	0.309	0.238	0.161	0.142	0.145	8.41	12.53
Fair Health	0.524	0.320	0.353	0.325	0.326	0.175	0.265	0.197	13.30	20.54
Poor Health	0.378	0.351	0.249	0.319	0.408	0.191	0.443	0.237	22.05	33.63
	Correlation Coefficients with Self-Reported Health									
	0.3611		0.3044		-0.4923		-0.5689		-0.2811	

However, the fact that self-reports seem to be meaningful does not imply we can blindly use them in our estimations without considering whether these measures should pass further tests before including them with confidence in our models. One important concern is about the potential endogeneity of these measures relative to the issue under study. Many previous researchers have suggested that self-assessed health measures might exaggerate individuals' actual health due to the tendency of respondents to use health problems as a convenient rationalization for difficulties in the labor market, putting in doubt the validity of these measures.[3] For example, in studies of the retirement decision, if the respondent's self-reported health status is merely a rationalization of the labor force decision (e.g., reporting being in poor health if they withdrew from the labor force), then unobserved factors affecting the labor supply decision will also affect self-reported health status. This implies that self-reported health is endogenous (i.e., correlated with unobservable factors affecting labor market decisions), biasing the coefficients of interest. Consequently, the large significant estimates of the impact of self-reported health may not indicate that this is a good measure of true health status, but merely a noisy measure of the dependent variable.

[3] For extended discussions of this "justification hypothesis" see Lambrinos (1981), Myers (1982), Parson (1982), Bazzoli (1985), Anderson and Burkhauser (1985), Stern (1989), Bound (1991), Kerkhofs and Lindeboom (1995), Blau et al. (1997), Kreider (1999), Bound et al. (1999), Kerkhofs et al. (1999), O'Donnell (1998), and Dwyer and Mitchell (1999).

Other researchers (e.g. Johnson (1977), Bazzoli (1985), and Bound et al. (1995)) criticize the use of self-reported health limitation in regression models of labor market participation, since health, measured as a condition limiting work, can be considered as an endogenous regressor, or even the same measure as the dependent variable. Hence, this measure may simply imply a tautological relationship between self-reported health and the retirement decision. Dwyer and Mitchell (1999) argue that additional problems can arise from the fact that subjective health measures may actually be assessments of leisure preferences, rather than true indicators of health status. That is, people who enjoy work tend to downplay health problems and postpone applying for DI benefits, while those who dislike work tend to apply soon after the onset of a sufficiently severe medical condition.

There is a substantial literature that provides some evidence in favor of these types of biases. Parson (1982) instruments self-reported health measures with future mortality and finds evidence supporting the justification hypothesis. He concludes that the use of self-reported health will cause significant biases in the coefficients of economic variables. Similar conclusions, using similar methods, were also reached by Anderson and Burkhauser (1985). However, Bound et al. (1998) have criticized the use of mortality as an instrumental variable, after finding evidence of endogeneity of mortality in these models stemming from measurement error. Bazzoli (1985) finds that self-reported health status affects the retirement decision differently depending on whether the variable is measured before or after the decision in question takes place. For example, self-reported health seems to have more significant effect when reported after retirement, lending support to the justification hypothesis. However, the time elapsed between the two measurements of health, which can be up to two years, may account for most of the difference. Finally, Blau et al. (1997) also find evidence of endogeneity of self-reported health in a retirement model, using the dummy variable indicating whether a person is in poor health.

In contrast to the studies reported above, there are many other studies which have found little evidence, or no evidence at all, of endogeneity in self-reported disability measures. Stern (1989) finds very weak evidence against the exogeneity of self-reported measures of disability in the labor force participation decision. Using data from the Netherlands, Kerkhofs et al. (1999) find some evidence of endogeneity of self-reported health limitation in the retirement decision, but little evidence in the disability application decision. Using the Health and Retirement Study (HRS) data, Dwyer and Mitchell (1999) conclude that self-rated health measures (including self-reported work limitations) are not endogenously determined with labor supply. Furthermore, they find no evidence to support the justification hypothesis. Using expectations indicators of labor supply, McGarry (2004) does not find much support for the justification hypothesis. Bound (1989) also notes that: "when outside information on the validity of self-reported measures of health is incorporated into the model, estimates suggest that the self-reported measure of health perform better than have been believed."

Another concern about the reliability of self-reported health status might stem from measurement error due to misreporting of respondents. However, the hypothesis that individuals systematically misreport their health and disability status in an anonymous, confidential survey does not seem highly plausible to us. Specifically, we found a high degree of internal consistency in responses to questions across the various sections of the HRS survey.[4] For this to be consistent with systematic misreporting, the respondents had to tightly coordinate their misreporting with other more "objective" reports, such as beginning and ending dates of jobs, dates of application, receipt of DI benefits, etc. If we were to believe that respondents are sophisticated enough to systematically misreport information in such a coordinated, internally consistent manner, we must question virtually all of their survey responses, including all "objective" health and functional status indicators. However, the literature rarely questions the validity of the "objective" health status measures.[5]

For the purposes of our study, it is hard to argue that health reports are used to rationalize expected longevity reports, especially because the longevity expectations questions were asked in the section after the health questions. However, self-reported health could still be problematic for a number of reasons explained in the next two subsections, which supports our empirical strategy of comparing different measures of health flows to better understand the consequences of using these

[4] See Benítez-Silva et al. (2004) for a discussion, as well as a study of the bias in self-reported disability.

[5] Bound (1991) is an exception to this sweeping statement. He argues that part of the problem is that the objective health variables measure health, rather than work capacity. Bound also notes that misreporting of variables tends to have counteracting effects. Baker et al. (2004) also question objective measures using a dataset that matches (objective) self-reports with medical records, they find evidence of considerable response error that might lead to attenuation bias in empirical work.

indicators in a variety of empirical models.

2.2 Longitudinal Validity of Self-Reported Health Measures

The economic literature has paid relatively little attention to the issue of how to measure health changes, or to analyze whether this type of variables belong in our models. Even if it is widely recognized that health is a dynamic concept, in most cases researchers have focused on measuring the stock of health. As we discussed in the introduction, some researchers have argued that self-reported health status by itself has a dynamic component, but it has not been shown in any convincing fashion. In the cases in which researchers have tried to measure health dynamics, they have either limited themselves to including lagged values of the self-reports in their specifications, or have included a particular measure without discussing the consequences of using different measures to capture health trajectories and health changes, or their properties. In other disciplines, however, measuring health changes is at least as important as measuring the level of health. This is mainly because they are concerned with measuring the consequences of some type of intervention that can affect a person's health.

A literature in the medical, and medical care fields, which has especially focused on studies of quality of life indicators, argues that the ability of a measure to detect changes (which is a characteristic defined as *responsiveness* or *sensitivity to change*), is an intrinsic part of the general validity of any self-reported measure of health. Therefore, it is concluded that researchers cannot evaluate the validity of a measure without paying attention to its ability to detect changes. The discussions in Guyatt et al. (1989), Hays and Hadorn (1992), Liang (2000), Erickson (2000), Patrick and Chiang (2000), Epstein (2000), and more recently in Terwee et al. (2003) are quite illuminating. They extend the classic assessments of the general criteria that a good health outcome measure must have, which is discussed in McDowell and Newell (1987), Streiner and Norman (1989), and Bowling (1991). Terwee et al. (2003) emphasize a very important distinction regarding the link between validity and responsiveness. They argue that many researchers have failed to distinguish between cross-sectional validity and longitudinal validity of the measures they use, with only the latter concept being linked to responsiveness. Notice that self-reported health status indicators, in the poor to excellent scale, seem to clearly be cross-sectionally valid but their longitudinal validity is still an open issue.

2.3 Dynamic Health Status: heterogeneous reporting, cut-point shifts, peer effects, and information loss

As argued in the last subsection, to understand how changes in self-reported health can be linked to changes in other measures of health is an essential part of an analysis of the validity of self-reported health. Furthermore, we need to understand whether reported changes in health are measures of health trajectories, and whether these trajectories can influence how individuals report a variety of useful and important variables.

In the HRS (and in many other surveys), self-rated health was asked as follows: *"Would you say your health is excellent, very good, good, fair, or poor?"* It is uncontroversial to state that true health status is a continuous (latent) variable, which when asked to individuals, is discretized in a particular way (in this case into five categories). Two general arguments can be given against the idea that health stocks (and its lags) alone should be used in empirical models that try to control for the dynamics of health. First, self-reported health status may be affected by unobserved heterogeneity. Second, using its changes over time to proxy for health flows could lead to information loss. We analyze each of these possibilities in turn:

Heterogeneity in reporting health status across individuals and over-time: gray areas, cut-point shifts, and peer-effects

Forcing individuals to discretize their reported health status can result in different self-reported health levels for individuals with the same latent health status. It could also lead to different self-reported health levels at different points in time for the same individual, due to the lack of clear cut differences between the ad-hoc health categories proposed to the respondents.

The effect of this lack of distinction between categories can result in a situation where even if an individual's real health did not change over time, her(his) self-rated health might change.[6] Figure 1, panels (a) to (d), illustrates this situation. Panel (a) in the upper left-hand corner, represents individuals who are able to make clear distinctions between those five health categories (bins) when reporting their health status. There is of course no reason to believe that the cut-points would be the same for everyone, but for the moment, and for this type of individual, we assume that they would be the same for a given individual over time.

Panel (b) in Figure 1, exemplifies the case in which even though individuals can clearly distinguish different health categories, the cut-points might be shifting over time for a given individual. This might be due to changing perceptions of what particular health levels mean, due to changes in a reference group, or even changes in their own understanding of their health. These shifts, can again be different for different individuals and overtime, which would suggest the possible existence of time-variant unobserved heterogeneity in reporting.

Panel (c) in Figure 1, represents individuals who do not have clear cut distinctions between each two consecutive health categories. Instead, these respondents have what we call *gray areas*, like the one represented by the distance between $H_3^{'}$ to H_3. The idea is that if their health status falls within these areas, individuals cannot clearly categorize it, and their actual reports could be a function of the context, and hence seem arbitrary to the econometrician. Again, there is no reason to believe that these gray areas would be the same for everyone. In fact, individuals could be heterogeneous in terms of the size of these areas.

Finally, panel (d) in Figure 1, depicts the case in which the gray areas might be changing over time for a given individual, due, in part, to the same reasons as in the shifts in panel (b); reference groups effects or informational changes with respect to their own health. This case can be understood as combining both time-invariant unobserved heterogeneity in reporting, and time-variant

[6]Crossley and Kennedy (2000) using repeated self-reported health questions (in the same scale as the one we use in this study) only separated by a short period of time during an interview process, found that almost 30% of the individuals actually changed their self-reports. This instability of the self-reported measure is a source of concern, especially for any study trying to use lag self-reports or changes of these self-reports in any empirical model. Their explanations of the findings are linked with the way the questions were administered, and the type of questions asked in between the two instants. It is hard to assess whether we would expect more or less stability if instead of a few minutes we wait literally a couple of years to ask the question again, but in any case this is another reason to be cautious when using self-reported health as a measure of health dynamics.

unobserved heterogeneity. As long as the shifts are not too large, and partially correlated with observables, an estimator that captures time-invariant unobserved heterogeneity might be subject to relatively small biases.

There are a number of researchers who have suggested in different applications the existence of cut-points shifts in assessing self-reported health. However, these authors have only discussed it in the context of differences across individuals, or in individuals' responses across countries, but not in the context of responses by the same individual over time, and they have not discussed the extension to the gray areas that we are presenting. The latter is key, since it allows us to reformulate the econometric problem as arising from unobserved heterogeneity. The terminology researchers have used has varied, but Kerkhofs and Lindeboom (1995), Groot (2000), Sadana et al. (2000), Murray et al. (2001), van Doorslaer and Jones (2003), and Lindeboom and van Doorslaer (2003) provide evidence of this possibility, and consider it a problem of measurement error.

As explained above, if we assume the gray areas are stable over-time, an estimation strategy that can account for time-invariant unobserved heterogeneity, can be successful in consistently estimating the effects of health changes, so we would think of the measurement error problem as having its source in unobserved heterogeneity in reporting due to the existence of these areas. However, there could still be cut-point shifts of different magnitudes, or gray area changes over time for a given person. The latter would be difficult to account for within any estimation strategy.

Notice that in all figures both types of individuals can easily identify which health category their health belongs to if their health falls into the white areas in the bins. Individuals will categorize their health as: (1). poor health, if the real health is worse than H_1 (or the corresponding label in the other figures); (2). fair health, if it is between H_1 and H_2; (3). good health, if it is between H_2 and H_3; (4). very good health, if it is between H_3 and H_4; and (5). excellent health, if it is better than H_4.

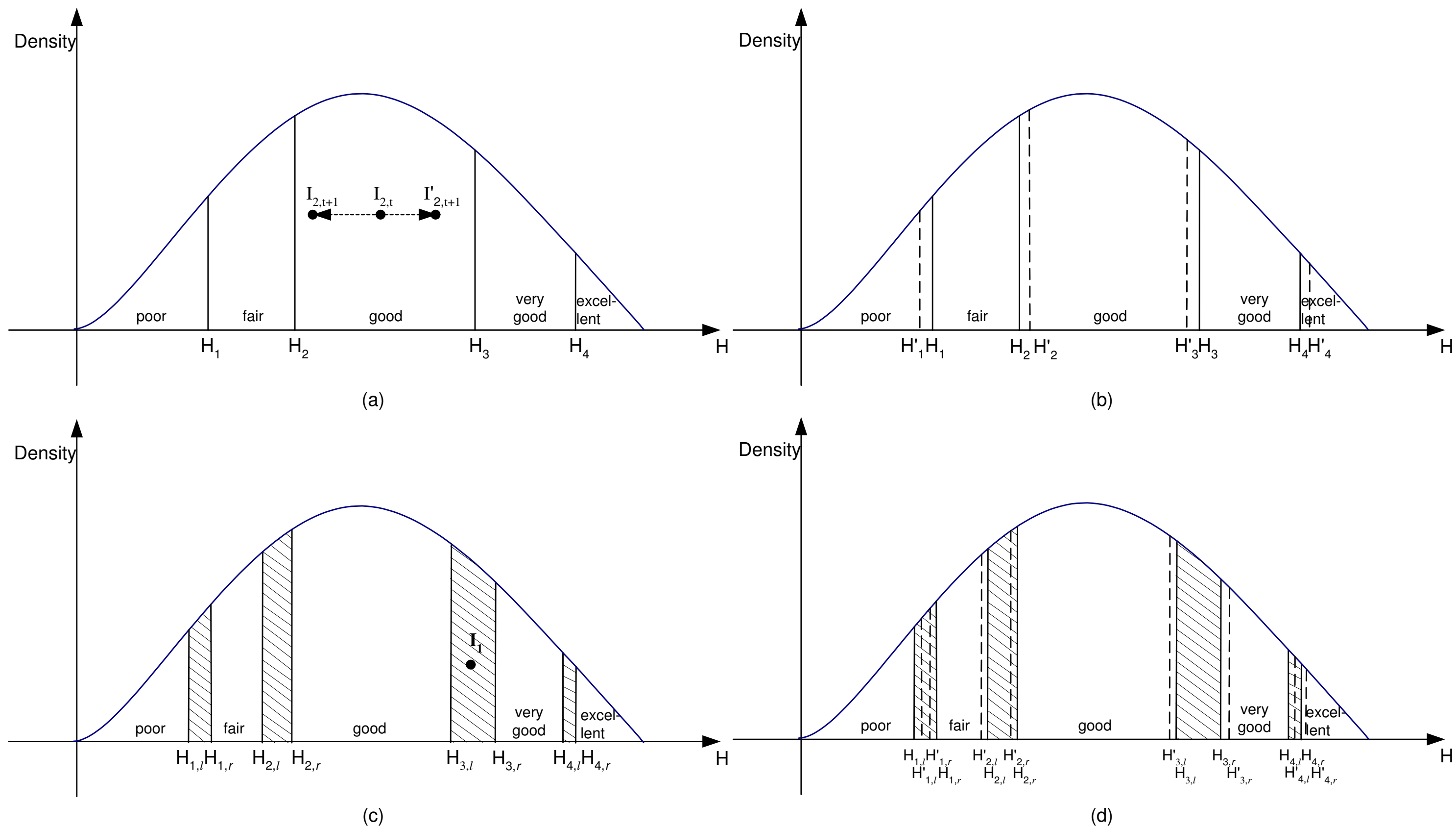

Figure 1: Heterogeneous Health Reports

Individuals in panel (a) and panel (b) would have little difficulty to differentiate their health between two consecutive categories, while in panel (b) the reports might change even if actual health does not change. However, individuals in panel (c) and (d) would find it rather difficult to decide what their health is if their health falls within one of those gray regions. Moreover, the difficulties increase with the thickness of those gray areas.[7] Hence, measurement error arises among those types of individuals. These areas are theoretical constructs, but they are a plausible explanations of the inconsistencies we will describe between the evolution of self-reported health measures and self-reported health flows, and the fact that there is no reason to believe that the cut-points will be the same for different people or will stay the same over time.

For example, let's assume individual I_1 is of the type portrayed by the figure in panel (c); her real health is the same in both period t and $t+1$, which falls in the gray area between $H_3^{'}$ to H_3. It is hard for her to decide whether she has good health or very good health. She could report either one with a positive probability, such that the following situation may happen. In period t, she reports her health as very good, but in period $t+1$, she reports her health as good. In this case, as econometricians, we observe that her health gets worse over time, while her real health has been I_1 all along. One of the possible reasons why she reported this way could be that in the first period she felt more optimistic than in the next period, so the measure is contextual.

Another explanation for these apparently inconsistent reports, and for the thickness of these gray areas, relies on peer effects. Self-reported health status is likely to be the result of an introspection that requires individuals to compare themselves either with a reference group or with themselves at previous points in time.[8] In the former case it is plausible that when reporting health status individuals are making a comparison with people of a similar age who they interact with or know about. This means that someone might report that they are in good health, while what they really

[7]By "thickness", we mean the horizontal distance between those two consecutive points at the edges of the gray areas, such as $H_3^{'}$ and H_3.

[8]The theory of reference groups has a long tradition in Sociology starting with Hyman (1942) and Kelley (1947), and more recently with some empirical applications, for example in Bank et al. (1990). In Economics the concept has been studied in a number of contexts (for example Pollak (1976) presents it as working through preferences), but rarely in the context of effects on reporting behavior. Manski (1993) provides an illuminating discussion of the difficulties that researchers can encounter to identify these types of effects in any empirical setting, and Alessie and Kapteyn (1991) and Kapteyn et al. (1997) present some empirical applications.

mean is that they are in good health in comparison with their particular reference group. In the following period, depending on the evolution of their health with respect to the reference group, they might report their health as being better or worse than before, even if there has been no change in their actual health, or even if their true health has evolved in the opposite direction. If these peer effects are present, it will be especially problematic to use changes in self-reported health status as measures of flows. Notice that this peer effects can affect the self-reports themselves, and could also affect the cut-points which could be shifting as a result of peer effects. Moreover, they could affect the width of the gray areas we have described, and even the possible changes of these areas over time.

In contrast, self-reported health changes were asked to individuals providing them with a clear reference point. Respondents were asked: *"Compared with your health when we talked with you in the previous wave (interview month-year), would you say that your health is better now, about the same, or worse?"* In this case, measurement errors and possible biases due to inconsistent replies over time as in the self-reported health status, are likely to be lower. The wording of the question forces individuals to provide a comparison with their own health at a different point in time, which minimizes the peer effects described above, and avoids the issue of cut-point shifts, since the assessment of the health change is not category-specific. Intuitively, however, we would expect this variable to be exposed to measurement error problems of its own, given the likelihood of observing recall problems among respondents. The latter is essentially equivalent to having people use different reference points in time when faced with this question. This is not a big problem as long as the measure actually captures the direction of health and its effects on contemporaneous reports of other variables and decisions.

Interestingly, in other surveys, most prominently in the National Health Interview Survey (NHIS) prior to 1997, as discussed, for example, in Cutler and Richardson (1997) and Cutler and Richardson (1998), and the British Household Panel Survey (BHPS), as described in Groot (2000) and Contoyannis et al. (2004), the self-reported health question was directly asked proposing individuals to compare their health with that of people their own age.[9] This indicates that any use of differences in

[9]Since 1997 the NHIS asks the self-reported health question without reference to an age group, and also asks the

this measure to proxy for health changes would for sure capture changes in their health as relative to their peers. In fact, Eriksson et al. (2001) compare self-reported health measures like the one in the HRS with another one similar to the one used in the BHPS, and conclude that measures which do not suggest a comparison group are more appropriate for longitudinal studies. They find that comparative measures lead individuals to overestimate their health in relation to others with increasing age. In the HRS, although the question does not provide any reference point, it is likely that some of these peer effects would still be present in the reporting by individuals. Therefore, if some of the individuals use this comparison when reporting their health, we are likely to see this overestimation in their responses, suggesting possible conflicting estimates for the variables measuring computed health changes and self-reported health changes.

Notice that we are expressing a concern regarding the existence of unobserved heterogeneity in reporting by individuals, with possibly many types of respondents, including individuals with heterogeneous distributions of the size of the gray areas in panels (c) and (d), individuals with clear and fixed cut-points as in panel (a), and individuals with shifting cut-points over time as in panel (b). Therefore, it will be critical in our estimations to account for this unobserved heterogeneity, and possibly also for the likely event that the unobserved components are correlated with the very reports of health and health changes. Properly accounting for these concerns will guide our empirical methodology.

Information loss

Even if individuals do not change the standards they use to evaluate their health, we may still lose information if we restrict attention to differences of health status to measure health flows. This could also be explained by turning to panel (a) in Figure 1. Assume individual I_2 has real health $I_{2,t}$ in period t, his health in period $t+1$ could be either $I_{2,t+1}$ or $I^{'}_{2,t+1}$. Although his self-rated health is good in both periods, unless we use a different measure of health flows, we would have thrown away information regarding his health change.

It is difficult to assess how large this information loss could be, given that the range of each of the self-reported health bins can vary widely in the population. But it is easy to see that it is something

self-reported change question.

empirical researchers should worry about. In fact, researchers in the medical field have described this type of information loss when analyzing how to assess change in health among severely ill patients. Bindman et al. (1990) present a case study in which the use of self-rated measures over time to assess changes in health, misses important information when dealing with people in very poor health. This is the case because once some individuals are in this lowest possible category they will keep reporting it even though they might actually be in worse health over time. These researchers call this effect the *floor phenomenon*, which is also described in Gold et al. (1996), and Testa and Simonson (1996). Baker et al. (1997), elaborating on this phenomenon, show that this loss of information can vary in size depending on the level of baseline health, with larger losses of information among people reporting worse health status. The self-reported measures in the HRS will naturally suffer from similar problems. Furthermore, since it can also happen in every possible health category, we can be facing a very important source of information loss, which might depend on how wide these categories are for given individuals, and the persistence of these health states among HRS respondents.

Given the likely presence of gray areas of reporting, cut-point shifts, inconsistencies, and loss of information, it is natural to argue in favor of a direct measure of health flows, like the self-reported health changes, which can minimize some of these problems, and better capture the dynamic effect of health.

3 The Data: Descriptive Analysis

The Health and Retirement Study is a nationally representative longitudinal survey of 7,700 households as of the first wave of interviews, headed by an individual aged 51 to 61 as of 1992-93. The primary purpose of the HRS is to study the labor force transitions between work and retirement with particular emphasis on sources of retirement income and health care needs. It is a survey conducted by the Survey Research Center (SRC) at the University of Michigan and funded by the National Institute on Aging.[10] Up to now, data from the first six waves of the survey are available. The last

[10]See Juster and Suzman (1995), also Gustman et al. (1994) and Gustman et al. (1995), or the HRS web page.

five waves of the data were conducted by phone using the computer assisted technology (CATI) which allows for much better control of the skip patterns and reduces recall errors.

Death and attrition have reduced the number of participants in the survey, and additional individuals, who have been included in our empirical work, have entered the survey later on mainly as spouses of previous respondents. In the empirical work we use all six available waves of the HRS. We construct a set of consistent variables on different sources of income, financial and non-financial wealth, health, and socio-demographic characteristics which will be assigned to each decision maker appropriately. The HRS collects information on both subjective and objective health measures on health stocks, such as self-rated health, chronic diseases, Activities of Daily Living (ADLs), and Instrumental Activities of Daily Living (IADLs).

3.1 Variables of Interest

The two main dependent variables in our empirical models are questions about individuals' subjective survival probability of living up to a certain age. First, "(What is the percent chance) that you will live to be 85 or more?" was asked to whoever is not older than age 75. And second, "(What is the percent chance) that you will live to be 75 or more?" was asked to those who are not older than 65.[11] Responses to these two questions are taken as measures of individuals' expected longevity.[12]

Regarding health variables, the self-rated health is a categorical variable which takes values from 1 to 5, where 1 represents poor health, 2 fair health, 3 good health, 4 very good health, and 5 excellent health. We also use an index for chronic disease, which incorporates information on

[11]These probability questions were asked in percentages in all waves except in wave 1, where respondents were asked to answer with a number between zero and ten. We have transformed those percentages or numbers to a value between zero and one, which can be interpreted as probabilities.

[12]The fact that a fairly large number of respondents replied to these questions giving potential focal points answers; either a zero probability (around 6% of the sample for the living to 75 question, but over 15% for the living to 85 question) a 50% probability (around 20% for both the living to 75 and living to 85 questions), or a 100% probability (also around 20% for the living to 75 question, but only around 10% for the living to 85 expectation), has been interpreted (see Hill et al. (2005)) as possibly indicating lack of information. This is rather controversial, and in any case it mainly creates the problem of non-normal regression errors, since the distribution of the dependent variable could be considered multi-modal. In general the results of conditional moments estimation, for example OLS or panel data regression models, are fairly robust to this problem, especially if the sample size is large. The most important properties of the linear estimators (that they are the best linear unbiased estimators and consistent, and that the variance estimator is unbiased and consistent allowing us to use conventional tests), survive the non-normality of the errors. However, there can be a loss of efficiency.

seven diseases; namely, high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, and arthritis. The chronic disease index is 1 if the individual has all the diseases, and 0 if he has none of them. The higher value of the chronic disease index has, the more chronic diseases he has. Notice that each chronic disease contributes $1/7$ to the index. Similarly, we have generated an index for ADLs and IADLs, which includes information on whether the individual has problems performing 23 daily activities. This measure takes the value 1 if the individual has difficulty in doing all these 23 activities, and 0 if he has no difficulty performing any of these 23, again each ADL or IADL contributes $1/23$ to the index. Functional limitation is measured as a binary indicator for whether the individual has any health problems which limit his ability to work. We also constructed an indicator for psychological problems, an indicator of self-reported memory ability, and an indicator of self-reported changes in memory ability.

The two main health flow variables we have generated are, first, the direct health flow measure, which is individuals' self-reports of their health changes compared to last interview. It is 1 if the individual reported his health is better, 0 if his health stays the same, and -1 if his health declines. Second, the computed health flow measure is an indicator of whether individuals' health gets better (a value of 1), same (0), or worse (-1), based on the calculated difference between the reports on his contemporaneous health and lagged health. These two flow measures are defined in such a way that their effects are comparable in our empirical models. Given the importance of these variables in our analysis, the next subsection goes into some detail in terms of what we could learn from analyzing, in an unconditional descriptive framework, the distribution of these indicators.

In order to incorporate them in our model of changes in the probabilities of living to certain ages, additional measures of health changes between each two consecutive waves have also been incorporated in the study. The computed changes in the chronic disease index, and that for the ADL index are calculated as the difference of these two indexes over waves. Binary indicators of new occurrence of working limitation in between waves, as well as changes in health insurance coverage have been constructed. Living habits such as smoking and drinking, and their changes between two consecutive waves have been measured as binary indicators. Notice that the latter are rather different from the rest of the variables, since in most cases the changing in these habits are the result of decisions by the individuals, and in most cases in the direction of quitting, which at these ages is likely to be linked to other health problems. This suggest that these variables might be collinear with other indicators of health changes, resulting in a loss of efficiency of the estimates. Additionally, a binary indicator for having no health insurance has been constructed based on the information on individuals' reported health insurance coverage status.

We have also included in our analysis information on individuals' demographic characteristics. Measures of age, gender, marital status, race and education have been included. Education is measured by the number of years the individual received formal education. Marital status takes the value one if the individual is married, 0 otherwise.

Parents', siblings' and kids' information has been collected in the HRS, which allows us to generate a series of variables that are likely to capture a number of possibly unobserved features correlated with longevity and expected longevity. Binary indicators for whether the individuals' mothers (fathers) lived up to age 75 and age 85 are constructed. Parents' education have been measured by the years of formal education they received. These six measures are mainly used to capture genetic effects on expected longevities. We also use a measure on number of living siblings from the supplement data provided by RAND which is constructed based on the HRS. Additionally, we construct an indicator for whether respondents have children living within 10 miles. The number of grandchildren is also used in our analysis.

Finally, variables on economic status and labor supply are constructed. Household net wealth measured in 10,000 dollars, which has been adjusted to dollars of 1992, includes savings and checking accounts, bonds, stocks, real estates, business owned, IRAs, trust, housing, and other assets net of debts. Individual income is taken from the RAND data which is a combination of labor income and social benefits. Self-reported retirement status is used as well, and it will be included as an exclusion restriction in the reporting equation since we believe it is likely to influence whether individuals have thought about their longevity.

3.2 Health flows

In the Health and Retirement Study, we are able to measure health flows in three different ways: self-reported health changes, computed health changes of self-rated health status, and onset of diseases. There are reasons, as discussed in Section 2, to believe that self-reported health changes could be the best measure among the three. Using the onset of diseases seems like a reasonable strategy too, however, this measure is not observed until a particular illness is diagnosed. This means that if we do not observe the onset of a disease it does not necessarily mean an individual is in better health. On the other hand, observing the onset of a disease is not necessarily a bad thing, it could mean that the disease has been recognized and under control. In a given period, individuals who have experienced the onset of a disease could be in an even better health status than those who have not been diagnosed, because the latter could actually have the disease but have not done much about it, since they do not know it and therefore are not treating it.

Table 2: Computed health changes and self-reported health changes

	SR-Health Changes			
Computed Health Changes	better	same	worse	Total
better	2,414	9,033	3,337	14,784
same	3,310	20,865	10,590	34,765
worse	1,288	9,111	7,752	18,151
Total Observations	7,012	39,009	21,679	67,700

Table 2 shows the inconsistency between the computed health changes and the direct reports on health changes. Individuals in the diagonal (in the oblique ellipse) are respondents with consistent reports, since their computed health changes are in line with their self-reported health changes. Also, those individuals whose computed health changes are zeros (in the horizontal ellipse), that is, their self-rated health falls into the same category as in the previous period, could be still considered as respondents with consistent reports. However, using only the computed health changes could lead to a loss of information, as explained in the previous subsection. Notice that this group comprises a bit more than 20% of our observations, which should raise a flag of caution for researchers relying solely on self-reported health status and its lags without exploiting the self-reported changes.

On the other hand, those with inconsistent reports are the individuals in the left bottom cells and right top cells, who are not in the two ellipses. These respondents with inconsistent reports account for 33.63% of the sample. The reason for these inconsistencies could be explained by the presence of gray areas, cut-point shifts, and also the peer effects that could potentially affect the self-reports in health status by the respondents in the HRS. It is not possible for us to assess the relative weight of these mechanisms in explaining these inconsistencies. But notice from the table, that in line with Eriksson et al. (2001), the group reporting to be in better health is considerably larger if we use the computed health changes than if we use the self-reported health changes, suggesting the possible overestimation of health changes when using a measure more open to inter-personal comparisons. In any case, given that it is debatable which measure is better without controlling for additional characteristics, we should be very careful in interpreting the effects of computed health changes in any empirical model that hopes to properly control for health status and health dynamics, in the presence of these inconsistencies

Table 3: Computed and Self-Reported Health Changes and Related Measures

	Pliv75		Pliv85		Index of Chronic Disease		Index of ADL		Doctor Visits	
Health Changes	Mean	Sd.	Mean	Sd.	Mean	Sd.	Mean	Sd.	Mean	Sd.
Better Comp. Health	0.660	0.287	0.479	0.323	0.236	0.177	0.150	0.171	7.91	13.00
Worse Comp. Health	0.638	0.294	0.447	0.321	0.237	0.178	0.178	0.193	9.77	17.40
Better Self-R Health	0.694	0.275	0.517	0.317	0.240	0.175	0.134	0.158	9.46	15.14
Worse Self-R Health	0.600	0.308	0.394	0.319	0.278	0.196	0.248	0.226	11.6	20.58

Table 3 shows that self-reported health changes seem to be better indicators of health trajectories on a variety of subjective and objective measures of health status than computed health changes. Notice that for all the measures, the difference between reporting better and worse health is much larger for the self-reported measure, meaning that the effect of the different trajectories is more explicit for this measure than for the computed measure. These tabulations gives us some initial evidence on the likely effects of these two variables in the conditional moments estimations we discuss in Section 5.

Table 4: Loss of Information and Longitudinal Validity

	Changes in Pliv75		Changes in Pliv85		Changes in Index of Chron. Dis.		Changes in Index of ADL		Changes in Doctor Visits	
Health Changes	Mean	Sd.	Mean	Sd.	Mean	Sd.	Mean	Sd.	Mean	Sd.
Comp-same;SR-better	0.004	0.26	0.019	0.29	0.016	0.08	-0.028	0.11	0.134	16.3
Comp-same;SR-worse	-0.011	0.28	-0.007	0.29	0.028	0.09	-0.019	0.14	2.026	19.9
	Correlation Coefficients with Computed Health Changes									
	0.0647		0.0718		-0.1020		-0.146		-0.081	
	Correlation Coefficients with Self-Reported Health Changes									
	0.0384		0.0588		-0.059		-0.046		-0.064	

Table 4 tackles the issue of loss of information and longitudinal validity. In those cases where computed self-reported health indicates that there has been no health change, but we observe changes in the self-reported health change variable, we can see that the variables capturing changes in objective and subjective measures of health status as more consistent with the self-reported measure than with the computed one. However, the evidence in these unconditional summary statistics is rather weak. The correlation coefficients (which in all cases are statistically different from zero) indicate that computed health changes capture meaningful health changes in objective and subjective indicators at least as effectively as self-reported health changes. From this unconditional evidence, we would conclude that self-reported health status seem to be longitudinally valid since it captures change in a significant way.

Table 5: Inconsistencies Between Computed and Self-Reported Health Changes

	Changes in Pliv75		Changes in Pliv85		Changes in Index of Chron. Dis.		Changes in Index of ADL		Changes in Doctor Visits	
Health Changes	Mean	Sd.	Mean	Sd.	Mean	Sd.	Mean	Sd.	Mean	Sd.
Comp-better;SR-better	0.036	0.27	0.065	0.30	0.014	0.08	-0.053	0.13	-1.99	23.6
Comp-better;SR-same	0.025	0.28	0.058	0.31	0.014	0.08	-0.021	0.11	-0.78	15.1
Comp-better;SR-worse	0.013	0.29	0.014	0.31	0.011	0.09	-0.051	0.15	0.62	14.8
Comp-worse;SR-better	-0.015	0.29	-0.011	0.31	0.028	0.09	-0.009	0.12	2.12	15.9
Comp-worse;SR-same	-0.019	0.28	-0.0004	0.30	0.028	0.08	0.011	0.12	2.04	13.9
Comp-worse;SR-worse	-0.042	0.29	-0.034	0.32	0.050	0.11	0.036	0.17	4.82	21.2

Table 5 explores the inconsistencies between self-reported health changes and computed health changes which account for more than 30% of our sample. The table allows us to observe whether the changes in related health indicators are in line with the changes in computed health changes or self-reported health changes, when they are inconsistent with each other. The table not only indicates that computed health changes capture fairly well the changes in a number of objective and subjective indicators, but also show that conditional on a given computed health change, the self-reported health change adds information. For example, conditional on being in better health by the computed health indicator, being in worse health by the self-reported indicator is correlated with comparatively lower increases in longevity expectations, and comparatively higher number of doctor visits. On the other hand, conditional on being in worse health by the computed health measure, those in better self-reported health have comparatively smaller decreases in expected life expectancy, and smaller increases in the more objective measures of health problems like ADLs and doctor visits. This evidence suggests that in the models that try to capture the effects of new information on changes in longevity expectation (which we will present and estimate in the next sections) we probably want to incorporate both types of measures since they seem to be providing complementary information.

Table 6: Summary Statistics

Variable	Obs	Mean	Sd.
Pliv75	22,361	0.662	0.279
Pliv85	22,218	0.454	0.312
Health			
Self-Reported Health Status	24,455	3.442	1.111
Better Computed Health Change	24,455	0.211	0.408
Same Computed Health Change	24,455	0.532	0.499
Worse Computed Health Change	24,455	0.257	0.437
Better Self-Reported Health Change	24,455	0.117	0.322
Same Self-Reported Health Change	24,455	0.521	0.500
Worse Self-Reported Health Change	24,455	0.362	0.480
Index of Chronic Disease	24,455	0.186	0.162
Index of ADLs and IADLs	24,455	0.130	0.154
Health Limitation	24,455	0.226	0.419
Psychological Problems	24,300	0.140	0.347
Number of doctor visits	24,455	8.007	14.389
Number of days hospitalized	24,455	0.300	1.476
Smoker	24,455	0.207	0.405
Drinker	24,455	0.543	0.498
No Health Insurance	24,455	0.180	0.384
Cognitive Ability			
Self-reported memory ability	23,176	3.224	0.953
Computed memory ability change	22,858	-0.119	0.707
Self-reported memory change	23,166	-0.112	0.416
Demographic Characteristics			
Age	24,455	57.946	4.982
Age squared	24,455	33.826	5.510
Male	24,455	0.401	0.490
Years of Education	24,455	12.674	2.880
Married	24,455	0.811	0.392
White	24,455	0.857	0.350
Economic Status			
Total net wealth (in $10,000)	24,455	35.223	-134.172
Total income (in $10,000)	24,455	2.149	2.880
Family Background			
Indicator of father lived up to 75	24,455	0.140	0.347
Indicator of mother lived up to 75	24,455	0.187	0.390
Indicator of father lived up to 85	24,455	0.458	0.498
Indicator of mother lived up to 85	24,455	0.601	0.490
Mother’s years of education	24,455	9.676	3.483
Father’s years of education	24,455	9.333	3.843
Indicator of kids living close	24,455	0.513	0.500
Number of grandkids	24,455	3.774	4.750
Number of siblings	24,455	2.934	2.417

3.3 Sample and sample restrictions

We use all six available waves of the HRS. But given the skip pattern of the longevity expectations variables we only include in our sample age eligible respondents who were asked the questions about their survival probabilities. Furthermore, those with missing information on their parents' mortality have been excluded from our sample, since these variables play an important role in our empirical analysis. We believe it is important to account for these genetically related health information in our empirical work. Additionally, individuals who died between our sample period (from 1992 to 2003) have automatically dropped from our sample after they died.

Some previous literature (see Hurd and McGarry (1995), Hurd and McGarry (2002)) have considered individuals with a higher probability of living up to age 85 than that of age 75 as lacking the capability of probabilistic thinking. However, it is not clear to us whether individuals thought about these two survival probability questions simultaneously and then reported them, or, they first thought of their probabilities of living up to age 75 and then conditional on surviving by age 75, they reported their probabilities of living up to age 85. Given the sequence of how these two probability questions were asked (individuals were firstly asked about their subjective survival probability of living up to age 75, and then the probability of living up to age 85), it is possible that some of them understood the second question as conditional on surviving to age 75. In fact, if the answer is conditional we can observe both higher and lower responses, therefore, it is hard to identify the percentage of individuals that might have understood the question in the conditional sense. We have decided to keep those people with higher probability of living up to age 85 in the sample, which represents only about 2% of the sample. In any case, further sensitivity analysis by eliminating those people does not change our empirical results in any significant way.

Table 7: Percentage of age eligible individuals who responded to the two subjective survival probability questions

	w1	w2	w3	w4	w5	w6
Pliv75	93.09	87.92	88.83	82.93	86.15	86.33
Pliv85	92.82	87.74	87.62	57.86	83.93	82.51

As a result, the sample used in the estimations includes more than 22,000 observations from more than 8,000 individuals. Summary statistics are presented in Table 6. 30% of the sample are male, and 87% are whites. The sample is 57 years old on average, with 12.70 years of education. 81.19% are married. More than 90% reported these two subjective survival probabilities as of the first wave of interviews, and we can observe in Table 7 how this percentage has evolved over time.

There are reasons to suspect that sample selectivity might be an issue in this study, given that we are only able to observe individuals' subjective survival probabilities for those who answered the questions. In the sample, we do observe people who are age eligible to answer these two probability questions did not give any answer. The response rates in the six waves are reported as below:

In Table A.1. in the appendix, we present summary statistics after dividing the sample between individuals who reported expected longevity and those who did not, for comparison. We observe that respondents with better health status (both better self-rated health, and less chronic diseases, or less functional limitations) are more likely to report; However, individuals with parents who died earlier than age 75 or 85 are more likely to report. This probably indicates that those events are likely to be correlated with having thought about longevity issues; Higher educated individuals, and whites, are more likely to report. Also, parents' education have a positive correlation with the likelihood of reporting. Females and those with health insurance coverage are more likely to report too; Finally, richer respondents are less likely to report their expected longevities.

4 Econometric Specifications

Our empirical strategy is to consistently estimate the effect of a variety of regressors, most prominently a number of health measures, on individual longevity expectations that are measured by the self-reported probabilities of living to age 75 and 85. To reach this goal, a few econometric questions need to be tackled. First of all, we want to control for observed heterogeneity. This is not a trivial task since omitted variable biases can be an important problem relevant to the reports of expected longevity. Second, we need to account for possible sample selection biases, since a non-trivial proportion of eligible individuals did not answer the longevity questions. It is reasonable to believe that respondents who answered the questions might not be a random sample from the population of interest, and those who are more likely to expect to live longer would also be more likely to respond. We will perform this correction with the standard techniques following Heckman (1979) and the suggestions by Hsiao (1986). Third, we want to take into account the unobserved heterogeneity potentially present in our characterization of the econometric model, which we suspect might have a lot to do with the way individuals report their health. If we do not control for the unobserved components we will be confounding partial and total effects of our variables of interest. Panel data sets allow us to model explicitly how those unobserved components enter the econometric specification, where we can choose to include them as a fixed effect or as a random variable and then test the different specifications.

We first estimate in a given cross-section or the pooled sample of observations, controlling for clustering (see Deaton (1997)), the following model

$$Pliv75_i = \alpha_1 + X_i'\beta + H_i'\lambda_1 + u_{1i}, \tag{1}$$

and

$$R_i = I\ (\alpha_2 + Z_i'\gamma + H_i'\lambda_2 + u_{2i} > 0), \tag{2}$$

where in (1) the set of individual characteristics X_i consists of various socio-economic and demographic variables, and H_i is a vector of health variables which include measures of health stocks and health flows. $Pliv75_i$ (or $Pliv85_i$ if that is the model we are estimating) represents the measure of expected longevity reported by individuals in the sample.

In (2), R_i is an indicator of whether the individual responded the expected longevity question, where it takes the value 1 when the $Pliv75 \geq 0$. Notice that u_{1i} and u_{2i} are not independent, and Z_i can be equal to X_i but will include additional variables in the empirical application, which will help in identifying the model. In this setting, the selection rule is potentially not independent of the behavioral function being estimated. It is fairly straightforward to estimate the full model by Maximum Likelihood or by standard Two-Step procedures.

The HRS provides us with repeated observations of the same individuals, and this allows us to control for potential unobserved components that could enter our econometric model. Our main equation of interest can now be written as

$$Pliv75_{it} = \alpha_1 + X_{it}'\beta_1 + H_{it}'\lambda + \mu_i + \nu_{it}, \tag{3}$$

where μ_i represents the unobserved heterogeneity component, and the ν_{it} are the idiosyncratic disturbances. This model can be estimated either assuming no correlation between observed explanatory variables and the unobserved effect (random effects), or allowing for arbitrary correlation between the unobserved effect and the observed explanatory variables (fixed effects). We can then test whether the random effects specification or the fixed effect specification is more appropriate, and whether the former is more appropriate than the pooled OLS regression.[13] In our case, even though we are also interested in estimating the effect of a number of time invariant regressors both on the selection equation and the equation of interest, the possibility of allowing for correlation between the unobserved component and the regressors, makes the fixed effect the most appropriate estimator. This decision will be supported by the specifications tests.[14]

We have, therefore, an unbalanced panel of respondents, where individuals are selected into the sample, or non-randomly drop from the sample that answered the expected longevity questions. To take this selection problem into account we can extend the classic Heckman (1979) sample selection correction approach where the first stage is a probit specification with no individual component and the second stage accounts for the individual component. This is suggested by Hsiao (1986), and it is in principle fairly straightforward. In the empirical results, we will present the estimates of both uncorrected and corrected regressions, with and without unobserved heterogeneity. The qualitative results of our empirical work do not change much from specification to specification,

[13]See Wooldridge (2002) and Baltagi (2005) for up to date and illuminating presentations of these issues.

[14]It has shown that, under some fairly realistic assumptions, the fixed effect model can exacerbate a possible measurement error problem. This means that a possible attenuation bias of the OLS or RE estimator (if measurement error was a problem to start with) might grow even larger under the fixed effect formulation. The common way to solve the problem is to construct some type of Instrumental Variable estimator. We have perform extensive sensitivity analysis of our model using Panel Data instrumental variables techniques and found no evidence of these problems with our FE specification. For example, a model that compares our preferred FE specification with the same specification but instrumenting self-reported health with its lagged, rejected the panel IV specification in favor of the traditional FE specification. Also, we have experimented with the Hausman and Taylor (1981) estimator which circumvents the possible worsening of the measurement error problem by using a variation of the techniques utilized in the estimation of instrumental variable models for panel data, and the results do not change in any significant way.

but the predicted effects of some of the variables do change in non-trivial ways once we move towards the consistent estimators of the parameters of interest.

The econometric specifications we have described so far only exploit the fact that individuals responded repeatedly to the same longevity expectations questions, but do not directly focus on why these reports might change from period to period. An additional econometric analysis of interest tries to directly explain how these reports have changed over time. With this objective in mind, we can estimate a model of changes in longevity expectations as a function of changes in a variety of exogenous regressors, including changes in the health variables of interest. We can think of this alternative model as trying to capture the source of the innovations to the self-reported longevity expectations. We follow Benítez-Silva and Dwyer (2005) in order to set up the appropriate econometric model to capture the role of new information. One conclusion from that work is that individuals integrate new and old information with weights attached to each of these sources when updating their expectations.

Assuming the linearity of the function that individuals use to process information regarding the variables of interest, and further assuming that the elements of the information set (Ω) used to formulate the expectations are jointly normally distributed, we can then write

$$Pliv75_{t+1,i} = Pliv75_{t,i} + \gamma(\omega_{t+1} - E[\omega_{t+1}|\Omega_t]), \tag{4}$$

where the term in parenthesis is the difference between the observed new information (ω_{t+1} will be represented in our empirical estimation by the changes in the exogenous characteristics) and the expectation of this information individuals had in the previous period. If we had all these elements, it would be possible to estimate (4) by a conditional moments estimator like OLS, and would expect a coefficient of one in front of the longevity expectations as of time t under the rational expectations hypothesis. Also, the coefficients in γ (which we will assume are time invariant) could be interpreted as the effect of the unanticipated components of new information on changes in expectations. However, we do not observe the second element in the brackets. Although it could be estimated by making further assumptions about the relationship between new and current information, this is likely to be a very noisy procedure. Instead, we estimate the following equation

$$Pliv75_{t+1,i} = \alpha + \beta Pliv75_{t,i} + \gamma \omega_{t+1,i} + \epsilon_{t+1,i}, \tag{5}$$

where there is no reason to believe that β should be equal to 1: since the unobserved expectation term from (4) would enter into the error term, leading to possible biases in the coefficients of interest.[15] Also, the estimates of the vector γ, are interpreted as the effects that changes in specific factors of Ω have on longevity expectations formation. A nonzero coefficient of a given change in a member of the information set can be interpreted as information not perfectly anticipated, and therefore not embedded in the individuals' longevity expectations as of period t.

It is also very common in a variety of literatures to use a formulation that implicitly assumes rational expectations, and proceeds by subtracting the self-reports of a given year from the self-reports of the following year. Given that Benítez-Silva and Dwyer (2005), and Benítez-Silva and Dwyer (2006) have shown that rationality cannot be rejected among this sample of HRS respondents regarding retirement expectations, we are fairly confident of the validity of the alternative specification we present below. Also, we will see that our own results from estimating (5) cannot reject the utilization of the specification shown below, at least on statistical grounds. The changes in expected survival probabilities between periods can be formulated as:

$$\Delta Pliv75 = \alpha + \Delta H_i \beta_1 + \Delta X_i \beta_2 + \epsilon_i, \tag{6}$$

where ΔH_i represents health changes between two particular periods, and ΔX_i are controls for time varying individual characteristics.

[15]We will treat these possible biases as a problem of measurement error, which can be ameliorated using IV techniques.

5 Empirical Results

5.1 Estimations of Expected Longevity

In Tables 8 to 10 we provide the results of estimating both cross-sectional and panel data expected longevity probability models with and without sample selection corrections. In all cases all the key variables, and most of the other regressors in the estimation are significant and have the intuitive signs. The results across specifications vary quantitatively, but they all tell essentially the same story in qualitative terms.[16] Our main result is that both health stocks and health flows significantly affect self-reported expected longevity, and most strikingly, which measure of health flows is used matters a great deal. In fact, we show that the self-reported health change is a more appropriate measure of health flows, compared with the computed health flow which uses the differences in self-reported health status. The self-reported change indicator for being in better health than in the previous period has a positive sign in the estimations, indicating that being in better health is correlated with an increase in the probabilities of living to age 75 and age 85. This is true even after controlling for the self-reported health level, which is also positively correlated with the probabilities of living to those ages.

On the other hand, and rather surprisingly for us at first, the computed health change indicator for better health has a very small negative (and in our preferred specification, insignificant) effect on the self-reported longevity measures. We have a number of possible interpretations for this result, with all of them suggesting caution to researchers in the use of self-reported health measures in longitudinal studies, and especially the use of computed self-reported health changes as proxies for health flows. In line with our discussion in section 2, we believe that the computed changes in self-reported health are very likely to be exposed to considerably more problems of unobserved heterogeneity in reporting due to the fact that many individuals might have a difficult time in distin-

[16] As discussed in Section 2, some of the variables included in the estimation can be considered either measures of health care utilization (like the number of doctor visits and hospitalizations, and even the index of chronic diseases) and therefore highly collinear with contemporaneous health indicators, or likely to be endogenous like an indicator for whether respondents have children living within 10 miles. We have estimated all the specifications we report below dropping these variables from the model, and the results do not change in any significant way. In particular the coefficients on the self-reported health measures are virtually unchanged.

guishing exactly between the imposed cut-points for the health levels in a consistent way over the period of analysis, leading to wide gray areas (in the terminology we have introduced). This version of the cut-point shifts problem that is described in some detail in the literature on cross-country comparisons of health, comes from the fact that there is nothing in the questionnaire suggesting to individuals that they should use the same reporting model as they did in the previous interviews, or even use the same cut-points between health levels. The more pervasive this problem is, the more likely that changes in computed self-reported health could capture the importance of the gray-areas problem (or the problem of shifting cut-points) that we presented in Figure 1. If that is the case, once we control for health levels, computed changes in health are more likely to be an expression of the difficulty to pinpoint their own health level, which could be correlated with lower expected life expectancies. This is true even after controlling for education and other socio-demographic measures, and in most cases its effects disappear once we account for unobserved heterogeneity that allows for correlation between the individual effect and the regressors of interest. These cut-points shifts or changing gray areas seem to be correlated with events we are not quite able to control for, and therefore they could be proxying for unobserved factors correlated with lower expected life expectancy.

A companion interpretation (and a fairly plausible one) is linked to the peer effects discussed in section 2. If many individuals report their health using a reference group as a comparison basis, then changes in health status might be more of an expression of the evolution of health compared with that reference group and less of an expression of their own health. And the latter is the one that we believe is more likely to be correlated with expected longevity. If computed improvements in self-reported health capture a *better than your peers* measure of health, there is no reason to expect them to have a significant positive correlation with expected longevity, which is more likely to be mainly driven by actual health, regardless of interpersonal comparisons. In fact, it is easy to believe that in general the information flow from your peers, among people this age, is more likely to be related to bad news about health events of people you know or your peers know. If individuals mostly receive bad news about other people's health, it is not surprising that they would tend to think they are doing better on a health measure open to interpersonal comparisons.

The negative and insignificant effect of the computed self-reported health improvement indicator, contrasts with the more intuitive, positive and significant effect of the self-reported health improvement variable. Remember that the latter is the result of a question that gives individuals a reference point by asking them to compare their current health to their own health two years prior. Even though we might expect quite a lot of recall errors on these responses, the formulation of the question basically assures that individuals will perform an intra-personal comparison when reporting their health change. Whether they take the exact health level two years ago (a level not given to them by the interviewer) as a reference point, or if they respond with respect to some other point in a more recent past, is not essential to our arguments. The fact is that this report is consistent with their answers regarding expected longevity in a completely different section of the survey.

Overall, the results do indicate that self-reported health levels in a cross-sectional sense, have the intuitive positive correlation with the self-reported expected longevity. However, the problems arise when using differences in health stocks over time as measures of health flows. Although we believe the differences in health reports over time do provide some information to researchers, they are more likely to be proxying for unobserved heterogeneity in reporting than in assessments of health. Therefore, it is more appropriate to use the self-reported health changes as measures of health flows to avoid the problems with the computed measure.

Table 8 shows the results of different possible econometric models of the longevity expectation without correcting for sample selection biases; equations (1) and (3). As it is the case in all the estimations we will show, the preferred econometric model is the Fixed Effect specification.[17] The presence of an individual specific unobserved heterogeneity component is supported by the standard

[17]One drawback of using the Fixed Effect estimator is that the effects of interesting variables like gender, race, and education cannot be estimated. For this we can analyze the results of the Random Effects estimator, or perform some sensitivity analysis with slightly different types of panel data models like those suggested in Hausman and Taylor (1981). As expected being a male is correlated with significantly lower probabilities of living to 75, other things equal (also if we perform an unconditional test, females report a significantly higher probability of living to 75), but rather unexpectedly being white reduces the self-reported probability of living to age 75 (in an unconditional test, whites do not report significantly different longevity expectations), maybe indicating that members of other races are relatively more optimistic in their reports, other things equal. On the other hand, higher levels of education are correlated with higher expected longevity. We have also performed some sensitivity analysis by dividing the sample by gender and re-estimating our preferred specification. The results do not change qualitatively, but the quantitative effects of most health variables (both subjective and objective measures) are larger for males, and some variables, like the cognition indicator (better cognition is predicted to increase expected longevity for females) or the net wealth measure (wealthier females report, other things equal, lower expected longevity) are only significant for females.

econometric tests, and further tests show that it is not appropriate to assume that this unobserved heterogeneity is uncorrelated with the variables of interest. The Fixed Effect estimator captures the type of heterogeneity we have been describing in relation to reporting. In order for the Fixed Effect estimates to be unbiased, we have to believe that even if the reporting of health might be varying over time due to the gray areas problem, each individual has a particular pattern in terms of the areas in which they distinguish the health categories. The Fixed Effect estimator is able to account for time-invariant unobserved heterogeneity, but it cannot remove any remaining time-variant biases created by omitted variables.

The use of the Fixed Effect estimator in Table 8, and in the rest of the tables, has two main effects on our results. First, it decreases considerably the effect of self-reported health status. For example, it reduces the effect of going from one health category to the next one (say from good to very good) on the average expected probability of living to age 75 by two thirds. In Table 8, this means that instead of increasing the average expected probability by about 10% as estimated in the OLS specification, or almost 7% as in the Random Effect specification, it increases it by only 3% in the Fixed Effect model. This means that not accounting for unobserved heterogeneity which correlates with the exogenous variables of interest can lead to serious biases and misleading inference. Second, it shows that computed health improvements have a very small negative and insignificant effect on the probability of living to age 75, instead of the much larger (and significant) negative effect estimated in the OLS specification. Notice that the coefficient on this binary indicator is reduced 15-fold from the OLS specification to the Fixed Effect estimator, and almost 12-fold from the Random Effects estimator to the Fixed Effect estimator. Even if the coefficient is not very precisely estimated, it is very close to zero, and it becomes almost the smallest coefficient in the table. However, the Fixed Effect estimator does not affect the coefficient of the self-reported health change much. The coefficient is reduced, but it is still fairly large, similar in magnitude to the effect of the health level, and very precisely estimated. The coefficient predicts that the better self-reported health indicator increases the average expected probability of living to age 75 by almost 3%. Table 9 shows the same specifications but correcting for possible selection bias due to lack of response to the longevity questions, therefore they jointly estimate equations (1) and (2), or (3) and

(2), depending on the column. The results do not vary much, and in the Fixed Effect specification the presence of selection bias is not supported.[18]

Table 10 provides some sensitivity analysis of different specifications, using subsets of the health indicators.[19] We only provide the results of the Corrected Fixed Effect estimator, which is our preferred specification and the one supported by our specification tests. We can see that the effect of Self-Reported health status is very stable, and increases the average expected probability of living to age 75 by around 3%. The effect of the self-reported health change is also stable and of similar magnitude. The last two columns are especially interesting to analyze since they estimate a model with lagged health status, which follows a fairly standard literature that uses this kind of specification to account for the dynamics of health. Notice that irrespective of whether we include the better self-reported health change indicator, the lagged health categorical variable is estimated to have a significant but very small effect on expected longevity. The effect is four times smaller than the effect of current health status, but more importantly, almost four times smaller than the effect of the better self-reported health indicator as well, in the last specification. The conclusion

[18]An additional source of selection into the sample used for estimation, which can result in biased coefficients, is survivorship bias. Given that our dependent variables are measures of longevity, it is natural to expect that those dying between the ways would be the ones reporting lower probabilities of surviving to certain ages, leaving our sample of survivors as a selected group of respondents that expect to leave longer. In fact, we have used the exit interview information from all available waves of the HRS to compare the expected survival probabilities of those we know survived to the next wave, with the expected probabilities of those we know die between waves (about 7% of the original wave 1 sample had died by wave 6). Our findings indicate that eventual survivors reported significantly higher probabilities of living to age 75 and 85, suggesting that survivorship bias could be an issue in our data. In order to assess the effects of this additional source of selection in our sample we have followed the recommendations in Portrait et al. (2004) to construct additional selection correction terms, and have included them in our preferred specifications. Even though in most cases these additional selection terms were significant, the effects on the results we report in this section were negligible. These results are available from the authors upon request. One possible problem of actually controlling for this additional selection concern, is that the correction terms we have constructed to account for this possible bias are likely to be correlated with the selection term regarding responses to the longevity questions. The reason for this concern is that those that ended up dying between waves were considerably less likely to report expected longevity probabilities. In fact, the difference in response rates is around 14 percentage points, and it is highly significant. If the probability to respond is serially correlated, it is likely that the unobserved components that make someone more likely to respond are correlated with the unobserved components that make that same individual more likely to survive. This would mean that the introduction of the additional correction terms would result in collinearity problems and additional biases of its own.

[19]We have also estimated the model allowing for serial correlation in the idiosyncratic component in the fixed effect model, and also have experimented with the Hausman and Taylor (1981) estimator which allows only a subset of the regressors in the fixed effect estimation to be correlated with the unobserved individual component (in our case we have assumed the correlation was with the health reporting variables, and also with education). Both types of specifications provide essentially identical results to the ones presented in our preferred specifications. These results are available from the authors upon request.

from this is twofold. Lagged health has a significant effect on expected longevity, but it is very small, and fails to capture the dynamic effect of health picked up by the self-reported health change indicator. The coefficient on the latter indicator is unaffected by the inclusion of the lagged health indicator.

Using our preferred specification, Table A.2. in the Appendix provides results of a model of expected probability of living to age 85. The results are qualitatively identical to the ones just presented for the Pliv75 variable. One difference is that here selection does matter, and the effect of the lagged self-reported health is estimated to be smaller and in this case insignificant. As discussed in Hurd and McGarry (2002) and Hamermesh (1985), the Pliv85 seems to be less in line with the life tables than Pliv75, with explanations linked to possible lack of adjustments by sex when forming longer term expectations, or the possibility of individuals assuming past improvements in life expectancy will continue and affect them favorably. We believe that an additional explanation is linked to the argument we discussed earlier about the possibility of the responses to the Pliv85 questions being taken by some individuals to mean that Pliv85 are conditional on Pliv75, which would bias these responses upwards for those respondents, as found by previous researchers. Given this evidence, we have chosen to use the estimation of the Pliv75 model as our benchmark, even if the differences in estimation results with the Pliv85 estimation are not qualitatively important.

Table A.3. in the appendix shows probit estimates of the selection equation used in the different specifications we have shown. It is worth noticing that the main exclusion restriction we have used to better identify the sample selection corrected models is an indicator of whether someone self-reports herself as retired. From the results, it is clear that this indicator is correlated with whether they answer the longevity expectations questions, but some sensitivity analysis we have performed shows that it is unlikely to be correlated with the error term in the main equation, since its effect on the expected longevity variables is estimated to be insignificant.

Finally, notice that the estimates of the other health variables have the expected signs, and are in line with the previous literature estimating both expected and actual longevity. For example, having more chronic diseases or more ADLs decreases the expected longevity probabilities, and the same is true for those that visited the doctor more often. On the other hand, those in better cognitive health (measured here by memory ability), expect to live longer, and the same it is true for married, wealthier individuals, and also for higher earners.

5.2 Changes in Expected Longevity

Table 11 estimates the model presented in (5) trying to capture the effect of new information on the changes in expected longevities. Following the recommendations of Benítez-Silva and Dwyer (2005) we use a sample-selection corrected Instrumental Variables procedure, without assuming rational expectations, in which variables with a significant effect would capture what has not been anticipated by individuals. Therefore, it is not surprising to see in the table that the computed health change indicator is significant in this context, since it can capture innovations either regarding the comparisons between the evolution of the person's health with that of their peers, or regarding cut-point shifts, which are more likely to come as a surprise than the information on their own health. The computed health change does play a significant role in this specification, which confirms the argument that although it is clearly a problematic measure to use to proxy for health dynamics, it indicates that the self-reported health status is longitudinally valid, since it captures unanticipated changes, even in a conditional moments estimation.

Notice that not too many variables are significant in this specification, in fact in the preferred specifications in columns 3 and 4, the only consistently significant variables are the self-reported health change indicators, the changes in performing activities of daily living, and the change in the number of grandchildren. Notice that, in line with the findings by Benítez-Silva and Dwyer (2005) on a very different context, the results seem to be consistent with the Rational Expectations Hypothesis. However, as those authors emphasize, this specification is not strictly speaking a rational expectations test, due to the informational requirements it imposes, assuming that individuals are forming expectations over the changes in the variables affecting the dependent variable. Column four, actually uses the fact that the coefficient on the lagged expected longevity is not significantly different from 1 to estimate the more common specification that has the difference in expected longevity probabilities as dependent variable, as presented in (6). The results are very similar. The main difference is that the effect of the self-reported health change indicator is now considerable larger, which we interpret as indicating that although not statistically incorrect, assuming that the coefficient on lagged Pliv75 is 1 can create some biases in the coefficients of variables that are correlated with the lagged expectation.

Finally, Table A.4. in the Appendix shows the very same specifications as Table 11 but using the Pliv85 as the dependent variable. The results do not change in any significant way, except for the fact that some additional variables become significant. For example, the recent death of the mother has a negative effect on expected longevity, and having better memory ability has a positive effect.

6 Conclusions

This is one of the first papers to directly study the use of subjective measures of health flows to capture the dynamic effect of health on both individuals' expected survival probabilities and how they evolve over time. Our findings are relevant to any behavioral or empirical study using self-reported health measures.

In our sample, we observe that for a large proportion of the sample, the difference between the self-reported health in two consecutive waves is not consistent with the self-reported health changes between the same two waves. This fosters our discussion that allows us to provide possible explanations for these differences, and at the same time illuminates the estimation results that suggest caution to researchers using self-reported health and computed self-reported health changes in many empirical applications. Our findings provide support for the use of self-reported health changes, whenever possible, to capture the dynamic component of health, as a complement to self-reported health status.

Moreover, we compare the effect of the two measures of health changes on the changes in the expected survival probabilities. Our results show that after controlling for the change of self-reported health and the change of other objective health measures, there is evidence that computed self-reported health changes are not perfectly anticipated and therefore have explanatory power in this kind of econometric specification. This suggests that self-reported health measures are longitudinally valid. However, the explanation for the reason behind their significance in these types of econometric models is likely to be linked with possible time-variant unobserved heterogeneity regarding reporting behavior of health status.

We hope our findings foster further research on the longitudinal properties of self-reported measures in household surveys, the dynamics of health, and the measures of health flows in general. Most of the efforts in the profession have concentrated in the cross-sectional properties of this type of variable. However, with the growing number of panel data sets that provide self-reported measures of a variety of variables, there is a pressing need to understand the pros and cons of using these measures in longitudinal econometric models. In this paper, we have paid particular attention to self-reported health given the wide use of these measures in just about any empirical application using household level data. However, our findings could be of use in many other contexts where self-reports are used, and the dynamics of self-reported measures are considered to be important.

References

Alessie, R. and A. Kapteyn (1991). Habit formation, interdependent preferences and demographic effects in the almost ideal demand system. *Economic Journal 101*(406), 404–419.

Anderson, K. H. and R. V. Burkhauser (1985). The retirement-health nexus: A new measure of an old puzzle. *Journal of Human Resources 20*, 315–330.

Baker, D., R. D. Hays, and R. H. Brook (1997). Understanding changes in health status: Is the floor phenomenon merely the last step of the staircase? *Medical Care 35*(1), 1–15.

Baker, M., M. Stabile, and D. Catherine (2004). What do self-reported, objective, measures of health measure? *Journal of Human Resources 39*(4), 1067–1093.

Baltagi, B. H. (2005). Econometric analysis of panel data. Third Edition. John Wiley and Sons, Ltd.

Bank, B. J., R. L. Slavings, and B. J. Biddle (1990). Effects of peer, faculty, and parental influences on students' persistence. *Sociology of Education 63*(3), 208–225.

Bazzoli, G. (1985). The early retirement decision: New empirical evidence on the influence of health. *Journal of Human Resources 20*, 214–234.

Benítez-Silva, H., M. Buchinsky, S. Cheidvasser, H. M. Chan, and J. Rust (2004). How large is the bias in self-reported disability? *The Journal of Applied Econometrics 19*(6), 649–670.

Benítez-Silva, H. and D. S. Dwyer (2005). The rationality of retirement expectations and the role of new information. *The Review of Economics and Statistics 87*(3), 587–592.

Benítez-Silva, H. and D. S. Dwyer (2006). Expectation formation of older married couples and the rational expectations hypothesis. *Labour Economics forthcoming*.

Bindman, A. B., D. Keane, and N. Lurie (1990). Measuring health changes among severely ill patients. the floor phenomenon. *Medical Care 28*(12), 1142–1152.

Blau, D. M., D. Gilleskie, and C. Slusher (1997). The effect of health on employment transitions of older men. manuscript.

Bound, J. (1989). Self-reported versus objective measures of health in retirement models. NBER Working Paper No. 2997.

Bound, J. (1991). Self-reported versus objective measures of health in retirement models. *Journal of Human Resources 26*(1), 106–138.

Bound, J., M. Schoenbaum, T. R. Stinebrickner, and T. Waidmann (1998). Modeling the effect of health on retirement behavior. manuscript.

Bound, J., M. Schoenbaum, T. R. Stinebrickner, and T. Waidmann (1999). The dynamic effects of health on the labor force transitions of older workers. *Labour Economics 6*(1), 179–202.

Bound, J., M. Schoenbaum, and T. Waidmann (1995). Race and education differences in disability status and labor force attachment. *Journal of Human Resources 30-S*, 227–267.

Bowling, A. (1991). Measuring health:a review of quality of life measurement scales. Philadelphia: Open University Press.

Burström, B. and P. Fredlund (2001). Self-rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes? *Journal of Epidemiology and Community Health 55*, 836–840.

Contoyannis, P., A. M. Jones, and N. Rice (2004). The dynamics of health in the british household panel survey. *Journal of Applied Econometrics 19*(4), 473–503.

Crossley, T. F. and S. Kennedy (2000). The stability of self-assessed health status. SEDAP Research Paper 26.

Cutler, D. M. and E. Richardson (1997). Measuring the health of the u.s. population. *Brookings Papers on Economic Activity. Microeconomics*, 217–282.

Cutler, D. M. and E. Richardson (1998). The value of health: 1970-1990. *American Economic Review: Papers and Proceedings 88*(2), 97–100.

Deaton, A. (1997). The analysis of household surveys. Baltimore: Johns Hopkins University Press.

Deeg, D., R. van Zonneveld, P. van der Maas, and J. Habbema (1989). Medical and social predictors of longevity in the elderly: Total predictive value and interdependence. *Social Science and Medicine 29*(11), 1271–1280.

Dwyer, D. and O. Mitchell (1999). Health problem as determinants of retirement: Are self-rated measures endogenous? *Journal of Health Economics 18*(2), 173–193.

Epstein, R. S. (2000). Responsiveness in quality-of-life assessment: Nomenclature, determinants, and clinical applications. *Medical Care 38*(9), II–91–II–94.

Erickson, P. (2000). Assessment of the evaluative properties of health status instruments. *Medical Care 38*(9), II–95–II–99.

Eriksson, I., A.-L. Undén, and S. Elofsson (2001). Self-rated health. comparisons between three different measures. results from a population study. *International Journal of Epidemiology 30*(2), 326–333.

Ferraro, K. F. and J. Kelley-Moore (2001). Self-rated health and mortality among black and white adults; examining the dynamic evaluation thesis. *Journal of Gerontology:Social Science 56B*(4), s195–s205.

Gan, L., G. Gong, M. Hurd, and D. Mcfadden (2004). Subjective mortality risk and bequests. NBER working paper No. 10789.

Gan, L., M. Hurd, and D. Mcfadden (2003). Individual subjective survival curves. NBER working paper No. 9480.

Gold, M., P. Franks, and P. Erickson (1996). Assessing the health of a nation: The predictive validity of a preference-based measure and self-rated health. *Medical Care 34*(2), 163–177.

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics 19*(3), 403–420.

Gustman, A. L., O. S. Mitchell, and T. L. Steinmeier (1994). Retirement research using the health and retirement survey. NBER Working Paper No. 4813.

Gustman, A. L., O. S. Mitchell, and T. L. Steinmeier (1995). Retirement measures in the health and retirement study. *Journal of Human Resources 30*, S57–S83.

Guyatt, G. H., R. A. Deyo, M. Charlson, M. N. Levine, and A. Mitchell (1989). Responsiveness and validity in health status measurement: A clarification. *Journal of Clinical Epidemiology 42*(5), 403–408.

Hamermesh, D. S. (1985). Expectation, life expectancy and economic behavior. *Quarterly Journal of Economics 100*(2), 389–408.

Hausman, J. A. and W. E. Taylor (1981). Panel data and unobservable individual effects. *Econometrica 49*(6), 1377–1398.

Hays, R. D. and D. Hadorn (1992). Responsiveness to change: an aspect of validity, not a separate dimension. *Quality of Life Research 1*(1), 73–75.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Hill, D., M. Perry, and R. J. Willis (2005). Estimating knightian uncertainty from survival probability questions on the hrs. manuscript, University of Michigan.

Hsiao, C. (1986). Analysis of panel data. Econometric Society Monographs. Cambridge University Press.

Hurd, M., D. Mcfadden, and A. Merrill (1999). Predictors of mortality among the elderly. NBER working paper No. 7440.

Hurd, M. and K. McGarry (1995). Evaluation of the subjective probabilities of survival in the health and retirement study. *Journal of Human Resources 30*, s7–s56.

Hurd, M. and K. McGarry (2002). The predictive validity of subjective probabilities of survival. *The Economic Journal 112*(482), 966–985.

Hyman, H. (1942). The psychology of status. *Archives of Psychology 269*.

Idler, E. L. and Y. Benyamini (1997). Self-rate health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior 38*(1), 21–37.

Johnson, W. (1977). The effect of disability on labor supply. *Industrial and Labor Relations Review 30*, 380–381.

Juster, F. and R. Suzman (1995). An overview of the health and retirement survey. *Journal of Human Resources 30*(Supplement), s7–s56.

Kapteyn, A., S. Van de Geer, H. Van de Stadt, and T. Wansbeek (1997). Interdependent preferences: An econometric analysis. *Journal of Applied Econometrics 12*(6), 665–686.

Kelley, H. (1947). Two functions of reference groups. In G. Swanson, T. Newcomb, and E. Hartley (Eds.), *Readings in Social Psychology*, pp. 410–414. Henry Holt and Co. New York.

Kerkhofs, M. and M. Lindeboom (1995). Subjective health measures and state dependent reporting errors. *Health Economics 4*, 221–235.

Kerkhofs, M., M. Lindeboom, and J. Theeuwes (1999). Retirement, financial incentives and health. *Labour Economics 6*, 203–227.

Kreider, B. (1999). Latent work disability and reporting bias. *Journal of Human Resources 20*, 734–769.

Lambrinos, J. (1981). Health: A source of bias in labor supply models. *Review of Economics and Statistics 63*(2), 206–212.

Lee, Y. (2000). The predictive value of self assessed general, physical, and mental health on functional decline and mortality in older adults. *Journal of Epidemiology and Community Health 54*, 123–129.

Liang, M. (2000). Longitudinal construct validity: Establishment of clinical meaning in patient evaluative instruments. *Medical Care 38*(9), II–84–II–90.

Lindeboom, M. and E. van Doorslaer (2003). Cut-point shift and index shift in self-reported health. Tinbergen Institute Discussion Paper 042/3.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies 60*(3), 531–542.

McDowell, I. and C. Newell (1987). Measuring health: A guide to rating scales and questionnaires. Oxford University Press.

McGarry, K. (2004). Health and retirement: Do changes in health affect retirement expectations? *Journal of Human Resources 39*(3), 624–648.

Mete, C. (2005). Predictors of elderly mortality:health status, socioeconomic characteristics and social determinants of health. *Health Economics 14*, 135–148.

Miilunpalo, S., I. Vuori, P. Oja, M. Pasanen, and H. Urponen (1997). Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *Journal of Clinical Epidemiology 50*(5), 517–528.

Mossey, J. and E. Shapiro (1982). Self rated health: a predictor of mortality among the elderly. *American Journal of Public Health 72*(8), 800–808.

Murray, C., A. Tandon, J. Salomon, and C. Mathers (2001). Enhancing cross-population comparability of survey results. GPE Dicussion Paper, 35. World Health Organization.

Myers, R. J. (1982). Why do people retire from work early? *Aging and Work 5*, 83–91.

O'Donnell, O. (1998). The effect of disability on employment allowing for work incapacity. Working Paper No. 98-13. University of Kent at Canterbury.

Parson, D. O. (1982). The male labor force participation decision: Health, reported health,and economic incentives. *Economica 49*, 81–91.

Patrick, D. L. and Y.-P. Chiang (2000). Measurement of health outcomes in treatment effectiveness evaluations. *Medical Care 38*(9), II–14–II–25.

Pollak, R. A. (1976). Interdependent preferences. *American Economic Review 66*(3), 309–320.

Portrait, F., R. Alessie, and D. Deeg (2004). Disentangling the age, period, and cohort effects using a modeling approach: An application to trends in functional limitations at older ages. Utrecht University, manuscript.

Sadana, R., C. Mathers, A. Lopez, C. Murray, and K. Iburg (2000). Comparative analysis of more than 50 household surveys on health status. GPE Dicussion Paper, 15. World Health Organization.

Siegel, M., E. H. Bradley, and S. V. Kasl (2003). Self-rated life expectancy as a predictor of mortality: evidence from the hrs and ahead surveys. *Gerontology 49*(4), 265–271.

Smith, V. K., D. H. Taylor, and F. A. Sloan (2001). Longevity expectations and death: Can people predict their own demise? *American Economic Review 91*(4), 1126–1134.

Stern, S. (1989). Measuring the effects of disability on labor force participation. *Journal of Human Resources 24*, 361–395.

Streiner, D. L. and G. R. Norman (1989). Health measurement scale: A practical guide to their development and use. Oxford University Press.

Terwee, C. B., F. W. Dekker, W. M. Wiersinga, M. F. Prummel, and P. M. M. Bossuyt (2003). On assessing responsiveness of health-related quality of life instruments: Guidelines for instrument validation. *Quality of Life Research 12*(4), 349–362.

Testa, M. A. and D. C. Simonson (1996). Assessment of quality-of-life outcomes. *New England Journal of Medicine 334*(13), 835–840.

van Doorslaer, E. and A. Jones (2003). Inequalities in self-reported health: validation of a new approach to measurement. *Journal of Health Economics 22*(1), 61–87.

Wooldridge, J. (2002). Econometric analysis of cross section and panel data. MIT Press.

Table 8: Estimation of Pliv75 without Sample Selection Correction[20]

Variable	Pooled OLS Coef. (S.E.)	Random Eff. Coef. (S.E.)	Fixed Eff. Coef (S.E.)
Self-Reported Health Status	0.060 (0.002)	0.045 (0.002)	0.020 (0.003)
Better Computed Health Change	-0.031 (0.004)	-0.024 (0.004)	-0.002 (0.004)
Better Self-R Health Change	0.030 (0.005)	0.022 (0.004)	0.018 (0.005)
Index of Chronic Disease	-0.089 (0.012)	-0.099 (0.014)	-0.056 (0.022)
Index of ADLs and IADLs	-0.122 (0.016)	-0.152 (0.016)	-0.125 (0.021)
Health Limitation	-0.006 (0.005)	-0.010 (0.005)	-0.010 (0.006)
Psychological problems	-0.010 (0.005)	-0.016 (0.006)	-0.016 (0.009)
Number of doctor visits	0.00002 (0.0001)	0.0002 (0.0001)	0.0003 (0.0001)
Number of days hospitalized	-0.002 (0.001)	-0.002 (0.001)	-0.003 (0.001)
Self-Reported Memory Ability	0.021 (0.002)	0.014 (0.002)	0.005 (0.002)
Smoker	-0.045 (0.004)	-0.037 (0.005)	0.002 (0.009)
Drinker	0.015 (0.003)	0.010 (0.004)	0.000 (0.005)
No Health Insurance	-0.006 (0.004)	-0.009 (0.004)	-0.010 (0.005)
Age	0.000 (0.004)	-0.002 (0.005)	0.002 (0.009)
Age squared	0.002 (0.004)	0.004 (0.004)	0.006 (0.007)
Male	-0.036 (0.004)	-0.038 (0.005)	-
Education	0.003 (0.001)	0.005 (0.001)	-
Married	-0.001 (0.004)	0.007 (0.005)	0.022 (0.009)
White	-0.057 (0.005)	-0.054 (0.007)	-
Total net wealth (in $10,000)	0.00002 (0.00001)	-0.000001 (0.00001)	-0.00002 (0.00001)
Total income (in $10,000)	-0.001 (0.001)	0.0003 (0.0006)	0.0002 (0.0008)
Indicator of father liv75	0.045 (0.003)	0.040 (0.004)	0.003 (0.018)
Indicator of mother liv75	0.045 (0.003)	0.038 (0.004)	0.007 (0.012)
Mother's education	0.004 (0.001)	0.004 (0.001)	-
Father's education	0.000 (0.001)	0.001 (0.001)	-
Indicator of kids living close	-0.010 (0.003)	-0.002 (0.003)	0.008 (0.004)
Number of grandkids	0.002 (0.000)	0.001 (0.000)	0.001 (0.001)
Number of siblings	0.001 (0.001)	0.001 (0.001)	-0.001 (0.002)
Wave Indicator	0.008 (0.001)	0.007 (0.001)	-0.009 (0.008)
Constant	0.287 (0.113)	0.400 (0.129)	0.265 (0.313)
Observations	25,173	25,173	25,173
R^2	0.175	0.1726	0.0754
Average Pliv75	0.664	0.664	0.664

[20]Breusch-Pagan Test: $\chi^2 = 1204.05$, p-value: 0.0000.
[20]Hausman Test: $\chi^2 = 286.94$, p-value: 0.0000.

Table 9: Estimation of Pliv75 with Sample Selection Correction[21]

Variable	Heckman Coef. (S.E.)	Corrected RE Coef. (S.E.)	Corrected FE Coef. (S.E.)
Self-Reported Health Status	0.057 (0.003)	0.044 (0.003)	0.021 (0.003)
Better Computed Health Change	-0.029 (0.004)	-0.023 (0.004)	-0.003 (0.004)
Better Self-Reported Health Change	0.029 (0.005)	0.021 (0.005)	0.018 (0.005)
Index of Chronic Disease	-0.095 (0.013)	-0.103 (0.014)	-0.055 (0.022)
Index of ADLs and IADLs	-0.120 (0.016)	-0.150 (0.016)	-0.125 (0.021)
Health Limitation	-0.008 (0.005)	-0.011 (0.005)	-0.009 (0.006)
Psychological problems	-0.011 (0.005)	-0.017 (0.006)	-0.016 (0.009)
Number of doctor visits	0.00001 (0.0001)	0.0002 (0.0001)	0.0003 (0.0001)
Number of days hospitalized	-0.002 (0.001)	-0.002 (0.001)	-0.003 (0.001)
Self-Reported Memory Ability	0.022 (0.002)	0.014 (0.002)	0.005 (0.002)
Smoker	-0.047 (0.004)	-0.038 (0.005)	0.002 (0.009)
Drinker	0.011 (0.004)	0.008 (0.004)	0.001 (0.006)
No Health Insurance	-0.005 (0.004)	-0.008 (0.004)	-0.011 (0.005)
Age	0.000 (0.004)	-0.003 (0.005)	0.003 (0.009)
Age squared	0.002 (0.004)	0.004 (0.004)	0.006 (0.007)
Male	-0.035 (0.004)	-0.037 (0.005)	-
Education	0.001 (0.001)	0.004 (0.001)	-
Married	-0.002 (0.004)	0.006 (0.005)	0.022 (0.009)
White	-0.067 (0.007)	-0.060 (0.008)	-
Total net wealth (in $10,000)	0.00003 (0.00001)	0.000004 (0.00001)	-0.00002 (0.00002)
Total income (in $10,000)	-0.0006 (0.0007)	-0.0003 (0.0007)	-0.0002 (0.0008)
Indicator of father liv75	0.046 (0.003)	0.041 (0.004)	0.003 (0.019)
Indicator of mother liv75	0.044 (0.003)	0.038 (0.004)	0.007 (0.012)
Mother's education	0.003 (0.001)	0.003 (0.001)	-
Father's education	0.000 (0.001)	0.001 (0.001)	-
Indicator of kids living close	-0.010 (0.003)	-0.003 (0.003)	0.008 (0.004)
Number of grandkids	0.002 (0.000)	0.001 (0.000)	0.001 (0.001)
Number of siblings	0.001 (0.001)	0.000 (0.001)	-0.001 (0.002)
Wave Indicator	0.011 (0.002)	0.008 (0.002)	-0.010 (0.008)
Constant	0.332 (0.116)	0.431 (0.131)	0.254 (0.314)
Inverse Mills' ratio	-0.132 (0.056)	-0.077 (0.057)	0.029 (0.074)
Observations	26,273	25,173	25,173
R^2	-	0.173	0.074
Average Pliv75	0.673	0.664	0.664

[21] Hausman Test: $\chi^2 = 297.70$, p-value: 0.0000.

Table 10: **Sensitivity Analysis of Pliv75 estimation using Corrected Fixed-Effect Model**

Variable	Corrected FE Coef. (S.E.)	Corrected FE Coef (S.E.)	Corrected FE Coef (S.E.)	Corrected FE Coef. (S.E.)	Corrected FE Coef (S.E.)
Self-Reported Health Status	0.020(0.003)	0.021(0.003)	0.020 (0.003)	0.021 (0.003)	0.020 (0.003)
Lagged S-R Health	-	-	-	0.004 (0.002)	0.005 (0.002)
Better Computed Health Change	-	-0.002 (0.004)	-	-	-
Better S-R Health Change	-	-	0.018 (0.005)	-	0.018 (0.005)
Index of Chronic Disease	-0.054 (0.022)	-0.052 (0.022)	-0.057 (0.022)	-0.049 (0.022)	-0.052 (0.022)
Index of ADLs and IADLs	-0.127 (0.021)	-0.128 (0.021)	-0.125 (0.021)	-0.127 (0.021)	-0.124 (0.021)
Health Limitation	-0.009 (0.006)	-0.009 (0.006)	-0.009 (0.006)	-0.009 (0.006)	-0.008 (0.006)
Psychological problems	-0.016 (0.009)	-0.016 (0.009)	-0.016 (0.009)	-0.015 (0.009)	-0.016 (0.009)
Number of doctor visits	-0.0003 (0.0001)	-0.0003 (0.0001)	-0.0003 (0.0001)	-0.0003 (0.0001)	-0.0003 (0.0001)
Number of days hospitalized	-0.003 (0.001)	-0.003 (0.001)	-0.003 (0.001)	-0.003 (0.001)	-0.003 (0.001)
S-R Memory Ability	0.005 (0.002)	0.005 (0.002)	0.005 (0.002)	0.005 (0.002)	0.005 (0.002)
Smoker	0.001 (0.009)	0.001 (0.009)	0.002 (0.009)	0.001 (0.009)	0.002 (0.009)
Drinker	0.001 (0.006)	0.001 (0.006)	0.001 (0.006)	0.001 (0.006)	0.001 (0.006)
No Health Insurance	-0.010 (0.005)	-0.010 (0.005)	-0.011 (0.005)	-0.011 (0.005)	-0.011 (0.005)
Age	0.003 (0.009)	0.002 (0.009)	0.003 (0.009)	0.002 (0.009)	0.003 (0.009)
Age squared	0.006 (0.007)	0.006 (0.007)	0.006 (0.007)	0.006 (0.007)	0.006 (0.007)
Married	0.022 (0.009)	0.022 (0.009)	0.022 (0.009)	0.022 (0.009)	0.022 (0.009)
Total net wealth (in $10,000)	-0.00002 (0.00001)	-0.00002 (0.0002)	-0.00002 (0.00001)	-0.00002 (0.00002)	-0.00002 (0.00002)
Total income (in $10,000)	-0.0003 (0.0008)	-0.0003 (0.0008)	-0.0002 (0.0008)	-0.0003 (0.008)	-0.0003 (0.0008)
Indicator of father liv75	0.004 (0.018)	0.004 (0.018)	0.004 (0.018)	0.004 (0.018)	0.003 (0.018)
Indicator of mother liv75	0.006 (0.012)	0.007 (0.012)	0.006 (0.012)	0.007 (0.012)	0.007 (0.012)
Indicator of kids living close	0.008 (0.004)	0.008 (0.004)	0.008 (0.004)	0.008 (0.004)	0.008 (0.004)
Number of grandkids	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Number of siblings	0.000 (0.002)	-0.009 (0.008)	-0.011 (0.008)	-0.009 (0.008)	-0.009 (0.008)
Wave Indicator	-0.010 (0.008)	-0.001 (0.002)	0.000 (0.002)	-0.001 (0.002)	-0.001 (0.002)
Constant	0.245 (0.314)	0.268 (0.314)	0.232 (0.314)	0.258 (0.314)	0.242 (0.314)
Inverse Mills' ratio	0.019 (0.074)	0.016 (0.074)	0.030 (0.074)	0.026 (0.074)	0.040 (0.075)
Observations	25,217	25,177	25,207	25,177	25,173
R^2	0.071	0.075	0.069	0.083	0.082
Predicted Pliv75	0.664	0.664	0.664	0.664	0.664

Table 11: Estimation of changes in Pliv75 with Sample Selection Correction[22]

Variable	Corrected IV Coef. (S.E.)	Corrected IV Coef. (S.E.)	Corrected IV Coef. (S.E.)	Corrected FE Coef. (S.E.)
Pliv75 in previous Wave	0.960 (0.038)**	0.954 (0.038)**	0.959 (0.038)**	-
Better Computed Health Change	0.027 (0.006)**	-	0.026 (0.006)**	0.022 (0.008)**
Better S-R Health Change	-	0.015 (0.007)**	0.012 (0.007)*	0.035 (0.012)**
Changes in Index of Chronic Disease	-0.041 (0.032)	-0.049 (0.032)	-0.043 (0.032)	-0.018 (0.043)
Changes in Index of ADLs and IADLs	-0.079 (0.025)**	-0.088 (0.025)**	-0.078 (0.025)**	-0.099 (0.037)**
Onset of a Health Limitation	-0.013 (0.010)	-0.013 (0.010)	-0.013 (0.010)	-0.010 (0.014)
Onset of Psychological problems	-0.004 (0.016)	-0.005 (0.016)	-0.004 (0.016)	-0.013 (0.023)
Start to Smoke	0.022 (0.022)	0.022 (0.022)	0.021 (0.022)	-0.044 (0.033)
Start to Drink	0.004 (0.010)	0.004 (0.010)	0.004 (0.010)	0.010 (0.016)
Loss of Health Insurance	-0.014 (0.007)*	-0.014 (0.007)*	-0.014 (0.007)*	-0.010 (0.011)
Change in doctor visits	-0.0002 (0.0002)	-0.0003 (0.0002)	-0.0003 (0.0002)	-0.0002 (0.0002)
Change in days hospitalized	-0.001 (0.001)	-0.002 (0.001)	-0.002 (0.001)	-0.002 (0.002)
Computed Memory Ability Change	0.005 (0.006)	-	0.005 (0.006)	-0.008 (0.008)
S-R Memory Ability Change	-	-0.011 (0.015)	-0.011 (0.015)	0.015 (0.023)
New widowhood	-0.001 (0.020)	0.002 (0.020)	-0.001 (0.020)	0.035 (0.029)
Changes in total net wealth (in $10,000)	-0.00003 (0.00001)	-0.00003 (0.00002)	-0.00003 (0.00002)	-0.0004 (0.00002)**
Changes in total income (in $10,000)	0.00003 (0.001)	-0.00003 (0.001)	0.00003 (0.001)	-0.0005 (0.002)
Recent father's death	0.004 (0.012)	0.004 (0.012)	0.004 (0.012)	0.016 (0.020)
Recent mother's death	-0.002 (0.011)	-0.001 (0.011)	-0.001 (0.011)	0.016 (0.016)
Indicator of kids moving close	0.005 (0.007)	0.005 (0.007)	0.005 (0.007)	0.012 (0.011)
Changes in number of grandkids	0.002 (0.001)**	0.002 (0.001)**	0.002 (0.001)**	0.002 (0.001)**
Changes in number of siblings	0.000 (0.003)	0.001 (0.003)	0.001 (0.003)	0.001 (0.004)
Constant	0.017 (0.028)	0.026 (0.028)	0.017 (0.028)	-0.057 (0.016)**
Inverse Mills' ratio	0.007 (0.033)	0.015 (0.033)	0.009 (0.033)	0.438 (0.149)**
Observations	13,135	13,127	13,125	13,125
Test of Weak Instruments	Reject, p=0.000	Reject, p=0.000	Reject, p=0.000	-
Test of Over-id. Restrictions	Cannot Rej., p=0.11	Cannot Rej., p=0.098	Cannot Rej., p=0.117	-
R^2	-	-	-	0.0019

[22] One $*$ indicates significance at the 10% level, two $**$ indicate significance at the 5% level or better.

Table A.1.: Summary Statistics by Subjective Survival Probabilities Reporting Behavior

Variable	Reported Pliv		Did not report Pliv	
	Obs	Mean (Sd.)	Obs	Mean (Sd.)
Pliv75	22,361	0.662 (0.28)	-	-
Pliv85	22,218	0.454 (0.31)	-	-
Health				
Self-Reported Health Status	22,376	3.482 (1.09)	2,079	3.007 (1.18)
Better Computed Health Change	22,376	0.21 (0.41)	2,079	0.212 (0.41)
Same Computed Health Change	22,376	0.536 (0.49)	2,079	0.498 (0.50)
Worse Computed Health Change	22,376	0.254 (0.43)	2,079	0.291 (0.45)
Better Self-Reported Health Change	22,376	0.120 (0.32)	2,079	0.088 (0.28)
Same Self-Reported Health Change	22,376	0.521 (0.50)	2,079	0.519 (0.50)
Worse Self-Reported Health Change	22,376	0.359 (0.48)	2,079	0.393 (0.49)
Index of Chronic Disease	22,376	0.184 (0.16)	2,079	0.208 (0.17)
Index of ADLs and IADLs	22,376	0.127 (0.15)	2,079	0.165 (0.20)
Health Limitation	22,376	0.221 (0.41)	2,079	0.290 (0.45)
Psychological problems	22,236	0.139 (0.35)	2,064	0.157 (0.36)
Number of doctor visits	22,376	7.936 (14.31)	2,079	8.775 (15.12)
Number of days hospitalized	22,376	0.279 (1.27)	2,079	0.518 (2.85)
Smoker	22,376	0.206 (0.40)	2,079	0.219 (0.41)
Drinker	22,376	0.554 (0.49)	2,079	0.431 (0.49)
No Health Insurance	22,376	0.176 (0.38)	2,079	0.227 (0.42)
Self-Reported Memory Ability	22,369	3.237 (0.95)	807	2.874 (0.99)
Computed Memory Ability Change	22,075	-0.119 (0.71)	783	-0.110 (0.72)
Self-Reported Memory Change	22,362	-0.111 (0.41)	804	-0.134 (0.44)
Socio-Demographic				
Age	22,376	57.881 (5.00)	2,079	58.644 (4.65)
Age squared	22,376	33.753 (5.53)	2,079	34.607 (5.19)
Male	22,376	0.382 (0.49)	2,079	0.603 (0.49)
Year of Education	22,376	12.796 (2.769)	2,079	11.361 (3.62)
Married	22,376	0.807 (0.39)	2,079	0.854 (0.35)
White	22,376	0.865 (0.34)	2,079	0.767 (0.42)
Total net wealth (in $10,000)	22,376	34.14 (-121.74)	2,079	46.78 (228.25)
Total income (in $10,000)	22,376	2.144 (2.77)	2,079	2.201 (3.83)
Family Background				
Indicator of father lived up to 75	22,376	0.455 (0.49)	2,079	0.484 (0.50)
Indicator of mother lived up to 75	22,376	0.605 (0.48)	2,079	0.564 (0.49)
Indicator of father lived up to 85	22,376	0.137 (0.34)	2,079	0.169 (0.37)
Indicator of mother lived up to 85	22,376	0.189 (0.39)	2,079	0.168 (0.37)
Mother's years of education	22,376	9.782 (3.42)	2,079	8.532 (3.88)
Father's years of education	22,376	9.434 (3.80)	2,079	8.247 (4.11)
Indicator of kids living close	22,376	0.509 (0.50)	2,079	0.556 (0.49)
Number of grandkids	22,376	3.707 (4.68)	2,079	4.497 (5.38)
Number of siblings	22,376	2.901 (2.39)	2,079	3.293 (2.65)

Table A.2.: Sensitivity Analysis of Pliv85 Estimation using Corrected FE model

Variable	Corrected FE Coef. (S.E.)	Corrected FE Coef. (S.E.)
Self-Reported Health Status	0.018 (0.004)	0.019 (0.003)
Lagged S-R Health	-	0.004 (0.003)
Better Computed Health Change	-0.001 (0.005)	-
Better S-R Health Change	0.012 (0.006)	0.013 (0.006)
Index of Chronic Disease	-0.073 (0.024)	-0.071 (0.024)
Index of ADLs and IADLs	-0.104 (0.024)	-0.103 (0.024)
Health Limitation	-0.013 (0.007)	-0.012 (0.007)
Psychological problems	-0.012 (0.010)	-0.011 (0.010)
Number of doctor visits	-0.0004 (0.0001)	-0.0004 (0.0001)
Num of days hospitalized	-0.001 (0.001)	-0.001 (0.001)
S-R Memory Ability	0.010 (0.003)	0.010 (0.003)
Smoker	-0.008 (0.010)	-0.008 (0.010)
Drinker	-0.007 (0.006)	-0.007 (0.006)
No Health Insurance	-0.003 (0.005)	-0.003 (0.005)
Age	-0.056 (0.009)	-0.056 (0.009)
Age squared	0.031 (0.007)	0.031 (0.007)
Married	0.016 (0.010)	0.016 (0.010)
Total net wealth (in $10,000)	0.00001 (0.00002)	0.00001 (0.00002)
Total income (in $10,000)	-0.002 (0.001)	-0.002 (0.001)
Indicator of father liv85	0.014 (0.017)	0.014 (0.017)
Indicator of mother liv85	0.006 (0.011)	0.006 (0.011)
Indicator of kids living close	0.001 (0.005)	0.001 (0.005)
Number of grandkids	0.000 (0.001)	0.000 (0.001)
Number of siblings	0.002 (0.002)	0.074 (0.009)
Wave Indicator	0.073 (0.009)	0.002 (0.002)
Constant	2.382 (0.345)	2.373 (0.345)
Inverse Mills' ratio	-0.149 (0.068)	-0.145 (0.068)
Obs	24,915	24,915
R^2	0.042	0.044
Average Pliv85	0.465	0.465

Table A.3.: Selection Equations of different Corrected FE Pliv75 estimations

Variable	Coef. (S.E.)	Coef. (S.E.)	Coef.(S.E.)	Coef. (S.E.)
Self-Reported Health Status	0.130 (0.020)	0.149 (0.022)	0.128 (0.020)	0.147 (0.022)
Better Computed Health Change	-	-0.101 (0.039)	-	-0.104 (0.039)
Better S-R Health Change	-	-	0.032 (0.055)	0.041 (0.055)
Index of Chronic Disease	0.222 (0.118)	0.260 (0.121)	0.213 (0.119)	0.251 (0.121)
Index of ADLs and IADLs	-0.044 (0.138)	-0.036 (0.138)	-0.042 (0.138)	-0.033 (0.139)
Health Limitation	0.110 (0.047)	0.120 (0.047)	0.109 (0.047)	0.119 (0.047)
Psychological problems	0.048 (0.047)	0.050 (0.047)	0.046 (0.047)	0.049 (0.047)
Number of doctor visits	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Numbers of days hospitalized	-0.011 (0.004)	-0.011 (0.004)	-0.012 (0.004)	-0.012 (0.004)
Self-R memory ability	-0.048 (0.018)	-0.051 (0.018)	-0.048 (0.018)	-0.051 (0.018)
Smoker	0.078 (0.041)	0.080 (0.042)	0.078 (0.041)	0.081 (0.041)
Drinker	0.184 (0.034)	0.186 (0.034)	0.183 (0.034)	0.186 (0.034)
No Health Insurance	-0.032 (0.040)	-0.028 (0.040)	-0.034 (0.040)	-0.030 (0.040)
Age	0.004 (0.044)	0.007 (0.044)	0.004 (0.044)	0.007 (0.044)
Age squared	-0.019 (0.039)	-0.021 (0.039)	-0.019 (0.040)	-0.021 (0.039)
Male	-0.067 (0.041)	-0.059 (0.041)	-0.067 (0.041)	-0.059 (0.041)
Years of Education	0.065 (0.007)	0.064 (0.007)	0.065 (0.007)	0.064 (0.007)
Married	0.038 (0.042)	0.038 (0.042)	0.039 (0.042)	0.039 (0.042)
White	0.392 (0.043)	0.386 (0.044)	0.393 (0.043)	0.387 (0.044)
Total net wealth (in $10,000)	-0.0003 (0.0001)	-0.0004 (0.0001)	-0.0003 (0.0001)	-0.0004 (0.0001)
Total income (in $10,000)	0.004 (0.010)	0.003 (0.010)	0.004 (0.010)	0.003 (0.010)
Indicator of father liv75	-0.072 (0.034)	-0.073 (0.034)	-0.072 (0.034)	-0.074 (0.034)
Indicator of mother liv75	0.036 (0.035)	0.037 (0.035)	0.036 (0.035)	0.036 (0.035)
Mother's education	0.025 (0.007)	0.025 (0.007)	0.025 (0.007)	0.024 (0.007)
Father's education	-0.003 (0.006)	-0.003 (0.006)	-0.003 (0.006)	-0.003 (0.006)
Indicator of kids living close	0.012 (0.033)	0.012 (0.033)	0.013 (0.033)	0.014 (0.033)
Num of grand kids	0.005 (0.003)	0.004 (0.003)	0.005 (0.003)	0.004 (0.003)
Num of siblings	0.005 (0.007)	-0.134 (0.012)	-0.133 (0.012)	0.005 (0.007)
Wave Indicator	-0.134 (0.012)	0.005 (0.007)	0.005 (0.007)	-0.133 (0.012)
Self-Reported as retired	0.182 (0.043)	0.176 (0.043)	0.182 (0.043)	0.176 (0.043)
Constant	0.732 (1.242)	0.644 (1.232)	0.739 (1.247)	0.654 (1.237)
Obs	26,325	26,278	26,314	26,273
R^2	0.108	0.109	0.1079	0.1087
Log-likelihood	-4,098.91	-4,073.84	-4,095.53	-4,070.73

Table A.4.: Estimation of Changes in Pliv85 using Corrected Sample Selection IV and FE models

Variable	Corrected IV Coef. (S.E.)	Corrected IV Coef. (S.E.)	Corrected IV Coef (S.E.)	Corrected FE Coef (S.E.)
Pliv85 in previous wave	1.063 (0.047)**	1.062 (0.047)**	1.061 (0.047)**	-
Better Computed Health Change	0.035 (0.007)**	-	0.034 (0.007)**	0.027 (0.009)**
Better S-R Health Change	-	0.012 (0.008)	0.008 (0.008)	0.035 (0.013)**
Changes in Index of Chronic Disease	-0.072 (0.037)*	-0.078 (0.037)**	-0.072 (0.037)*	-0.060 (0.049)
Changes in Index of ADLs and IADLs	-0.066 (0.030)**	-0.079 (0.030)**	-0.065 (0.030)**	-0.112 (0.042)**
Onset of Health Limitation	-0.002 (0.011)	-0.003 (0.011)	-0.002 (0.011)	-0.002 (0.016)
Onset Psychological problems	-0.009 (0.017)	-0.009 (0.017)	-0.009 (0.017)	0.021 (0.027)
Start to Smoke	0.001 (0.024)	0.001 (0.024)	0.001 (0.024)	-0.022 (0.037)
Start to Drink	0.002 (0.012)	0.002 (0.012)	0.002 (0.012)	0.005 (0.018)
Loss of Health Insurance	-0.031 (0.008)**	-0.031 (0.008)**	-0.031 (0.008)**	-0.031 (0.012)**
Change in doctor visits	-0.0004 (0.0002)	-0.0004 (0.0002)	-0.0004 (0.0002)	-0.0001 (0.0002)
Change in days hospitalized	-0.004 (0.001)**	-0.004 (0.002)**	-0.004 (0.001)**	-0.007 (0.002)**
Computed Memory Ability Change	0.013 (0.007)*	-	0.013 (0.007)*	0.006 (0.009)
S-R Memory Ability Change	-	0.0001 (0.016)	-0.002 (0.015)	0.029 (0.026)
New widowhood	0.014 (0.022)	0.017 (0.023)	0.014 (0.022)	0.047 (0.033)
Changes in total net wealth (in $10,000)	-0.000005 (0.00002)	-0.000006 (0.00002)	-0.000005 (0.00002)	-0.00001 (0.00002)
Changes in total income (in $10,000)	-0.002 (0.001)	-0.002 (0.001)	-0.002 (0.001)	-0.002 (0.002)
Recent father's death	-0.013 (0.016)	-0.013 (0.016)	-0.013 (0.016)	-0.014 (0.022)
Recent mother's death	-0.028 (0.012)**	-0.029 (0.012)**	-0.028 (0.012)**	-0.022 (0.018)
Indicator of kids moving close	0.0004 (0.009)	0.000 (0.009)	0.000 (0.009)	-0.003 (0.013)
Changes in num. of grandkids	0.002 (0.001)**	0.002 (0.001)**	0.002 (0.001)**	0.004 (0.001)**
Changes in num. of siblings	0.004 (0.003)	0.005 (0.003)	0.004 (0.003)	0.005 (0.005)
Constant	-0.030 (0.025)	-0.023 (0.024)	-0.030 (0.024)	-0.071 (0.019)**
Inverse Mills' ratio	0.068 (0.034)**	0.085 (0.035)**	0.071 (0.034)**	0.587 (0.143)**
Observations	12,927	12,919	12,917	12,917
Test of Weak Instruments	Reject, p=0.000	Reject, p=0.000	Reject, p=0.000	-
Test of Over-id. Restrictions	Cannot Rej., p=0.207	Cannot Rej., p=0.154	Cannot Rej., p=0.195	-
R^2	-	-	-	0.0029