The issue discussed in this paper is an important one and the motivation is nicely done. The utility formulation in (3) is also quite intuitive. At this stage, the authors specialize to determining variables education, age and income. While this choice is quite common and data availability is good, it should be mentioned that this is one particular but not the only possible choice.

But that is a minor matter.

[This comment is well taken. Of course this is not the only possible one. There is another paper we are working on to describe the type of variables and their implications for identification. In fact, identification can come from any of, or combination of the variables mentioned.]

A more critical, and in my opinion restrictive, assumption is that life cycle income has a deterministic trajectory. This essentially reduces the model to a static one and creates part of the identifiability problem. I would have been happier with a stochastic income trajectory with the mean and variance modelled in terms of initial income and elapsed time.

[The stochastic income trajectory is possible in principle but requires convolution of the stochastic component in income (or in general if a variable that is stochastic is used in the utility specification) with the error term defined in the utility specification and the stochastic component in the individual discount rate. For instance, when we look at equation (10) in the paper, there are components including the interactions of variables (x) and the stochastic component (u) in individual discount rate. With the inclusion of stochastic income (or any other stochastic variable), the interactions will be more involved with interactions terms of these stochastic components. The naïve estimator we are proposing in this paper enables us to replace discount rate with its prediction, therefore instead of proceeding with equation (10), we obtain equation (12). The estimator on the other hand becomes a two stage estimator where in the second stage an estimated quantity (the individual discount rates) is used. This is why we correct the standard errors by applying a bootstrap methodology. A full estimation framework that directly estimates the form in equation (10) is a topic of another paper. This form can be estimated using a heterogeneous ordered probit specification (This as the reviewer suggests includes modelling of the variance term of the ordered probit model). This is an ongoing research agenda for us. The type of variables that are included in this paper in fact are either fixed (like education) or have deterministic trends (like age), or have stochastic but fixed growth parameters (like income, the growth parameter is unknown but once it is known (estimated), the trajectory of income is deterministic). As stated it can be extended to include other types of variables with stochastic components as the reviewer suggests, but requires careful studying of the interaction terms and making sure of the conditions for identification. The estimator that is required to do all these is an ongoing project.]

The identification strategy of introducing individual specific discount rates in (8) is nice but the specification eventually becomes (12) using (14) for discount rate. Now (14) has the same determining variables as the lifetime utility formulation. So the final form actually involves many squares and cross products of these three variables and problems of co-linearity and correlated error may well arise. A clearer discussion is needed to convince the reader that the estimation strategy successfully avoids or solve these issues. The authors indeed comment that "…final estimation form may not be trivial to estimate…" but the estimation strategy is not yet convincing.

[As the reviewer points the identification and estimation are two different things. Identification is secured with the functional forms, distributions assumed in the paper. The functional form based on the structural model gives the identification in our methodology, so it is not required to augment

the discounting equation with extra variables that are not present in the main utility equation. However, we know this type of identification can encounter problems in the estimation due to the problems the reviewer rightfully listed, such as "..has the same determining variables as the lifetime utility formulation. So the final form actually involves many squares and cross products of these three variables and problems of co-linearity and correlated error may well arise**". We definitely agree with this, as any two step estimation methodology that depends on functional form identification (i.e. Heckman Selection Model is also functionally identified, but it is always recommended to include an exclusion restriction to estimate it in practice) can be hard to estimate in practice and securing identification with exclusion restrictions can be a safer choice. In the revised version, we will include an exclusion restriction (other variable(s) in the discount rate equation (14) that are not included in the main utility equation) to estimate the model]**

The concern is also propagated in the results. For instance, all λ estimates reported in Table 1 are negative (albeit insignificant). This would imply a negative (or zero) growth rate for income which is highly counterintuitive. On the other hand, if they are taken as 0 then we are back to a purely static specification. I believe the authors should report the discount rate estimation results and diagnostics for all the models in more detail.

[The specific coefficient estimates are not discussed in the paper. The reviewer is right about the fact that the λ estimates reported in Table 1 are insignificant (that will be discussed below). However, the main aim of the paper is to show the negative coefficient we generally obtain in classical SAH estimates for the female coefficient disappears if the methodology proposed in the paper is applied. Already the most pronounced and arguably important finding relates to the sign of the female coefficient in the paper. The negative sign of female coefficient in the base model (ordered probit), meaning that females are less likely to report good health as compared to males, turns to be positive yet insignificant for the discount rate specifications where there is no reporting bias. Therefore, the utility framework combined with the individual discount rate assumption removes the female lower health reporting in SAH.

Further, the coefficients are smaller in magnitude in the discount rate specifications. The marginal effects presented in Table 2 also indicate that females reported very good health level are more likely have a higher discount rate as compared to males which means that they are less-future oriented. These findings support our hypothesis that the reason for females to report generally poorer health compared to their male counterparts can be related to the heterogeneity in individual discount factors.

The standard errors of the individual coefficients of λ estimates are most likely to be suffering from the multicolinearity as the reviewer suggested in the previous comment. Therefore, as a result higher standard errors leads to insignificant results. On the other hand, the negativity of the coefficients does not mean that the income trajectory has a decreasing pattern for higher income individuals. The income variable is also a discrete variable with "Very poor, Poor, Medium, Rich and Very rich" are the discrete categories. The omitted group is the "Very Poor" category, therefore the remaining λ estimates indicates that there is no significant difference of the other income level growths from the "Poor" group. This itself can be discussed of course whether this true in general, however it is not panel data we identify this parameter in our estimation. The identification comes from the theoretical justification for the individual behavior regrading health utility. In that respect what we estimate as the income growth can capture other labor market differences among these different income groups.]

Minor comment:

Third line in 3.2 on p 9: "…this theoretically novelty…" should be "…this theoretical novelty…"

[This will be fixed in the revised version]

Figure 1 on P 11: the legend "Graphs by female" is confusing as both male and female histograms are presented – but which one is which should be clearly mentioned. I assumed that the order is male, female.

[Your assumption is right. 0 corresponds to Male (Female = 0), and 1 corresponds to Female. Graphs by the variable Female (which is equal to 0 if the individual is Male, and equals to 1 if the individual is Female) is meant in the Figure. However, we agree that it might be confusing. We will remove "Graphs by female" statement and replace 0 and 1 on top of the graphs by "Male" and "Female" respectively.]

Footnote 6 on p 12 seems to be erroneous as the values specified for smoking intensity are discrete.

[Yes. The footnote should be as follows:" The variable is constructed as a measure that takes a value of 5 for smoking regularly, 4 for smoking sometimes, 2 for previously smoked but quitting, and 1 for never smoking." This will be fixed in the revised version of the paper.]