No 159

The Precautionary Saving Motive and Wealth Accumulation

Mauro Mastrogiacomo and Rob Alessie

CPB discussion paper

CPB Discussion Paper

No 159
October, 2010

The Precautionary Saving Motive and Wealth Accumulation

Mauro Mastrogiacomo and Rob Alessie

CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14
P.O. Box 80510
2508 GM The Hague, the Netherlands

Telephone +31 70 338 33 80
Telefax +31 70 338 33 50
Internet www.cpb.nl

ISBN 978-90-5833-475-6

Abstract in English

We quantify the relative importance of the precautionary saving motive in determining wealth accumulation. Puzzling results have appeared on the relative importance of the precautionary motive when this is derived either using a self reported measure of uncertainty about future income rather than observed life-cycle income variation. In this study we show that if one takes into account explicitly the uncertainty of the second income earner results converge using both methods. Precautionary savings account for about 30\% of wealth accumulation. However we also claim that obtaining converging results does not necessarily answer the question on the empirical relevance of precautionary savings, as the amounts being saved largely differ among studies due to the country specific incentives to save and to the measure of wealth accumulation.

Key words: precautionary savings, income uncertainty

JEL code: D12, D91, E21

Abstract in Dutch

Wij kwantificeren hoe belangrijk het voorzorgmotief is in de vrije besparingen van gezinnen. Tegenstrijdige resultaten zijn verschenen in de literatuur . Dit heeft te maken met twee verschillende benaderingen van inkomensonzekerheid: zoals consumenten die zelf ervaren, of via de variatie van huishoudinkomen over de tijd gemeten.

Wij laten in onze studie zien dat, als men ook de onzekerheid van de tweede verdiener in het gezin meeneemt, de resultaten via deze twee benaderingen ongeveer aan elkaar gelijk zijn. Enerzijds zijn besparingen uit voorzorg ongeveer gelijk aan 30% van het opgebouwde vermogen, anderzijds is het feit dat wij convergerende resultaten vinden niet persé een indicatie dat de vraag over het belang van deze besparingen is beantwoord. Men moet zowel landspecifieke institutionele prikkels, alsook de verschillende maatstaven van vermogensopbouw in acht nemen.

Steekwoorden: vrije besparingen, inkomensonzekerheid

Een uitgebreide Nederlandse samenvatting is beschikbaar via www.cpb.nl.

The Precautionary Saving Motive and Wealth Accumulation

Mauro Mastrogiacomo (CPB, VU University Amsterdam, Netspar)
Rob Alessie (University of Groningen, Netspar)

October 1, 2010

Abstract

We quantify the relative importance of the precautionary saving motive in determining wealth accumulation. Puzzling results have appeared on the relative importance of the precautionary motive when this is derived either using a self reported measure of uncertainty about future income rather than observed life-cycle income variation. In this study we show that if one takes into account explicitly the uncertainty of the second income earner results converge using both methods. Precautionary savings account for about 30% of wealth accumulation. However we also claim that obtaining converging results does not necessarily answer the question on the empirical relevance of precautionary savings, as the amounts being saved largely differ among studies due to the country specific incentives to save.

Keywords: precautionary savings, income uncertainty
JEL codes: D12, D91, E21

1 Introduction

The identification of the impact of saving motives on saving decisions has created several problems to applied researchers in recent years. This is definitely the case for the precautionary motive. While there is theoretical consensus that a broad formulation of the intertemporal allocation problem allows the identification of 2 or 3 different motives to save, it is far from clear what the empirical relevance is of these motives. Life-cycle savings (like those for retirement or purchasing a house), precautionary savings (due to income uncertainty) and bequests, are the motives that can be identified if one is able to estimate structural models that may need up to the third derivative of the utility function and possibly add some non testable assumptions to the model. In this paper we focus on precautionary savings, therefore we show in the appendix a standard identification strategy of the precautionary motive within a life cycle framework when future income is uncertain.

Those applied studies who have attempted these cumbersome estimations have been rewarded with results that are far apart from each other. Kennickell and Lusardi (2004) survey the results for precautionary savings and lament the large range of conclusions derived. In our view comparing these studies is also difficult, as empirical strategies change across studies as well as the data being used, which sometimes refer to completely different countries with very different saving incentives.

We return to the dispute that since the 90's has set the scene for a big part of this literature: the relation between earnings uncertainty and precautionary savings. We will contend that seemingly puzzling results can be resolved, but at the same time we will also show that converging empirical results do not necessarily answer the question on the empirical relevance of the precautionary motive.

Most authors seem to have agreed that different empirical approaches in the definition of income risk are destined to depict precautionary savings as being marginal (Guiso, Jappelli, and Terlizzese 1992), relevant (Lusardi 1997) or extremely important ((Carroll and Samwick 1998) and (Ventura and Eisenhauer 2006)). This means a wealth accumulation for precautionary reasons of about 2% (Guiso, Jappelli, and Terlizzese 1992), about 20% (Lusardi 1997) or around 30-40% ((Carroll and Samwick 1998) and (Ventura and Eisenhauer 2006)). While the first two studies use subjective short term income uncertainty as a proxy of income risk (also referred to as subjective earnings variance (SEV), see Section 3), the latter two use observed life-cycle income variation (LIV) . This means that while in the first approach respondents are asked to report the probability of a household income change within the next year, in the second approach reported income over a number of years and/or cohorts is being used. The emerged consensus points in the direction that subjective questions relative to next year personal income do not incorporate enough variation to represent life cycle income risk, which is indeed better picked up by the second approach. Estimates using the self reported probability distribution of next year income changes are therefore far lower than all other estimates. Given the high quality of the studies and the data being used, these differences are disturbing to any applied researcher. We also believe that income uncertainty over a longer period should be larger then the uncertainty that can be derived by looking at the next year. However it has not been shown that this would motivate larger immediate savings, rather than revealing even more puzzles related to time consistency and procrastination.

In this study we show empirical evidence that tries to reconcile these two positions in the empirical debate. We do not play with the time horizon on which uncertainty should be questioned. Our strategy is to show that the study of Guiso, Jappelli, and Terlizzese (1992) underestimates subjective earning variance not specifically, or at least not only, due to the short term horizon of the income question being asked, as noted by Lusardi (1997). The other reason, which was not researched yet, is the internal inconsistency of considering only the answer of the head of the household when uncertainty is being asked about household income rather than individual income. The second income earner in the household (whose income is typically more at risk) is not being taken into the analysis in this way. We intend therefore to fix this. In addition we follow

Lusardi's suggestion that income risk better being instrumented by unemployment risk, in order to take care of the measurement error that arises due to the short term nature of the subjective expectations question (Lusardi 1997). However contrary to Lusardi's study we don't only instrument the head of the household income risk but explicitly both income earners' in the household.

As a preview of our results, we show that our estimates reproduce closely the results in Lusardi (1997) and Guiso, Jappelli, and Terlizzese (1992). However when we include the uncertainty of the second income earner the share of precautionary savings due to subjective earning variance increases to about 30%. Similar results are also derived using the other empirical approach, based on observed income variation. This magnitude is thus comparable to results reported by Carroll and Samwick (1998) and Ventura and Eisenhauer (2006), whose studies are also partly replicated on our data.

The reason to estimate different series of models using the different empirical approaches is that we want to make sure that results are not driven by the population and the time period being studied. There is indeed empirical evidence that shows that Italians and Americans may differ in saving behavior if not due to different preferences at least because of very different institutions (as example for all: the compulsory saving system present in Italy and largely absent in the US). In this sense the Netherlands (the country we focus on in this study) is more similar to Italy. The need for precautionary savings in the Netherlands may actually be even lower, due to the developed employment protection legislation .

Though we will show converging results of the different empirical approaches, we will propose one important distinction. The concepts of precautionary savings or precautionary wealth are also different within each country, and not only non-comparable between countries. Converging results do not necessarily imply a clear cut answer to the question on the empirical relevance of the precautionary motive.

Such a study is relevant in the current policy debate that is going on in many western countries. Policy makers in countries with an extensive unemployment and disability insurance system often propose policies that imply more responsibility of individuals into insuring themselves against income and health risks. But are individuals motivated to such savings already? The effectiveness of these policy proposals will evidently depend on whether one can or will be able to increase the share of private savings that is meant to insure these risks.

The data are introduced in Section 2. Next we review the empirical approaches mentioned above. Section 4 concludes.

2 Data and descriptive analysis

In this study we use the DNB household survey (DHS). The DHS is administered by CentERdata, which is associated with Tilburg University, the Netherlands. The survey is sponsored by De Nederlandsche Bank (DNB), the Dutch central bank. The aim of the DHS is, among others, to furnish information on both economic and psychological determinants of savings. The survey is conducted annually, starting 1993/1994. In this study, we use the waves up to and including 2008. Each year, the survey contains approximately 1500 households (well over 2500 individuals) and is an unbalanced panel . Wealth questions are asked in 5 separate sub-questionnaires that are released at different points in time over the year. This feature tends to increase non response.

The models that we are going to estimate will be based on different subsamples. This because the questions on subjective earnings variance are only asked in the period 1994-2002. The main characteristics of the sample, like age, family size or education are available for a larger groups of respondents. After selecting out time inconsistent observations for age and year of observation we end up with a basic sample of about 5700 households, which are good for about 18500 point observations where information on active savings is available.

Active savings is one of the dependent variables in the multivariate analysis. We will therefore illustrate in detail the construction of the variable "active savings". The DHS provides very detailed information on households' assets and liabilities, which enables us to calculate an approximation of active household savings. The main source of information for the definition of active savings comes from a specific question concerning the amount of money put aside in the last 12 months. The question is formulated as follows: "About how much money has your household put aside in the past 12 months?". This question is answered by a sub-sample (14948 point observations) who fills in the psychological questionnaire. Answers to this question come in 7 categories, where the first interval is "less than € 1,500" and the last "€ 75,000 or more". We assign to each respondent an amount of active savings equal to the middle point of the interval chosen, or to the lower bound if the category chosen is the last. Evidently those who do not save or even dissave are not accounted for in this way. In order to solve this problem we have to combine different variables present in our data. The first is the answer to the question "Did your household put any money aside in the past 12 months?" which can be answered with a yes or no. The second is question "How is the financial situation of your household at the moment?" which allows the following 5 answers: 1) there are debts, 2) need to draw upon savings, 3) it is just about manageable, 4) some money is saved, 5) a lot of money can be saved.

Those who answer that no money were put aside and that they just about manage with their financial situation, are imputed zero savings (2803 observations). Those who did not put aside money, and either are in debt or drawing upon their savings are imputed a (negative) measure of active savings (673 observations), which we will describe below. Finally those who answer that they did put money aside in the last 12 months, but did not answer the question on active savings, are imputed a (positive) measure of active savings, if they claim that some money or a lot of money can be saved (141 observations, which increases the sample to about 18500 observations). The active saving measure used in the imputation is identical in spirit to the one used for the PSID data in the American literature (Bosworth and Anders 2008). It is based on the first difference of net financial wealth and isolates passive savings in the form

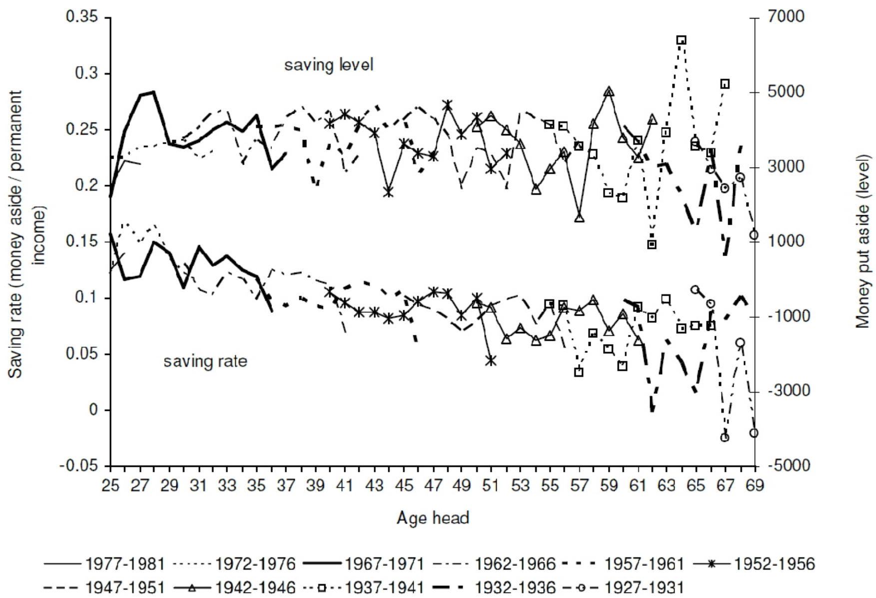

Figure 1: Active savings and saving rate over age and cohort.
The saving rate is plotted at the median and outliers due to too low permanent incomes are removed. Saving levels are at the mean, as the median of a (mostly) categorized variable is not informative.

of capital gains (Berben, Bernoth, and Mastrogiacomo 2006). Our definition is therefore a refinement relative to previous studies (Alessie and Teppa 2010) who impute active savings only looking at the first difference of net financial wealth.

Active savings is further used to compute the individual saving rate. This is the ratio between active savings (as defined above) and permanent income. The variable permanent income is imputed by using the method put forward by (Kapteyn, Alessie, and Lusardi 2005). We use the variable saving rates only for our descriptives.

Permanent income has about 4000 missing values, which reduces the sample to little more than 14000 observations. We compute the saving rate only for a sub-sample of the population, in the sense that outliers are removed (for instance those with permanent income slightly above zero, or with unreliably high saving rates, this reduces the sample further to 13639 observations).

In Figure 1 we plot the development of active savings and the saving rate over age and cohort. Each segment represents a cohort. The graph shows for instance that give age the saving rate of the cohort born between 1967 and 1971 is higher than the saving rate of the cohort born between 1972-1976 (that is at about age 33). The differences are not large. A movement along the segment depicts an age-time effect, while the vertical distance between the segments

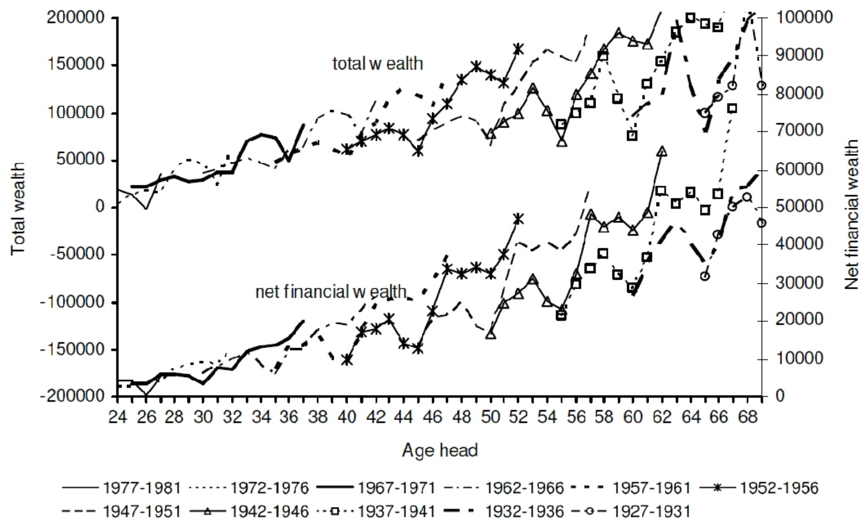

Figure 2: Net financial wealth and total wealth by age and cohort. Figures are plotted by lowess smoothing.

represent cohort-time effects. The figure shows in general that a decrease in savings over age is revealed when looking at saving rates (left vertical ax), not so clearly looking at saving levels (right vertical ax). For most cohorts we notice that, given age, the saving rate of younger cohorts is somewhat higher. This cohort-time effect is less visible when only looking at levels. This suggests that these cohort differentials are (permanent) income related. The figure shows that on average the variation in saving levels is small (between 1 and 5 thousands euro each year). The higher saving rates of the youth (about 10-15%) is of course due to the low permanent income of this cohort.

In Figure 2 we plot net financial wealth and total wealth by age and cohort. The last two items are defined as follows. Net financial wealth sums up the balance of checking accounts, saving accounts, deposits or certificates, business accounts and balances of stocks, bonds and mutual funds, plus some other minor assets. It subtracts several items such as checking account overdrafts, consumer or study debt and other debts. Total wealth also accounts for non financial assets, such as housing wealth, durables and secondary properties, using a self reported measure relative to the current market value of the home. The residual mortgage on all properties is subtracted. Notice that the first difference over time of net financial wealth is being used (cleaned of capital gains) to impute active savings for missing item responses, as reported above.

Figure 2 shows the accumulation process of net financial wealth and total worth. It shows both a positive cohort-time effect and age-time effect. This means that given age younger cohorts tend to be wealthier. This is graphically represented by the vertical distance between the segments. Also given the cohort wealth increases when aging, that is when we move along each separate segment. The cohort differentials are somewhat larger when we look at financial wealth(left vertical ax, notice that the scale is there half than on the other ax) of older cohorts. These differentials are smaller for older cohorts' total worth. This indicates that the net value of their dwellings is higher, possibly due to larger increases in the value of their real estate.

We also extract from the data information about subjective expectations of income risk. This information is only available in some years (1994-2002). This will reduce the sample by about 40%, that will be reduced further by minor item on responses. Part of the multivariate analysis, depending on the method being tested will exclude certain cohorts, or individuals for which some instruments are not observed. Estimating different models will therefore imply using different estimating samples, varying in dimension between about 2500 to 5.000 observations. In Table 1, we show descriptive statistics of the most relevant variables for the estimating sample used in the models below.

The descriptive statistics show some interesting differences among the 3 samples. These differences are larger between Sample1 on the one side and Samples 2 and 3 on the other. This difference is due to the shorter time period of the latter samples. In particular the sample reporting expectations of future earnings is somewhat older, has higher income but lower net worth relative to Sample 1. The summary statistics do not suggest any evident selection in the samples.

3 Income risk and wealth accumulation

We first replicate some studies that have appeared in the literature in order to understand how our data relate to these studies. In this section we argue that it is possible to reconcile empirical findings based on subjective earnings variance (Guiso, Jappelli, and Terlizzese 1992) and observed life-cycle income variation.

3.1 Subjective earnings variance

We now extend the use of data on subjective earnings variance, up to including the second income earner. Due to data limitations we need to simplify slightly the setting chosen by Guiso, Jappelli, and Terlizzese (1992) who uses data for the SHIW survey. Our data only report the percentage growth rate of nominal earnings (z), and not the percentage growth of prices. Due to the historical low level of inflation of the Netherlands during our sample period, we think that this shortcoming is not too serious. This is not the only difference between the Dutch data and the SHIW data. In the DHS respondents are asked "We would like to know a bit more about your expectations of the next 12 months. Below we have presented a number of possible changes in income (y). Please indicate with any of those changes, how likely you think it is that the total income of your household will change by that percentage in the next 12 months", which needs some manipulation to be used. Therefore individuals are asked to report the likelihood $p(z_{i,t})$ such that $z_a < z_{i,t} < z_b$, where in turn $z_b = (\infty, 0.15, 0.1, 0.05, 0, -0.05, -0.1, -0.15)$ and $z_a = (0.15, 0.1, 0.05, 0, -0.05, -0.1, -0.15, -\infty)$.

In order to treat this information empirically we have taken the middle point of each intermediate category such that $z_{i,t} = \frac{z_a + z_b}{2}$ and the lower (upper) bound of the upper (lower) extreme category, such that the 7 possible income changes amount to the following:

category	z
1	0.15
2	0.125
3	0.075
4	0
5	-0.075
6	-0.125
7	-0.15

Other then in the study of Guiso, Jappelli, and Terlizzese (1992), $p(z_{i,t}|y)$ is not revealed directly. We follow largely the approach in Hochguertel (2003), which we formalize hereafter.

Instead of being confronted with a probability distribution, the Dutch respondent is asked about a likelihood scale k_j , ($j = 1, ..., 7$). These seven possible answers stem from completely unlikely, to completely likely. In order to translate the k categories to a probability level, we adapt the notation slightly and refer now to $p(z_{i,t}|y, k_j)$, rather than $p(z_{i,t}|y)$ only. We have assumed values of $p(z_{i,t}|y, k_j)$ such that the following 3 conditions are fulfilled

1. $p(z_{i,t}|y, k_{j-1}) < p(z_{i,t}|y, k_j)$
2. $p(z_{i,t}|y, k_j) - p(z_{i,t}|y, k_{j-1}) \simeq p(z_{i,t}|y, k_{j-1}) - p(z_{i,t}|y, k_{j-2})$
3. $0 < p(z_{i,t}|y, k_j) < 1$

The first condition indicates that these probabilities have consistent ordering with the likelihood scale. We assume out all possible inconsistencies in ordering probabilities. The second condition resembles the approach by Hochguertel (2003) as we also assume that probabilities are (almost) equally spaced. This means that the difference between "highly likely" and "likely" is the same as the distance between "highly unlikely" and "unlikely". The third condition states that also the extreme cases allow a minimal level of uncertainty, and that the probability distribution is not degenerated.

Table 1: Summary statistics for 3 sub samples

	Sample 1: log wealth		Sample 2: active savings		Sample 3: expectations	
	mean	sd	mean	sd	mean	sd
Number of cohabiting children	0.86	1.13	0.94	1.16	1.00	1.17
Age	47	10.27	46	9.85	49	11.76
Net worth (WEALTH)	207138	168031	161662	142340	189084	154901
Education head	2.60	0.78	2.56	0.80	2.60	0.84
Family size	2.63	1.35	2.75	1.36	3.01	1.18
Hours worked head	40.42	8.45	41.95	9.23	43.70	8.73
Hours worked partner	17.45	16.85	20.81	16.92	26.05	15.38
Log permanent income	10.07	0.44	10.09	0.47	10.18	0.45
Log total assets household	4.82	1.33				
Log total assets head	11.47	1.62				
Log wealth rate	1.66	1.25	1.18	1.59	1.49	1.21
Cohabiting partner	0.76	0.43	0.79	0.41	1.00	0.00
Motive : bequest money to children	2.45	1.65	2.42	1.61	2.54	1.62
Motive : to supplement future social security	4.85	1.75	4.65	1.90	4.44	1.96
Active savings	3962	4205	3690	4136	4158	4526
Household with self employed	0.29	0.46	0.26	0.44	0.07	0.26
Gross income head	43020	21343	45009	20626	48650	21325
Gross income partner	7901	12594	9466	13343	12521	14340
Head employed	0.91	0.29	0.84	0.37	0.78	0.42
Partner employed	0.33	0.47	0.39	0.49	0.56	0.50
Current year	2001	5	1997	2	1996	2
Year of birth head	1954	11	1950	10	1947	12
N	5472		3527		2444	

Explanatory note: Sample 1 is observed in the period 1994-2008, samples 2 and 3 in period 1994-2002. The different sample sizes are the result of keeping the most possible observations when selecting on available information

These three conditions result empirically in the following schedule[1]

$p(z_{i,t}\|y, k_j) = .002$	if	$j = 1$	highly unlikely
$p(z_{i,t}\|y, k_j) = .20$	if	$j = 2$	somewhat unlikely
$p(z_{i,t}\|y, k_j) = .35$	if	$j = 3$	unlikely
$p(z_{i,t}\|y, k_j) = .50$	if	$j = 4$	not likely nor unlikely
$p(z_{i,t}\|y, k_j) = .65$	if	$j = 5$	likely
$p(z_{i,t}\|y, k_j) = .80$	if	$j = 6$	somewhat likely
$p(z_{i,t}\|y, k_j) = .998$	if	$j = 7$	highly likely

As all 7 questions are independent, the sum of all probabilities is not bounded to be equal to 1. We therefore compute the following conditional probabilities

$$p_j = \frac{p(z_{i,t}|y, k_j)}{\sum_{j=1}^{7} p(z_{i,t}|y, k_j)} \tag{1}$$

the probabilities p_j are then used as components of the probability distribution of each individual. Say for instance that an individual reports all income changes to be somewhat unlikely (probability of 20%) with exception of a decrease between 5% and 10% (category 5), which is felt to be somewhat likely (probability of 80%), then the denominator of equation 1 is equal to 2 and all income changes have a p_j of 10%, with exception of category 5, which then has a p_j of 40%.

Net household income from the income questionnaire is then used to proxy for y where missing values are replaced by the answer to the categorial question INKHH (The total net income of your household consists of the income of all members of the household, after deduction of taxes and premiums for social insurance policies, taken as the sum total over the past 12 months. Into which of the categories mentioned below did the total net income of your household go in the past 12 months?), which is asked again in the psychological questionnaire.

Therefore

$$\sigma^2 = \sum_{j=1}^{7} \left[p_j * (z_j - \mu_z)^2\right] * y^2$$

where $\mu_z = \sum_{j=1}^{7} p_j * z_j$

The only level of certainty that is allowed is derived from the question INKZEKER (How certain do you feel about this change of income?) when respondents answer "totally certain". When we organize all this information we end up with the following distribution of the subjective standard deviation over current earnings.

[1] In the estimation of the empirical model we have experimented with schedules that were somewhat different, but this did not affect our results in any interesting way.

Table 2: Frequency distribution of the ratio of the subjective standard deviation to the mean of earnings in selected classes sigma/Y)

	Partner (%)									
Head (%)	0	0 - 1.5	1.5 - 2.5	2.5 - 3.5	3.5 - 4.5	4.5 - 6.5	6.5 - 10	10 - 15	Total	N
0	11.2	5.2	1.1	1.3	1.2	2.3	2.6	0.4	25.2	1107
0 - 1.5	3.4	9.3	1.5	1.3	1.3	1.5	1.5	0.5	20.4	896
1.5 - 2.5	0.5	0.9	0.8	0.6	0.5	0.7	0.6	0.3	4.8	212
2.5 - 3.5	1.0	1.7	0.5	1.2	0.8	1.3	1.1	0.2	7.8	344
3.5 - 4.5	1.1	1.6	0.6	0.7	1.5	1.5	1.1	0.2	8.4	368
4.5 - 6.5	1.7	2.2	0.8	1.3	1.9	4.6	3.3	0.6	16.4	719
6.5 - 10	1.2	1.7	0.6	0.8	1.0	2.9	5.2	1.0	14.3	629
10 - 15	0.2	0.3	0.2	0.2	0.2	0.4	0.6	0.5	2.6	116
Total	20.3	22.9	6.1	7.4	8.5	15.1	16.0	3.7	100	
N	890	1005	269	324	374	664	702	163		4391
Sum diagonal		34.3								
Sum above diagonal		35.4								
Sum below diagonal		30.3								

Explanatory note: The mean of the ratio of the subjective standard deviation to the mean of earnings (σ/Y) is about 3%.

Table 2 reports the head's and partner subjective standard deviation of mean earnings. If we only look at the head, like in the Italian studies, we notice very similar findings (Guiso, Jappelli, and Terlizzese 1992). We also have a large fraction of heads of the household (about 25%) perceiving no uncertainty. In the Dutch case the ratio between the subjective standard deviation and mean income is on average similar to the one reported for Italy by Lusardi (1997) (about 3%).

Table 2 shows also that while individual uncertainty tends to be lower, household uncertainty is much larger. Those households expecting no change in income in the opinion of both cohabiting respondents is about 11% (this is 1/3 relative to the Italian studies) and the household in which both members fall in the same uncertainty category (the sum of the main diagonal in table 2) is only 34%. Further the sum of the above and below diagonal frequencies is very similar. This means that while the average income uncertainty for the household is very similar to the one of the single respondent (above and below diagonal cells compensate each other), at the household level the variation is much higher. Evidently the disagreement within the couple returns a higher level of income uncertainty relative to the one reported by the head of the household alone. As consumption and saving decisions depend on both members of the household, it is essential to account for this household level heterogeneity.

We assume therefore that both households members uncertainty matters to the saving decision. We adapt the notation slightly and rewrite $z_{i,t}$ as $z_{\omega,t}$ in order to take into account the different positions and household dimensions. Therefore $z_{\omega,t}$ is the percentage nominal household earnings change, and $\omega = f, h, p$. Where f denotes household h denotes head and p denotes partner. Notice that according to the question "By what percentage do you think the total net income of your household will increase/decrease in the next 12 months?" both respondents report their own opinion about the common household income change. We assume that percentage nominal household earnings change of both household members should be averaged out, that is to say: $z_f = \frac{z_h + z_p}{2}$.The reason we take the average of both uncertain changes is to acknowledge that household saving is the results both of income and expenditures within the household. The working partner needs of course to be accounted for, but also the non working partner has a share of the responsibility in the consumption decisions and her opinion about future income uncertainty is relevant to the saving decision. One could argue that a weighted average, depending on household bargaining power, could be a better measure than a simple average. In the simple case in which bargaining power solely depends on individual income though we would attach no power to the non working partner. In essence we would implicitly be assuming exclusively private consumption, or the irrelevance of household production within the household, which is not justified in this contest. Clearly our approach is largely in line with assuming public consumption within the household. This is also justified on empirical grounds. Our data contain from 2004 onwards the following question: "Now we would like to ask you how your household is organized and how financial decisions are taken. Which of the following statements represents the situation in your household

Table 3: How is your household is organized about taking financial decisions?

	n	%
All our money belongs to both of us, no distinction between mine and yours	1165	77%
Part of the money is private, the other part is mutual money	203	13%
The money we earn individually is one's own	43	3%
I control the finances, my partner receives an allowance	11	1%
My partner controls the finances, I receive an allowance	10	1%
I get part of the household money, my partner controls the rest	6	0%
My partner receives part of the household money, I control the rest	11	1%
Another settlement	26	2%
The above is not applicable for my situation/I do not have a partner	2	0%
don't know	33	2%
N	1510	

Explanatory note: These figures are related to the survey 2005. Source: DHS, own computations

most?". Answers to this question are reported in Table 3. Evidently a very small fraction of households takes money related decisions on the base of their exclusive personal income. Non the less in the empirical analysis we will both use a weighted average in order to compute z_f and a row average. The latter example is now pursued further to explain our computations.

The variables z_h and z_p are not independent. The variance computation must take this into account, thus

$$\sigma^2_{z_f} = \frac{\sigma^2_{z_h}}{4} + \frac{\sigma^2_{z_p}}{4} + \rho\frac{\sigma_{z_h}\sigma_{z_p}}{2}$$

The next cases are relevant:

$$\sigma^2_{z_p} = 0 \text{ and } \sigma^2_{z_h} = 0, \text{ implies that } \sigma^2_{z_f} = 0. \quad (2)$$

$$\sigma^2_{z_p} > 0 \text{ and } \sigma^2_{z_h} = 0, \text{ implies that } \sigma^2_{z_f} = \frac{\sigma^2_{z_p}}{4} \quad (3)$$

$$\sigma^2_{z_p} = 0 \text{ and } \sigma^2_{z_h} > 0, \text{ implies that } \sigma^2_{z_f} = \frac{\sigma^2_{z_h}}{4} \quad (4)$$

$$\sigma^2_{z_p} > 0 \text{ and } \sigma^2_{z_h} > 0, \text{ implie sthat } \sigma^2_{z_f} = \frac{\sigma^2_{z_h}}{4} + \frac{\sigma^2_{z_p}}{4} + \rho\frac{\sigma_{z_h}\sigma_{z_p}}{2}. \quad (5)$$

The frequencies of these 4 cases can be derived using Table 2. As in the last case we need to determine the level of ρ, we assume $\rho = cov(\sigma^2_{z_h}, \sigma^2_{z_p}) = 0.35$, as derived in our sample[2]. This means that in some cases $\sigma^2_{z_f}$ will be lower than either $\sigma^2_{z_h}$ and $\sigma^2_{z_p}$.

[2] We have carried out some sensitivity analysis to this parameter as this value may seem high. We report some results below.

Notice that the measure of household earnings variance derived in expressions 2 to 5 is statistically meaningful but not theoretical. Beside that in our data the value of $\sigma^2_{z_f}$ appeared to be lower than $\sigma^2_{z_h}$ (the measure used previously in the literature). This could imply a mechanically lower value of the variable and therefore an artificially higher elasticity of savings to $\sigma^2_{z_f}$. In order to tackle these two problems we use also a second definition of $\sigma^2_{z_f}$, that is

$$\sigma^2_{z_f} = \frac{\sigma^2_{z_p}}{2} + \frac{\sigma^2_{z_h}}{2} \tag{6}$$

This measure is theoretically defendable, as it is in line with a simple unitary model with no consumption sharing. Beside it will also imply $\sigma^2_{z_f} > \sigma^2_{z_h}$ as in our data it appears that $\sigma^2_{z_p} > \sigma^2_{z_h}$.

3.2 Multivariate analysis

We take as starting point the same specification as in Guiso, Jappelli, and Terlizzese (1992) and Lusardi (1997) we estimate the following model:

$$\begin{aligned} \ln \frac{W_{i,t}}{Y_i} &= \beta_0 + \beta_1 Age_{i,t} + \beta_2 Age^2_{i,t} + \beta_3 Age^3_{i,t} + \beta_4 Fsize_{i,t} \\ &\quad + \beta_5 Nchild_{i,t} + \beta_6 \ln(Y_i) + \beta_7 \frac{\sigma^2_{z_{\omega,t}}}{Y_i} + u_{i,t} \end{aligned} \tag{7}$$

where $W_{i,t}$ is wealth[3], Y_i is permanent income (Kapteyn, Alessie, and Lusardi 2005), which is cleaned of outliers. Further $Fsize$ represents family size and $Nchild$ the number of cohabiting children. Differently from the Italian studies we use a panel dataset. We therefore run first pooled OLS regressions (in order to correct the standard errors for within group dependence) and also a random effect model. Table 4 collects the estimation results.

The estimates in Table 4 show that the results of the pooled OLS[4] and of the random effect model both return an effect for precautionary accumulation of about 4% when we only look at the head of the household. This finding is perfectly in line with results from Guiso, Jappelli, and Terlizzese (1992) and Lusardi (1997) that proposed and estimated these models already. They find accumulations of about 2.8%. Relative to their sample we have used an additional selection criterion, by excluding singles. This is due to the need of comparability with the other specifications. However when we included singles we had almost the same findings (results available from the author).

Lusardi (1997) noticed also that a question about next year income changes could actually measure with error the level of life time income uncertainty, which in the theoretical model is supposed to motivate precautionary savings. Lusardi

[3] About 5% of our sample who has negative wealth.

[4] We have also experimented with median regression. The proportion of precautionary savings decreased somewhat in all models, however the impact relative to the other specifications did not change qualitatively.

(1997) claims that Italians face income risk not so much due to short term income changes (labor contracts typically cover a longer horizon, leaving little room to uncertainty for the next year). Unemployment risk is thus a better measure of income uncertainty. This is also true for the Netherlands, where labor contracts are actually binding at national level. However while Lusardi (1997) uses regional unemployment as an instrument to correct this measurement error, we use the age of entrance into the labor market (results are very similar when we experiment with and indicator for sector of employment). This variable is created using information on the amount of years that one has contributed into the pension system and does not use information about education. There are two reasons to prefer this instrument in the Dutch context. The first is that regional unemployment in the Netherlands is not interesting. Contrary to Italy, in the Netherlands unemployment does not differ much by region, and interregional commuting is extremely common, due to the geographic characteristics of the country. Further all Dutch employees are covered by an unemployment benefit. The largest component in income risk depends on the duration of the unemployment benefit, which is proportional to experience. Therefore the age of entrance into the labor market is related to unemployment-based income risk, rather than the general level of unemployment.

The results of the IV regression, where age of entrance of both spouses are used as instruments, reveal that precautionary accumulation rises to 25% of total accumulation. This is again in line with results from Lusardi (1997) who noticed an increase of accumulation to 25% also in her IV results. At the bottom of the table we also report the value of the Sargan test statistic for over-identification and of a F-test for the joint significance of the instruments in the first stage equation. The first test statics are very low and therefore not statistically significant. This implies that we can reject the hypothesis of the model being misspecified (for instance because the instrument was to be included in the main equation). The F statistic suggests that the instrument is statistically relevant. This is also the case for all other specifications that we show. So far our data reproduced the basic findings and the reassessed results appeared in the literature. However these rates of accumulation are still substantially lower of those presented by Carroll and Samwick (1998) and Ventura and Eisenhauer (2006).

Table 4: Estimation results for reported earnings variance and wealth accumulation

	Head only			Head and partner		
	OLS1	IV	RE	OLS2	IV1	IV2
Head civil servant						0.0745
Age	0.2505**	0.1983*	0.4971***	0.2408*	0.2112	0.2506*
Age head square	-0.0042*	-0.0029	-0.0091***	-0.004	-0.0034	-0.0023
Age head cube	0	0	0.0001***	0	0	0
Education						0.0632
Family size	-0.0322	-0.161	-0.0555	-0.0377	-0.1994	-0.2771
Log permanent income	0.2566***	0.2814***	0.1917**	0.2593***	0.3396***	0.2792***
Motive : bequest money to children						0.0409**
Motive : to supplement future social security						0.0288**
Household with self employed	0.2212**	0.0132	-0.0109	0.2218**	-0.0108	0.0757
Subjective earnings variance head	0.0005***	0.0033**	0.0004***			
Subjective earnings variance (mean head and partner)				0.0008***	0.0056***	0.0056**
Year 1996						0.1653***
Year 1997						0.0959
Year 1998						0.024
Year 1999						-0.8111***
Year 2000						-0.6224**
Year 2001						0.2765
Year 2002						0.6256**
Year of birth head						0.0974***
Constant	-5.7808***	-5.6459***	-9.0145***	-5.6543***	-6.1686***	-199.3928***
R^2	0.0317	0.0131	0.0259	0.0328	0.0148	0.03
N	2444	2444	2444	2444	2444	2444
Mean earnings variance	87	87	87	64.73	64.73	64.73
$\left(1-1/\exp(\beta_7 * \sigma_w^2/Y)\right)$	3.9%	24.9%	3.4%	5.1%	30.3%	30.3%
Sargan test statistic ${\chi^2}_{(1)}$		0.059			0.0001	0.172
F-Test instruments first stage		10.9			10.7	8.9

Explanatory note: The marginal effect of the subjective earnings variance on wealth accumulation $(1-1/\exp(\beta_7 * \sigma_w^2/Y))$ is reported as a percentage effect. The instrument used in the IV is the age of entrance into the labor market of both spouses. We report results of test statistics on the validity of the instruments in the first stage regression, and the results of the over-identification test. ***=1%, **=5%, *=10% statistically significant.

Table 4 includes therefore four more specifications where now household earning variance, as computed in equations 2 to 5, replaces the one reported by the head of the household. The first is again a pooled OLS. It reveals that precautionary accumulation is about 5% (see column OLS2). This is twice as much as in Lusardi's study, though still very low.

When we instrument it using the age of entrance into the labor market of both adults in the household we notice that precautionary accumulation increases to 30.3% (columns IV1). Notice that now we have taken into account both the problem related to the short term horizon of the subjective question (thus the measurement error problem, by instrumenting as in Lusardi 1997) and the problem of the household income related question being confronted only with the head of the household answer in previous studies.

So far we have kept the model specification identical to that of Lusardi (1997). The last panel IV specification in the right panel of the table departs from the previous specifications in the literature and extends the set of regressors (column IV2). We add some additional usual suspects, such as education and sector of employment. We also add year of birth, in order to account for possible spurious correlations and a dummy for self employment. In this specification we also more explicitly account for the panel structure by adding time effects. However as we already included age and year of birth we cannot include simple time dummies (this is due to the identity between current year and the sum of age and year of birth). We use thefore the transformation proposed by Deaton and Paxson, where all time effects add up to zero (Deaton and Paxson 1994). Due to these transformations, there are no real time effects. All transitory time effects are assumed to be business cycle shocks instead of, for instance, changes in preferences.

We have also added two more controls for the bequest motive and the life cycle savings. These are the answers to questions about the importance of bequeathing money to children, or to save in order to supplement social security in the future. The positive coefficients indicate that those who find these motives important also have higher accumulation. We don't interpret this as a causal effect, but we find it interesting that also adding these controls did not affect our findings about the precautionary motive.

These findings indicate that, once the time horizons and household composition issues are taken into account at the same time, the empirical results derived using subjective earnings variance return considerable rates of accumulation.

3.3 Sensitivity analysis

In the empirical approach above we have made some important assumptions. The first is that the computation of the variance of the correlated variables needed to follow an exclusively statistical definition as in expressions 2 to 5. We relax this assumption, as anticipated above, by using a theoretically defendable proxy, namely $\sigma^2_{z_f} = \frac{\sigma^2_{z_p}}{2} + \frac{\sigma^2_{z_h}}{2}$. Results for this specification are listed in Table 5 under columns IV3 and IV4. The second assumption is that the household expected change in earnings depends equally on the expectation of both household members, namely $z_f = \frac{z_h + z_p}{2}$. We have experimented also with another definition: $z_f = z_h(g) + z_p(1-g)$, where g is a weight that depends on personal income. Results of this specification are reported in model IV5. Finally we have set $\rho = 0$ in the estimation reported under model IV6.

All these estimates confirm the magnitude of the result above, and do not alter the main conclusions. Such a larger impact of precautionary savings may seem odd in a European context. One remark is that Americans are exposed to more risk while Europeans, especially in The Netherlands, have less need for insurance. While this remark is definitely true for the level of accumulated personal saving, this might not hold for the allocation of those savings to different motives. Indeed personal savings in The Netherlands are lower in level relative to those in countries with less generous retirement systems.

In order to remove this kind of concerns, we will next move to estimating a model that uses income variation in the determination of precautionary savings. In this way we eliminate the concern of comparing Dutch results with previously obtained results on US data, that testify of completely different institution-driven saving incentives. This step is necessary as our results on SEV have not been replicated on US data, therefore we replicate the study on LIV on data from The Netherlands.

3.4 Life-cycle income variation

In order to estimate precautionary accumulation using observed income rather then subjective income expectations we replicate the model estimated by Ventura and Eisenhauer (2006), which is also closely linked to the theoretical model presented in the appendix and has the main advantage of being estimated on the same Italian data, like the studies quoted above. By estimating such a model we can remove the suspect that our previous findings on subjective earning variance depend on the some hidden peculiarity in our data. We also re-estimate the model presented by (Carroll and Samwick 1998), that aims to explain the log of wealth and is generally more often referred to in the literature. In order to apply these two methods, we use our data as if these were repeated cross sections rather then a panel.

Table 5: Estimation results for reported earnings variance and wealth accumulation

	OLS3	IV3	IV4	IV5	IV6
Head civil servant			0.0753	0.0922	0.07
Age	0.2423*	0.227	0.2615*	0.3316**	0.2540*
Age head square	-0.0041	-0.0037	-0.0024	-0.0036	-0.0023
Age head cube					
Education			0.0631	0.0675	0.0685
Family size	-0.038	-0.1995	-0.2828	-0.3432	-0.2815
Log permanent income	0.2594***	0.3411***	0.2812***	0.2625**	0.2628***
Motive : bequest money to children			0.0408**	0.0424**	0.0420**
Motive : to supplement future social security			0.0272*	0.0281**	0.0287**
Household with self employed	0.2206**	-0.0261	0.0671	-0.0189	0.0808
Subjective earnings variance (mean head and partner)				0.0054**	0.0068**
Subjective earn. var unitary model	0.0003***	0.0021***	0.0021**		
D&P Year 1996			0.1569***	0.1219***	0.1672***
D&P Year 1997			0.0911	0.0522	0.0933
D&P Year 1998			0.0202	0.023	0.0252
D&P Year 1999			-0.8063***	-0.8377***	-0.8082***
D&P Year 2000			-0.6307**	-0.5477**	-0.6067**
D&P Year 2001			0.277	0.2895	0.2601
D&P Year 2002			0.6333**	0.6019**	0.6258***
Year of birth head			0.0997***	0.1099***	0.0971***
Constant	-5.6803***	-6.4585***	-204.1668***	-225.2260***	-198.7225***
R2	0.0329	0.0148	0.0294	0.0262	0.033
N	2444	2444	2444	2444	2444
Mean earnings variance	177	177	177	73.16	51.9
$\left(1 - 1/\exp(\beta_7 * \sigma_w^2/Y)\right)$	5.2%	31%	31%	32.6%	29.7%
Sargan (score) $\text{chi}^2(1)$		0.006	0.114	0.114	0.147
F-Test instruments first stage		12.09	9.97	8.32	9.74

Explanatory note: The marginal effect of the subjective earnings variance on wealth accumulation $\left(1 - 1/\exp(\beta_7 * \sigma_w^2/Y)\right)$ is reported as a percentage effect. The instrument used in the IV is the age of entrance into the labor market of both spouses. We report results of test statistics on the validity of the instruments in the first stage regression, and the results of the over-identification test. ***=1%, **=5%, *=10% statistically significant.

3.4.1 A model for active savings

In their model Ventura and Eisenhauer (2006) regress savings (SAVE) on a number of variables. The proxy for bequest motives (BEQUEST) is the saving rate of a corresponding elderly cohort, the variance of future income is derived using the income variance (VAR) of the young cohorts. The latter is therefore a proxy of the precautionary motive. Life cycle savings are introduced using the wage gap between cohorts (WDIFF) of old and young respondents with similar education[5]. It is not clear why these proxies should correspond to the theoretical variables implied by the model, however it is still interesting for our case to replicate this method. Their model is:

$$\begin{aligned} SAVE &= \alpha_0 + \alpha_1 WDIFF_{i,t} + \alpha_2 VAR_{i,t} + \alpha_3 BEQUEST_{i,t} + \alpha_4 WEALTH_{i,t} + \\ &+ \alpha_5 Fsize_{i,t} + \alpha_6 Male_{i,t} + \alpha_7 Educ_{i,t} + \alpha_8 IMILLS_{i,t} + \varepsilon_{i,t} \end{aligned}$$

The remaining variables identify male respondents ($Male$), educational level ($Educ$), and the probability of a given group featuring a positive level of savings ($IMILLS$), that should account for any possible selection or truncation bias. The exclusion restriction being used is the variable 'head employed' and is significant in the first stage regression. Table 6 contains the estimation results for this model.

In Table 6 precautionary accumulation for Model 1 is computed as $s * \frac{\sum \frac{\alpha_2 VAR_{i,t}}{SAVE_{i,t}}}{N}$. The sample is smaller relative to Model 2, as in this model wealth data are needed that have higher non response and also because the sample is split in cohorts. Model 1 beside is based on the sample were the dependent variable is present. We report a precautionary accumulation that is equal to 36%. This is extremely close to the results presented above, for subjective earning variance of both adults in the household.

3.4.2 A model for log wealth

Carroll and Samwick (1998) estimate the following model:

$$\log(W_i) = a_0 + a_1\omega_i + a_2\log(Pi) + a_3^{'}Z_i + a4\xi_i + \varepsilon_i \tag{8}$$

In their empirical application W is wealth (we have about 5% negative wealth outcomes) the term ω denotes the log of the variance of log income[6], P is permanent income, Z is a vector of taste shifters, and ξ represents risk aversion (we will use a dummy for self employed to capture this). We report in Table 6 results of this specification under Model 2 and 3. These are IV regressions where ω is instrumented (due to measurement error) using education, sector and employment. In order to compute the share of precautionary savings into total

[5] We refer to the study of Ventura et al for a more precise definition of these variables.

[6] We compute the within group variance, that is to say the variance per household as we observe income over time.

Table 6: Active savings model

	Model 1	Model 2	Model3
	Estimate	Estimate	Estimate
Age		0.0153	0.0411**
Net worth (WEALTH)	0.0076***		
Proxy for bequest motive (BEQUEST)	-0.0124		
Education head	353.7009**		
Family size	-489.0531***		
Probability to be a saver (IMILLS)	3976.0808***		
Log variance of log income household			0.4107***
Log variance of log income head		0.1738***	
Log permanent income		0.9024***	0.9373***
Household with self employed		0.1286	0.1183
Male	-236.199	0.0178	0.1830***
Variance of future income (VAR)	0.0075***		
Wage gap between cohorts (WDIFF)	0.0693***		
Constant	-1038.3651*	-1.983	-14.2901***
N	3527	5472	5472
Share of precautionary savings	36%	18%	31%

Explanatory note: the marginal effect (share) in the last row for Model 1is computed as $s * \frac{\sum \frac{\alpha_2 VAR_{i,t}}{SAVE_{i,t}}}{N}$. The marginal effect in Model 2 is computed by comparing the actual distribution of wealth with the distribution that would prevail if all households faced the same, small amount of income uncertainty, as explained in the text.

wealth according to this model we compare the actual distribution of wealth with the distribution that would prevail if all households faced the same, small amount of income uncertainty. We do that by first regressing the predicted values in 8 on our empirical proxies listed above for 8, by OLS. In this way we derive a $\hat{a}_1$ that is then employed to compute a new measure of log wealth:

$$\log(W_i^*) = \log(W_i) - \hat{a}_1 (\hat{\omega}_i - \omega^*) \tag{9}$$

where $\hat{\omega}_i$ is predicted in the first stage regression of the log of variance of log income on the instruments set, while $\omega^* = \min(\hat{\omega}_i)$.The marginal effect reported in the table is then computed as

$$\frac{\frac{1}{N}\sum_{i=1}^{N} W_i - \frac{1}{N}\sum_{i=1}^{N} W_i^*}{\frac{1}{N}\sum_{i=1}^{N} W_i} * 100\% \tag{10}$$

The results in Model 3 are comparable to the results in Carroll and Samwick (1998), though the estimated marginal effect of 31% is somewhat smaller. Nonetheless this magnitude also is in line with the rest of our computations.

In Model 2 we have also replicated the analysis of Carroll and Samwick (1998), but now substituting the income of the head of the household in the variance computation (dropping therefore the income of the partener). In line with the findings above it appears that when only the head is accounted for the share of precautionary savings appears smaller. In this setting this should depend on the lower value of the variance of the head's income relative to household income.

4 Conclusions

We show different empirical strategies aiming to quantify the empirical relevance of the precautionary saving motive. Puzzling results in the past had created doubts on this matter as, depending on the way in which income uncertainty was defined, precautionary savings were found as being a negligible or extremely relevant determinant of saving accumulation. In the first case the subjective probability distribution of next year earnings was being used. In the second observed life cycle earnings variance. Those who had tried to reconcile these opposing findings had noted that short term income uncertainty may be too low to motivate immediate savings. We should better instrument this variable as it measures life cycle income uncertainty with error. Alternatively one could look explicitly at a longer time horizon, where uncertainty over a longer time period should indeed be larger. But the longer the period taken into account, the more relevant all worries related to individual time consistency and procrastination become.

In this study we point out that subjective earnings variance over the next year, as reported by the head of the household, may be low. Household uncertainty, including therefore the secondary earner, is much larger due to disagreement about future outcomes within the household. When we take this into account the puzzle is resolved, in the sense that we obtain converging results. Our estimates based on subjective earnings variance and observed life-cycle income variation deliver a very similar picture. In both cases the share of precautionary saving is about 30-40%. However the first approach is based on wealth rates, while the second on active savings and the third on the log of wealth. The converging predictions imply therefore a dissimilar impact of the precautionary motive, as accumulated wealth is not only the sum of per period savings.

This indicates that comparisons between methods are difficult also due to the dependent variable that is being analyzed. Considering that the definition of personal wealth differs by country (pension savings are excluded in Europe), we believe that comparing empirical finding between countries will not contribute to answering further the question on the empirical relevance of precautionary savings.

References

Alessie, R., and F. Teppa (2010): "Saving and Habit Formation: Evidence from Dutch Panel Data," *Empirical Economics*, 38(2), 385–407.

Berben, R. P., K. Bernoth, and M. Mastrogiacomo (2006): "Households' Response to Wealth Changes, Do Gains or Losses make a Difference?," *DNB working paper*, 90.

Bosworth, B. P., and S. Anders (2008): "Saving and wealth accumulation in the PSID, 1984-2005," *CRR WP*, 2008-2, 43.

Carroll, C., and A. Samwick (1998): "How Important is Precautionary Saving?," *The Review of Economics and Statistics*, 80(3), 410–419.

Deaton, A., and C. Paxson (1994): "Intertemporal Choice and Inequality," *The Journal of Political Economy*, 102(3), 437–467.

Guiso, L., T. Jappelli, and D. Terlizzese (1992): "Earnings uncertainty and precautionary saving," *Journal of Monetary Economics*, 30(2), 307–337.

Hochguertel, S. (2003): "Precautionary motives and portfolio decisions," *Journal of Applied Econometrics*, 18, 61–77.

Kapteyn, A., R. Alessie, and A. Lusardi (2005): "Explaining the Wealth Holdings of Different Cohorts: Productivity Growth and Social Security," *European Economic Review*, 49(5), 1361–1391.

Kennickell, A., and A. Lusardi (2004): "Disentangling the Importance of the Precautionary Saving Motive," *NBER working papers series*, 10888, 1–64.

Lusardi, A. (1997): "Precautionary Saving and Subjective Earnings Variance," *Economics Letters*, 57(3), 319–326.

Mazzocco, M. (2004): "Saving, Risk Sharing, and Preferences for Risk," *The American Economic Review*, 94(4), 1169–1182.

Ventura, L., and J. G. Eisenhauer (2006): "Prudence and precautionary saving," *Journal of Economics and Finance*, 30(2), 155–168.

Appendix A: Savings and replacement rates

In this section we aim to show the theoretical relevance of precautionary saving motives within an intertemporal utility maximization framework when a policy parameter, the retirement replacement rate, is called into play. In order to do so we abandon the main line of the paper, that is the intrahousehold interaction. We use a constant relative risk aversion (CRRA) utility function first as suggested by Carroll and Samwick (1998), who argue that CARA utility functions, that we treat next, have the drawback of not accounting for the relative impact of income changes at different quantiles of the income distribution. The first utility function is:

$$U(C) = \frac{C^{1-\gamma}}{1-\gamma}, \gamma > 0 \tag{11}$$

as in our data we do not have consumption one strategy to handle this equation is to use $C = w - s$, where w is wage and s is savings.

We assume a two-period model. In the first period individuals work for their wage. In the second period they receive a retirement benefit with replacement rate α . We also allow for a mean preserving spread k to income as in Ventura and Eisenhauer (2006), such that $E(k) = 0$. This implies that

$c_2 = (\alpha w - k + s)$

We now aim to identify the role of the retirement system (here exemplified by the retirement replacement rate) and uncertainty in determining savings. The Euler equation implied by this model is:

$$(w - s)^{-\gamma} = \beta E_t \left\{ R_{t+1} \left(\alpha w - k + s\right)^{-\gamma} \right\} \tag{12}$$

where R_{t+1} is the return on assets or gross interest rate in period $t + 1$ and β is the individual discount factor, or impatience parameter. This can be rewritten as

$$1 = R_{t+1} \beta E_t \left\{ \left(\frac{(\alpha w - k + s)}{(w - s)} \right)^{-\gamma} \right\} \tag{13}$$

The Euler equation should hold in expectations, as we don't know w_2 nor C_2.

We take the log of 13

$$0 = r_{t+1} + \log \beta + \log \left\{ E_t \left[\left(\frac{(\alpha w - k + s)}{(w - s)} \right)^{-\gamma} \right] \right\} \tag{14}$$

Using polynomial expansions this can be reduced (see Appendix 1) to the following expression:

$$s = \frac{(1-\alpha)w}{2} + \frac{1}{4} \gamma Var \left((\alpha - 1) w - k + 2s \right) \tag{15}$$

Uncertainty (variance term) increases savings. This is achieved through precautionary savings. Also a higher risk aversion parameter increases precautionary savings. $\frac{(1-\alpha)w}{2}$ embodies the intertemporal saving motive (that decreases when the replacement rates increases) and the bequest motive together. This is evidently a simple intertemporal smoothing mechanism, as it takes the average of first and second period consumption. This allows us to derive some testable implications. The first is that wealth accumulation and expected replacement rates are negatively related. The second is that uncertainty about future income increases wealth accumulation. Notice also that if $\alpha = 1$, that is to say if there is no drop in income upon retirement, savings solely depend on the term $\frac{1}{4}\gamma Var\left(2s-k\right)$, that is to say all savings are precautionary and depend on the uncertainty about future consumption. In the empirical approach later on we will proxy this variance term using different definitions depending on the interaction between household members. Evidently the interaction between the variances of the different income earners in the household does not emerge from the solution in 15. A plausible expectation however is that when the household members disagree about the development of future income, precautionary savings should be higher. This results is proven by Mazzocco (2004).

For the moment we find it relevant to show that this intuition does not strictly depend on the choice of the utility function. If we modify the utility function by assuming a CARA function:

$$U\left(C\right) = -\frac{e^{-\eta c}}{\eta}$$

and assume that innovation to income follow a random walk with normally distributed shocks, then $c_t\tilde{}Normal, \sigma^2$.

Expectation in period 1 of marginal utility in period 2, assuming $k = 0$, for simplicity, is then[7]:

$$E\left(-\frac{e^{-\eta(\alpha w+s)}}{\eta}\right) = e^{E(-\eta(\alpha w-k+s))+\frac{1}{2}Var(-\eta(\alpha w+s))} = e^{-\eta(w-s)}$$

where the last equality depends on the Euler condition where $e^{-\eta(w-s)}$ is the marginal utility of current period.

By taking logs and rearranging we get to a closed form solution:

$$s = \frac{w\left(1-\alpha\right)}{2} + \frac{1}{2}\eta\sigma^2.$$

All the implications discussed above in the CRRA case are also valid in the CARA case, including the fact that we do not explicitly consider the household members separately in the theoretical discussion.

Empirical estimation of a consumption model is less interesting in our setting, as different authors have already noticed that the residual definition of C

[7] If $x\ \tilde{}N(E(x), \sigma^2)$, then $E(e^x) = e^{E(x)+\frac{1}{2}\sigma^2}$

could be troublesome, due to the fact that s is typically imputed using wealth data (as it is the case in our data), rather then directly observed. We have verified the implications sketched above empirically with a series of reduced form models, where the dependent variables used above are confronted with the expected retirement replacement rate. Due to the limited amount of observations none of these computations was reliable.

Appendix B: formal derivation

Taking the log of 13 we had derived

$$0 = r_{t+1} + \log\beta + \log\left\{E_t\left[\left(\frac{(\alpha w - k + s)}{(w - s)}\right)^{-\gamma}\right]\right\}$$

The second order Taylor expansion of $\log\left(\frac{\alpha w - k + s}{w - s}\right)^{-\gamma}$ around its unconditional mean is, simplifying the notation with $C^* = \frac{\alpha w - k + s}{w - s}$:

$$\log\left(C^*\right)^{-\gamma} \simeq \log E_t\left(C^*\right)^{-\gamma} + \left(\frac{C^{*-\gamma} - E_t\left(C^*\right)^{-\gamma}}{E_t\left(C^*\right)^{-\gamma}}\right) - \frac{1}{2}\left(\frac{C^{*-\gamma} - E_t\left(C^*\right)^{-\gamma}}{E_t\left(C^*\right)^{-\gamma}}\right)^2 \tag{16}$$

Notice that $\left(\frac{C^{*-\gamma} - E_t(C^*)^{-\gamma}}{E_t(C^*)^{-\gamma}}\right)$ could be simplified with a first order Taylor expansion [8], this implies that equation 16 can be rewritten as:

$$\log\left(C^*\right)^{-\gamma} \simeq \log E_t\left(C^*\right)^{-\gamma} + \left(\frac{\left(C^*\right)^{-\gamma} - E_t\left(C^*\right)^{-\gamma}}{E_t\left(C^*\right)^{-\gamma}}\right) + \tag{17}$$

$$-\frac{1}{2}\left(\log\left(C^*\right)^{-\gamma} - \log E_t\left(C^*\right)^{-\gamma}\right)^2 \tag{18}$$

Taking the expected value of 17:

[8] Take the function
$f(x) = \frac{e^x - E_t X_{t+1}}{E_t X_{t+1}} = \frac{e^x}{\nu X_{t+1}} - 1$, around $\bar{x} = \log\left(E_t X_{t+1}\right)$, where $x = \log\left(X_{t+1}\right)$ .Since $f(x) \simeq f(\bar{x}) + f'(\bar{x})(x - \bar{x})$,
in this case $\frac{X_{t+1} - E_t X_{t+1}}{\nu E_t X_{t+1}} = \underbrace{\frac{e^{\log\left(X_{t+1}\right)}}{E_t X_{t+1}} - 1}_{\text{function}} \simeq$

$$\underbrace{\frac{e^{\log\left(E_t X_{t+1}\right)}}{E_t X_{t+1}} - 1 + \frac{e^{\log\left(E_t X_{t+1}\right)}}{E_t X_{t+1}}\left(\log\left(X_{t+1}\right) - \log\left(E_t X_{t+1}\right)\right)}_{\text{Taylor Expansion}} = \log\left(X_{t+1}\right) - \log\left(E_t X_{t+1}\right)$$

$$E_t \log (C^*)^{-\gamma} \simeq E_t \log E_t (C^*)^{-\gamma} + E_t \left(\frac{(C^*)^{-\gamma} - E_t (C^*)^{-\gamma}}{E_t (C^*)^{-\gamma}} \right) + \tag{19}$$

$$-\frac{1}{2} E_t \left(\log (C^*)^{-\gamma} - \log E_t (C^*)^{-\gamma} \right)^2 \tag{20}$$

Notice that the expected value of the deviation from the unconditional mean is equal to zero and that $E_t \log E_t (C^*)^{-\gamma} = \log E_t (C^*)^{-\gamma}$. We can therefore rewrite 19 and rearrange the terms such that:

$$\log E_t (C^*)^{-\gamma} \simeq E_t \log (C^*)^{-\gamma} + \frac{1}{2} E_t \left(\log (C^*)^{-\gamma} - \log E_t (C^*)^{-\gamma} \right)^2 \tag{21}$$

Substituting for $\log E_t (C^*)^{-\gamma}$in the RHS of 21 from the same 21:

$$\log E_t (C^*)^{-\gamma} \simeq E_t \log (C^*)^{-\gamma} + \frac{1}{2} E_t \left[\begin{array}{c} \log (C^*)^{-\gamma} - E_t \log (C^*)^{-\gamma} \\ -\frac{1}{2} \left(\log (C^*)^{-\gamma} - \log E_t (C^*)^{-\gamma} \right)^2 \end{array} \right]^2 \tag{22}$$

We neglect the terms of 3rd or higher order as we are taking a second order Taylor expansion:

$$E_t \left\{ \left[\begin{array}{c} \log (C^*)^{-\gamma} - E_t \log (C^*)^{-\gamma} + \\ -\frac{1}{2} \left(\log (C^*)^{-\gamma} - \log E_t (C^*)^{-\gamma} \right)^2 \end{array} \right] \right\}^2 \simeq E_t \left[\left(\log (C^*)^{-\gamma} - E_t \log (C^*)^{-\gamma} \right)^2 \right] \tag{23}$$

Using 23 into 22:

$$\log E_t (C^*)^{-\gamma} \simeq E_t \log (C^*)^{-\gamma} + \frac{1}{2} E_t \left[\left(\log (C^*)^{-\gamma} - E_t \log (C^*)^{-\gamma} \right)^2 \right] \tag{24}$$

that is also better as:

$$\log E_t (C^*)^{-\gamma} \simeq E_t \log (C^*)^{-\gamma} + \frac{1}{2} Var \left(\log (C^*)^{-\gamma} \right) \tag{25}$$

We now replace back $\frac{\alpha w - k + s}{w - s} = C^*$. This implies that:

$$\log E_t \left(\frac{\alpha w - k + s}{w - s} \right)^{-\gamma} \simeq E_t \left[-\gamma \left((\alpha w - k + s) - (w - s) \right) \right] + $$

$$+ \frac{1}{2} \gamma^2 Var \left((\alpha w - k + s) - (w - s) \right)$$

If we now return to 14 we can rewrite:

$$0 = r_{t+1} + \log\beta + E_t\left[-\gamma\left((\alpha w - k + s) - (w - s)\right)\right] + \quad (26)$$

$$+\frac{1}{2}\gamma^2 Var\left((\alpha w - k + s) - (w - s)\right) \quad (27)$$

Rearranging:

$$E_t\left(\alpha w - k + s\right) - (w - s) = \sigma r_{t+1} + \sigma\log\beta + \frac{1}{2}\gamma Var\left((\alpha w - k + s) - (w - s)\right) \quad (28)$$

thus

$$E_t\left(\alpha w - k + s\right) - (w - s) = \sigma r_{t+1} + \sigma\log\beta + \frac{1}{2}\gamma Var\left((\alpha - 1)\, w - k + 2s\right) \quad (29)$$

We assume without loss of generality that $E_t(k) = 0$, as this is a mean preserving spread of income. And that the two discounts cancel out each other. Further notice that $E_t\left(\alpha w - k + s\right) - (w - s) = \alpha E_t(w) + E_t\left(s\right) - w + s = \alpha w + s - w + s = (\alpha - 1)\, w + 2s$

Rearranging:

$$(\alpha - 1)\, w + 2s = \frac{1}{2}\gamma Var\left((\alpha - 1)\, w - k + 2s\right) \quad (30)$$

where:

$Var\left((\alpha - 1)\, w - k + 2s\right) = w^2 Var\left(\alpha - 1\right) + \underbrace{Var(k)}_{\sigma_k^2} + 4Var(s) + \underbrace{Cov((\alpha - 1)\, w, k)}_{0} +$
$2wCov((\alpha - 1)\, , s) + \underbrace{Cov(k, 2s)}_{0} =$

$w^2 Var\left(\alpha - 1\right) + \sigma_k^2 + 4Var(s) + 2wCov((\alpha - 1)\, , s) = w^2 Var\left(\alpha - 1\right) + \sigma_k^2 + 4E(s - \bar{s})^2 + 2wCov((\alpha - 1)\, , s)$

thus:

$$s = \frac{(1 - \alpha)w}{2} + \frac{1}{4}\gamma Var\left((\alpha - 1)\, w - k + 2s\right) \quad (31)$$