Analysis of the Precautionary Saving Motive Based on a Subjective Measure (SAVE 2005-2007)[1]

Michael Ziegelmeyer[†]

February 2009

[1] The following text was written as my diploma thesis at the beginning of 2008. The author gratefully acknowledges the excellent support and advice of Dr. Michela Coppola in supervising this diploma thesis. The author is also grateful to the MEA seminar participants and Professor Börsch-Supan for their valuable comments. Katharina Flenker and Armin Rick provided excellent work in proof-reading this thesis.

[†] Mannheim Research Institute for the Economics of Aging (MEA); University of Mannheim; L13,17; 68131 Mannheim; Germany; e-mail: Ziegelmeyer@mea.uni-mannheim.de

Abstract:

The importance of the precautionary saving motive for households' saving behaviour is unquestioned in the literature of the last two decades. However, the magnitude of precautionary savings and its influencing factors could not be satisfactorily determined. A subjective measure of the desired amount of precautionary savings in the German SAVE study 2005-2007 allows for the evaluation of these questions on a new basis without relying on a specific definition of wealth. This study supports the view of a low or at most moderate contribution of precautionary savings to wealth accumulation ranging from 6% for total net wealth to 26% for financial wealth. Carroll's buffer-stock model (1992, 1997) is introduced since it establishes the theoretical foundation of this study. In a multivariate data analysis, the panel structure and the wide range of theoretically relevant variables available in the SAVE datasets allows investigating the factors which influence the subjective measure of the desired amount of precautionary savings. Education, current income, variables which reflect a precautionary attitude, bequest and old-age provision motives have strong and significant effects in the expected way. The variables for individual income risk have the expected sign over all specifications. However, they remain insignificant at the usual significance levels, which can be due to the construction of these variables or the German social insurance and welfare system.

Contents

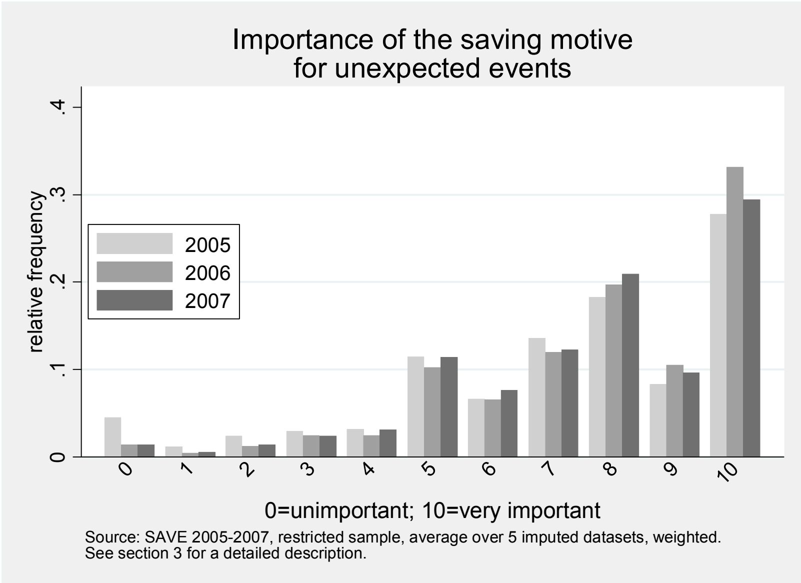
Tables

Figures

1 Introduction

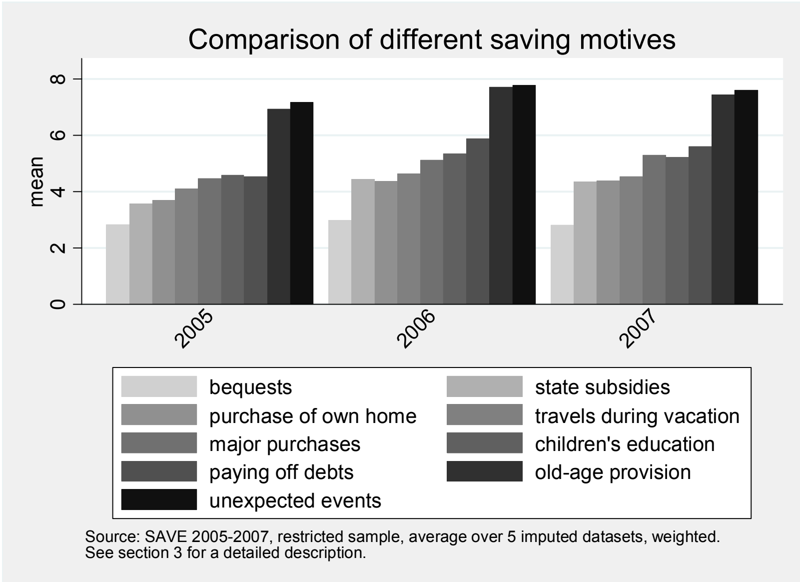
The decision on how much to save and therewith how much to consume is one of the most important economic decisions households have to face. However, this decision is still not very well understood. The complexity of this decision is founded in the intertemporal horizon and in the forward-looking behaviour of agents, as well as the different motives to save and the interaction among these motives. Generally, today's savings are tomorrow's consumption which causes utility in one way or another. The range of how the savings could be consumed is wide, whether invested in daily needs, in the purchase of a home, in the children's education, in bequests to children for their consumption, or in travelling. Another saving motive, the precautionary saving motive, has gained more and more attention in the last 20 years, and its fundamental role has been emphasised in a wide range of economic literature. The idea of precautionary savings is summarised in the additional savings accumulated for preparing for unforeseen events. Precautionary savings as any other kind of savings can be increased in two ways: first, a reduction in consumption at constant labour supply or, second, a rise in the labour supply at constant consumption (Carroll & Kimball, 2007, p. 2).

Figure 1: Importance of the saving motive for unexpected events



But how important is the precautionary saving motive? One possible way to answer this question is a comparison of the importance of different saving motives. Looking at the German SAVE study, which is explained in detail in chapter 3, figure 1 shows that most heads of households assign a relatively high importance to the precautionary saving motive given the range from 0 (unimportant) to 10 (very important). Figure 2 compares the mean of the importance over different saving motives. It can be seen that together with the old-age provision motive, the motive for savings for unforeseen events is the most important saving motive.

Figure 2: Comparison of different saving motives



Similar results were found in the Survey of Consumer Finances (SCF) for the US (Carroll, 1997, p. 1). Understanding the saving process serves not only a self purpose, but rather it is needed for a well-directed economic policy. In the case of precautionary savings, public social insurances have an impact on certain risks like health, unemployment (Kazarosian, 1997, p. 246) or the risk of poverty in old age, which influence the amount of precautionary savings and automatically the personal saving rate.[2] The personal saving rate

[2] As Krueger and Kubler (2006, pp. 737-738) state, a pay-as-you go social security system crowds out private savings, reduces the capital stock and diminishes the growth of the economy of future generations compared to an investment-based social security programme. On the other side it

itself is linked to the capital stock and the capital stock to economic growth. Further on, there is a link between precautionary savings and the impact of government debt as Barsky & Mankiw & Zeldes (1986) examined.[3] From these points of view, understanding the precautionary saving motive, its influencing factors, its role in the saving process, and the linkage to other saving motives over the life-cycle is of fundamental relevance for understanding the saving behaviour of households.

This study deals with the analysis of the precautionary saving motive in the Germany using a subjective measurement of the desired amount of precautionary savings of the SAVE study 2005-2007. The aim of this study is to answer the following research question: "Which factors influence the subjective measure of the desired amount of precautionary savings and to what extent?" To answer this question, the existent theoretical models and empirical work are investigated first. Afterwards, the findings are used to prepare and analyse the SAVE dataset from 2005-2007. For this analysis, the SAVE dataset offers a wide range of appealing characteristics to answer this question:

allows for intergenerational risk sharing. This comes into play if there are aggregate shocks to labour and capital income and the returns to wages and capital are imperfectly correlated. This again reduces savings via the precautionary channel.

3 If future income is uncertain, a tax cut today, which is offset by an increase in future taxes, reduces uncertainty about future wealth since the tax cut provides certain wealth today. This reduced uncertainty leads to lower precautionary savings and to higher consumption today.

- Different definitions of wealth variables have been used as the explained variable to proxy precautionary savings. The new subjective measure of the desired amount of precautionary savings allows using the variable directly as the explained variable preventing the circumstance that the proxies are not a one-to-one map of precautionary savings. This procedure was chosen by Kennickell & Lusardi (2005), who were the first to use a direct measure of precautionary savings as an explained variable.
- All studies dealing with precautionary savings face the problem of creating a measurement of risk on the right hand side which is both observable or constructible and exogenous (Browning & Lusardi, 1996, p. 1835). In addition to income risk, this study implements measures for different kinds of risk that should influence precautionary savings from a theoretical point of view. Recent studies account for longevity risk (Palumbo, 1999; Lusardi, 1998; Kennickell & Lusardi, 2005), health risks (Kennickell & Lusardi, 2005; Kong & Lee & Lee, 2007), business risk (Kennickell & Lusardi, 2005), and pension uncertainty (Murata, 2003) in addition to the well known income risk. The SAVE survey offers proxies for most of the factors mentioned.
- To estimate consistent coefficients of the variables of interest, SAVE makes it possible to include controls which have been found to be relevant from a theoretical and empirical perspective.

- Since the omitted variable bias is always a problem, the panel structure of the SAVE survey can control time invariant, unobserved variables. This increases the accuracy of the estimates.
- For each year five multiple imputed SAVE datasets are available. Inference, which takes not only the variance within each dataset into account but also the variance between the datasets themselves, can be drawn on the basis of these five multiple imputed datasets.

After the clarification of the research question and the possibilities offered by the SAVE dataset, definitions of the subject of interest are provided next.

1.1 Definitions related to the precautionary saving motive

To make the subject of this course clear, definitions are provided in this section. It is helpful to call to attention the distinction between precautionary saving and precautionary savings. As Carroll & Kimball (2007) point out in their introduction, precautionary saving is a flow figure, whereas precautionary savings is the size of stock of precautionary saving accumulated in the past. After this clarification, I use precautionary savings and precautionary wealth as synonyms. Leland (1968), who was the first to set up a two period model that included the precautionary saving motive, defines "precautionary demand for saving … as the extra saving caused by future income being random rather than determinate" (Leland, 1968, p. 465). What Leland meant by being random is not clear. In economics, a distinction between risk and uncertainty is made (Menz, 2007, pp. 9-10). The economic understanding of risk is that an individual can assign certain probabilities to different states of nature, whereas uncertainty cannot be quantified (Dictionary of economics, 2003, pp. 338, 390). Since risk contains measurable probabilities, insurances can step in. Carroll & Kimball (2006, p. 2) define that "precautionary saving is additional saving that results from the knowledge that the future is uncertain." My interpretation of this definition is that precautionary saving results from perceptional uncertainty. Further on, they expand future uncertainty beyond income uncertainty.

To define precautionary savings in a more satisfactory way for this study, “background uncertainty” is defined first. In this study “background uncertainty” is defined as the perceptional risk and uncertainty influencing present or future consumption of an individual, whereby, first, the individual is not able or, second, not willing to insure against the risk. The first point comes from the fact that there is a lack of insurance based on the absence of full information or control of the insurance company, leading to adverse selection or moral hazard of the individuals insured. Moreover, an insurance contract cannot be feasible, because the estimation of the damage distribution function is not known such as for natural disasters or terrorist attacks. Second, individuals may not be well informed about available insurance contracts or information and evaluation costs for insurances against every insurable risk may be too high. Precautionary savings are thus defined as the additional amount of savings of an individual facing background uncertainty compared to an individual with the same expected income path without facing any background uncertainty. This definition of precautionary savings is roughly similar to the definition of Gourinchas & Parker (2002, pp. 75-76). They split total wealth in buffer and life-cycle wealth, whereby the buffer wealth corresponds closely to the definition of precautionary savings used here. This discussion and the quotes of definitions of precautionary savings should help to better understand what precautionary savings stand for, and it should raise awareness of the complexity of this topic. As Essig (2005, p. 6) mentioned, the exact definition of precautionary savings, the question of which risks it should insure against and in which time horizon, is still not clear.

1.2 Empirical evidence and magnitude of precautionary savings

Much work has been done related to the precautionary saving motive based on different approaches. This study is classified in the micro-empirical work on precautionary savings, which tries to link wealth and uncertainty directly. In contrast to simulation approaches like Zeldes (1989b) or Carroll (1997), micro panel data allow for a direct test whether people change their behaviour due to changes in risk according to theoretical predictions. Compared to structural estimation approaches (Gourinchas & Parker, 2002; Cagetti, 2003), it is less restrictive in its assumptions. However, the estimation of model implied parameters is not possible since no well defined model underlies the estimation.

Empirical evidence on the importance of precautionary savings is mostly based on reduced form regressions of net worth or financial assets on proxies for income risk. The regression results range from no significant influence of the precautionary saving motive to the accumulation of different measures of wealth (Skinner, 1988), to a small influence between 1%-4.5% (Guiso & Jappelli & Terlizzese, 1992; Lusardi, 1997, 1998; Arrondel, 2002), and to results up to more than 50% (Carroll & Samwick, 1998). Two relatively recent studies quantified the contribution of precautionary savings to financial wealth in Germany. Whereas Bartzsch (2006, p. 15) found a share of precautionary savings relative to financial wealth of 20.6% on average, Fuchs-Schündeln & Schündeln (2005, p. 1101) identified a contribution of 22.1% for East Germany and 12.9% for

West Germany. For an overview of studies investigating the importance of the precautionary saving motive see table A.3 in the appendix. This table provides information about the authors, the country, the dataset and the investigated time period, the dependent and independent variables, and the main results. This overview is not limited to regression based studies, but also includes examples of structural estimation approaches and Euler equation methods like that one of Dynan (1993). This wide range of results offered in table A.3 is due to different dependent variables caused by different definitions of wealth, unequal controls and risk variables, as well as varying estimation techniques and functional forms in addition to different datasets in varying time periods.

The variety of empirical results makes an ongoing investigation necessary. After the importance of the precautionary saving motive has been clarified, definitions have been provided and the research question of this study and its new possibilities have been introduced; the remainder of this study is organised as follows:

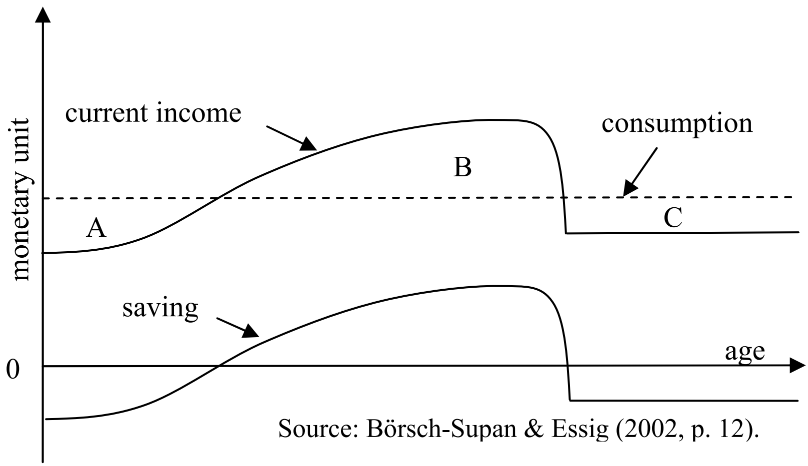
Section 2 gives an overview of the theoretical modelling of precautionary savings; section 3 introduces the dataset of the German SAVE Survey of 2005-2007 and discusses the applied imputation procedure in more depth; section 4 presents descriptive statistics of precautionary savings, whereas section 5 establishes the underlying empirical equation for the estimation; section 6 outlines the estimation techniques adapted, and section 7 describes and discusses the results; section 8 concludes and gives a perspective for further research.

2 Theory

2.1 The standard life-cycle model

Before this study goes into the theoretical details of models incorporating the precautionary saving motive, the origin of these models, the standard life-cycle model, is introduced first. In contrast to the traditional Keynesian consumption function in which consumption is a function of current disposable income, the permanent income hypothesis of Friedman (1957) argues that consumption is a function of permanent income, where permanent income is defined as the agent's total lifetime resources divided by the number of periods the agent is going to live (Romer, 2006, pp. 347-348). In addition, Modigliani & Brumberg (1954) showed that rational agents try to smooth their consumption over their entire life-cycle. Hence, the neoclassical life-cycle-permanent-income hypothesis was born (hereinafter referred to as the standard life-cycle model). The result of the standard life-cycle model derives from an intertemporal utility maximisation problem of a rational, forward looking agent who is maximising his life-time utility by choosing the optimal amount of consumption in each period. Assuming that the life-cycle utility is the sum of discounted values of future utility, where the utility function is additively separable over time and the function is the same over all periods and has a positive but decreasing marginal utility, the agent tries to smooth consumption over his life-cycle. The instruments to do this are borrowing and saving. The agent will borrow, if his current income is below his permanent income, and he will save, if his current income is above his permanent income. Moving life-cycle resources from one period to another is undertaken until the marginal utility of consumption is constant over time. Figure 3 pictures the typical development of consumption, current income, and saving over the life-cycle.

Figure 3: Consumption, income, and saving over the life-cycle



Source: Börsch-Supan & Essig (2002, p. 12).

As shown in the figure, young households should borrow money to finance their consumption level (phase A), whereas households should pay off their accumulated debts and accumulate wealth in phase B, from which they can consume in phase C. Thus, the standard life-cycle model induces an old-age provision motive (Schunk, 2007b, p. 9). The more interested reader is referred to the articles of Browning & Lusardi (1996) and Rodepeter (1999), which give a detailed overview of the theories and the assumptions of the standard life-cycle model and its extensions.

However, the standard life-cycle model cannot explain a wide range of empirical puzzles. Only two puzzling facts are mentioned here. First, as Börsch-Supan et al. (2003) point out, in Germany elderly people do not dissave on average, since median or mean savings and also saving rates by cohort or cross section remain positive in old age despite a very generous pensions and health systems. This result is in contrast to the predictions of life-cycle models. To introduce the bequest motive can help to explain the so-called "German Saving Puzzle", but as mentioned by Börsch-Supan et al. (2003, p. 15), bequests can be the result of an accident, a strategy, or decreased consumption caused by weak health. A result which underpins this argumentation is that bequests do not seem to be related to the number of children in a household (Börsch-Supan & Essig, 2002, pp. 16-17; Hurd, 1987).

The second puzzle, which is strongly related to the importance of the precautionary saving motive in the life-cycle process, is the evidence Carroll & Summers presented in their article in 1991. They found that consumption growth tracks income growth very closely over the life-cycle in different occupational groups (pp. 318-327) and that consumption growth is high in countries in which income growth is high and vice versa (pp. 308-315). These stylized facts contradict important results of the rational expectation version of the permanent-income hypothesis for consumption. To name only one, there should be no parallel movement of income and consumption, since the anticipated growth rate of income should not be correlated to the growth rate of consumption (p. 306).

For the inverted u-shape of income and consumption within a working life, known as the income/consumption parallel in low frequency data, researchers investigated several explanations (see Browning & Crossley (2001, pp. 12-14) for a summarisation of the explanations). One of the most promising explanations (Carroll, 1997, pp. 34-38) was enabled through the extension of the standard life-cycle model with the introduction of a precautionary motive (Browning & Lusardi, 1996, p. 1798). This saving motive allows possible explanations for other empirical observations too (Guiso et al., 1992, p. 308). The next section of this study summarises the development of life-cycle models extended by the precautionary saving motive until they were able to explain the inverted u-shape of income and consumption mentioned above. Subsequently, one of the most famous models including the precautionary saving motive is introduced from which an explanation for the income/consumption parallel in low frequency data can be derived.

But before the theory of precautionary saving is expanded, other extensions of the life-cycle model and new approaches must also receive special attention and should be briefly mentioned; that is liquidity constraints and new approaches through behavioural economics (Browning & Lusardi, 1996, pp. 1846-1848). Liquidity constraints appear if households are not able to borrow. Binding as well as not currently binding liquidity constraints reduce consumption today (Zeldes, 1989a, p. 314). Deaton's (1991) interaction of liquidity constraints and an uncertain labour income induces precautionary savings.

The new approaches through behavioural economics question the general framework of life-cycle models. Thaler (1994, pp. 186-188) addressed three criticisms to life-cycle models. First, to find the optimal consumption rule, the household has to solve an intertemporal dynamic maximisation problem. This is even hard for an economist, since these problems can be quite complex. In addition to that, the household has to make assumptions about, for example, its utility function and has to build expectations about future income flows, the likelihood of dying, and risk of medical expenditures. In addition, the possibility of learning is practically very limited. Only learning from others or following good rules of thumb may help to get close to the optimal consumption path. Rodepeter & Winter (1999) evaluated different rules of thumb compared to the optimal solution. Not only bounded rationally is a concern, which is founded in limited mental capabilities, but also the lack of self-control some households face, which is the second criticism of Thaler. Today households are seduced into consumption. Their preferences are time-inconsistent. A relative new model, which is able to incorporate time-inconsistent preferences and other phenomena, is the "dual-self" model of Fudenberg & Levine (2006). The third point mentioned by Thaler is the assumption of fungibility. Households do not treat every source of wealth equally for their consumption purposes. Their changing propensity to consume out of wealth is assumed to be founded in different “mental accounts” (Thaler, 1990). Other points which should be added are the influences of social norms on an individual’s behaviour (Akerlof, 2007) or the effect of social context, which is paraphrased as “keeping up with the Jones” (Lavoie, 2004, pp. 641, 647).

2.2 Development of models including uncertainty

In 1968, Leland was the first who introduced the effect of uncertainty on savings in a formal analysis of a two-period model of consumption. The setup for the two-period model is a known labour income y_1 in period 1 and the distribution of Y_2, the uncertain labour income in period 2. Assuming additive utilities ($U_G = U_1+U_2$), the consumption in period 1 (c_1) should be chosen in a way that $E[U(c_1,(1+r)(y_1-c_1)+Y_2]$ is maximised. His article already points out that the concept of risk aversion alone is not able to capture higher saving by increasing wage uncertainty. Precautionary demand for saving is the result of concavity and a positive third derivative of the utility function. Therefore, certainty equivalence models like that famous one of Hall (1978), which uses a quadratic utility function, are not able to deal with income uncertainty, since the income variance is not included in the consumption function if the third derivative is zero. Leland as well as Sibley (1975, pp. 76-78) and Miller (1976), who expanded Leland's two-period model to a multi-period one, conclude that with a positive third derivative of the utility function, precautionary savings will rise in the presence of income uncertainty. A positive third derivative is, for example, the result of utility functions with decreasing absolute risk aversion.

Based on these insights, Kimball (1990a) introduced the concepts of absolute and relative prudence parallel to the Arrow-Pratt measure of absolute and relative risk aversion. "Prudence", which is a property of utility functions, leads to higher savings with increasing background uncertainty. In contrast to risk aversion, which is based on the negative second derivative of a utility function, the third derivative of a utility function needs to be positive in the presence of prudence. Expanding the life-cycle model through precautionary savings broadens the model's horizon. It is by far less restrictive than the certainty equivalence models and allows for the modelling of more complex behaviour (Browning & Lusardi, 1996, pp. 1798, 1808).

A further step was undertaken by Zeldes (1989b). Zeldes was the first to use numerical methods to approximate the optimal consumption rules of some multi-period problems based on constant relative risk aversion utility functions. These numerical methods are necessary since there are no closed-form decision rules for optimal consumption functions with more realistic assumptions than the certainty-equivalence models, which do not allow for precautionary savings. In particular, this is valid for constant relative risk aversion utility functions (excluding closed-form solutions for constant absolute risk aversion utility functions [which are implausible, see Kimball, 1990b, for arguments]).

Computer simulations were also used by Carroll to develop the so-called buffer-stock model of saving. Since the econometric equation developed in this paper is influenced by this model, the buffer-stock model and its implications are introduced in the next section.

2.3 The buffer-stock model of saving

The buffer-stock model of saving was developed by Carroll (1992, 1997) and is closely related to the models of Zeldes (1989b) and Deaton (1991). The decision-unit in this model is an impatient and prudent household. Impatience is the preference to consume today compared to consuming tomorrow in the sense that the household would like to borrow against future income if the income is constant over time and there is no uncertainty. Furthermore, the household raises savings in the presence of increasing uncertainty, which is implemented by a "prudent" utility function with a positive third derivative. While prudence induces higher consumption in future periods, impatience drives consumption to increase in the presence. This tension leads to an optimal and unique wealth to income ratio if the parameters of the model are chosen in a plausible range as done by Carroll (1992, 1997). If the individual is below the optimal stock of wealth, prudence dominates impatience and increases savings until the optimal stock of wealth is reached. If the opposite is the case, impatience dominates prudence and will decrease savings up to the optimal level. This result is attractive, since it is similar to a rule of thumb behaviour suggested by financial planning guides (Carroll, 1997, p. 46). The behaviour suggested by the financial planning guides is to hold a certain multiple of (permanent) income as buffer-stock for unforeseen events. However, this advice does not answer the question what the optimal wealth to permanent income target should be. Exactly this gap is filled by the buffer-stock model. In the following, the model structure of the buffer-stock model is explained and the main results are presented. Afterwards, the power and limitations of the model are mentioned. The next section concludes with implications of the buffer-stock model for the research question being investigated.

2.3.1 Theoretical foundation of the buffer-stock model

To ease the determination of a solution for the decision problem, the buffer-stock model structures the decision problem rigorously via several assumptions. The consumer faces the following stylized intertemporal maximisation problem (Carroll, 1992, pp. 72-73; Carroll, 1997, pp. 5-6):

$$\max_{C_t} E\sum_{t=0}^{T}\left(\frac{1}{1+\delta}\right)^t u(C_t)$$

$$\begin{aligned}\text{subject to}: W_{t+1} &= (1+r)(W_t + YL_t - C_t)\\ YL_t &= P_t V_t\\ P_{t+1} &= (1+g)P_t N_{t+1}\end{aligned}$$

where the subscript t denotes the time period and T is the time of death, W_t is net wealth, YL_t is non-capital income (labour income), P_t is permanent income, C_t is consumption, V_t is a multiplicative transitory and N_t a multiplicative permanent shock, g is the growth rate of permanent income, δ is the discount rate, and r is the interest rate. Furthermore, a CRRA utility function ($u(C) = C^{1-\rho}/1-\rho$) is assumed, where ρ is the coefficient of relative risk aversion, which equals the coefficient of relative prudence minus one in the case of a CRRA utility function. In addition, the current resources or cash on hand X_t is defined as $X_t = W_t + YL_t$.

Setting $c_t = C_t/P_t$ and $x_t = X_t/P_t$ allows for reducing the number of variables from three to two, and the optimal consumption rule can be written in the form $c_t(x_t)$. In the next step the parameters of the model (r, g, δ, ρ, $\sigma^2(\ln V_t)$, $\sigma^2(\ln N_t)$, $p(YL_t=0)$) have to be estimated or defined. Since the way to model uncertainty is essential for the resulting buffer-stock behaviour, Carroll's procedure is briefly outlined (Carroll, 1992, pp. 64-72). Using non-capital income YL_{it} for each household i over a period from 1976 to 1985 of the Panel Study of Income Dynamics (PSID), a measure of permanent income $\overline{YL_i}$ for each household was calculated as the average of YL_{it} over all time periods t. A $YL_{it}/\overline{YL_i}$ variable was constructed leading to ten observations for each household. The distribution of this variable suggests that there is a certain probability of earning zero non-capital income, which was not due to measurement error. Carroll estimated the probability to be around 0.5% in each period. The rest of the distribution of $YL_{it}/\overline{YL_i}$ indicates that both the transitory and the permanent income shocks follow a lognormal distribution. After a decomposition of the transitory and permanent shock, both were estimated. To obtain conservative estimates, Carroll finally assumed that the values are $\sigma_{\ln N} = \sigma_{\ln V} = 0.1$.

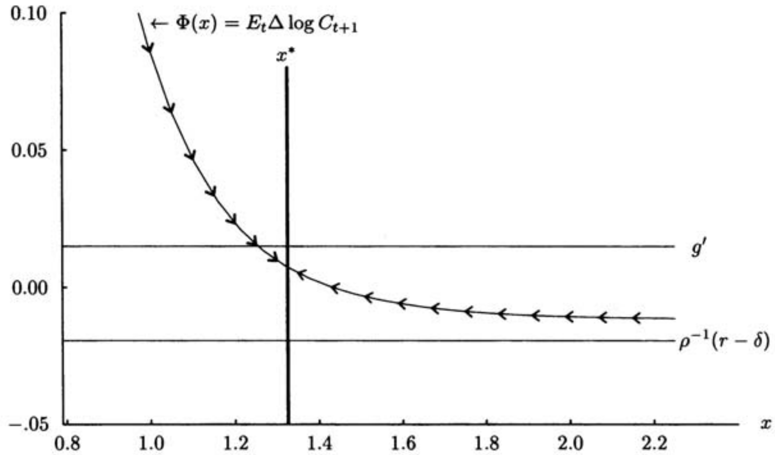
Further on, attention should be paid to the implementation of impatience (Carroll, 1992, p. 74). To be impatient, the consumer must fulfil the condition $\rho^{-1}(r-\delta) < g$, where $\rho^{-1}(r-\delta)$ is the growth rate of consumption under certainty. It is derived from the fact that the expected present value of consumption and income should be equal over the life-cycle. As consumption growth is smaller than permanent income growth g, current consumption is greater than current income and thus the consumer must be impatient.

After estimating and defining all the parameters and their relationship, the optimal consumption rule is the result of a backward induction (Carroll, 1992, pp. 74-76, 128-130). In the case of certainty about the time of death and the absence of a bequest motive, it is optimal to consume all resources left over in the last period. Knowing the optimal behaviour in the last period, the consumer can solve backward the period-by-period Euler equations. Carroll found that the optimal consumption rules converge rapidly if the remaining life time is longer than ten years (Carroll, 1992, p. 75). The buffer-stock behaviour of consumers can be explained best by the resulting log-linearised consumption Euler equation of the form (Carroll, 1997, p. 10):

$$E_t \Delta \ln C_{t+1} \approx \rho^{-1}(r-\delta) + \rho/2 \, \mathrm{var}_t(\Delta \ln C_{t+1}) + e_{t+1},$$

where $E_t \Delta \ln C_{t+1}$ is the expected consumption growth rate, $\rho^{-1}(r-\delta)$ is the consumption growth rate under certainty, and $\rho/2 \, \mathrm{var}_t(\Delta \ln C_{t+1})$ is additional consumption growth influenced by the coefficient of risk aversion or prudence ρ and the uncertainty in the labour income process. Figure 4 summarises some of the key properties of buffer-stock saving. For a more detailed discussion and the theoretical foundations see Carroll 1997 (pp. 11-15) and Carroll 2004:

Figure 4: Expected consumption growth as a function of cash on hand



Source: Carroll, 1997, p. 11: Figure 1a.

- There exists a unique target of the cash on hand to permanent income ratio denoted by x*.
- Households below the optimal target have a higher expected and declining consumption growth rate, since they save more to reach the optimal target value x*, and they reduce their savings when they get closer to the target value x*. Households above the optimal target value have a lower and increasing expected consumption growth rate, since they consume more to reach the optimal target value x*, and they reduce consumption when they get closer to the optimal target x*.
- As cash on hand approaches infinity, the expected consumption growth rate converges to the consumption

growth rate under certainty. As cash on hand converges to zero, the consumption growth rate increases to infinity.

2.3.2 Power of the buffer-stock model

After the setup and the implications of the buffer-stock model have been presented, the power of the buffer-stock model to solve empirical puzzles is investigated next. Carroll (1997, pp. 32-47) demonstrates how the buffer-stock model is able to explain different empirical puzzles. In his analysis, the benchmark of the buffer-stock model is the standard life-cycle model and a Keynesian model ($C = a_0 + a_1Y + u$). The most prominent empirical puzzle and its solution are introduced as follows:

One result of the simulations of the buffer-stock model over different occupational groups is that income tracks consumption relatively closely up to an age of 45 or 50. In other words, if a target cash on hand to permanent income ratio exists, the growth rate of income and consumption must converge in aggregate (Carroll, 1992, p. 91). As outlined above, this corresponds to the stylized fact known as the income/consumption puzzle (Carroll & Summers, 1991). Whereas the Keynesian model can explain this puzzle very easily assuming a_0 to be close to zero and a_1 to be close to one, the prediction of the standard life-cycle model is at odds with this empirical observation (see section 2.1). This result was enforced by other simulations of consumption models. For example, the models of Gourinchas & Parker (2002) and Cagetti (2003) predicted the income/consumption parallel in the early life-cycle too.

Other empirical puzzles the buffer-stock model is able to explain like the consumption/ income divergence in high frequency data (pp. 38-42) and household age/wealth profile (pp. 42-47) are described in detail by Carroll (1997).

2.3.3 Limitations of the buffer-stock model

As indicated, the buffer-stock model is able to provide explanations for several empirical puzzles. But still there are empirical findings that the buffer-stock model is not able to explain. Some of these findings are mentioned, before the first article testing the main results of the buffer-stock model is briefly discussed.

As described above, the buffer-stock behaviour developed by Carroll is a result of the tension between impatience, prudence, and the chance of zero earnings. However, the chance of zero earnings is not a plausible assumption in Germany because of the existing social welfare payments. This leads to serious problems for buffer-stock behaviour in the model, since this behaviour relies on the strong preference caused by the utility function to prevent zero consumption resulting in savings in the previous periods. However, one can argue that instead of the prevention of zero consumption, the household wants to maintain a certain living standard above the living standard which can be guaranteed by social welfare payments. Further on, there seems to be evidence that buffer-stock behaviour is only engaged in by households up to 50 years of age (Carroll, 1997, pp. 3, 36; Carroll & Samwick, 1998, p. 414) or up to age 40 or 45 as estimated by Gourinchas & Parker (2002, p. 49).

Still puzzling is the behaviour of very wealthy households (Carroll, 1997, p. 50). Their enormous financial assets are not consistent with the implications of the buffer-stock model. Furthermore, only financial assets are modelled in the buffer-stock model. Thus, the largest fraction of overall wealth, the housing wealth, (Carroll, 1997, p. 47) is not taken into consideration so far.

Jappelli et al. (2006) were the first to test the implications of the buffer-stock model. The main prediction of the buffer-stock model is that households below their optimal wealth target should save and households above should dissave until the optimal wealth to income ratio is reached. To test this buffer-stock behaviour, the covariance between the gap of current and target wealth and consumption should be in the range between 0.4 and 0.7 (Jappelli et al., 2006, pp. 8-11). The sample was restricted to households with a household head between the age of 20 and 50, since the hypothesis should be tested on the group where the buffer-stock behaviour is most likely to emerge. The data of the 2002 and 2004 Italian Survey on Household Income and Wealth (SHIW) were analysed. This data set provides a subjective measure of the amount of desired precautionary savings similar to SAVE. Since the optimal wealth stock is required for the analysis, they declared the desired amount of precautionary savings as the target wealth. Working with different kinds of robustness checks, the estimated covariances are far below the simulated range resulting from buffer-stock behaviour (Jappelli et al., 2006, pp. 14-21). Further on, they found that the wealth to income ratio of younger households increases with age, which provides indirect evidence against buffer-stock behaviour in

Italy (pp. 21-25, 40) since the buffer-stock model predicts a constant relation between wealth and income, at least for young individuals.

2.3.4 Implications of the buffer-stock model

From the discussion of the theoretical modelling of precautionary savings, several valuable hints were collected for the empirical modelling. Firstly, from a theoretical point of view model relevant parameters are identified. Carroll (1997, p. 19) investigated the degree of the influence of these parameters on the amount of buffer-stock savings. He found that the degree of uncertainty (unemployment risk, transitory and permanent income shocks) and the coefficient of relative risk aversion have a strong influence, whereas permanent income growth, the interest rate, and the discount factor have a weaker influence on buffer-stock savings. These results are volatile with respect to different occupational and educational groups. For the empirical model developed in this study, measures of these variables are constructed, and control variables for different occupational and educational groups are added. Secondly, the variables like consumption or cash on hand were divided by permanent income to get a ratio of these two variables. Statements and implications of the buffer-stock model are provided related to these ratios. Thus, the subjective measure of precautionary savings in this analysis is related to permanent income as well. Thirdly, the theory emphasises the importance of permanent income compared to current income. An aim should be to construct an appropriate measure for permanent income, too. Fourthly, households seem to engage in buffer-stock behaviour only in the first half of the life-cycle. Therefore, it is controlled for different ages of the household head, and other saving motives related to the precautionary motive are added. Fifthly, saving behaviour was found to be different for very wealthy households. Thus, adding wealth dummies helps to absorb different saving behaviours in different wealth groups. See section 5 for a detailed description of the included variables and the motivation behind these variables. But before this issue is addressed, a description of the SAVE dataset and descriptive statistics are provided in sections 3 and 4.

3 The SAVE dataset 2005-2007

3.1 The SAVE survey and its focus

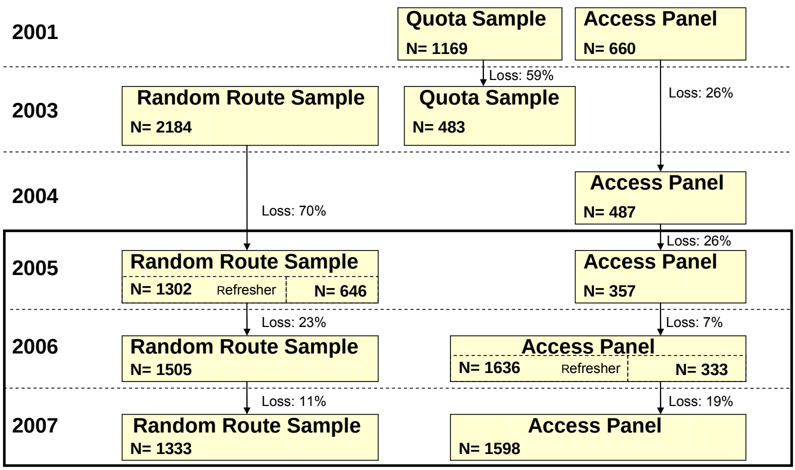
As emphasised by most researchers (Rodepeter, 1999, p. 119-120) and also in the conclusion of the “Household Saving” review of Browning & Lusardi (1996), the puzzling results of saving and consumption behaviour in general and of the precautionary saving motive especially makes more detailed, reliable, and extensive datasets necessary. According to Browning & Lusardi (1996, p. 1849):

> *“To assess whether behavior is consistent with the theory we need information on, for example, health status, subjective perceptions of mortality risk, and the situation of any children.“*

The SAVE dataset introduced in 2001 tries to fill this gap for Germany. In addition to individual health status and life expectancy, a wide range of necessary questions are included to assess risks, uncertainties, and liquidity constraints. It is a hard task to get reliable measures of variables like the personal discount rate, risk aversion or prudence. Attempts have been done in SAVE, but these theoretical concepts are somehow difficult to elaborate. Taking aspects of behavioural economics also into account, variables of self-control and the ability of planning and calculation of future income, consumption or saving streams should also be included to set an extensive framework for a detailed analysis.

Since the question about the subjective measure of precautionary savings was first asked in 2005, this study takes the SAVE survey from 2005-2007 as the base for this analysis. Figure 5 shows the SAVE survey over the years with all its sub-samples. The years 2005 to 2007 consists of two sub-samples. The Random Route sample is a multiple stratified multistage random sample, whereas the Access Panel is a quota sample (Schunk, 2006; Börsch-Supan et al., 2008). Even though the sample technique is different, comparisons of these two different sub-samples show that sample statistics of different variables are quite similar even if the Access Panel might cause a bias through non-randomness based on the latitude of judgement of the interviewer. To increase the available number of observations, both sub-samples are analysed.

Figure 5: Sample design of SAVE



Source: Börsch-Supan et al. (2008).

3.2 The subjective measure of the desired amount of precautionary savings

The question about the subjective amount of precautionary savings (hereinafter simply precautionary savings), on which this work is based, was introduced in the 2005 SAVE survey and was also requested in 2006 and 2007. The question printed in the German questionnaire is translated best in the following manner:

> *"55. About how much do you think you and your family need to have in savings for unanticipated emergencies and other unexpected things that may come up?"*[4]

The first questionnaire to ask this question was the Survey of Consumer Finances (SCF) for the US in 1995.[5] The aim of this question was to measure the desired amount of "buffer-stock" savings, which is a stock figure (Kennickell & Lusardi, 2005, pp. 10-11). For a better understanding of this question, it was placed after a sequence of questions concerning savings (this was done in the SCF as well as in the SAVE survey). In all three years the question before the precautionary savings question was to asses the importance of different saving motives (SAVE: question 54 in

[4] In the German SAVE Study of 2005 this question was an exact translation of the corresponding question in the SCF. The question was phrased in German as follows: "Wie viel Ersparnis benötigen Sie und Ihre Familie zur Vorsorge vor unvorhergesehenen Ereignissen?"

[5] This question was also included in two other surveys: the "Italian Survey of Household Income and Wealth" from 2002 on and the "Dutch CentER Panel" in 2005.

2005). Before that question, the amount of household savings of the previous year was asked (SAVE: question 53 in 2005). Since question 53 asks for a flow figure, it is discussable whether or not the interviewees understood question 55 in the intended way. However, the amounts in the answers on precautionary savings suggest that most of the interviewees gave a stock figure as an answer (see table 6). In addition, the item non-response over all three years (table 1) shows that question 55 was somehow hard to answer.

Table 1: Missing values of the question about precautionary savings

year	# obs	missing values	in %
2005	2298	504	21.9%
2006	3462	894	25.8%
2007	2924	757	25.9%

Source: SAVE 2005-2007.

In contrast to the SCF95 with around 3%, the item non-response is fairly high at above 20% in SAVE. Since this question cannot be seen as a critical one which the respondent does not want to answer, this high item non-response could be the result of that a significant number of the interviewees did not understand the question or they had no idea what their precautionary savings should be. Thus, two typical advantages of subjective measures, the low cognitive burden and the low item non-response rate, are not the case here. Table 2 shows different characteristics of household heads who answered the question and household heads who did not. The argument that the question is somehow hard to understand is confirmed in this table, since missing rates are higher for older household heads, household heads with a basic education, and low income households. A t-test of the differences in the item non-response rates between each of these groups and the rest of the sample is significant at the 1% level.

Table 2: Comparison of some sample characteristics between the answered and not answered questions about precautionary savings

	answered	not answered
age		
< 25	4.29%	7.15%
25 - 64	74.78%	62.65%
≥ 65	20.93%	30.21%
education		
basic (9/10 years)	7.59%	15.97%
higher (13/14 years)	59.18%	60.22%
undergraduate (16/17 years)	15.98%	11.74%
graduate (18/19 years)	17.25%	12.06%
net monthly income		
< €1300	21.38%	31.06%
€1300 - €2599	44.61%	44.54%
≥ €2600	34.01%	24.40%
# obs	6529	2155

Source: SAVE 2005-2007, average from 5 imputed datasets, unweighted, outlier excluded.

Looking at the response behaviour over the years (table 3) gives a puzzling result. Instead of an expected learning process, there were even more respondents who answered the question the year before and had a missing value one year later than the other way around. One typical disadvantage of subjective measures is that the respondents give answers which are socially desirable (Schunk, 2007b, p. 4). This however should play a minor role, since there is no socially desired attitude towards precautionary savings. Maybe there is a tendency that higher precautionary savings are deemed better.

Table 3: Response behaviour over the years

		2006			
		answered	missing value	not participated	total
2005	answered	**1,227**	**268**	306	1,801
	missing value	**197**	**146**	161	504
	not participated	1,156	480	0	1,636
	total	2,580	894	467	3,941

		2007			
		answered	missing value	not participated	total
2006	answered	**1,859**	**334**	387	2,580
	missing value	**315**	**423**	156	894
	not participated	0	0	467	467
	total	2,174	757	1010	3,941

Source: SAVE 2005-2007, unweighted.

It is important to note that the question asked for the “desired” amount of precautionary savings. Thus, the level of precautionary savings which were reported should be unaffected by economic cycles the household has to face.

Based on the fact that deleting interviewees who did not answer the questions included in the analysis may lead to a selection bias, the unanswered questions are imputed using the structure of the dataset and the information given by the person him-/herself and interviewees who answered the question of concern and their characteristics. In my opinion, it is debatable to impute values of precautionary savings to households who did not understand the concept behind this question. However, it seems to be the best method to handle this problem. The imputation procedure and its results are described in more detail in the next section.

3.3 Role of imputation

The SAVE dataset was imputed every year from 2003 onwards using a “Markov Chain Monte Carlo multiple imputation procedure” to fill the missing values with plausible substitutes. For a detailed description of the whole procedure and the implementation see Schunk’s paper about this topic in 2008. This procedure has two important advantages: first, as already mentioned, the missing value of a variable might not be random resulting in biased estimates. Second, if an observed-case analysis is done, this can lead to a serious loss of efficiency due to the sometimes drastically reduced sample sizes.

To allow that the relationships between the observed variables are estimated first, and estimates of these relationships are used to predict the missing values, the missing data must fulfil the “ignorable” criteria (Caneron & Trivedi, 2005, pp. 925-927). For that, two assumptions have to hold: first, the MAR (missing at random) assumption makes sure that the probability of a missing value does not depend on the missing value itself after controlling for the other observed variables, which are correlated to the missing value; second, the parameters for the missing values must be unrelated to the parameters which a researcher wants estimate from the data. The MAR assumption is normally not testable, whereas the second assumption is satisfied in the most cases. Therefore, the imputation procedure should include all relevant variables to estimate the missing values and to conserve the correlation structure of the dataset.

Variables with low missing rates (mainly socio-demographic variables) are imputed first. Then other variables are imputed making use of the additional information of the already imputed variables. After all gaps are filled, the procedure is repeated for most of the variables to impute these variables with as much information as possible, because now all variables can be included in the analysis based on the fact that there are no missing values left. This part is related to the "Markov Chain Monte Carlo imputation" in the name of the imputation procedure above. The procedure is repeated five times to fulfil convergence criteria. After five loops, the procedure stops and one complete dataset is obtained. The overall procedure is repeated five times generating five datasets with different imputed values. This refers to the "Multiple" in the name above. From these five datasets the coefficients and standard errors are calculated according to Rubin's rules (Rubin, 1996, pp. 467, 477; Schunk, 2007a, pp. 37-38). Whereas the new coefficient is the average over the coefficients of the 5 datasets, the new standard errors take not only the within-imputation variance into account but also the between-imputation variance between the 5 imputed datasets. For more details related to this topic, the interested reader is referred to Rubin & Schenker (1986), Rubin (1987, 1996) and Little & Rubin (2002) amongst others. In this study, means and medians of the descriptive statistics are calculated over all five datasets.

The most important question of this analysis, the question about the desired amount of precautionary savings, has amongst others the highest missing rates (see table A.2 for the missing rates of the variables included in this analysis). From this point of view a proper understanding of the result of every imputation step is necessary. To make clear why different imputation procedures were chosen between 2005 and 2006/2007, table 4 is shown first.

Table 4: Zero precautionary savings over the years

year	# obs	# zero obs	in %
2005	1794	168	9.4%
2006	2568	0	0.0%
2007	2167	0	0.0%

Source: SAVE 2005-2007, only observed values, unweighted.

Whereas in 2005, 9.4% of all respondents gave a zero value, the answers for 2006 and 2007 were always positive. This was the reason to add a probit in 2005. The imputation procedure for each year is summarised as follows (table 5):

Table 5: Imputation procedure of precautionary savings

2005	2006	2007
	step 1	
probit	-	-
	step 2	
regression of positive values	regression of all values	regression with trimming (+/- 1.96 standard deviations around the mean, in this case values above €180,000 are excluded for the estimation of the coefficients)
adding a random variable	adding a random variable	adding a random variable
shooting	shooting	shooting
rounding off to hundred	rounding off to hundred	rounding off to hundred

The procedure is described as follows for continuous or quasi-continuous variables (Schunk, 2007a, pp. 13-15). After a probit (only in 2005) and a linear regression, the conditional expected values are estimated and a random variable is added to the estimated values. This added random variable is a normal distributed random variable with a mean equal to zero and a standard deviation equal to the root mean squared error of the regression before. The added random variable is censored to the maximum or minimum of +/- one standard deviation. It should ensure the variability of the imputed data. Since the regression and the added random variable provides estimates that are out of range of the observed values, the shooting process ensures that all values are between the minimum and maximum of the observed data by adding an always newly generated random variable with the appropriate sign, constructed in the same way as described above, to the imputed and out of range value (a negative random variable is added if the predicted value is below the minimum of the observed values, and a positive random variable is added if the predicted value is above the maximum of the observed values). This procedure ceases if the value is located in the observed range.

This method of adding a random variable and shooting is also done by the imputation of other datasets like the Spanish Survey of Household Finances (EFF) (Barceló, 2006, pp. 24-25) and the SCF (Kennickell, 1998). However, this alone is no argument for the validity of this method.

Schunk (2007a, pp. 25-32) compares descriptive statistics of the median and the mean of the observed and the five imputed datasets in 2003/2004. The complete SAVE dataset, which obtains the observed and the imputed data, shows a consistent pattern for financial and saving variables. He found that the mean of the observed data is notably smaller than for the imputed values. Schunk refers to the imputation results of Hoynes et al. (1998), who arrived at similar results. However it may be wrong to conclude that "for most financial asset items, the included conditioning variables shift the distribution to higher values for financial wealth on average …" (Schunk, 2007a, p. 28) is also valid for the amount of precautionary savings. The following thoughts and data evidence related to the amount of desired precautionary savings make this point clear. Table 6 illustrates the situation and compares descriptive statistics of the observed and the imputed data:

Table 6: Descriptive statistics for the observed data and the 5 imputed implicates

	obs. data	imp. data 1	imp. data 2	imp. data 3	imp. data 4	imp. data 5
2005						
# obs	1801	504	504	504	504	504
mean	10655	52094	53729	53255	51503	51018
median	3000	44200	45900	50200	48250	43650
std. err.	88802	39434	39674	37944	37408	40145
min	0	0	0	0	0	0
max	3000000	241900	228100	224000	176100	211900
2006						
# obs	2580	894	894	894	894	894
mean	13673	36852	34852	35324	34951	35274
median	5000	34050	31200	32750	32700	31800
std. err.	56746	23371	24167	22648	22798	23450
min	5	300	0	200	0	0
max	1000000	128600	115000	106000	99600	136000
2007						
# obs	2174	757	757	757	757	757
mean	13989	12987	12758	12502	12792	12685
median	5000	11400	10900	10800	11400	10900
std. err.	85380	8962	8989	9072	8862	9233
min	1	0	0	100	0	100
max	3000000	55200	56600	54000	60500	56700

Source: SAVE 2005-2007, unweighted, no correction for outlier.

Table 6 displays that the mean of precautionary savings is five times higher for the imputed than for the observed values in 2005 and around three times higher in 2006. In 2007 the means are relatively equal. The results may imply that the inclusion of the outlier in 2005 and 2006 increased the mean and the median drastically. This conclusion is overhasty as the subsequent investigation shows.

For this investigation, the imputation procedure has been followed step by step in all three years.[6] After the regression and the probit (only in 2005), the means of the imputed data are relatively close to the observed means. However, the regression leads to negative estimated values (2005: 30%; 2006: 12%; 2007: 6%). The added error term flattens the estimated distribution for the unobserved values. As a consequence, the amount of negative and the mean of the positive values strongly increase, whereas the overall mean remains roughly constant. The positive values are kept and many trials are needed to make the negative values positive by adding additional error terms according to the shooting process described above. Thus, adding the error term and the subsequent shooting process are responsible for this substantial increase in the mean of the imputed values. The value finally imputed often differs remarkably from the value originally estimated by the first regression.[7]

[6] Tables with distributions and summarised statistics can be requested from the author.

[7] A similar pattern is observed for the question about the annual savings the year before. Other imputed variables have to be inspected first.

Excluding outliers in 2007 reverses the positive effects on the mean of the imputed data. The problems of the imputation procedure itself remain. In conclusion, the drawbacks in the imputation procedure and its changes over the years make it necessary to rely only on the observed values of precautionary savings. Therefore, the household who did not answer the question about their precautionary savings are excluded in the subsequent analysis.

3.4 Inflation, outlier, and weights

The euro amounts are not inflation adjusted in this analysis. The gain of an inflation correction is very limited in this case, and the additional assumption that, e.g., the inflation is equal over different groups categorised by region, consumption behaviour, or assets is very strong. Besides, the correction for inflation should not change any estimated empirical evidence below.

The handling of outliers is always critical for the estimation results. If they are not corrected properly, this can result in a serious bias of the estimated coefficients. Since the subjective measure of precautionary savings is the explained and most critical variable in this analysis, using the logarithm as functional form of the explained variable may reduce the problem of outliers. Nevertheless, if the amount of precautionary savings was higher than €200,000 (42 values ≈ 0.64% of the restricted sample 2005-2007), the values were compared over the years for the same household making use of the panel structure. After looking at the panel structure, a correction was made if a decimal place mistake occurred (8 values corrected). Observations were kept if a decimal place mistake could not be assumed and precautionary savings in the one year was not more than ten times larger than in another year (6 values kept). If the household only participated in one year, observations were deleted only if the values were more than €300,000 (2 values deleted). All the other observations above €200,000 were deleted. To sum up, correcting for outliers is always a process including a certain degree of arbitrariness. However, the procedure chosen should ensure that this arbitrariness is kept to a minimum.

The exclusion of the missing values of the question about precautionary savings leads not only to a painful reduction of the number of observations but also to a strong decrease in the fraction of households participating in all three years. Further on, the exclusion of missing values properly leads to the selection bias described at the beginning of section 3.3. In spite of these serious disadvantages, the exclusion seems more appropriate, since the bias caused by the implemented imputation procedure seems to be higher than the bias caused by the limitation on the observed values of precautionary savings. After the correction of outliers and the limitation to observed values of precautionary savings (hereinafter called restricted sample), the following panel structure is obtained (table 7).

Table 7: Panel structure

panel structure	**all values (after outlier correction)**			**restricted sample**		
05 / 06 / 07	# obs	in %	acc. in %	# obs	in %	acc. in %
X X X	1619	41.1%	41.1%	934	27.9%	27.9%
0 X X	1297	32.9%		911	27.2%	
X X 0	205	5.2%	38.3%	281	8.4%	39.2%
X 0 X	5	0.1%		121	3.6%	
X 0 0	469	11.9%		458	13.7%	
0 X 0	341	8.7%	20.6%	442	13.2%	32.9%
0 0 X	3	0.1%		201	6.0%	
	3939	100%	100%	3348	100.0%	100%

Source: SAVE 2005-2007.

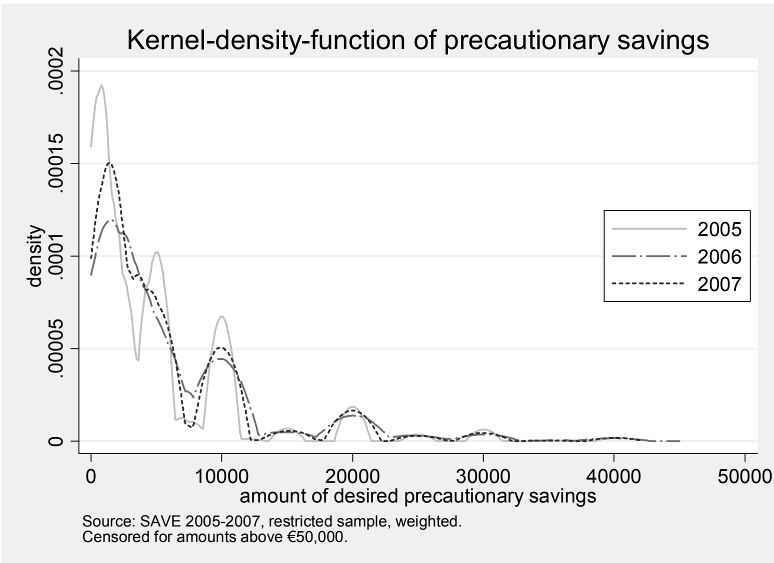
Weighting is a useful operation to eliminate survey bias due to over or under-sampling of certain parts of the population. In the next section the descriptive statistics of the restricted sample are calculated using weights based on the Mikrozensus 2004 for the SAVE dataset of 2005, on the Mikrozensus 2005 for the SAVE dataset of 2006, and Mikrozensus 2006 for the SAVE dataset of 2007. The Mikrozensus is a yearly repeated statistical survey by the German Federal Statistical Office which covers around 1% of all households in Germany. The households are chosen in a way that they are a representative mapping for the German population.

Nine different categories of households are constructed from the restricted sample. They differ in three different net income classes (below €1300, €1300 to €2599, €2600 and above) and three age classes (under 35 years of age, 35 to 54, 55 and above). To calculate the weights the relative frequency of households in the Mikrozensus is divided by the relative frequency of households in the restricted SAVE sample, both part of the same category of households. Despite the fact that the restrictive sample is used, the weights calculated for each year are in a reasonable range (not smaller than 0.55 and not greater than 2.19). Thus, the post stratification weights make sure that the restricted sample is a representative image of the population in the dimensions the weights were calculated. Whereas the necessity of weights is established for descriptive statistics, weights are not unquestioned for the regression based inference. Radbill & Winship (1994, pp. 242-247) demonstrated that unweighted OLS should be preferred over weighted OLS if the weights are only a function of the explanatory variables as it is the case here. If the model is correctly specified, both weighted and unweighted OLS will produce unbiased and consistent estimates. Unweighted OLS yields smaller standard errors and is therefore more efficient. As a result, weights are used in the descriptive analysis. The regression based inference is done without weights.

4 Descriptive Statistics

In this section descriptive statistics are introduced. The focus is on bivariate analyses of the relationship between precautionary savings and explanatory variables, which are found to be important from an empirical and theoretical point of view. The number of descriptive statistics has to be limited, and only some of the most relevant figures are displayed. Before starting with the bivariate analysis, the subjective measure of precautionary savings itself is graphically analysed. Figure 6 plots the kernel-density-function for each year from 2005 to 2007.

Figure 6: Kernel-density-function of precautionary savings



A concentration of observations in the area between €0 and €5,000 inclusively is noticed, whereas the most observations were concentrated between $5,000 and $10,000 in the article of Kennickell & Lusardi (2005, p. 12). Around 67% of all answers are in this range over the whole time period. The amount of the desired precautionary savings is concentrated around the so-called focal points, e.g., €1,000 with 12.6%, €2,000 with 10%, €5,000 with 17.6%, and €10,000 with 14.8% over the complete time period. The following table shows the mean, median, and standard deviation over the years for the restricted sample:

Table 8: Descriptive statistics of precautionary savings

year	# obs	mean	median	std. err.	min	max
2005	1794	6194	3000	14684	0	500000
2006	2568	9330	5000	19400	5	300000
2007	2167	8834	3500	24528	1	1000000

Source: SAVE 2005-2007, restricted sample, weighted.

Whereas the H_0 hypothesis of similarity of the yearly distributions is rejected for 2005/2006, 2005/2007, and 2006/2007 using a two-sample Kolmogorov-Smirnov test of the equality of distributions based on the weighted restricted sample, H_0 cannot be rejected for 2006/2007 with a p-value of $p=0.621$ for the unweighted restricted sample. However, as pointed out above, weights are a useful tool to get a representative sample at least for the two chosen characteristics of income and age. Therefore, weights are used for the ongoing descriptive analysis.

One issue, which has always been of great interest from a macro perspective, is the question about the contribution of precautionary savings to wealth accumulation. The traditional way to get an answer in micro-empirical studies is to calculate the difference between the observed wealth, which includes automatically precautionary savings, and a predicted measure of wealth without precautionary savings. The first step is to regress total or financial wealth on several households' characteristics plus some risk measures. Setting the variables of risk to zero or to their minimal observed values and holding all the other characteristics of the household constant, wealth is predicted using the estimated coefficients (e.g. Carroll & Samwick, 1998, pp. 416-417; Bartzsch, 2006, pp. 13-15). In a second step, the sum of the differences between the observed wealth and the predicted wealth over all households in the sample is taken and divided by the sum of observed wealth over all households: the obtained quotient is assumed to be the average fraction of wealth which is due to precautionary savings. This is the procedure chosen mainly to get the estimates of precautionary savings described in section 1.2. This calculation method has several drawbacks, e.g. the risk measures could be only poor proxies of the real risks the household is facing or they may not capture the whole range of risks the households might want to cover with their precautionary savings.

A completely different way to estimate the contribution of precautionary savings to wealth accumulation is to use a subjective measure of precautionary savings. A first possibility consists in calculating the mean over the whole sample of both precautionary savings and a measure of wealth and dividing the first one by the second one. A second method is to calculate the ratio of the subjective measure of precautionary savings to different measures of wealth for each respondent, taking the median afterwards. The mean is not a useful measure in this case, because of negative values for total wealth and a high fraction of zero financial wealth. Table 9 summarises the results:[8]

Table 9: Contribution of precautionary savings to different measures of wealth

	SAVE				Kennickel & Lusardi[1]	Jappelli et al.[2]
	2005	**2006**	**2007**	**2005-2007**	**1995, 1998**	**2002, 2004**
mean precuationary savings	6194	9330	8834	8298	20855	55137
mean total net wealth	141053	148380	141486	144079		
mean financial wealth	31648	32117	31081	31645		
∑ precuationary savings/ ∑ total net wealth	4.4%	6.3%	6.2%	5.8%	8%	
∑ precuationary savings/ ∑ financial wealth	19.6%	29.1%	28.4%	26.2%	20%	
median ratio of						
precuationary savings/ total net wealth	4.0%	5.7%	5.4%	5.0%		31%
precautionary savings/ financial wealth	30.6%	45.5%	38.5%	38.5%		332%

Source: SAVE 2005-2007, restricted sample, average over 5 imputed datasets, weighted.
[1]Kennickell & Lusardi (2005, pp. 1, 14, 16): the mean in USD is calculated out of the information of three subsamples.
[2]Jappelli et al. (2006, p. 12): the sample was restricted to household heads aged between 20 and 50.

The results are compared to the results in two other articles working with a subjective measure of precautionary savings: Kennickell & Lusardi working with the SCF of 1995 and 1998 and Jappelli et al. working with the SHIW of 2002 and 2004. The results of the SHIW differ remarkably from the SCF, although the question in the SHIW is similar to the question in the SCF. In contrast, the results of the

[8] The definitions of the wealth categories correspond to the definitions chosen in Börsch-Supan et al. (2008).

SAVE study are much more in line with the SCF. In the 2005-2007 SAVE survey precautionary savings account for 5.8% of total net wealth and 26.2% of financial wealth (1st calculation method). It is not clear how the ratios reported in the article of Kennickell & Lusardi (2005, p. 16) are calculated. The author assumes that the calculation was done using the first method, whereas Jappelli et al. (2006, p. 12) interpreted the ratio to be calculated after the second method.

But how does the size of precautionary savings develop over different age classes? In the case of this study, one has to keep in mind that figures 7-9 are the result of both age and cohort effects, which cannot be separated using such a short panel.

Figure 7: Mean of precautionary savings over age classes (2005-2007)

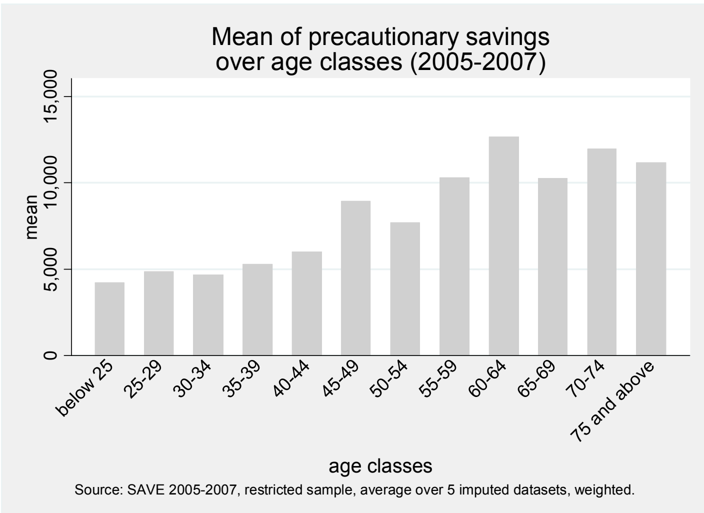
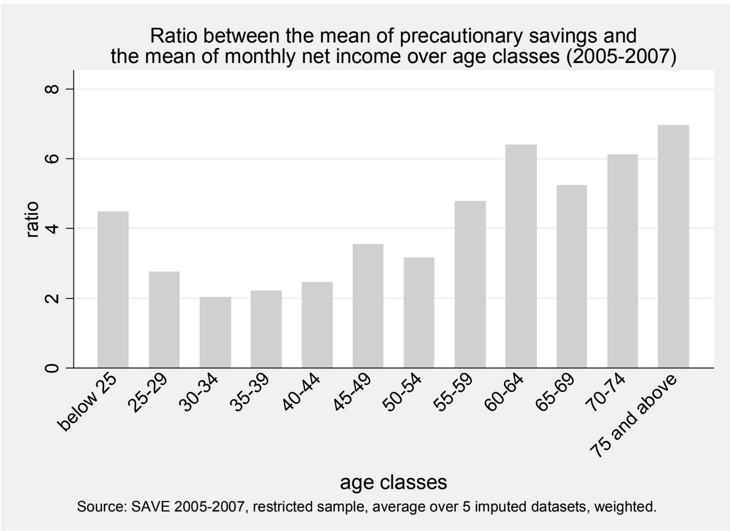


Figure 7 shows that on average precautionary savings increase up to approximately 60 years of age and thereafter remain on a high level. Setting precautionary savings in relationship to monthly net income and financial wealth alters the picture. Whereas the stock of precautionary savings is more than four times higher than monthly net income for households below 25 years of age, the ratio decreases to about twice times the monthly net income of households between 30 and 34 years of age, and increases up to 6 times the monthly net income for households between age 60 and 64 (figure 8; the median ratio of precautionary savings to monthly net income displays a u-pattern too).

Figure 8: Ratio between the mean of precautionary savings and the mean of monthly net income over age classes (2005-2007)



Compared to financial assets, the fraction of precautionary savings is above 70% for households below 25 years of age and quickly decreases to a level of around 20% between 30 and 54, before it slightly increases again to above 30% (figure 9; the median ratio of precautionary savings to financial wealth results in a similar pattern). This u-pattern can also be found for the median ratio of precautionary savings to total net wealth. Thus, precautionary savings seems to be by far more important for young and old households.

Figure 9: Ratio between the mean of precautionary savings and the mean of financial wealth over age classes (2005-2007)

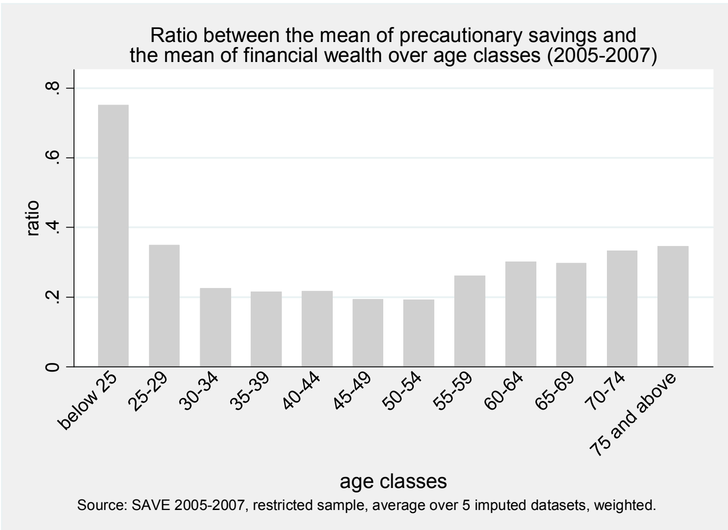
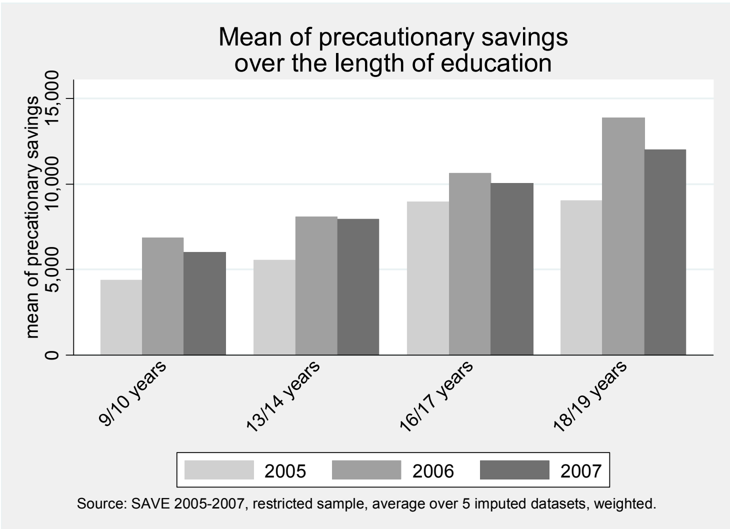


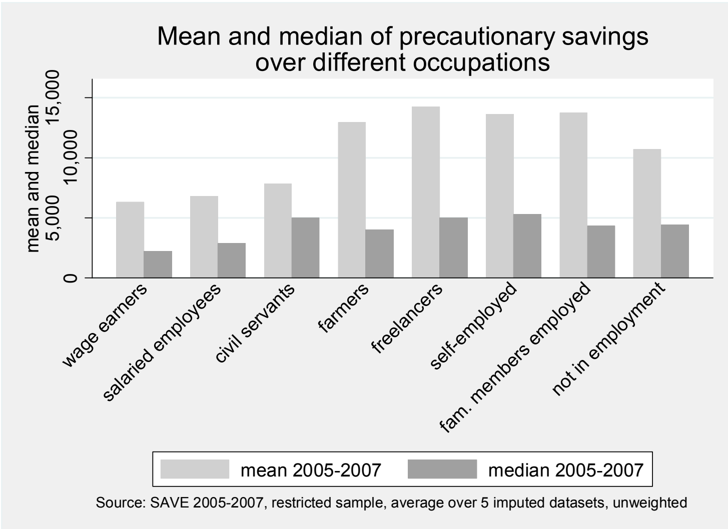
Figure 10 shows the mean of precautionary savings over different educational groups. The groups are constructed from the length of educational and professional training (see table A.1 in the appendix). As can be seen, there is a positive relationship between the length of education and the mean of precautionary savings. The figure plotting the median precautionary savings over the length of education leads to the same result (not shown here).

Figure 10: Mean of precautionary savings over the length of education



Controlling for self-selection into different occupations is important for this investigation, as the self-selection process may be driven by different risk attitudes, which influence the precautionary savings too. Section 5.2.4 gives deeper insights. Figure 11 presents the mean and the median precautionary savings over different occupations. Since civil servants face the lowest income risk, one could expect civil servants to have the lowest mean or median precautionary saving. That self-selection into different occupations comes into play here, can be seen in figure 11.

Figure 11: Mean and median of precautionary savings over different occupations (2005-2007)



There the civil-servants have a similar median amount of precautionary savings than the self-employed and freelancers. Using a mean comparison t-test, the H_0 hypothesis of equal means between the self-employed workers and freelancers, farmers, family members employed in the family business, and individuals who are currently not in paid employment cannot be rejected in each case. Additional controls like a measure of risk aversion may help to obtain the partial effect of being self-employed. The multivariate analysis in the following sections will exploit the possibility to control for a wide range of variables.

As Carroll (1997, p. 19) points out, risk aversion together with the degree of uncertainty is the most influential determinant of buffer-stock savings: the higher the risk aversion, the higher the buffer-stuck savings, holding all other explanatory variables constant. The critical point in micro-empirical studies is to find an appropriate measure for risk aversion. Börsch-Supan & Essig (2002, pp. 87-91) constructed a measure of risk aversion out of the average of five questions about the willingness to take risks with respect to the respondent's own health, career, money, leisure time and sport, and driving.

Figure 12: Median of precautionary savings over different measures of risk aversion (2005-2007)

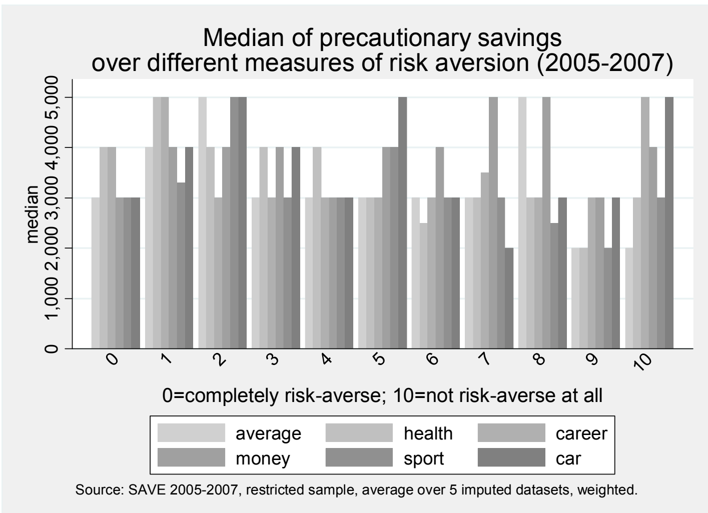


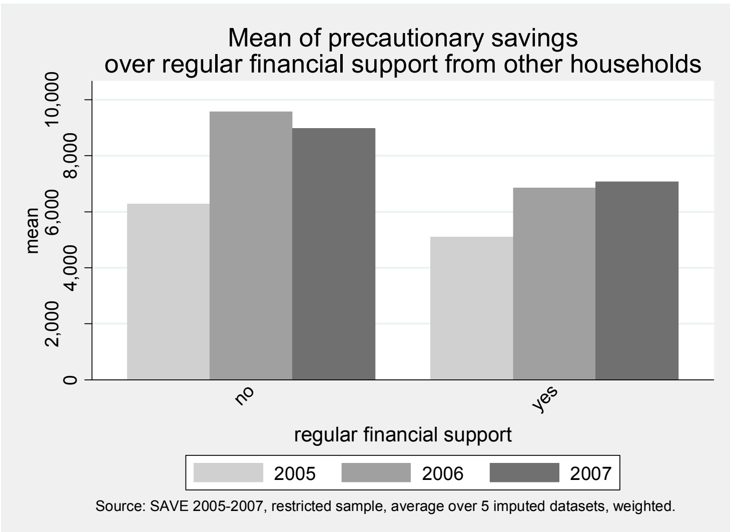
Figure 12 pictures the median precautionary savings over each of the five categories and the newly constructed average measure of risk aversion (using the mean leads to a similar pattern). Opposite to the expectations, there is no substantial increase in precautionary savings with a higher degree of risk aversion. This is the case for all 6 measures. Börsch-Supan & Essig (2002, pp. 88-90) presented a similar puzzling result. In the SAVE study of 2001, there is an increase in the saving rate for households who are less risk averse. Other questions concerning the risk attitude such as placing a whole day's income on a bet (question 123 in 2005) were not asked again in 2007. Also, the financial decision questions (questions 104-107 in 2005) were only asked in 2005. Since there is not a better measure available for the risk attitude, the average over the five categories is used in the multivariate analysis.

Another important control variable is the regular support by other households.[9] Support from other households is similar to an insurance. If a household gets into financial distress, the financial support from another household can step in and help out. In all three years the mean of precautionary savings of households not receiving

[9] Question 73 (in 2005) regarding irregular support of more than €25 per month from persons in another household together with question 71 (in 2005) regarding regular support would have been the preferred variable, since financial distress is for most households a transitory situation and only irregular support is needed. However, there was a mistake in 2005. Only households who answered that they received support on a regular basis were also asked to answer the question about receiving support on an irregular basis. This was not the case in 2006 and 2007.

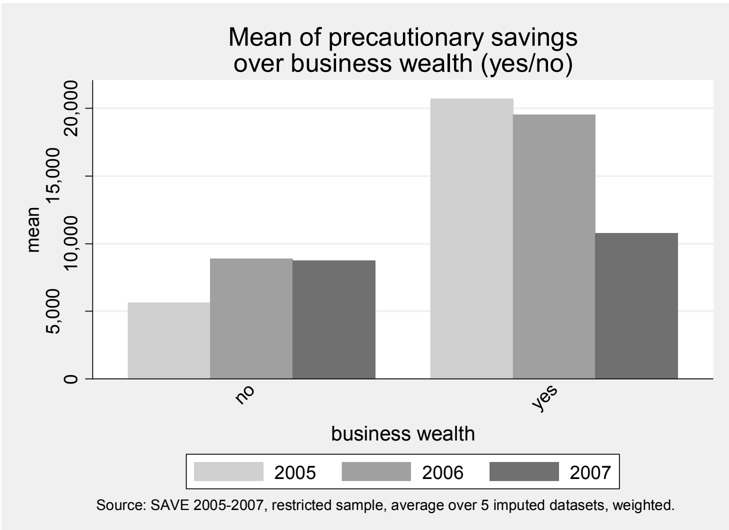
regular financial support is higher than the mean of precautionary savings of households receiving regular support (figure 13). However, the only year in which this result is statistically significant is 2006. This result holds for the median of precautionary savings as well (not shown here). Thus, regular financial support from other households may tend to decrease households' precautionary savings. Other insurance variables are not explicitly discussed here. They are included in the multivariate analysis.

Figure 13: Mean of precautionary savings over regular financial support from other households



Hurst & Kennickell & Lusardi (2005) investigated the differences in saving behaviour of individuals who were business owners and those who were not. They concluded that one has to properly account for the differences in the saving behaviour between business owners and non-business owners to take different environments they face and characteristics they have into account. Figure 14 plots the mean of precautionary savings over business ownership. In all three years, the possession of business wealth is connected to higher precautionary savings. The pattern is the same for an inspection of the median of precautionary savings over business wealth.

Figure 14: Mean of precautionary savings over business wealth (yes/no)



The final issue addressed in this section is a measure of income uncertainty. This variable is gained from the following question (questionnaire of 2005):

76. During the last five years, did your income ...

□ *fluctuate significantly;* □ *fluctuate slightly;* □ *not fluctuate at all?*

Figure 15: Mean of precautionary savings over income fluctuations

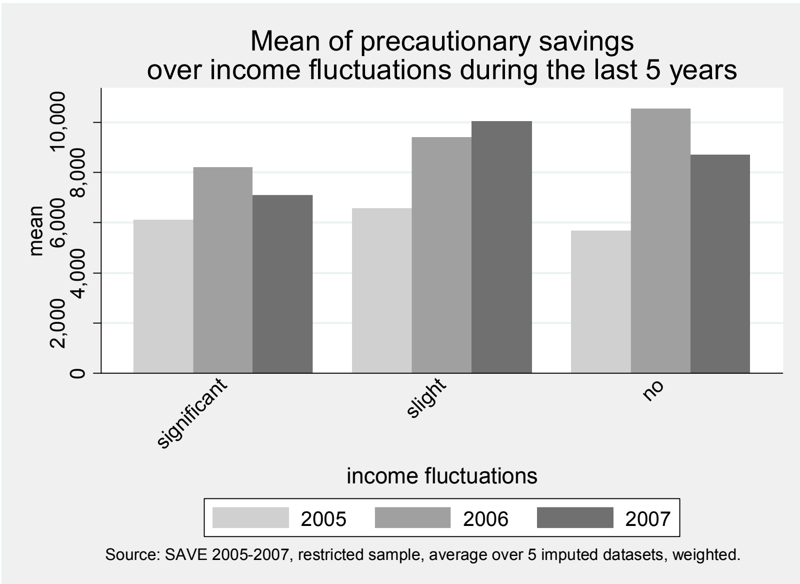


Figure 15 shows the mean precautionary savings over households categorised according to their income fluctuations during the last five years. The prediction of the theory is that increasing income fluctuations involve a rise in precautionary savings. In contrast to that, figure 15 and the equivalent graph presenting the mean of precautionary savings suggest that there is no such relationship in a bivariate analysis.

To summarise the results of the bivariate analysis, it was found that especially very young (below 25 years) and elderly households (from age 55 to 60 onwards) hold more precautionary savings relative to their monthly net income or to their financial wealth. In addition, a positive relationship between the length of education and precautionary savings was discovered. Wage earners and salaried employees have the lowest mean and median, freelancers and self-employed are amongst the highest. Moreover, risk aversion has no influence on the amount of precautionary savings. Regular financial support from other households reduces as expected the amount of precautionary savings and business owners have substantially more precautionary wealth. In contrast to the theory, there is no relationship between the amount of precautionary saving and income fluctuations over the past five years. Since the results seen in the bivariate analysis might be driven by variables we had not controlled for, it is time to turn to the multivariate empirical analysis, which is discussed in detail in the following three sections.

5 Development of an empirical equation for the SAVE dataset

5.1 Reduced form regression equation

Almost all micro-empirical studies investigating the precautionary saving motive[10] are based on the estimation of the following equation:

$$\text{I.}\ \frac{W_h}{Y_h^P} = f(\text{risk}_h, X_h) \text{ or II. } g(W_h) = f(\text{risk}_h, Y_h^P, X_h),$$

where W_h is a measure of wealth held by household h, Y_h^P is the permanent income, risk_h is a vector of measures of different kinds of risks, and X_h is a vector of control variables (e.g., age, gender, education, occupation). This wealth to permanent income equation is based on the life-cycle models introduced above. An implication of these models is the estimation of the ratio of a measure of wealth to permanent income (equation I.).

To allow for non-homothetic preferences, the functional form of the logarithm makes it possible to add the log of permanent income on

[10] For instance, Guiso et al. (1992, p. 324), Starr-McCluer (1996, pp. 289-290), Kazarosian (1997, p. 242), Lusardi (1997, p. 323; 1998, p. 449), Carroll & Samwick (1998, p. 413), Engen & Gruber (2001, p. 560), Arrondel (2002, p. 188), Murata (2003, pp. 5, 10), Hurst & Kennickell & Lusardi (2005, p. 6), Kennickell & Lusardi (2005, p. 3), Bartzsch (2006, p. 4).

the right hand side (equation II.). Now the influence of explanatory variables on the ratio of precautionary savings to permanent income must not be constant, and the influence is allowed changing with increasing permanent income. This is done in the subsequent multivariate analysis.

5.2 Problems related to the empirical assessment of the precautionary saving motive

The estimation of such an equation is a difficult task. Several problems are related to the empirical estimation of the precautionary saving motive. Kennickell & Lusardi (2005, pp. 4-9) give a good overview. In the next subsections several issues are explained, and the implementation using the SAVE dataset is presented. Since not all variables are discussed in detail, table A.1 in the appendix contains the definition of every variable used in this analysis of the SAVE dataset and table A.2 provides information about the item non-response.

5.2.1 The explained variable

Total net wealth or financial wealth were mainly used as explained variables in micro-empirical studies so far (see table A.3 in the appendix). However, it is difficult to define a measure of wealth that satisfies the needs of precautionary savings. The time required for accessibility of money plays a major role and depends on the risks and uncertainties a household has to face. Whereas longevity risk has a very long time-frame, a broken car or washing machine must be replaced more quickly. Even though financial wealth is more liquid than housing wealth, real estate can serve as a security for credit. Thus, considering only financial wealth may be defined too narrowly and taking total net wealth may be defined too widely. Therefore, liquidable wealth has been introduced, but it is difficult to draw the line (Kennickell & Lusardi, 2005, p. 16). Further on, different measures of wealth are the results of a wide range of saving motives. Thus, different motivations and related backgrounds are mixed in these measures. The desired amount of precautionary savings offers another possibility as an explained variable and partly circumvents the problems of commonly used wealth variables. Since this measure has only been used in the article of Kennickell & Lusardi (2005) as the dependent variable, the knowledge about the validity of this measure is still very limited. This study may help to shed more light on this issue.

5.2.2 Income risk and other risks over the life-cycle

Using the precautionary saving motive in empirical work, leads to many difficulties. The most notable difficulty is to find measures of risk and uncertainty that are correlated with precautionary savings and vary strongly enough across the population (Lusardi, 1998, p.1).

Income risks

For a long time, the focus has been exclusively on income risk defined as earnings plus transfers risk excluding non-capital income in most cases. It was implicitly assumed that income risk is the most important risk an individual has to face. Estimating the variance of income, which was the method chosen by Carroll & Samwick (1998) and Kazarosian (1997) using panel data, is one possibility.

Skinner (1988) used occupation dummies as a proxy for income risk. His results revealed that in contrast to the theory individuals with occupations related to riskier income flows saved less than people in other occupational groups. This points in the direction of self-selection into jobs according to the individual's risk aversion. Thus, the results state that occupation dummies are normally not a good proxy for income risk but they are necessary controls. Another way to approach this problem is to construct a subjective measure of income uncertainty as Guiso et al. (1992) and Lusardi (1997) did using the 1989 SHIW. Their measure was based on a question about the change of expected nominal earnings of a household one year after the household was examined (Guiso, 1992, pp. 311-317). The variance derived cannot be seen as a measure of earning variance over a whole life-time since it is the forecast for only one period in the future. In addition, the low variance which was constructed can be the result of the labour market in Italy having only a low percentage in short-term contracts (Lusardi, 1997, p. 322). Nevertheless, both Guiso et al. and Lusardi found a small but significant influence of the subjective measure of income risk.

As pointed out by many authors (e.g., Carroll, 1992; Engen & Gruber, 2001), unemployment is likely to be one of the main sources of uncertainty about future income. Lusardi (1998, p. 451) used a subjective measure based on an individual's evaluation of the probability of becoming unemployed in the next year. Another measure of unemployment risk are regional unemployment rates (Lusardi, 1997, pp. 323-325; Kennickell & Lusardi, 2005, p. 18; Essig, 2005, pp. 17, 31). This exogenous source of income variation can reflect the general, not individual specific probability of becoming unemployed. Along with an increase in both measures, precautionary savings should increase to buffer against a higher general or individual unemployment risk.

To sum up, it is of utmost importance to find appropriate measures of risk. The drawback of the constructed measures of risk out of panel data might be that it could be a difficult task to eliminate measurement error to obtain the "real" transitory and permanent income for the variance calculation (Kennickell & Lusardi, 2005, p. 5). Aggregated risk measures constructed from time series lose the individual information which is so important in this case (Guiso et al., 1992, p. 308). Further on, the household could already be insured against this risk (Caballero, 1991, pp. 862-863; Browning & Lusardi, 1996, pp. 1803, 1821), which is investigated in a further subsection. Since the interviewee might better know the kinds and magnitudes of risk he or she has to face, subjective measures of risk or a subjective question about the amount of precautionary savings are the better method of collecting these data. But misreporting due to misunderstanding or desirability of a certain answer may lead as well to a bias (Schunk, 2006, p. 4; Essig, 2005, p. 7; Alessie & Kapteyn, 2001).

Three different measures of income risk are applied in this study to meet the concerns of the outlined difficulties: one individual specific backward looking, one individual specific forward looking, and a general measure of income risk.

The individual specific backward looking measure is constructed from the question already presented at the end of section 4 (question 76 in the SAVE survey of 2005). Two dummies are constructed using no income fluctuation as the base group. The intuition behind this question is that households having significant fluctuations over the past five years should have higher precautionary savings than the other two groups on average. Those households with no fluctuations at all should have the lowest precautionary savings, holding other factors constant.

The forward looking measure of income risk is the measure already used by Lusardi (1998, p. 451) or Essig (2005, pp. 15-17). In the SAVE questionnaire the respondent was asked to rate the likelihood of job loss in the year of the survey for the household head (p_H) and his/her partner (p_P). The range is from 0% (very unlikely) to 100% (very likely). Under the assumption that the replacement rate is a=60% for all households[11] and the knowledge about the individual specific income of the household head (Y_H) and the partner (Y_P), the income variance is calculated from the formula $p_H(1-p_H)(1-a)^2Y_H + d_P\, p_P(1-p_P)(1-a)^2Y_P$, where d_p is a dummy of having a partner. The likelihood of becoming unemployed of the household head and the partner is assumed to be independent, and p_H or p_P are set to zero for civil servants or individuals who are not working, which includes both unemployed and retired individuals. The included measure, which is the calculated standard deviation of income divided by

[11] The earned rate is 67% if there is a child in the sense of § 32 paragraph 1, 3 to 5 Einkommensteuergesetz (EStG). Since the age of children and their status are not known in SAVE, this distinction cannot be made. Moreover, the claim for unemployment money and the duration of availability depends on the age of the individual and the kind, length, and times worked in the past. For more information, the homepage of the "Bundesagentur für Arbeit" is a helpful guide.

current income, is the variation coefficient of income in the near future.

The last measure of income risk are the unemployment rates in the 16 federal states to cover the general unemployment risk in the federal state the respondent is living (Essig, 2005, pp. 17, 31).[12]

Health risk

As Börsch-Supan (2005) describes, there are additional risks influencing individuals over the life-cycle. Many of them influence the current and the future path of the ability to consume in one way or another. The risks he mentioned stretch from economic and political risks to biometric and family risks. From this broad range health risks are considered next. Health risks are critical from two perspectives. On the one side a poor health can influence the labour income, on the other side it may raise medical expenses and reduce expenses in other categories in the future. Moreover, it can already influence the consumption behaviour today through large investments in health today (Picone et al., 1998). This means that a higher risk in a drastically worsening health situation should lead to higher precautionary savings today. Guiso et al. (1996, pp. 163-168) used the number of ill days as a proxy for health uncertainty. Kennickell & Lusardi (2005, p. 19, appendix) included the state-specific level of out-of-pocket health costs and a more individual specific measure on expected health care expenses in the next 5 to

[12] Information are taken from the homepage of the "Statistische Ämter des Bundes und der Länder". The unemployment rates for each federal state are the average unemployment rates for each year.

10 years, but only the expected future health care expenses were significant and had the expected sign. Kong & Lee & Lee (2007) calculated the variance in the health status using a four year panel of Korean data. This measure had the correct sign and was significant for a sample restricted to household heads over 64 years of age. Instead of reducing precautionary savings, households covered by a private health insurance had higher levels of wealth than uninsured households (Starr-McCluer, 1996). Thus, private health care insurance coverage must be strongly endogenous, since it mirrors a common attitude towards risk.

In this study, health risk is constructed from a question asking for the expectation about the future development of the own and the partner's health situation. The dummy constructed is equal to one if the health development is expected to be poor for at least one household member (household head or the partner) (see table A.1 in the appendix). This poor expected health development should induce higher precautionary savings today.

Longevity risk

The risk that life expectancy is underestimated by individuals and that resources actually needed are higher than the individual originally accounted for is called longevity risk. Precise measurements of longevity risk are hard to quantify. Medical progress and its impact on life expectancies is difficult to measure. Kennickell & Lusardi (2005, pp. 19-20) included a proxy for this risk using the variation coefficient of longevity (a similar procedure was chosen by Palumbo (1999, p. 406)). For that procedure, life expectancy was calculated for each person based on mortality probabilities categorised by age, gender, and race. From the distribution of life expectancy the standard deviation was derived. The ratio between this standard deviation and the difference between life expectancy and current age was their variable for longevity risk. This procedure does not pay any attention to the information each individual has about his/her living and health condition (Hurd & McGarry, 2002, p. 966).

The SAVE study uses this information asking each individual to estimate his/ her own and his/ her partner's life expectancy. Smith & Taylor & Sloan (2001) as well as Hurd & McGarry (1995, 2002) elaborated that individuals are quite good at their subjective evaluation of their own longevity and their survival probabilities respectively. Consequently, these variables can be useful tools in micro-empirical equations of life-cycle models. As Börsch-Supan & Essig (2005, pp. 7-8) presented, individuals tended to underestimate their life expectancy in the SAVE study of 2004. The reason is that individuals are misled since they have already survived certain risks up to their present age, and they take life expectancy of newborns as their own. Since these mistakes are made and it is not possible to identify "right" and "wrong" estimates, it is a difficult task to construct an appropriate measure of longevity risk.

For this study only a measure of the expected years left to live (maximum of the household head and his/her partner) is included in the analysis. A plausible hypothesis seems to be that if longer years left to live are correlated, e.g., with an increasing health risk, an increase in this variable should result in an increase in precautionary savings. Lusardi (1998, p. 452) included a variable which measured the subjective probability of living to age 75 with a similar intention.

Business risk

Kennickell & Lusardi (2005, p. 20) constructed the failure rate of businesses categorised by type, age, and state of business and included this measure as a proxy of business risk. In this study, only a dummy for the possession of business assets is included to control for the different behaviour of business owners. In examining the different behaviour of business owners and non-business owners, Hurst & Kennickell & Lusardi (2005) found that after splitting the sample into these categories, the contribution of precautionary savings to total net wealth accumulation dropped from around 50% to less than 10% for non-business owners and to less than 12% for business owners. The results support the findings of a low or moderate contribution of precautionary savings to wealth accumulation and points out the importance of controlling for business owners, which was explicitly or implicitly done by all the studies finding a low or moderate contribution of precautionary savings. Compared to the pooled SCF of 1995 and 1998 with 11%, business owners account for only 4.7% in the restricted SAVE sample from 2005-2007 (average over all five imputed datasets).

Pension risk

Many people are uncertain about the level of their future public pension (Murata, 2003). This uncertainty should lead to a higher private old-age provision via higher precautionary savings. Therefore, a dummy is added, which is one if at least for one, the household head or the partner, there is no guess possible about the expected level of the public pension. This variable is not available for the year 2005.

Finally, the connection between the buffer-stock model and the empirical equation introduced in section 5.1 is made by Carroll & Samwick (1998, pp. 412-413). They present an almost linear relationship between the optimal cash on hand to permanent income ratio and the measures of income uncertainty. Similar to Bartzsch (2006, p. 4) it is assumed that a linear relationship exists for the uncertainty and risk measures used in this study if no dummy variable is used.

5.2.3 Estimation of permanent income

Following the argumentation of the permanent income hypothesis, households should consume out of permanent income and not out of current income. The purpose of precautionary savings is to smooth consumption around permanent income to maintain the living standard in the face of risk and uncertainty. Considering this aspect, precautionary savings should be set into relation to permanent income (Engen & Gruber, 2001, p. 561).

Permanent income normally is a variable constructed from panel data as done by many authors (Kazarosian, 1997, p. 242; Fuchs-Schündeln & Schündeln, 2005, p. 1098; Bartzsch, 2006, pp. 5, 9). There are three reasons why constructing permanent income out of panel data does not seem to be a passable way using the SAVE dataset: first, the panel analysed consists of only three observations over time; second, the time distance between each observation in time is only one year; thus we have only two years between the first and the third observations in time; third, observations over all three years can only be obtained for less than 50% of all households. If sufficient panel data are not available, Kennickell & Lusardi (2005, p. 17) use a subjective measure of permanent income, which asks for the "normal" income of households. A similar question is not available in the SAVE dataset. This leads to the method of constructing permanent income out of a cross section developed by King & Dicks-Mireaux in 1982 (pp. 254-257) and applied by many authors, e.g., Starr-McCluer (1996, pp. 292-294), Murata (2003, pp. 21-24), and Essig (2005, pp. 8-10). Subsequently, this method is described in more detail, followed by criticism and the introduction of the method applied in this study.

Permanent income of individual i is denoted by y_i^p and depends on a vector of observable characteristics X_i (gender, education, occupation, …) with a coefficient vector β, unobservable characteristics s_i (ability, motivation, early family upbringing), and a correction term for different cohorts $c(age_i)$, since younger cohorts normally are better off due to capital accumulation and technical advancements. This can be expressed in the following functional equation:

$$\text{I. } f\left(y_i^p\right) = \beta X_i + s_i - c(age_i)$$
$$\text{with } E(s_i) = 0 \text{ and } var(s_i) = \sigma_s^2$$

However, what is observed is current income y_{it}^c, which is not equal to permanent income ($y_i^p \neq y_{it}^c$) for two reasons: first, the income-age profile over the life-cycle $h(age_{it} - \overline{age})$, and, second, the transitory income shocks u_{it}. Taking both points together results in the equation below, which reflects the relationship between transitory and permanent income.

$$\text{II. } f\left(y_{it}^c\right) = f\left(y_i^p\right) + h(age_{it} - \overline{age}) + u_{it}$$
$$\text{with } E(u_{it}) = 0,\ var(u_{it}) = \sigma_u^2,\ cov(u_{it}, s_i) = 0$$

Plugging equation I in equation II and making additional assumptions, the next equation is obtained:

$$\text{III. } f(y_{it}^c) = \beta X_i + g(age_{it}) + \eta_{it}$$
$$\text{where } g(age_{it}) = h(age_{it} - \overline{age}) - c(age_{it})$$
$$\text{and } \eta_{it} = s_i + u_{it} \text{ and } E(\eta_{it}) = 0,\ var(\eta_{it}) = \sigma_s^2 + \sigma_u^2$$

Now permanent income can be estimated in different ways. The easiest way is that permanent income is equal to the age-adjusted earnings profile $\hat{\beta} X_i$ of the individual i ignoring the individual's specific effect s_i and the cohort effect $c(age_i)$ (see equation I; chosen by Carroll & Dynan & Krane (1999, pp. 18-19); Arrondel (2002, p. 191)). A more complex way is to take these two effects into account:

$$\text{IV.}\ f(\hat{y}_i^p) = \hat{\beta}X_i + \hat{s}_i - \hat{c}(age_{it}),\ \text{where}\ \hat{s}_i = \hat{\alpha}\left[f(y_{it}^c) - f(\hat{y}_{it}^c)\right]$$

There are some critical assumptions in the whole procedure, which are not outlined in this study, but by far the most important and also most critical one is the estimation of the individual specific component s_i. This part is further developed. At first, s_i has to be estimated. Neglecting this term can lead to a serious bias in the estimation of the coefficients β. Since s_i is a fraction of the total error term η_{it} in equation III, this fraction can normally be estimated using panel data. Given that only a cross section is available, estimates from other panel studies have to be used. The estimate of $\hat{\alpha}$ implemented in the studies of King & Dicks-Mireaux (1982, p. 255), Starr-McCluer (1996, p. 294), and Essig (2005, pp. 8-9) is 0.5 based on the studies of Lillard & Willis (1978, p. 991) and Lillard (1977, p. 45). Since these studies were written more than 30 years ago and rely on US data, to assume that $\hat{\alpha}$=0.5 seems inappropriate for actual German data. Moreover, the size of $\hat{\alpha}$ depends crucially on explanatory variables already included in X_i. Even more important is that for each individual the time constant unobservable effect is treated as being half of the difference between observed and imputed current income. This treatment makes things easier than they really are.

These concerns brought up another method to deal with this issue. The log of current income along with other income specific control variables is included at the right hand side (see the equation on page 88). Controls are the level of employment (e.g., full-time, part-time, low level, not employed), number of income sources, probability of receiving a high inheritance, and a dummy for retirement and past unemployment. Accounting for past unemployment is necessary since this often lowers the permanent income and the stock of wealth (Arrondel, 2002, p. 189). Additional to these controls, the controls included also for precautionary savings can hold income relevant variables constant. Age, gender, marital status, foreigner, education, and occupational group are some examples (see table A.1 for more details). Since panel techniques are applied to estimate the coefficients, it can be controlled for time invariant unobservable variables, which was the main concern above. With this method it is not possible to obtain an estimate of the coefficient of permanent income. However, it is possible to hold constant all the variables that may influence permanent income. This may allow estimating the other coefficients properly.

5.2.4 Necessary controls and other theoretically important variables

Even though from a theoretical point of view, the interest focuses mainly on the different measurements of risk, control variables play an important role in achieving unbiased estimates of the coefficients. If there is a correlation between an explanatory variable and an omitted variable, which has an influence on the explained variable, the result will be an upward or downward bias of the coefficient of interest depending on the sign of the covariance of the omitted variable and the explanatory variable and the coefficient of the omitted variable. Control variables are added to the equation since several risk variables and the explained variable seem to be correlated with other variables. However, not only the consistent estimation of the coefficients of risk variables should be achieved, but also the effects of additional variables which are important as outlined in the theoretical discussion earlier are qualified and quantified. For a detailed list and how the variables are constructed, the reader is referred to table A.1 in the appendix. Most common are controls for age, gender, education, and marital status.[13]

Household composition

A dummy of whether the household has children or not and a dummy of whether an additional person except for the partner and the children lives in the household's home is added. It is not possible to specifically identify the additional person(s). The intention behind the inclusion of this dummy is that additional persons living in the household's home who are not the children of the household head or the partner are mainly older people, maybe the parents, who have to be cared for in many cases. This home care may not only be time but also money extensive, since medicine and outpatient care is needed. The parents being cared for often have their own income sources. However, if savings and pension payments of the parents are not enough, their children have to pay depending on their income and wealth. Thus, precautionary savings should increase in the presence of additional person(s) in the household. The effect of children in the household is not clear. On

[13] Different specifications for age were implemented and the p-values from the F-tests of the age terms were compared. The lowest p-value was obtained form age alone. However, the coefficient was still insignificant. Thus, only age is included in the specifications.

the one side children can raise additional needs for precautionary saving, on the other side children may limit the possibility to do so. However, the last point made should be irrelevant since it was asked for the “desired” amount of precautionary savings.

Occupation dummies

Moreover, occupation dummies are included since not controlling for self-selection leads to a serious underestimation of precautionary savings as Lusardi already stated in 1997 (p. 320). Fuchs-Schündeln & Schündeln (2005) confirmed the results of Skinner (1988) using the German reunification as a natural experiment. Making use of the German civil servants in the German SOEP and the fact that most of the civil servants in Eastern Germany did not select their jobs in consideration of income risk, since income risk in the German Democratic Republic was not a matter of concern, they identified the importance of self-selection into different jobs and quantified the effect of self-selection on precautionary savings.

Foreigner dummy

The article of Piracha & Zhu (2007) emphasised the importance of the distinction between natives and immigrants. They used a change in the German nationality law in 2000 as a natural experiment to determinate the change in the saving behaviour of immigrants. Since the new nationality law reduced uncertainty for immigrants, they found a reduction of precautionary savings for immigrants of around 13% after the law came into effect. Thus, a dummy representing a foreigner is included to partly control for a different saving behaviour of natives and immigrants. Since foreigners face additional risks (e.g. more difficulties to become employed) compared to natives, it is expected that being a foreigner has a positive influence on precautionary savings.

Expectation about income growth

According to the buffer-stock model, expectations about future income growth are included. If the expectation of future income growth drops, the result should be a rise in the optimal wealth to income target (Carroll, 1997, pp. 13-15). Two measures are available; that is, a variable asking for the expectations about the future development of the own financial situation for the longer time horizon as well as a variable reflecting the likelihood of an income increase one year in the future both ranging from 0 (very negative/ unlikely) to 10 (very positive/ likely).

Impatience

The importance of a measure of impatience was pointed out at the end of section 2. To get a measure of the time preference, a dummy for the smoking status is included. The idea behind this proxy is that smokers evaluate the current utility of smoking higher than negative influences on their future health status. Recently Khwaja & Sloan & Salm (2006, pp. 674-676) confirmed the positive relationship between the degree of impatience and being a smoker. Moreover, they found that smokers have a higher risk tolerance than non-smokers (pp. 676-678). Further on, smoking status seems to be a time-invariant characteristic of the individual (pp. 674, 678). Thus, an individual who smoked in the past should have the same preference structure. It is the opinion of the author that this must not be the case, because the alteration of an individual's smoking behaviour might result from a change in the preference structure. In this study the better fit is obtained using a dummy equal to one if the individual is a current smoker ignoring smoking behaviour in the past. This dummy of the smoking status is used as a proxy for impatience and risk tolerance in this paper. In the context of precautionary savings, smoking behaviour as a proxy for impatience was already used by Lusardi (2003, p. 12) and Kennickell & Lusardi (2005, p. 20).

Besides that, a dummy variable is added to control for the time preference from another perspective. The underlying question asked the respondents to place themselves on a scale from 0 (I am easy going and take each day as it comes. I don't think or worry much about the future) to 10 (I think about the future a lot and have a pretty good idea of where I want to be and want to do in the future). The dummy is equal to 1 if the respondent answered a value lower than 5 (easy going type) and the base group corresponds to individuals who think a lot about the future.

Other preferences

The attitude towards risk is not only important for self-selection into different jobs. If risk aversion is also positively correlated with prudence, which is the case for the CRRA utility function, higher risk aversion leads to higher precautionary savings. The constructed variable used as a measure of risk aversion was already introduced in detail in section 4.

SAVE offers a wide range of variables which may reflect general attitudes towards saving. A dummy variable is constructed out of the question asking for whether or not households hold a certain minimum amount in their current account. Since the current account is besides cash and credit cards the most liquid money, to limit spending not to fall below this minimum amount should reflect a precautionary attitude towards risk and should therefore raise precautionary savings on average. Moreover, a dummy is included for whether or not the household is not a regular saver type.

Health

As current health status has an influence on future precautionary savings based on a relation to future health problems and the perception of expected health expenditures (Arrondel, 2002, p. 189), controls are included for current health status.

Liquidity constraints

Liquidity constraints influence the ability of households to get credit when in financial distress. If households in financial distress want to prevent consumption from falling below a minimum amount, precautionary savings are necessary. Conversely, households who can borrow need not as much wealth to protect them against a drop in current income. But it is difficult to construct an appropriate measure of the household's possibilities to borrow. Kennickell & Lusardi (2005, p. 21) summarise some of the established measures in other studies. Different variations in constructing a liquidity index have been tried in this study. The result is a dummy variable being one if a credit was refused, not granted in the full amount, or if the household did not apply for credit because they thought that the credit would not be granted. The time horizon for these questions is the last five years. Moreover, the overdraft limit is included. A higher limit should reduce the necessary precautionary savings at least in the short run. Other and additional investigated possibilities are a dummy for low income families (e.g. monthly income lower than €900 (€1200), which corresponds to the 10th (20th) percentile) and the ratio of outstanding debt over total net wealth.

Insurance

Precautionary savings are a buffer against risks the household is not insured against. The private and public insurance market therefore plays a major role in the magnitude of precautionary savings. The German social insurance system is one of the most fully developed social security systems in the world dating back to the years 1883, 1884, 1889, and 1912 during the German Empire and the year 1927 during Weimar Republic. The mainly obligatory unemployment, health, accident, old-age, and long-term care insurance system in addition to social benefits are well developed compared to other countries like the US. This strong insurance system may reduce the influence of certain long-term and short-term life-cycle risks mentioned above. The past years brought strong changes about the social security system, and it started a still ongoing process towards increasing self-responsibility. The so-called "Hartz IV Law", the pension and health care reforms are only some of the ever returning vocabulary in the last few years. Whether this process is observable in the data in changing buffer-stock savings is an interesting question, since the uncertainty about the social security system is raised in the general awareness, and the objective economic risk in the transitional period could be higher (Börsch-Supan, 2005, p. 1). However, the private insurance market could have taken over, which is only observable in a limited way in the SAVE dataset. Social insurance coverage is compulsory for the largest portion of the German population (for the exemptions see the Sozialgesetzbuch [SGB]). Nevertheless, not every individual who is not subject to the compulsory insurance coverage holds a voluntary insurance. Distinctions and nuances cannot be identified with SAVE in 2005 and 2006. In 2007 respondents had to indicate whether their employment relationship underlies the compulsory social insurance coverage. The questions about the ownership of a private or social long-term care insurance plan were only asked in 2007.

For all three years two dummies for private insurance coverage (occupational disability and liability insurance) are included. In the case of private occupational disability insurance as well as private liability insurance, precautionary savings should decrease. As introduced and discussed already in section 4, regular support from other households is included too.

Interaction between different motives to save

As Hurst & Kennickell & Lusardi (2005, p. 7), Kennickell & Lusardi (2005, pp. 8-9), or (Jeppelli et al., 2006, p. 22) point out, wealth accumulated for bequest purposes and the old-age provision can be used additionally to the precautionary savings as buffer-stock savings if unforeseen events occur and more money is needed than saved for unforeseen events. If the precautionary savings are enough, then savings for bequests and old-age provision can serve their original purpose. To control for these two saving motives, two variables for the importance of the bequest motive and the importance of the old-age provision motive are included ranging from 0 (unimportant) to 10 (very important).

Wealth dummies

In empirical studies it is often emphasised that the saving behaviour of wealthy households is different than the saving behaviour of less wealthy households (see section 2.3.4). In the case of rich households, wealth accumulation has to serve additional purposes and cannot be explained alone to finance own future consumption or the consumption of heirs. Carroll (2000) investigates alternative models. He clarified that wealth itself can cause utility whether through power or social status. Therefore, four wealth dummies corresponding to the quartiles of the total net wealth distribution are constructed to control for other purposes that savings can serve.

Other specifications

Specifications with a wide range of other variables and different functional forms have been accomplished.[14] The variables included

[14] Some of the additional variables investigated are: interactions of all risk variables with different age groups; interactions between age and educational groups; interaction between educational and occupational groups; social welfare payments; past smoker; household size; number of children; dummies for one, two, …, more than five children; income variance out of expected unemployment probabilities with and without replacement rate; unemployment probabilities; dummy for low income

in the final version are selected, first, to meet the theoretical requirements and, second, to incorporate previous empirical findings. If more variables in SAVE seemed to be promising, the specification showing the best fit was chosen. The adjusted R^2, the Akaike Information Criterion, and the Schwarz Information Criterion served as a reference point (Gujarati, 2003, pp. 536-538). Further, including additional variables allows for the checking of the robustness of the applied specification.

5.2.5 Functional form

Using the logarithm of the desired amount of precautionary savings or wealth is a typical approach in estimating higher amounts of monetary units because of the skewed distribution. The logarithm normally leads to less serious effects of outliers as mentioned above and reduces the problem of heteroskedasticity. In addition, robust standard errors are used in all estimated equations.

Using the log of precautionary savings makes it necessary to exclude zero precautionary savings. Only in 2005 168 households with zero precautionary savings were observed (see table 4). Hence, a selection model like the Heckman two-step procedure (Heckman, 1976) seems unnecessary in this case (at least not for 2006 and 2007). Since the exclusion of observations at the left tail side may

families; a ratio of outstanding debt over total net wealth; permanent or temporary working position; change in the income situation during the last five years; save goal; optimism; self-assurance; record keeping; other variables out of the health category.

lead to a bias, another procedure is implemented. A small and positive amount (plus €1) is added to ensure that all values of precautionary savings are bigger than zero. Thus, this leads to a small and negligible right shift in the distributions of precautionary savings.

As described at the beginning of this section, using the logarithm allows adding the variable for permanent income (in this study current income) on the right-hand side to permit non-homothetic preferences.

5.3 The basic empirical equation for SAVE

After considering all the points made above, this leads to the following empirical equation, which is estimated separately for every year:

$$\log(precautionary\_savings_i) = \beta_0 + \beta_I \begin{pmatrix} \log(current\_income_i) \\ income\_controls_i \end{pmatrix} + \beta_R(risk_i) + \beta_C(controls_i) + \eta_i$$

where the form of the equation was described in sections 5.1 and 5.2.5, the dependent variable was introduced in section 5.2.1, the permanent income challenge was clarified in section 5.2.3, the risk variables were presented in section 5.2.2, and the control variables were discussed in section 5.2.4. Since 5 imputed datasets exist for every year (in this study every dataset is restricted as outlined above), the normal procedure is to run the regression on every dataset and compare the results to analyse the variation between the datasets (Caneron & Trivedi, 2005, p. 934). After that Rubin's rules are applied to combine the estimates out of all five imputed datasets (see section 3.3). Table 10 shows the results of the regression of each year using Rubin's rules. Since the results of these regressions do not lead to stable results over the years for most of the explanatory variables, only the coefficients and standard errors of the risk variables are shown here justified in the importance of the risk variables. The specification is the same over all years except for theoretically important variables which are only available in 2006 or 2007. These variables are added in the regression of 2006 and 2007.

Table 10: Robust OLS estimation for each year from 2005-2007

	2005			2006			2007		
log(precautionary_savings+€1)	Coef.		Std. Err.	Coef.		Std. Err.	Coef.		Std. Err.
std(net income)/net income	-0,008		0,749	0,240		0,346	0,505		0,386
significant earnings fluctuations	0,125		0,170	0,068		0,078	-0,016		0,080
slight earnings fluctuations	0,130		0,145	0,072		0,062	-0,088		0,068
provincial unemployment rate	-0,004		0,029	-0,019		0,013	-0,012		0,015
expected years left to live	-0,017	**	0,008	-0,004		0,003	-0,009	**	0,004
poor development of hh health	0,072		0,195	0,082		0,084	0,021		0,092
uncertain pension				-0,092		0,062	-0,166	**	0,069
# obs	1794			2568			2167		
average R-sq	0,224			0,266			0,282		

Note: * : 10% significance level; ** : 5% significance level; *** : 1% significance level.
Source: SAVE 2005-2007, restricted sample, unweighted, application of Rubin's rules.

Comments:
The control variables used are the same as in the RE and FE estimations in section 7.
The variable pension uncertainty of a household is only available in 2006 and 2007.
In 2007 the following controls have additionally entered the regression: obligation to contribute to social insurance system; private long-term care insurance; additional long-term care insurance.
For a detailed description of the variables see table A.1 in the appendix.

There is strong evidence for heterogeneity among the three years. A first hint for the heterogeneity is the changing evidence of the impact of explanatory variables over the years. To formally test the heterogeneity between the years, a pooled regression for the years 2005 and 2006 was run. An additional time dummy was included and all variables were interacted with a time dummy for 2005 and one for 2006. The time intercept for 2006 and the interaction terms for 2006 were found to be jointly significant from zero at a 1% significance level using a Wald test. The same was found for the 2006/2007 and the 2005/2007 combinations. The insignificance of the risk variables and the strong heterogeneity make more convenient empirical procedures necessary. Thus, to control for endogeneity is one major concern. In connection with these more appropriate empirical procedures a detailed discussion of the results is done. Methods, applications, and results are discussed in section 6 and 7.

6 Estimation technique

6.1 Pooled cross sections and macro shocks

Since the variables of interest, mainly the risk variables, are not significant analysing each cross section separately (table 10), which could be based on the insufficient sample size of each cross-section, one way to increase the sample size is to use independently pooled cross sections. This can be done by using the cross sections in SAVE from 2005-2007. The observations might not be identically distributed over time since the effects of independent variables may change. Further on, as Kennickell & Lusardi (2005, p. 7) emphasised, macro shocks can change the distribution of variables in different years. For example, a household with higher unemployment risk is more likely to be hit by a shock and a resulting reduction of wealth during a recession. Possible solutions are different intercepts for each year and different slope coefficients for variables, which are expected to change (Wooldridge, 2003, pp. 408-413).

To deliver consistent and efficient estimates, the same assumptions apply compared to a normal OLS-estimation. This means that correlation in the error term under the assumption of independently sampled observations is ruled out. But this can only be true having no panel data, because in panel data the error term of a household is normally serially correlated over the years. In SAVE, there is a non negligible panel component (see table 7).

To sum up, one advantage is that pooling the cross sections of 2005-2007 raises the observations and therefore the efficiency of our estimates. The drawback is the serial correlation of the error term for the most part of the total sample, which is based on the panel structure. Moreover, endogeneity may cause biased inference. This issue is addressed next.

6.2 Endogeneity problems

When explanatory variables are correlated with the error term in a regression, the researcher has to deal with the challenge of endogeneity. One possible way for endogeneity to occur is whenever the direction of causality of explanatory variables is ambiguous. On the one side, wealth or home equity can be the result of precautionary savings. On the other side, the precautionary savings behaviour itself can be strongly influenced by the level of wealth and home equity as explained above. To circumvent this challenge, the following estimations are done twice for robustness reasons: one time with possible endogenous variables, the other time without. Estimating the equation without possible endogenous variables might induce an omitted variable bias if the eliminated variables are correlated with other regressors. The following variables have been ex ante identified as two-way causality variables: wealth dummies and a dummy for home equity. This procedure was also chosen by Kennickell & Lusardi (2005, p. 20) and Schunk (2007b, p. 12).

Endogeneity can also occur for two other reasons: first, measurement error in the regressor(s) and, second, omitted variable bias. In these cases instruments can be a useful tool as used in the studies of Carroll & Samwick (1998, p. 414) and Lusardi (1997, pp. 323-325). However, to find valid instrument variables (IV) is normally a hard task because they should fulfil three properties (Wooldridge, 2003, pp. 461-500): First, the instruments must be exogenous; second, the instruments have to be correlated with the endogenous variable; third, the instruments are not allowed being explanatory variables for the explained variable, which is often subsumed under property one. Poor instruments are easily obtained if the instruments and the endogenous variable are weakly correlated. The problem with poor instruments is not just that the variance of the IV estimator is much larger than the variance of the OLS estimator. A more serious problem is that the IV estimator can have a larger asymptotic bias even if the instrument and the error term are only modestly correlated.

The difficulty in finding appropriate instruments and the drawbacks of poor instruments resulted in the method to control for endogeneity applied in this study. This method makes use of the panel structure and applies the random effects (RE) and fixed effects (FE) model to control for time-constant omitted variables. But there is no possibility to control for time-varying omitted variables that are correlated with the dependent variable as instruments can do.

6.3 Panel estimation techniques

This section reviews the panel estimation techniques applied in this study (Wooldrige, 2002, pp. 248-291). The topic is motivated, the mechanics introduced, and the most essential assumptions are discussed.

y and $x \equiv (x_1, x_2, \ldots., x_k)$ are observable random variables and s is an unobservable or omitted random variable, often called unobserved effect. s can stand for unobserved variables like early family upbringing, skills, and motivation. s is assumed to be time constant. h=1, …, N denotes the household and t=1, …, T is the actual time period. Supposing a linear model gives the following relationship:

$$(1)\ y_{it} = \beta_0 + x_{it}\beta + s_i + u_{it}$$

where β is the K·1 vector of interest and u_{it} is the idiosyncratic error or idiosyncratic disturbance. As already mentioned in section 5.2.4, omitting a relevant variable, in this case the individual time constant error s, can lead to biased estimates if s has an influence on y_{it} and $cov(x_{jt}, s_i) \neq 0$ for some j. In this procedure, s_i is treated as a random variable as is normally the case in the modern view. The mechanics introduced here are for a balanced panel, which means that for each time period the same number of not changing households are available. In an unbalanced panel this is not the case. However, the mechanics are similar to the balanced case but not explicitly discussed in this study. The statements are made with a focus on the asymptotic properties, where N grows to infinity and T is fixed.

The first assumption discussed is the strict exogeneity assumption, which the explanatory variables have to fulfil for both the RE and the FE model. The interpretation of this assumption is that after x_{it} and s_i is controlled for, x_{it} has no partial influence on y_{ik} for all $k \neq t$. Stated in the form of the idiosyncratic error, gives

(2) $E(u_{it}/ x_{i1}, \ldots, x_{iT}, s_i)=0$ for all t=1, 2, ..., T.

This assumption implies that $cov(u_{it}, s_i)=0$ and that

(3) $E(x'_{ik}u_{it})=0$ for t,k=1, 2, …, T.

Assumption (3) is much stronger than the so called zero contemporaneous assumption, which is $E(x'_{it}u_{it})=0$ for t=1, 2, …, T. However, assumption (3) allows that $cov(x_{it}, s_i) \neq 0$. The last statement is important since it leads to the modern distinction between RE and FE models. One assumption the RE model has also to include is that $cov(x_{it}, s_i)$ must be equal to zero. This does not have to be the case for the FE model.

The FE model uses relationship (4):

(4) $y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)\beta_{FE} + (u_{it} - \bar{u}_i)$

Since it uses only the variation of a variable over time for each individual, the FE estimator is called the within estimator. Under the assumption of strict exogeneity and the rank condition, the FE model is unbiased and consistent. It is more robust than the RE model since the additional assumption $E(x'_{it}s_i)=0$ is not necessary. A serious drawback is that only time varying explanatory variables can be included in the analysis. The rank condition is not fulfilled for a time constant variables and therefore they have to be excluded.

The within estimator is also efficient, if $E(u_{it}u'_{it} \mid x_{it}, s_i) = \sigma_u^2 I_T$ (where I_T is the identity matrix), which means that the idiosyncratic errors u_{it} are homoskedastic across t and are not serially correlated. If this is not the case, the estimator is not efficient and a robust variance matrix estimator should be used to calculate test statistics.

The RE model takes advantage of the information in the correlation structure of the composite error term using a generalised least square (GLS) method. For the GLS estimation to be consistent, the usual rank condition for GLS must be assumed (Wooldridge, 2002, p. 258). It can be shown that the RE estimator is the weighted sum of the between and the fixed effects estimator (Wooldridge, 2002, pp. 286-288), resulting in the next equation:

$$(5)\quad y_{it} - \vartheta \bar{y}_i = (x_{it} - \vartheta \bar{x}_i)\beta_{RE} + (u_{it} - \vartheta \bar{u}_i)$$

$$\text{with } \vartheta = 1 - \sqrt{\frac{\sigma_u^2}{T \cdot \sigma_s^2 + \sigma_u^2}}$$

The estimation procedure is unbiased if $E(x'_{it}s_i)=0$ for t=1, 2, …, T and efficient under the additional assumptions $E(u_{it}u'_{it} \mid x_{it}, s_i) = \sigma_u^2 I_T$ and $E(s_i^2 \mid x_{it}) = \sigma_c^2$. Without the last two assumptions, a robust variance matrix should be used (Wooldridge, 2002, pp. 262-263). For large N and a fixed T, the price for standard errors and test statistics that are robust to heteroksedasticity and serial dependence is negligible even if both last named assumptions hold.

One advantage of the RE model over the FE model is that if the assumption E($x'_{it}s_i$)=0 is fulfilled, the variances of the RE estimators have much smaller variances than the FE estimators, because more information contained in the error term structure are exploited. In addition, the FE model is not able to produce estimates for time constant variables and can only produce imprecise estimates for variables that do not have much variation over time (Wooldridge, 2002, p. 286).

It could be that the model does not contain individual unobserved effects. This can be tested under H_0: $\sigma_s^2 = 0$, which is the statistical equivalent to the non-existent unobserved effect (Wooldridge, 2002, pp. 264-265). The Breusch and Pagan Lagrangian multiplier test statistic for RE offers a way to test H_0. If H_0 is rejected, panel techniques should be applied.

To test whether the RE or the FE model is appropriate, the E($x'_{it}s_i$)=0 must be proved (Wooldridge, 2002, pp. 288-291). The Hausman test formulates the H_0 and H_1 hypothesis under the assumption of strict exogeneity. Under H_0: E($x'_{it}s_i$)=0 both the RE and the FE model should produce consistent estimates. The difference between the estimated coefficients weighted with the variance covariance matrix of the estimates should be small. H_1 states that E($x'_{it}s_i$)≠0. In this case only the FE model is consistent, and the difference between the estimated coefficients should be large. If H_0 is rejected, the FE model should be the preferred choice.

But what happens if the assumption of strict exogeneity fails? Then not only the Hausman test is inappropriate, but also the coefficients of the RE and FE models are not consistently estimated. To reject or not reject the strict exogeneity assumption, the following test is suggested (Wooldridge, 2002, p. 285). A regression based test for strict exogeneity is the estimation of the following expanded version of the FE model.

$$(6)\ y_{it} = x_{it}\beta + w_{i,t+1}\delta + s_i + u_{it}\ ,\ t=1, 2, \ldots, T\text{-}1$$

where $w_{i,t+1}$ contains variables of x_{it}, which may violate the strict exogeneity assumption. One time period is lost through the inclusion of $w_{i,t+1}$. The strict exogeneity assumption is rejected if deviations of δ from 0 are significant since this would violate assumption (2) or (3).

7 Results

7.1 Application of RE and FE models

After the theoretical discussion of the RE and the FE models, they are applied. The Breusch and Pagan Lagrangian multiplier test statistic for RE supports this application. The H_0 hypothesis of no random effects is rejected at a 1% significance level in all five datasets.

Since both the RE and FE models are critically based on the assumption of strict exogeneity, the test for strict exogeneity should be implemented as suggested at the end of section 6. The unbalanced panel of this study, which stretches over only three time periods, may easily lead to significant elements of δ. This fact questions the explanatory power of two elements of δ which were found to be significant at a 10% level. It is the opinion of the author that the assumption of strict exogeneity cannot be properly tested under these circumstances.

The further procedure begins with the estimation of the RE model. In most cases it is a good idea to include a separate time period intercept (Imbens & Wooldridge, 2007, p. 1). This is done by including two dummy variables, one for 2005 and one for 2007. The reported intercept is the intercept for 2006 and must be corrected by the dummy coefficient for the corresponding year to obtain the intercept for 2005 or 2007. The RE model is the preferred model since it is more efficient if the outlined assumptions hold and allows for the estimation of time-constant factors, which cannot be estimated using a FE model. To decide whether the additional assumption for the RE model holds, a FE model is estimated with the same coefficients and the Hausman test is applied next. The H_0 hypothesis of the Hausman test is rejected at a 1% significance level in all five datasets.[15] Thus, the coefficients of the RE and the FE effect model are too different taking the standard errors into account. If this is the case, some researchers infer that a FE model must be preferred to a RE model.

Instead of rejecting the RE model in favour of the fixed effect model, another procedure is chosen to guarantee that the assumption $E(x'_{it}s_i)=0$ is fulfilled. This assumption is violated if missing covariates are correlated with observed covariates. Skrondal & Rabe-Hesketh (2004, pp. 52-53) and Berkhof & Snijders (2007, pp. 142-147) suggest including the mean over time $\overline{x}_i$ for every individual and for each variable as additional covariates. These terms will absorb the bias of the coefficients if the assumption $E(x'_{it}s_i)=0$ is violated. Thus, the RE model expanded by $\overline{x}_i$ as additional covariates combines the advantages of both RE and FE models. It produces consistent estimates and permits the inclusion of time constant variables as regressors. The results are shown in the last column of table 11. As can be seen very easily, the difference in the coefficients of the FE model and the expanded RE model has

[15] The female dummy is dropped since gender does not change over time.

almost vanished. The H_0 hypothesis of the Hausman test cannot be rejected with a p-value of 1.00 in all five datasets.

However, the expanded RE model as well as the FE model can fail to estimate coefficients of variables accurately if they do not vary very much over time. Column 6 of table 11 shows the frequency of changes in each variable over time. Variables show less frequent changes for preference variables, marital status, educational and occupational status, for example. Moreover, the imputation procedure could be responsible for part of the variation of these variables. The applied imputation procedure makes no use of the panel structure, and the hotdeck procedure is mainly applied to the relative stable variables over time. The hotdeck imputation could have led to an increase in the changes of these variables.

For robustness reasons, table 12 shows the results of other specifications using the RE model extended by the inclusion of $\overline{x}_i$ as additional covariates such as: estimation without possible two-way causality variables (specification 1.); exclusion of zero precautionary savings caused by not adding one to the amount of precautionary savings (specification 2.); restriction of the sample to a balanced panel to learn more about changing behaviour over time (specification 3.); restriction of the sample size to household heads not older than 50 years, since in this part of the life-cycle buffer-stock behaviour is most likely to emerge (specification 4.); restriction of the sample to non-business owners, since, as mentioned in section 5.2.2, saving behaviour varies significantly between business owners and non-business owners (specification 5.); and changes of the covariates and their functional form as already mentioned at the end of section 5.2.4 (not shown). The construction of all variables and the base group of each dummy variable is outlined in table A.1 in the appendix.

Table 11: Random effects, fixed effects, and expanded random effects estimation

	RE		**FE**		changes	**expanded RE**	
log(precautionary_savings+1)	Coef.	Std. Err.	Coef.	Std. Err.	over time	Coef.	Std. Err.
log(net income+1)	0.171 ***	0.050	0.125 *	0.067	89.6%	0.124 **	0.060
work full	-0.167 *	0.097	-0.061	0.225	8.0%	-0.059	0.178
work part	0.002	0.096	0.146	0.172	4.6%	0.147	0.141
work little	0.016	0.092	0.210	0.155	10.2%	0.210 *	0.125
# income sources	0.063 **	0.028	0.019	0.046	48.0%	0.020	0.038
high heritage probability	0.258 **	0.110	0.187	0.164	17.3%	0.187	0.143
unemployed	-0.102	0.112	-0.009	0.182	8.1%	-0.008	0.150
past unemployment	0.055	0.052	-0.005	0.125	9.9%	-0.005	0.108
retired	-0.134	0.096	-0.023	0.193	2.4%	-0.017	0.162
std(net income)/net income	0.300	0.269	0.310	0.396	51.3%	0.309	0.327
significant earnings fluctuations	0.090	0.064	0.108	0.100	25.6%	0.109	0.084
slight earnings fluctuations	0.055	0.052	0.082	0.079	40.7%	0.081	0.066
provincial unemployment rate	-0.020	0.013	-0.093	0.091	97.0%	-0.092	0.075
expected years left to live	-0.010 ***	0.003	-0.010 *	0.006	96.5%	-0.010 **	0.005
age	0.011 **	0.004	0.013	0.028	100.0%	0.013	0.023
Eastern Germany	0.182	0.128	-0.042	0.346	0.3%	-0.112	0.307
female	-0.086	0.064		0.000	0.0%	-0.114 *	0.067
foreigner	-0.014	0.180	0.342	0.268	1.6%	0.342	0.225
single	-0.090	0.084	-0.195	0.260	2.5%	-0.196	0.220
separated or divorced	-0.376 ***	0.094	-0.110	0.234	4.1%	-0.111	0.193
widowed	-0.358 ***	0.125	-0.233	0.307	0.9%	-0.245	0.267
children in hh	-0.095	0.060	0.121	0.152	7.9%	0.124	0.124
additional person in hh	0.172 *	0.096	0.155	0.163	5.7%	0.156	0.142
basic education	-0.220 *	0.114	-0.071	0.237	3.6%	-0.072	0.195
undergraduate education	0.187 ***	0.068	0.465 ***	0.157	8.2%	0.465 ***	0.127
graduate education	0.223 ***	0.085	0.291	0.206	5.9%	0.291	0.178
civil servant	0.080	0.107	0.184	0.309	0.8%	0.188	0.267
selfemployed or freelancer	0.156	0.105	-0.087	0.236	4.1%	-0.086	0.188
good state of health	-0.134 ***	0.052	-0.046	0.081	8.7%	-0.047	0.068
poor state of health	-0.129	0.092	-0.127	0.149	21.9%	-0.127	0.123
poor development of hh health	0.068	0.067	-0.016	0.095	17.0%	-0.015	0.081
future income situation	0.007	0.009	0.005	0.013	54.4%	0.005	0.011
develop. own economic sit.	0.025 **	0.013	-0.004	0.022	77.1%	-0.004	0.018
smoker	-0.149 **	0.062	-0.132	0.180	5.8%	-0.131	0.147
easy going	-0.188 **	0.079	-0.028	0.121	15.1%	-0.028	0.100
risk aversion	0.024 *	0.012	0.015	0.021	91.7%	0.015	0.018
no min. amount in cur. account	-0.238 ***	0.050	-0.215 ***	0.077	29.8%	-0.215 ***	0.064
no regular saver	-0.199 ***	0.049	-0.097	0.079	26.9%	-0.097	0.065
no liability insurance	-0.216 ***	0.083	-0.082	0.136	11.7%	-0.083	0.115
pr. occup. disability insurance	-0.099 *	0.056	0.082	0.101	13.4%	0.082	0.083
regular support	0.138	0.090	0.097	0.153	8.1%	0.096	0.132
liquidity index	-0.075	0.089	-0.053	0.159	10.1%	-0.053	0.132
overdraft limit	0.027 ***	0.006	0.020 *	0.011	64.5%	0.020 **	0.008
motive bequest	0.036 ***	0.008	0.030 **	0.015	70.6%	0.030 **	0.012
motive old-age	0.057 ***	0.010	0.027 *	0.014	76.7%	0.027 **	0.012
business owner	0.208 *	0.110	0.139	0.252	3.8%	0.133	0.200
wealth_1	-0.321 ***	0.069	-0.047	0.108	16.3%	-0.047	0.089
wealth_3	0.181 **	0.077	-0.016	0.126	21.3%	-0.014	0.106
wealth_4	0.364 ***	0.093	-0.058	0.161	13.5%	-0.057	0.134
homeowner	0.042	0.076	-0.007	0.188	6.8%	-0.007	0.154
d2005	-0.699 ***	0.052	-0.582 ***	0.074		-0.582 ***	0.062
d2007	-0.109 ***	0.041	-0.225	0.187		-0.222	0.153
constant	6.286 ***	0.514	7.665 ***	1.976		6.416 ***	0.697
average R-sq within	0.064		0.078			0.078	
average R-sq between	0.290		0.156			0.303	
average R-sq overall	0.242		0.142			0.260	

Note: * : 10% significance level; ** : 5% significance level; *** : 1% significance level.
Source: SAVE 2005-2007; restricted sample; unweighted; # obs 6527; # groups 3348;
obs per group: min=1; avg=1.9; max=3; Rubin's rules were used to calculate the coefficients and standard errors.
The coefficients of the additional covariates in the expanded RE model are not shown.

Table 12: Specifications 1.) - 5.) of the expanded random effects estimation

expanded RE	**1.)**		**2.)**		**3.)**		**4.)**		**5.)**	
log(prec. savings+1)	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
log(net income+1)	0.124 **	0.061	0.136 ***	0.044	0.171 *	0.091	0.088	0.076	0.104 *	0.060
work full	-0.057	0.178	-0.001	0.120	0.029	0.281	0.283	0.206	-0.057	0.182
work part	0.149	0.141	0.054	0.105	0.216	0.206	0.303 *	0.181	0.154	0.145
work little	0.212 *	0.125	0.091	0.098	0.274	0.168	0.204	0.170	0.194	0.129
# income sources	0.020	0.038	-0.021	0.028	0.043	0.048	-0.030	0.055	0.027	0.040
high heritage prob.	0.185	0.144	0.163	0.121	0.117	0.204	0.329	0.241	0.199	0.147
unemployed	-0.008	0.150	0.062	0.106	0.107	0.220	0.171	0.182	-0.024	0.154
past unemployment	-0.004	0.108	-0.118	0.082	0.014	0.163	0.254	0.166	0.010	0.112
retired	-0.017	0.164	-0.122	0.121	0.046	0.257	-0.749 *	0.391	0.029	0.173
std(net income)/net income	0.299	0.326	0.198	0.260	0.486	0.417	0.152	0.403	0.248	0.335
sig. earnings fluct.	0.107	0.084	0.083	0.064	0.126	0.115	0.040	0.115	0.094	0.087
slight earnings fluct.	0.080	0.066	0.047	0.048	0.056	0.088	-0.002	0.092	0.074	0.067
pr. unemployment rate	-0.090	0.075	-0.031	0.049	-0.034	0.129	-0.060	0.083	-0.091	0.076
exp. years left to live	-0.010 **	0.005	0.000	0.003	-0.008	0.007	-0.012 **	0.006	-0.009 *	0.005
age	0.013	0.023	0.006	0.019	-0.035	0.050	-0.023	0.040	0.014	0.023
Eastern Germany	-0.120	0.287	-0.177	0.268	-0.275	0.385	-0.197	0.392	-0.145	0.346
female	-0.139 **	0.068	-0.056	0.047	-0.072	0.086	-0.094	0.091	-0.116 *	0.069
foreigner	0.350	0.225	0.126	0.205	0.626 *	0.367	0.371	0.284	0.397	0.247
single	-0.192	0.223	-0.240	0.164	-0.198	0.297	-0.432 **	0.213	-0.180	0.229
separated or divorced	-0.109	0.192	-0.148	0.130	-0.028	0.289	0.010	0.234	-0.124	0.201
widowed	-0.244	0.267	-0.041	0.244	-0.140	0.373	-0.311	0.431	-0.263	0.275
children in hh	0.127	0.124	0.121	0.094	0.134	0.206	-0.121	0.173	0.175	0.126
add. person in hh	0.159	0.140	0.274 **	0.111	0.089	0.222	-0.028	0.202	0.141	0.148
basic education	-0.072	0.195	0.053	0.140	-0.014	0.316	-0.066	0.317	-0.033	0.196
undergr. education	0.463 ***	0.127	0.283 ***	0.098	0.518 **	0.179	0.239	0.192	0.465 ***	0.133
graduate education	0.286	0.179	0.253 **	0.111	0.266	0.216	0.058	0.186	0.249	0.202
civil servant	0.191	0.274	0.248	0.248	-0.089	0.331	0.160	0.357	0.220	0.293
selfem. or freelancer	-0.086	0.189	0.117	0.129	-0.051	0.264	0.298	0.269	-0.053	0.207
good state of health	-0.046	0.068	-0.035	0.053	0.015	0.094	-0.031	0.104	-0.064	0.068
poor state of health	-0.126	0.122	-0.026	0.091	-0.154	0.185	0.448 **	0.215	-0.111	0.127
poor develop. of hh health	-0.014	0.081	-0.080	0.066	-0.067	0.116	-0.029	0.146	-0.004	0.083
future income situation	0.005	0.011	0.005	0.008	0.010	0.015	0.018	0.013	0.008	0.011
develop. own eco. sit.	-0.004	0.018	0.000	0.013	-0.017	0.026	-0.029	0.025	0.000	0.019
smoker	-0.133	0.147	-0.080	0.099	-0.044	0.213	-0.234	0.185	-0.127	0.150
easy going	-0.030	0.100	0.032	0.068	-0.062	0.147	0.294 **	0.141	-0.031	0.100
risk aversion	0.015	0.018	0.002	0.013	0.019	0.027	0.014	0.026	0.016	0.018
no min. amount	-0.215 ***	0.064	-0.071	0.046	-0.257 ***	0.098	-0.183 **	0.084	-0.218 ***	0.066
no regular saver	-0.099	0.065	-0.012	0.050	-0.112	0.094	-0.048	0.095	-0.071	0.066
no liability insurance	-0.083	0.115	-0.110	0.087	-0.145	0.153	0.192	0.164	-0.080	0.116
pr. occup. disability ins.	0.082	0.082	0.103	0.068	0.011	0.115	0.174 *	0.106	0.085	0.086
regular support	0.101	0.132	0.033	0.102	-0.208	0.225	0.055	0.154	0.110	0.132
liquidity index	-0.052	0.131	0.041	0.085	0.051	0.199	0.045	0.160	-0.047	0.135
overdraft limit	0.020 **	0.008	0.012 *	0.006	0.022 *	0.013	0.029 *	0.018	0.020 **	0.009
motive bequest	0.030 **	0.012	0.024 ***	0.008	0.011	0.017	0.035 **	0.018	0.031 **	0.012
motive old-age	0.027 **	0.012	0.005	0.009	0.029 *	0.017	0.060 ***	0.020	0.024 **	0.012
business owner	0.131	0.197	0.151 *	0.083	0.045	0.316	0.243	0.222		
wealth_1			-0.051	0.075	-0.063	0.128	0.008	0.104	-0.032	0.089
wealth_3			0.041	0.082	0.039	0.130	0.112	0.130	-0.030	0.111
wealth_4			0.125	0.102	0.075	0.190	0.058	0.179	-0.090	0.139
homeowner			-0.086	0.114	-0.121	0.205	-0.033	0.196	-0.003	0.159
d2005	-0.583 ***	0.062	-0.119 ***	0.044	-0.521 ***	0.092	-0.528 ***	0.090	-0.594 ***	0.063
d2007	-0.220	0.153	-0.111	0.103	-0.063	0.266	-0.162	0.174	-0.213	0.156
constant	6.125 ***	0.723	6.303 ***	0.464	5.148 ***	0.966	5.553 ***	0.862	6.483 ***	0.707
average R-sq within	0.078		0.032		0.078		0.088		0.079	
average R-sq between	0.282		0.301		0.413		0.267		0.297	
average R-sq overall	0.242		0.273		0.265		0.218		0.255	
# obs	6527		6359		2802		3410		6222	
# groups	3348		3271		934		1778		3237	

Note: * : 10% significance level; ** : 5% significance level; *** : 1% significance level.
Source: SAVE 2005-2007; restricted sample; unweighted; Rubin's rules were used to calculate the coefficients and standard errors. The coefficients of the additional covariates in the expanded RE model are not shown.

Specification:				
	1.)	without ex ante endogenous variables	4.)	restricted to age<=50
	2.)	restricted to pos. precautionary savings	5.)	restricted to non-business owners
	3.)	restricted to a balanced panel		

7.2 Discussion of the results

In the next section the results are discussed. Since the risk variables are of special interest, this section begins with the description of the results related to the risk variables. The rest of the section addresses variables which have been found to be significant. If the results are not further specified, they refer to the expanded RE estimates of table 11.

Income risk

The results of table 11 and the robustness checks of table 12 show that the forward looking and backward looking measures of income risk have the right positive sign (with only one exception, namely for slight income fluctuations in the age<= 50 specification, which however is close to zero and highly insignificant). Moreover, an increase of precautionary savings can be observed from the base group with no income fluctuations to the group with slight income fluctuations and even stronger to the group with significant income fluctuations. The p-value of significant income fluctuations is 20%, and the p-value of slight income fluctuations is 22% for the expanded RE model. Precautionary savings increase by 10.9% for a household with significant income fluctuations compared to a household with no income fluctuations, holding all other factors constant. The p-value for the forward looking risk measure is 35% in the expanded RE model. Thus, as it was already the case in the cross-sectional analysis, the forward and backward looking income risk variables remain insignificant at the usual significance levels.

The reasons for insignificance of the coefficients are multi-layered. The questions and the variables constructed from these questions could be the first reason. The backward looking risk measure was constructed using the question about income fluctuation in the past five years. This question does not distinguish between expected and unexpected income fluctuations. Only unexpected income fluctuation should increase precautionary savings. Seasonal workers, freelancers, and self-employed workers may face strong but expected income fluctuations. Furthermore, fluctuation or variation in income must not be that great of a concern, as long as labour income does not fall below a certain income that the household wants to maintain in every case. Considering this, fluctuations must not be an appropriate measure of labour income risk. The shortfall probability can be a better measure of risk. However, such a measure is even harder to construct since the income bound the household wants to maintain has to be known, positive and negative shocks must be distinguished, and the magnitude of the shocks should be determined.

The forward looking measure faces the challenge of the very short time horizon the households were asked to take into account (on average, slightly more than half a year from the point the question was asked until the end of that year). If the household heads were able to answer this question within the time frame requested, longer term labour contracts and zero settings for retirees, civil servants, and individuals who were not working may result in many answers of no unemployment risk at all (Lusardi, 1997, p. 322), which is actually the case (2005: 76.5%; 2006: 71.77%; 2007: 70.31 % for the household head; 2005: 71.82%; 2006: 66.67%; 2007: 65.78% for the partner [average over 5 datasets]).

The general measure of income risk is insignificant too. The coefficient is negative and has the opposite sign compared to the theoretical prediction. Thus, more individual specific measures of risk seem to be more reliable variables to measure the risk exposure of an individual.

Another reason why the measurements or proxies for income variation have mainly insignificant results is probably due to the German social insurance system. Public unemployment insurance may reduce the effects of income uncertainty compared to other countries like the US. A restriction of the sample to household heads not older than 50 years, in which the buffer-stock behaviour is most likely to emerge, yields even weaker significance levels.

Health risk

Poor expectations about the future health status do not change precautionary savings significantly. The coefficient is highly insignificant for most of the specifications. The reason for that could be based on the German health insurance system, which is obligatory for a high proportion of the German population. From 2004 on, additional payments for medicines and nursing, the quarterly medical practice fee, and so on increased. Nevertheless, the maximum amount a household has to pay is restricted to 2% of its yearly gross income after taking tax exemption limits for families into account. An exemption is the restriction to 1% of the gross income for individuals who have a chronic disease (Merten, 2003).

As noted earlier, the discussion and reformation of the German health care system are still in progress. It would be interesting to investigate how the effect of this variable will develop in future surveys.

Longevity risk

The proxy variable for longevity risk, the expected years left to live (maximum of the household head or the partner), shows negative and significant results for most of the specifications. Already the cross-sectional estimates point in this direction. A one year increase in the expected years left to live reduces precautionary savings by 1%. This result is somehow puzzling, since it was expected that with increasing years left to live precautionary savings would increase due to a longer time horizon for instance in which the household can face different kinds of risks such as health risks. A possible explanation could be that households having more years left to live are healthier households. However, control variables have been included to hold factors related to health constant. The expectation of more years left to live could be related to a more optimistic view, and more optimistic households will also hold less precautionary savings. But controlling for optimism does not alter this result. Another possibility is that more years left to live are an indicator of a less risky life, and this leads to a decrease in precautionary savings. To prevent such a relationship, risk aversion was already included in the specifications. Moreover, the effects are not significant if the sample is restricted to household heads aged 50 or below. The coefficient is almost equal to zero if households with no precautionary savings are excluded. This result may indicate that households with no precautionary savings do not care much about their future and may tend to overestimate their remaining life expectancy due to the lack of information about factors that may reduce their life expectancy.

Pension risk

Since this variable was not available in the 2005 dataset, only the cross-sectional analysis of 2006 and 2007 can be used to draw some inference (table 10). The variable of pension uncertainty has a p-value of 13% in 2006 and a p-value of 2% in 2007. On the first viewing, the negative sign of pension uncertainty may cause amazement. The variable used was taken out of a set of questions in which the household head had to guess the percentage of his/her pension and that of the partner compared to their anticipated last wage. There was also the possibility of an answer of "no guess possible". If there is no guess possible, uncertainty about the future pension level must be high. The positive effect could be, however, dominated by another effect pointing in the other direction. The household head may be less informed about his future pension level, since he does not care that much about it. If he does not seriously care about it, it is likely that the household head will not care about future uncertainty and therewith precautionary savings as well. Even if the attempt was done to control for such preference structures, it could be that the controls are not able to fully assimilate this effect. In addition, especially for younger households it might be more difficult to estimate their future pension level than for older households. This is confirmed in the negative correlation coefficient of -0.31 between the discussed dummy variable and the age of the household head.

Current Income

If current monthly net income increases by 10%, precautionary savings will increase by 1.24%. The effect is significant throughout all specifications (the significance level drops to 25% in only one specification). The following calculation should be allowed, although only current income and not permanent income is available. If a household's monthly net income, which corresponds to the mean monthly net income over the sample, is increased by €1,000 (all other factors are held constant by the sample means), average precautionary savings will increase by €405 (calculations are done using the unweighted sample means over all imputed datasets). Evaluated at the 1995 population mean, Kennickell & Lusardi (2005, p. 23) found a precautionary savings increase of $221 after a "normal" income raise of $1,000.

Education

As found in other studies (Arrondel, 2002; Engen & Gruber, 2001; Lusardi, 1998) dealing with the precautionary saving motive, education has a positive influence on precautionary savings. The effect for undergraduate education is highly significant over all except for one specification. Graduate education reaches the 5% or 10% significance level or comes close to the 10% significance level in almost all specifications. Whereas the sign of basic education is negative compared to the base group (high education), undergraduate education results in 47% more precautionary savings compared to the base group. This effect is reduced to a still very high increase of 29% for graduate education compared to the base group.

Household composition

The variables of household composition show a robust pattern over all specifications (exception: specification 4). The marital status of single, separated or divorced, and widowed lead to at least a 10% reduction of precautionary savings compared to the base group of married and living with husband/wife. Thus, the protection against unforeseen events of the husband or wife living with the household head induces higher precautionary savings, holding other influences constant. Children and an additional person in the household have positive coefficients on precautionary savings as expected. Admittedly, all the coefficients are not significant at the usual significance levels due to too high standard errors.

Foreigner

That foreigners may face additional risks compared to natives is indicated by the positive coefficient, which is close to the 10% significance level for almost all specifications. A household without German citizenship has 34% more precautionary savings than a German household with the same characteristics.

Preferences

The dummy for holding a certain minimum amount in a current account is highly significant if households with zero precautionary savings are included in the sample. This variable seems to reflect quite well a precautionary attitude. If a household does not try to ensure a minimum amount in his current account, precautionary savings are reduced by 22% for unchanged characteristics of the household.

In addition, precautionary savings of households with irregular saving behaviour or households with a smoking household head have the expected negative coefficients in all specifications but are not significant at the usual significance levels.

Moreover, the result of the coefficient of risk aversion is puzzling. Even if the coefficient is not significant over different specifications, the coefficient unexpectedly depicts the opposite sign. An indication for that could already be found in the descriptive analysis. In most of the specifications, a one point increase towards the direction of not risk-averse at all on a scale from 0 to 10 increases precautionary savings between 1% and 2%.

Liquidity index

Only the overdraft limit in this category shows a significant effect over all specifications. In contrast to the expectation, a higher overdraft limit of €1,000 increases precautionary savings by 2%. Instead of a dominating effect of the possibility to borrow in the short-run, this variable may reflect again a precautionary attitude. Through this precautionary attitude both the overdraft limit and precautionary savings may rise.

Insurance

Private insurance coverage (occupational disability and liability insurance) increases precautionary savings at around 8% for each of the two private insurances. The signs of the insignificant coefficients are stable over the specifications with only one exception for the private liability insurance in specification 4. This finding suggests that private insurance coverage must be endogenous. Starr-McCluer (1996) drew the same conclusion from her investigation about the effect of private health insurance coverage.

Other saving motives

The linkage between different saving motives can be seen in the significant coefficients of the bequest and old-age provision motives. This effect was also found by Lusardi (1998). A higher importance of one point on a scale ranging from 0 to 10 results for both motives in an increase of precautionary savings of around 3%. Thus, households who tend towards leaving bequest or care about their old-age provision accumulate a larger amount of precautionary savings, holding other factors constant.

Further comments

Before this sections ends, some comments are made about the results of variables which have not been mentioned. As presented in section 4, the influence of being a business owner is always positive, however only once significant at a 10% level. Both proxy variables of the expectation about future income growth have small standard errors. However, the coefficients are close to zero which leads to insignificant t-values. As can be seen from the intercepts, there is no statistically significant difference between 2006 and 2007. In contrast to this finding, the dummy coefficient of 2005 is negative and highly significant. This confirms the observations already made at the beginning of section 4 namely that the unweighted distribution of years 2006 and 2007 are similar, whereas both years 2006 and 2007 differ significantly compared to 2005. Even if the sample is restricted to a balanced panel (specification 3), the dummy for 2005 remains highly significant. Since the question remained the same over all three years, this fact is puzzling.

8 Conclusion

This study has investigated the effects of different factors on a subjective measure of precautionary savings in a qualitative and quantitative way using the German SAVE dataset 2005-2007. In a first step, the development of theoretical models incorporating the precautionary saving motive was described with a special focus on Carroll's buffer-stock model (1992, 1997). This dispute helped to identify important theoretical variables and their mechanisms to influence precautionary savings. The drawbacks of this study may be seen in the discussible meaning of the question about the desired amount of precautionary savings and the imputation procedure of this question, which was inconsistent over the years and not appropriate in the author's opinion. Even if the imputation procedure was only a side issue of this study at first, the high item non-response rate of the question of concern and the strong effect of the imputation of this question on the complete sample made a deeper investigation necessary. This investigation concluded with a restriction of the sample size to observed values of precautionary savings only.

The following descriptive analysis showed amongst other things that precautionary savings with a mean of €8,298 (2005-2007, weighted) account for 6% of total net wealth and 26% of financial wealth, which supports the evidence for a low or at most moderate contribution of precautionary savings to wealth accumulation. The problems associated with the empirical estimation of a reduced form equation of precautionary savings were discussed in depth.

After the use of panel estimation techniques and the application of Rubin's rules, estimations on the restricted sample present mainly insignificant coefficients on the risk variables. Explanations for this insignificance range from drawbacks of the constructed variables to the reduced risk caused by the German social insurance and welfare system. As expected, precautionary savings increase with current income. The length of education and professional training has very strong effects on the amount of precautionary savings. Undergraduate education leads to 47% and graduate education to 29% higher precautionary savings compared to high education in the base group, other relevant factors being equal. The relationships between precautionary savings and the old-age as well as the bequest motive are also significant and result in an increase in precautionary savings caused by an increase in the importance of these motives. Behaviour towards a precautionary attitude also has a very strong and significant effect. The dummy for holding a certain minimum amount in a current account and the amount of the overdraft limit point in this direction. There is also some evidence that foreigners have additional precautionary savings.

Even if the insignificance of the risk variables are to a certain extent disappointing, this study found significant explanatory variables even after a rich set of controls were included and sensitivity tests were applied. Furthermore, this study contributes by implementing a method to prove the qualitative and quantitative influence of variables on a subjective measure of precautionary savings with the German SAVE dataset. To improve the further examination of the research question, the following advices should be considered: first, a more precise and better understandable question about the desired amount of precautionary savings in the ongoing SAVE surveys may help to reduce the item non-response and the precision of the answers themselves; second, a new imputation procedure making use of the panel structure of the survey and other estimation techniques to allow the estimated values to be only in a defined range given by the answers of the respondents may raise the overall data quality and prevent a selectivity bias additional to the increase in the sample size; third, better specified questions about different kinds of risks should help to achieve significant coefficients.

The realisation of the above mentioned improvements would allow for investigating this research question again and for researching other interesting questions. For instance, different measurements of wealth have mainly been used as a proxy for precautionary savings so far. How extensive is the bias caused by these proxy variables compared to the new subjective measure of precautionary savings? This can be investigated by comparing the coefficients and standard errors of the explanatory variables estimated with the same regression specification only with different explained variables. Depending on the definition, wealth can be negative for quite a large portion of households. Since the explained variable in my analysis is the logarithm of the desired amount of precautionary savings, an inverse hyperbolic sine transformation of wealth (Pence, 2006, pp.

5-7; Essig, 2005, pp. 10-11, 48; Carroll & Dynan & Krane, 1999, pp. 14-16) with a transformation parameter of $\theta=1.05$[16] would offer a way to compare the different measures.

Maybe it is in the precautionary saving motive's nature that many questions, e.g., about the strength of the precautionary saving motive, its influencing factors, and the magnitude of precautionary savings cannot be satisfactorily answered. Thus, the engagement of economic research has to continue to better understand this important part of households' saving behaviour.

[16] The inverse hyperbolic sine function will be almost equal to the logarithm function for a transformation parameter of $\theta=1.05$ and strictly positive values.

9 References

Akerlof, G. (2007): "The missing motivation in macroeconomics." American Economic Review, 97, 5-36.

Alessie, R. and A. Kapteyn (2001): "New Data for Understanding Saving." Oxford Review of Economic Policy, 17(1), 55-69.

Arrondel, L. (2002): "Risk management and wealth accumulation behavior in France." Economics Letters, 74, 187-194.

Bannock, G., R. E. Baxter, and E. Davis (2003): "Dictionary of economics." London: Economist [et al.].

Barceló, C. (2006): "Imputation of the 2002 wave of the Spanish Survey of Household Finances (EFF)." Occasional Paper No. 0603, Bank of Spain.

Barsky, R. B., N. G. Mankiw, and S. P. Zeldes (1986): "Ricardian Consumers with Keynesian Propensities." American Economic Review, 76(4), 67-91.

Bartzsch, N. (2006): "Precautionary saving and income uncertainty in Germany – new evidence from microdata." Deutsche Bundesbank Discussion Paper Series 1, 44/2006.

Berkhof, J. and T. A. B. Snijders (2007): "Diagnostic Checks for Multilevel Models." In: Leeuw, J. de and E. Meijer (Ed.): "Handbook of Multilevel Analysis." 139-173, New York: Springer.

Börsch-Supan, A. (2005): "Risiken im Lebenszyklus - Theorie und Evidenz." Mea-Discussion-Paper 69-2005, MEA-Mannheim Research Institute for the Economics of Aging, University of Mannheim.

Börsch-Supan, A., M. Coppola, L. Essig, A. Eymann, and D. Schunk (2008): "The German SAVE study. Design and Results." Mea Studies 06, MEA-Mannheim Research Institute for the Economics of Aging, University of Mannheim.

Börsch-Supan, A. and L. Essig (2002): “Sparen in Deutschland: Ergebnisse der ersten SAVE-Studie.” Köln: Deutsches Institut für Altersvorsorge.

Börsch-Supan, A. and L. Essig (2005): “Personal assets and pension reform: How well prepared are the Germans?“ Mea-Discussion-Paper 85-2005, MEA-Mannheim Research Institute for the Economics of Aging, University of Mannheim.

Börsch-Supan, A. and A. Lusardi (2003): “Saving: A Cross-National Perspective.” In: Börsch-Supan, A. (Ed.): “Life-Cycle Savings and Public Policy: A Cross-National Study in Six Countries.” 1-32, New York: Academic Press.

Browning, M. and T. F. Crossley (2001): “The Life-Cycle Model of Consumption and Saving.“ The Journal of Economic Perspectives, 15(3), 3-22.

Browning, M. and A. Lusardi (1996): “Household Saving: Micro Theories and Micro Facts.” Journal of Economic Literature, 34(4), 1797-1855.

Bundesagentur für Arbeit (2007): “Unemployment money.” http://www.arbeitsagentur.de/nn_25650/Navigation/zentral/Buerger/Arbeitslos/Alg/Alg-Nav.html, last internet access 22.2.2008.

Caballero, R. J. (1991): “Earnings Uncertainty and Aggregate Wealth Accumulation.” The American Economic Review, 81(4), 859-871.

Cagetti, M. (2003): “Wealth Accumulation Over the Life Cycle and Precautionary Savings.“ Journal of Business & Economic Statistics, 21(3), 339-353.

Caneron, A. C. and P. K. Trivedi (2005): “Microeconometrics. Methods and Applications.” Cambridge [et al.]: Cambridge University Press.

Carroll, C. D. (1992): “The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence.” Brookings Papers on Economic Activity, 2, 61-135.

Carroll, C. D. (1997): “Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis.” Quarterly Journal of Economics, 112, 1-56.

Carroll, C. D. (2000): “Why Do the Rich Save So Much?” In: Slemrod, J. B. (Ed.): “Does Atlas Shrug? The Economic Consequences of Taxing the Rich.” Harvard University Press.

Carroll, C. D. (2004): “Theoretical Foundations of Buffer Stock Saving.” Economics Working Paper Archive 517, The Johns Hopkins University, Department of Economics.

Carroll, C. D., K. E. Dynan, and S. D. Krane, (1999): “Unemployment Risk and Precautionary Wealth: Evidence from Households' Balance Sheets.“ Economics Working Paper Archive 416, The Johns Hopkins University, Department of Economics.

Carroll, C. D. and M.S. Kimball (2007): “Precautionary Saving and Precautionary Wealth.“ Entry for The New Palgrave Dictionary of Economics, 2nd Ed.

Carroll, C. D. and A. A. Samwick (1998): “How Important is Precautionary Saving?” Review of Economics and Statistics, 80(3), 410-419.

Carroll, C. D. and L. H. Summers (1991): "Consumption Growth Parallels Income Growth: Some New Evidence." In: Bernheim, B. D. and J. B. Shoven (Eds.): “National Saving and Economic Performance.” 305-343, Chicago: Chicago University Press for NBER.

Deaton, A. (1991): “Saving and Liquidity Constraints.“ Econometrica, 59(5), 1221-1248.

Dynan, K. E. (1993): “How Prudent are Consumers?” The Journal of Political Economy, 101(6), 1104-1113.

Engen, E. M. and J. Gruber (2001): “Unemployment insurance and precautionary saving.” Journal of Monetary Economics, 47, 545-579.

Essig, L. (2005): “Precautionary saving and old-age provisions: Do subjective saving motive measures work?“ Mea-Discussion-Paper 84-2005, MEA-Mannheim Research Institute for the Economics of Aging, University of Mannheim.

Friedman, M. (1957): “A Theory of the Consumption Function.” Princeton: Princeton University Press.

Fuchs-Schündeln, N. and M. Schündeln (2005): “Precautionary Savings and Self-Selection: Evidence from the German Reunification “Experiment”.” Quarterly Journal of Economics, 120(3), 1085-1120.

Fudenberg, D. and D. K. Levine (2006): “A Dual Self Model of Impulse Control.” American Economic Review, 96(5), 1449-1476.

Gourinchas, P. O. and J. Parker (2002): “Consumption Over the Life-Cycle.” Econometrica, 70, 47-89.

Gruber, J. and A. Yelowitz (1999): “Public Health Insurance and Private Savings.” The Journal of Political Economy, 107(6), 1249-1274.

Guiso, L., T. Jappelli, and D. Terlizzese (1992): “Earnings Uncertainty and Precautionary Saving.” Journal of Monetary Economics, 30, 307-337.

Guiso, L., T. Jappelli, and D. Terlizzese (1996): “Income Risk, Borrowing Constraints, and Portfolio Choice.“ The American Economic Review, 86(1), 158-172.

Gujarati, D. N. (2003): “Basic Econometrics.” 4th Ed., Mc Graw Hill.

Hall, R. E. (1978): “Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence.“ The Journal of Political Economy, 86(6), 971-987.

Heckman, J. J. (1976): “The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models.” Annals of Economic and Social Measurement, 5(3), 475-492.

Hoynes, H., M. Hurd, and H. Chand (1998): “Household Wealth of the Elderly under Alternative Imputation Procedures.” In: Wise, D. A. (Ed.): “Inquiries in the Economics of Aging.” 229-257, Chicago: The University of Chicago Press.

Hubbard, R. G., J. Skinner, and S. P. Zeldes (1995): “Precautionary Saving and Social Insurance.” The Journal of Political Economy, 103(2), 360-399.

Hurd, M. D. (1987): “Savings of the Elderly and Desired Bequests.” The American Economic Review, 77(3), 298-312.

Hurd, M. D. and K. McGarry (1995): "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study." Journal of Human Resources, 30(0), S268-S292.

Hurd, M. D. and K. McGarry (2002): “The Predictive Validity of Subjective Probabilities of Survival.” The Economic Journal, 112(482), 966-985.

Hurst, E., A. Kennickell, and A. Lusardi (2005): “Precautionary Savings and the Importance of Business Owners." NBER Working Papers 11731, National Bureau of Economic Research, Inc.

Imbens, G.W. and J. M. Wooldridge (2007): “Linear Panel Data Models.“ Lecture Notes.

Jappelli, T., M. Padula, and L. Pistaferri (2006): "A Direct Test of the Buffer-Stock Model of Saving." CEPR Discussion Papers 6576, C.E.P.R. Discussion Papers.

Kazarosian, M. (1997): “Precautionary Saving – A Panel Study.” Review of Economics and Statistics, 79, 241-247.

Kennickell, A. B. (1998): “Multiple Imputation in the Survey of Consumer Finances.” Proceedings of the 1998 Joint Statistical Meetings, Dallas TX.

Kennickell, A. and A. Lusardi (2005): “Disentangling the Importance of the Precautionary Saving Motive.” Working Paper. Board of Governers of the Federal Reserve System. Washington D.C.

Khwaja, A., F. Sloan, and M. Salm (2006): ”Evidence on preferences and subjective beliefs of risk takers: The case of smokers.“ International Journal of Industrial Organization, 24(4), 667-682.

Kimball, M. S. (1990a): “Precautionary Saving in the Small and in the Large.“ Econometrica, 58(1), 53-73.

Kimball, M. S. (1990b): "Precautionary Saving and the Marginal Propensity to Consume." Working Paper 3403, NBER, Cambridge, MA.

King, M. A. and L-D. L. Dicks-Mireaux (1982): "Asset Holdings and the Life-Cycle." The Economic Journal, 92(366), 247-267.

Kong, M. K., J. Y. Lee, and H. K. Lee (2007): "Precautionary Motive for Saving and Medical Expenses under Health Uncertainty: Evidence from Korea." Economics Letters, doi: 10.1016/j.econlet.2007.11.004.

Krueger, D. and F. Kubler (2006): "Pareto-Improving Social Security Reform when Financial Markets Are Incomplete!?" American Economic Review, 96(3), 737-755.

Lavoie, M. (2004): "Post Keynesian consumer theory: Potential synergies with consumer research and economic psychology." Journal of Economic Psychology, 25, 639-649.

Leland, H. E. (1968): "Saving and Uncertainty: The Precautionary Demand for Saving." Quarterly Journal of Economics, 82, 465-473.

Lillard, L. A. (1977): "Inequality: Earnings vs. Human Wealth." The American Economic Review, 67(2), 42-53.

Lillard, L. A. and R. J. Willis (1978): "Dynamic Aspects of Earning Mobility." Econometrica, 46(5), 985-1012.

Little, R. J. A. and D. B. Rubin (2002): "Statistical Analysis with Missing Data." New York: Wiley.

Lusardi, A. (1997): "Precautionary saving and subjective earnings variance." Economics Letters, 57, 319-326.

Lusardi, A. (1998): "On the Importance of the Precautionary Saving Motive." American Economic Review, 88(2), 449-453.

Lusardi, A. (2003): "Planning and Savings for Retirement." Mimeo, Dartmouth College.

Menz, J.-O. (2007): "Precautionary Saving in a macroeconomic perspective: A comparison between New Keynesian and Post Keynesian consumption theory." Conference Paper, University of Warwick, United Kingdom.

Merten, M. (2003): „Die neuen Zuzahlungsregeln.“ Deutsches Ärzteblatt, 100(51-52), A-3360 / B-2799 / C-2617. http://www.aerzteblatt.de/v4/archiv/artikel.asp?id=39870, last internet access 24.2.2008.

Miller, B. L. (1976): “The Effect on Optimal Consumption of Increased Uncertainty in Labor Income in the Multiperiod Case.” Journal of Economic Theory, 13, 154-167.

Modigliani, F. and R. Brumberg (1954): “Utility analysis and the consumption function: An interpretation of cross-section data.” In: Flavell, J. H. and L. Ross (Eds.): “Social Cognitive Development Frontiers and Possible Futures.” Cambridge, NY: University Press.

Murata, K. (2003): “Precautionary Savings and Income Uncertainty: Evidence from Japanese Micro Data.“ IMES discussion paper series No. 2003-E-9.

Palumbo, M. G. (1999): “Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle.“ The Review of Economic Studies, 66(2), 395-421.

Pence, K. P. (2006): “The Role of Wealth Transformations: An Application to Estimating the Effect of Tax Incentives on Saving.“ Contributions to Economic Analysis & Policy, 5(1), article 20, 1-24.

Picone, G., M. Uribe, and R. M. Wilson (1998): “The effect of uncertainty on the demand for medical care, health capital and wealth.“ Journal of Health Economics, 17, 171-185.

Piracha, M. and Y. Zhu (2007): “Precautionary Savings by Natives and Immigrants in Germany.“ SOEPpapers on Multidisciplinary Panel Data Research, DIW Berlin, 1-25.

Radbill, L. and C. Winship (1994): “Sampling Weights and Regression Analysis.” Sociological Methods & Research, 23(2), 230-251.

Rodepeter, R. (1999): “Konsum- und Sparentscheidungen im Lebenszyklus. Theorie, Evidenz und Simulation unter klassischen und alternativen Ansätzen.” Dissertation, University of Mannheim.

Rodepeter, R. and J. K. Winter (1999): “Rules of thumb in life-cycle savings models.“ Sonderforschungsbereich 504 Publications 99-81, Sonderforschungsbereich 504, Universität Mannheim & Sonderforschungsbereich 504, University of Mannheim.

Romer, D. (2006): “Advanced macroeconomics.” 3rd ed., Boston [et al.]: McGraw-Hill.

Rubin, D. B. (1987): “Multiple Imputation for Nonresponse in Surveys.” New York: Wiley.

Rubin, D. B. (1996): “Multiple Imputation After 18+ Years.” Journal of the American Statistical Association, 91(434), 473-489.

Rubin, D. B. and N. Schenker (1986): “Multiple Imputation for Interval Estimation from Simple Random Samples with Ignorable Nonresponse.” Journal of the American Statistical Association, 81(394), 366-374.

Schunk, D. (2006): “The German SAVE Survey 2001 - 2006. Documentation and Methodology.“ Mea-Discussion-Paper 109-2006, MEA-Mannheim Research Institute for the Economics of Aging, University of Mannheim.

Schunk, D. (2007a): “A Markov Chain Monte Carlo Multiple Imputation Procedure for Dealing with Item Nonresponse in the German SAVE Survey.“ Mea-Discussion-Paper 121-2007, MEA-Mannheim Research Institute for the Economics of Aging, University of Mannheim.

Schunk, D. (2007b): “What Determines the Saving Behavior of German Households? An Examination of Saving Motives and Saving decisions.” Mea-Discussion-Paper 124-2007, MEA-Mannheim Research Institute for the Economics of Aging, University of Mannheim.

Schunk, D. (2008): “A Markov Chain Monte Carlo Algorithm for Multiple Imputation in Large Surveys.” Advances in Statistical Analysis, 92, 101-114.

Sibley, D. S. (1975): “Permanent and Transitory Effects of Optimal Consumption with Wage Income Uncertainty.” Journal of Economic Theory, 11, 68-82.

Skinner, J. (1988): “Risky Income, Life-Cycle Consumption and Precautionary Saving.” Journal of Monetary Economics, 22, 237-255.

Skrondal, A. and S. Rabe-Hesketh (2004): “Generalized Latent Variable Modeling. Multilevel, Longitudinal, and Structural Equation Models.” London [et al.]: Chapman & Hall/CRC.

Smith, V. K., D. H. Taylor Jr., and F. A. Sloan (2001): “Longevity Expectations and Death: Can People Predict Their Own Demise?” The American Economic Review, 91(4), 1126-1134.

Starr-McCluer, M. (1996): “Health Insurance and Precautionary Savings.“ The American Economic Review, 86(1), 285-295.

Statistische Ämter des Bundes und der Länder (2007): “Erwerbstätigkeit: Monatlicher Zahlenspiegel der Länder.“ http://www.statistik-portal.de/Statistik-Portal/de_inhalt02.asp, last internet access 23.2.2008.

Thaler, R. H. (1990): “Anomalies: Saving, Fungibility, and Mental Accounts.“ The Journal of Economic Perspectives, 4(1), 193-205.

Thaler, R. H. (1994): “Psychology and Savings Policies.“ The American Economic Review, 84(2), 186-192.

Wooldridge, J. M. (2002): “Econometric Analysis of Cross Section and Panel Data.” Cambridge, Mass. [et al.]: MIT Press.

Wooldridge, J. M. (2003): “Introductory Econometrics. A Modern Approach.” 2nd ed., Mason, Ohio [et al.]: Thomson.

Zeldes, S. P. (1989a): “Consumption and Liquidity Constraints: An Empirical Investigation.” The Journal of Political Economy, 97(2), 305-346.

Zeldes, S. P. (1989b): “Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence.” Quarterly Journal of Economics, 104(2), 275-298.

10 Appendix

Table A.1: Construction of the included variables in the estimated models

Variable	Description
precautionary savings	
log(precautionary _savings+1)	Logarithm of the amount of desired precautionary savings in € plus €1.
income and controls	
log(net income+1)	Logarithm of monthly household net income in € plus €1.
work_full work_part work_little	Dummies for categorical variable "extent of employment" of the household head: work_full = 1 if the household head has full-time employment. (at least 35 hours per week); 0 otherwise. work_part = 1 if the household head has part-time employment. (at least 15 hours, but less than 35 hours per week); 0 otherwise. work_little = 1 if the level of employment is low or casual. (less than 15 hours per week); 0 otherwise. Base group: not in paid employment.
# income sources	Number of income sources.
highheritage_prob	Subjective probability that the household will receive an inheritance or gift in the next two years, which significantly improves the household's financial situation. Scale from 0 to 1 with 0.1 increments.
unemployed	Dummy = 1 if the household head is currently unemployed.
past unemployment	Dummy = 1 if the household head has ever registered as unemployed.
retired	Dummy = 1 if the household head is retired.
income uncertainty	
std(net income)/net income	Using the question about the subjective probability to loose the current place of employment of the household head and the partner in the year of the survey, the standard deviation is calculated out of current net income and a replacement rate of 60%. It is assumed that there is no correlation of the employment status of the household head and his/her partner. The standard deviation is divided by the current household net income.
significant earnings fluctuations slight earnings fluctuations	Dummies for subjective ordinal variable "income fluctuation over the last 5 years": significant earnings fluctuations = 1 if income fluctuates significantly. slight earnings fluctuations = 1 if income fluctuates slightly. Base group: no earnings fluctuations.
provincial unemployment rate	Yearly average unemployment rate of the Federal State the household lives in.
longevity risk & controls	
expected years left to live	Maximum of the subjective life expectancy minus current age of the household head and the partner.
age	Age in years of the household head.
socio demographic controls	
east	Dummy = 1 if the household head lives in Eastern Germany.
female	Dummy = 1 if the household head is female.

foreign	Dummy = 1 if the household head is not a German citizen.
separated or divorced single widowed	Dummies for categorical variable "marital status": Base group: married and living together.
children in hh	Dummy =1 if there lives at least one child inside the household.
additional person in hh	Dummy =1 if there lives at least one additional person inside the household, which is not the partner or a child.
basic education undergraduate ed. graduate education	Dummies for categorical variables "highest general school or college leaving certificate" and "highest completed training for a professional qualification" of the household head: basic education = 1 if household head had 9 to 10 years educational training. undergraduate ed. = 1 if household head had 16 to 17 years educational training. graduate education = 1 if household head had 18 to 19 years educational training. Base group: household head had 13 to 14 years educational training.
civil servant selfemployed or freelancer	Dummies for categorical variable "employment status" of the household head: civil servant = 1 if current type of employment is civil servant. selfemployed or freelancer = 1 if current type of employment is self-employed or freelancer. Base group: other jobs or not employed.
health & controls	
good state of health bad state of health	Dummies for the ordinal variable "subjective current health status" of the household head: good state of health = 1 if state of health is described as good or very good. bad state of health = 1 if state of health is described as bad or very bad. Base group: fair state of health = 1 if state of health is described as fair.
bad development of hh health	Dummy = 1 if the expectation about the own or the partner's health situation on a scale from 0 (very negative) to 10 (very positive) is smaller than 4.
expected income growth	
future income situation	Subjective probability of an increase in the household head's net income one year ahead on a scale from 0 to 10 with 1 increments.
develop. own economic sit.	Ordinal variable about the expected development of the own financial situation ranging from 0 (very negative) to 10 (very positive) with 1 increments.
impatience	
smoker	Dummy = 1 if the household head is a smoker.
easy going	Dummy = 1 if the household head marked less than a 5 on a scale from 0 (easy going and take each day as it comes) to 10 (exactly planning the future).
preferences	
risk aversion	Average out of five questions about the self-evaluation of taking risks with respect to "my own health", "my career", "in money matters", "with respect to leisure time and sport", and "when driving" ranging from 0 (does not apply at all) to 10 (applies very well).
no min. amount in cur. account	Dummy = 1 if the household head subjectively replies not to ensure that he or she has a certain minimum amount in the current account.

no regular saver	Dummy for categorical variable "regularity of saving" of the household head and the partner. Dummy = 1 if the answers "I/we do not save because we do not have enough scope financially to do so" or "I/we do not save because we would prefer to enjoy life now" are given. Base group: all the other categories.
insurance	
pr. occup. disability insurance	Dummy = 1 if the household head or his/her partner is privately insured against occupational disability.
no liability insurance	Dummy = 1 if the household head or his/her partner has no liability insurance.
regular support	Dummy = 1 if the household receives regular support payments exceeding €25 per month from persons outside the household.
liqudity constraints	
liquidity index	Dummy =1 if a credit has been refused or not granted for the full amount requested or the household did not apply for a credit because the household members believed that the credit would be refused.
overdraft limit	Continuous variable of the agreed amount of overdraft facility. If the account has no overdraft facility, the value is set to zero.
other saving motives	
motive bequest	Absolute importance of the bequest motive on a scale from 0 (totally unimportant) to 10 (very important) with 1 increments.
motive old-age	Absolute importance of the old-age provision motive on a scale from 0 (totally unimportant) to 10 (very important) with 1 increments.
business	
business owner	Dummy =1 if the household head or the partner own any business assets.
wealth controls	
homeowner	Dummy = 1 if the household owns a home.
wealth_1 wealth_3 wealth_4	Dummies for each total net wealth (i.e., savings investments, building society savings, whole life insurance policies, savings bonds, share- and real-estate bonds, occupational and private pension schemes, real estate, business wealth etc.) quartile of the household: wealth_1 = 1 if total net wealth <= 3,000. wealth_3 = 1 total net wealth > 63,000 & total net wealth <= 222,000. wealth_4 = 1 total net wealth > 222,000. Base group: wealth_2 = 1 if total net wealth > 3,000 & total net wealth <= 63,000.
year dummies	
d2005 d2007	Dummy = 1 if the year of the survey is 2005. Dummy = 1 if the year of the survey is 2007. Base group: d2006 = 1 if the year of the survey is 2007.
uncertainty of pension (only 2006 + 2007)	
uncertain pension	Dummy = 1 if the respondent was not able to give an estimate of the percentage of the anticipated last wage/salary the household head or the partner will receive as their pension.
insurance (only 2007)	
obligation to contr. to social ins.	Dummy = 1 if the household head has an obligation to contribute to the social insurance system.
private long-term care insurance	Dummy = 1 if the household head has a private long-term care insurance and no social long-term care insurance.
add. long-term care ins.	Dummy = 1 if the household head has additional to the social long-term care insurance a private long-term care insurance.

Table A.2: Item non-response rates for the variables included in the analysis

		missing values					
		abs. number			in %¹		
name of variable	abbr. in SAVE	**2005**	**2006**	**2007**	**2005**	**2006**	**2007**
precautionary savings							
log(precautionary_savings+1)	fes3o	504	894	757	21.9%	25.7%	25.8%
income & controls							
log(net income+1)	f54o1	367	-	-	15.9%	-	-
	f54o2	316	-	-	20.8%	-	-
	f55o	-	553	342	-	15.9%	11.6%
level of employment	f22s1	12	22	16	0.5%	0.6%	0.5%
# income sources	f53m1_`k'	134	130	11	5.8%	3.7%	0.4%
high heritage probability	f88g1	30	90	127	1.3%	2.6%	4.3%
	f88g2	12	49	95	0.8%	2.1%	4.8%
	f89s	197	706	585	41.8%	79.1%	74.1%
unemployed	f23s1	60	125	69	2.6%	3.6%	2.4%
past unemployment	f26s1	28	81	6	1.2%	2.3%	0.2%
retired	f60s	0	0	0	0.0%	0.0%	0.0%
	f61s	0	9	0	0.0%	0.3%	0.0%
income uncertainty							
std(net income)/net income	f87g1	29	88	59	1.3%	2.5%	2.0%
	f87g2	18	70	64	1.2%	3.0%	3.3%
	f55o1	-	133	129	-	5.7%	6.6%
earnings fluctuations	f59s	51	121	245	2.2%	3.5%	8.3%
longevity risk & controls							
expected years left to live²	f90o1	85	101	217	3.7%	2.9%	7.4%
	f90o2	92	118	199	4.0%	3.4%	6.8%
	f91s	83	118	110	3.6%	3.4%	3.7%
	f91o1	52	93	114	12.4%	13.6%	17.7%
	f91o2	54	54	30	9.8%	8.5%	6.0%
age partner	f11o	8	6	13	0.5%	0.3%	0.7%
age	f07o	16	1	0	0.7%	0.0%	0.0%
socio demographic controls							
Eastern/ Western Germany	bula	0	0	0	0.0%	0.0%	0.0%
female	f06s	0	0	0	0.0%	0.0%	0.0%
foreigner	f08s	9	3	1	0.4%	0.1%	0.0%
marital status	f09s	2	11	23	0.1%	0.3%	0.8%
children in hh	f14o	17	85	132	0.7%	2.4%	4.5%
additional person in hh	f17s	15	59	73	0.7%	1.7%	2.5%
education	f20s1	16	8	3	0.7%	0.2%	0.1%
	f21s1	153	9	1	6.6%	0.3%	0.0%
job	f24s1	65	146	68	2.8%	4.2%	2.3%
health & controls							
state of health	fg1s1	5	23	30	0.2%	0.7%	1.0%
development of hh health	f85g3	18	73	72	0.8%	2.1%	2.5%
	f85g4	7	41	42	0.5%	1.7%	2.1%
expected income growth							
future income situation	f86g1	19	92	93	0.8%	2.6%	3.2%
develop. own economic sit.	f86g2	13	85	80	0.6%	2.4%	2.7%
impatience							
smoker	f94s	6	25	26	0.3%	0.7%	0.9%
easy going	f100p1	2	61	39	0.1%	1.8%	1.3%

		missing values					
		abs. number			in %¹		
name of variable	abbr. in SAVE	**2005**	**2006**	**2007**	**2005**	**2006**	**2007**
preferences							
risk aversion	f98bg1	20	110	125	0.9%	3.2%	4.3%
	f98bg2	340	511	502	14.8%	14.7%	17.1%
	f98bg3	39	159	164	1.7%	4.6%	5.6%
	f98bg4	66	202	217	2.9%	5.8%	7.4%
	f98bg5	160	214	222	6.9%	6.2%	7.6%
min. amount in cur. account	f50s	85	125	102	3.7%	3.6%	3.5%
saver type	f37s	24	86	69	1.0%	2.5%	2.4%
insurance							
liability insurance	f97s	22	115	138	1.0%	3.3%	4.7%
pr. occup. disability insurance	f96s	39	185	188	1.7%	5.3%	6.4%
regular support	f57bs	30	6	3	1.3%	0.2%	0.1%
obligation to contr. to social ins.	f27s1	-	-	93	-	-	3.2%
long-term care insurance	fg10s	-	-	22	-	-	0.7%
add. long-term care ins.	fg11s	-	-	65	-	-	2.5%
liqudity constraints							
liquidity index	fes1s	39	99	67	1.7%	2.8%	2.3%
	fes2s	45	131	131	2.0%	3.8%	4.5%
overdraft limit	fes5o	243	369	261	13.9%	13.3%	11.1%
other saving motives							
motive bequest	f46ag4	101	475	441	4.4%	13.7%	15.0%
motive old-age	f46g4	63	315	2684	2.7%	9.1%	91.4%
motive precautionary	f46g2	58	171	269	2.5%	4.9%	9.2%
business							
business owner	f81s	69	183	159	3.0%	5.3%	5.4%
wealth controls							
homeowner	f66s	11	42	44	0.5%	1.2%	1.5%
wealth³							
uncertainty of pension							
uncertain pension	f65s1	-	no missing values available				
pension not uncertain	f65s2	-	no missing values available				
year dummies							
d2005; d2007	year	0	0	0	0.0%	0.0%	0.0%

Note: The list includes all variables to construct the variables included in the analysis.
When the variable was already listed for the construction of a variable before, the variable is not listed again.
¹Percentages of missings as a % of those who had to answer this question.
If the superior question was not answered, this was count as an missing of the following question too.
²Observed values were replaced in 2005 and 2006 to make answers consistent.
³Since of the multitude of variables, the variables from which the wealth variable is constructed are not listed.
See Börsch-Supan et al. (2008) for missing rates of selected assets.
⁴In 2007 only 255 households were asked to answer the question about the importance of savings for the old-age provision by mistake. To fill the gaps of this question a special imputation procedure was written to make use of the panel structure of the dataset.
Source: SAVE 2005-2007, unrestricted sample.

Table A.3: Literature overview

authors (year of publication)	*dataset*	*evaluation period*	*explained variables/ optimisation problem*	*explanatory variables (risk variables)*	*controls/ parameters*	*main results*
Skinner (1988)[1]	CES	1972-1973	two different definitions of the saving rate	occupation dummies	log of net income, family size, age, age2	self-employed and salespeople have lower savings rates, although their income is more uncertain compared to other occupational groups
Caballero (1991)	overlapping-generations model		solution to an intertemporal saving problem of an agent	uninsurable labour-income uncertainty	interest rate, coefficient of relative risk aversion, discount rate	precautionary savings can account for more than 60% of the observed net wealth
Guiso & Jappelli & Terlizzese (1992)[1]	SHIW	1989	ratio of nondurable consumption to permanent earnings	self-reported measure of uncertainty (expected variance of earnings of the household head one year in the future)	age, family size, education, permanent and transitory earnings, net worth, one income recipients	precautionary savings account for 2% of households' net worth
			log of the ratio of net worth to permanent earnings		age classes, number of children, family size, area of residence	
Dynan (1993)[2]	CES	1985 4 quarters	consumption growth	quarterly consumption growth variance		the estimated coefficient of relative prudence is too small to induce precautionary savings
Hubbard & Skinner & Zeldes (1995)[3]	PSID	1984	modelling of household's consumption over the pre-retirement period	uncertainty about earnings, uncertain out-of-pocket medical expenses, and random date of death	time preference rate, interest rate, curvature of the utility function, earnings process, consumption floor	life cycle models expanded by the precautionary saving motive and social insurance programs can explain the saving behaviour of households with virtually no wealth
Starr-McCluer (1996)[1]	SCF	1989	log wealth (liquid assets, financial assets, net worth)	unevenness in health insurance coverage	permanent income, age, age2, race, gender, marital status, education, self-employment, children, persistent health problems, past and expected inheritance	mixed evidence since health insurance coverage increases significantly the savings of households
Carroll (1997)[3]	PSID	1968-85	maximising of life time utility over consumption in each period	estimated transitory and permanent labour income shocks, probability of zero earnings	interest rate, growth rate of permanent income, coefficient of relative risk aversion, discount rate	the buffer-stock model is able to solve the income/ consumption puzzle and other empirical puzzles

authors (year of publication)	*dataset*	*evaluation period*	*explained variables/ optimisation problem*	*explanatory variables (risk variables)*	*controls/ parameters*	*main results*
Kazarosian (1997)[1]	NLS	1965-1981	ratio of total net family wealth to permanent income	estimated measures of income risk specifications: 1.) total sample response to total uncertainty 2.) response to total uncertainty by occupation 3.) response to permanent and transitory uncertainty by occupation	occupation, education, and health dummies, age, age2, permanent income, bequest dummy, race dummy, number of children, married dummy	significant and positive effects of uncertainty on wealth; the effect is large and varies in size depending on occupation
Lusardi (1997)[1]	SHIW	1989	log of the ratio of wealth over permanent income	self-reported measure of uncertainty (expected variance of earnings of the household head one year in the future)	age, age2, age3, number of children, family size, log of permanent income	precautionary savings account for 2.8% of households' net worth
Lusardi (1998)[1]	HRS [start 1992]	not mentioned	total net worth	subjective measure of uncertainty based on expectations about the probability of becoming unemployed during the next year, longevity risk	age, age2, gender, race, education, health status, marital status, risk aversion, planning horizon, income expectations, bequest motive, permanent income, past unemployment, interaction with a one income earner dummy	uncertainty in earnings accounts for 1-3.5% of total net wealth and 2-4.5% of financial wealth; the longevity risk is not significant
Carroll & Samwick (1998)[1]	PSID	1981-1987	log of wealth (total net wealth, liquid assets, non-business and non-housing wealth)	estimated measures of income risk (relative equivalent precautionary premium (REPP), variance of income (VARY), and the variance of the log of income (VARLY))	(estimation was done in different subsamples according to education, occupation and industry) log of permanent income, age, age2, married, white, female, kids, occupation, education	setting the measurement of uncertainty to the smallest value predicted leads to a reduction of 32% for liquid assets, of about 45% for total net worth and of 50% for net worth excluding housing and business assets for households under 50 years of age
Carroll & Dynan & Krane (1999)[1]	SCF CPS	1983, 1989, 1992 1983, 1989, 1992	different measures of wealth	estimated unemployment risk (interacted with permanent income)	log of permanent income, number of income earners, homeowner, age, race, gender, marital status, number of children, education, age*education, occupation, occupation*age, industry, retirement status, defined benefit pension, turned down for credit	only significant effects of risk variables for total wealth (includes housing wealth); no significant effect for the lowest income group

authors (year of publication)	*dataset*	*evaluation period*	*explained variables/ optimisation problem*	*explanatory variables (risk variables)*	*controls/parameters*	*main results*
Gruber & Yelowitz (1999)[1]	SIPP CEX	1984-1993 1983-1993	log of household's total net worth or consumption	constructed measure of available current and future dollors from a insurance of medical expenses	gender, age, age2, race, education of household head and partner, marital status	Medicaid entitlement reduces wealth holdings (in 1993 by around 16.3%) and increases consumption significantly
Palumbo (1999)[4]	PSID NMCES CES	1984-1986 1977 1982-1983	modelling of household's consumption over the retirement period	uncertainty in the form of modelled future health care expenses, longevity and death risk	the coefficient of relative risk aversion is estimated in a way that the consumption observed is closest to the consumption levels predicted	importance of precautionary savings for elderly households
Engen & Gruber (2001)[1]	SIPP	1984-1990	ratio of financial assets to permanent income	expected unemployment benefit replacement rates	age, age2, gender, marital status, race, education, wage, wage2, partner's education, number of children, interactions between wage and age, state and year specific dummies	a reduction of the unemployment benefit replacement rate by 50 % raises average gross financial assets by around 14%; this effects is less strong for older workers and stronger for workers facing high unemployment risk
Arrondel (2002)[1]	INSEE	1997	logarithm of the ratio of financial and total net worth to permanent income	subjective measure of income risk	risk aversion, piece-wise linear funct. of age, education, social status, marital status, number of children, bequest motive, inheritance, urban vs. rural residence, dummy past unemployment, dummy past health problems	5% of wealth accumulation is due to the precautionary saving motive
Gourinchas & Parker (2002)[4]	CEX PSID	1980-1993	modelling of mean consumption over the life-cycle	estimated transitory and permanent labour income shocks, probability of zero earnings	given preferences and the income process, they simulated the expected consumption over the life-cycle	for young households (below 40 years), precautionary savings accounts for 60 to 70% of non-pension wealth
Cagetti (2003)[4]	SCF PSID	1989, 1992, 1995 1984, 1989, 1994	modelling of median wealth over the life-cycle	stochastic component of earnings	after the determination of the income process, the initial distribution of wealth, the degree of altruism, the discount factor, and the degree of risk aversion, they estimate the distribution of wealth over the life-cycle	at the beginning of the life-cycle, the wealth accumulation is mainly the result of the precautionary saving motive; if a households comes close to retirement, savings for retirement will become significant

authors (year of publication)	*dataset*	*evaluation period*	*explained variables/ optimisation problem*	*explanatory variables (risk variables)*	*controls/ parameters*	*main results*
Murata (2003)[1]	JPSC	1994, 1996	total net worth and financial wealth to permanent income ratio	labour earnings uncertainty constructed from a question concerning economic prospects, measure of uncertainty over public pension benefits	age, age2, education, double-income household dummy, current income, income of the previous year, number of children, age gap between husband and wife	public-pension uncertainty accounts for one third of financial wealth of young and middle-aged nuclear-family households and households without support from transfer payments; no significant effect of labour uncertainty on wealth accumulation
Fuchs-Schündeln & Schündeln (2005)[1]	GSOEP	1992-2000 1998-2000 (main analysis)	log of financial wealth, log of financial wealth plus housing wealth	civil servant dummy	log of permanent income, age, age2, gender, marital status, home ownership, year dummies, number of adults and children	22.1% of financial wealth are due to precautionary savings in the East and 12.9% in the West
Hurst & Kennickell & Lusardi (2005)[1]	PSID	1981-1987 1991-1997	log of total net worth, log of the ratio of total net worth over permanent income	computation of the variances of permanent and transitory income shocks (IV estimation)	age, age2, race, gender, marital status, education, past income and wealth shocks, year dummies, unemployment dummy	after splitting the sample in business owners and non-business owners, the accumulation of total net worth due to precautionary savings decreases from around 50% to less than 10% for non-business owners and less than 12% for business owners
Kennickell & Lusardi (2005)[1]	SCF PSID	1995, 1998 1990-1997	log of a subjective measure of the desired amount of precautionary savings	constructed measures for health risk, longevity risk, business risk, state unemployment rate	log of permanent income, age, age2, race, gender, marital status, education, household size, dummies for the age of children, health controls, wealth dummies, different variables to measure preferences and attitudes towards saving and risk, controls for liquidity constraints, future expenses, and macro shocks	precautionary savings account for 8% of total net worth and 20% of total financial wealth; the paper emphasis the importance of the precautionary saving motive for business owners and elderly households
Bartzsch (2006)[1]	GSOEP	2002 (1998-2002)	log of net financial wealth or non-business net wealth	five different constructed measures for income uncertainty; used instruments: education, occupation, industry	log of permanent income, age, age2, gender, marital status, number of adults/ children in the household, region, saving for retirement, saving for bequests, saving for children's education, willingness to take risks in financial matters	on average 20% of net financial wealth is due to precautionary savings; income risk is not significant as a regressor of housing wealth

authors (year of publication)	*dataset*	*evaluation period*	*explained variables/ optimisation problem*	*explanatory variables (risk variables)*	*controls/ parameters*	*main results*
Kong & Lee & Lee (2007)[1]	KHPS	1993, 1994 1996, 1997	change in nondurable consumption, change in household medical expenditures	constructed measure of health uncertainty using the question about the general health status	controls for multinomial logit models of the health status: age, age2, gender, marital status, education, smoking, drinking	health uncertainty causes a significant effect according to the theory of precautionary savings for a sample restricted to household heads older than 64
Piracha & Zhu (2007)[1]	GSOEP	1998-2005	three different definitions of savings	change in the nationality law (Germany in 2000)	age, gender, education, ethnic origin, house ownership, labour market status, benefit status, household size	significant reduction in savings of immigrants of around 13% after the change in the nationality law

Note: The literature overview does not claim to be a complete overview of the literature of the precautionary saving motive. It is limited to articles which have been found relevant for this study. Moreover, the articles differ in their approaches. This turns the arrangement of key contents into a difficult task if these contents should be categorised as done here. The author excuses for all the inadequacies which are a result of the arrangement of this table and the necessary simplification of certain contents.

Datasets:

CES/CEX (US) = Consumer Expenditure Survey
CPS (US) = Current Population Survey
HRS (US) = Health Retirement Survey
INSEE (France) = INSEE Survey on Wealth
JPSC (Japan) = Japanese Panel Survey of Consumers
KHPS (Korea) = Korean Household Panel Study

NLS (US) = National Longitudinal Survey
NMCES (US) = National Medical Care Expenditure Survey
PSID (US) = Panel Study of Income Dynamics
SCF (US) = Survey of Consumer Finances
SHIW (Italy) = Survey on Household Income and Wealth
SIPP (US) = Survey of Income and Program Participation

1= regression evidence
2= Euler equation method
3= simulation approach
4= structural estimation approach