WORKING
PAPER

Intertemporal Consumption with Directly Measured Welfare Functions and Subjective Expectations

ARIE KAPTEYN, KRISTIN J. KLEINJANS, ARTHUR VAN SOEST

WR-535

November 2007

This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the NICHD funded RAND Population Research Center (R24HD050906).

RAND LABOR AND POPULATION

Intertemporal Consumption with Directly Measured Welfare Functions and Subjective Expectations

Arie Kapteyn, RAND, email: kapteyn@rand.org

Kristin J. Kleinjans, University of Aarhus and RAND, email: kkleinjans@econ.au.dk*

Arthur van Soest, Tilburg University and RAND, email: avas@uvt.nl

Abstract

Euler equation estimation of intertemporal consumption models imposes heavy demands on data and identifiability conditions. For example, one typically needs panel data on consumption, assumptions on expectations, and a parameterization of preferences. We aim at reducing some of these requirements, by using additional information on respondents' preferences and expectations. The results suggest that individually measured welfare functions and expectations have predictive power for the variation in consumption across households. Furthermore, estimates of the intertemporal elasticity of substitution based on the estimated welfare functions are plausible and of a similar order of magnitude as other estimates found in the literature.

JEL Codes: D91, D84, D12

Key words: Expectations, Consumption, Euler equations

* Corresponding author. We thank Rob Alessie for making available the cleaned income data we use in the paper.

1. Introduction

Modern empirical studies of intertemporal allocation of consumption usually rely on Euler equations. For the estimation of such models, one typically needs panel data on consumption, assumptions on how respondents form their expectations, and a parameterization of preferences (see, for example, Hall 1978, Browning and Lusardi 1996, Carroll 2001, Attanasio and Low 2004). In this paper, we aim at reducing some of these estimation requirements by using subjective data on income expectations and data on the income levels respondents say they need to attain a given satisfaction level. The latter are taken as points on a contemporaneous utility function of consumption. We investigate if directly measured utility functions and expectations have explanatory power for consumption and savings behavior. An affirmative answer is useful because it implies that preference distributions are easy to estimate, in principle requiring cross section data only and avoiding the need to make arbitrary assumptions about expectations.

Thus, in contrast to the conventional approach, both utility functions and expectations are measured directly by asking questions to respondents. By combining the information thus obtained with data on consumption (computed from income and saving), we are able to test if these directly measured utility functions and expectations have explanatory power for consumption behavior. Economists have long been skeptical of the use of subjective responses in questionnaires that do not refer to objective phenomena, and to which extent such responses help to explain behavior is an open question. We investigate whether these models can be used to explain consumption and to analyze the sensitivity of saving and consumption to the interest rate (i.e., the intertemporal rate of substitution).

As a specification of preferences, we adopt the individual welfare function, a concept introduced by Van Praag (1968) and operationalized in numerous papers since, including Van Praag (1971), Van Praag and Kapteyn (1973), Kapteyn and Wansbeek (1985), Groot et al. (2004), and Van Praag and Ferrer-i-Carbonell (2004). In particular, we will use the individual welfare function of income (henceforth WFI), which in our dynamic framework is interpreted as a welfare function of consumption. In a static context, the WFI represents the satisfaction an individual attaches to a certain income (or consumption) level, measured on a continuous scale from 0 to 1. In Section 3, we will describe in some detail how the WFI is constructed from answers to a set of relatively straightforward questions.

An Euler equation relates the marginal utility of current consumption to the expected marginal consumption of the next period (which, in this paper, is next year). Thus, writing down an Euler equation for intertemporal allocation of consumption does not only require knowledge of the utility function but also of expectations. Therefore, we do not only elicit the individual utility function of consumption directly, but we also use direct information on expectations, following the approach pioneered by Dominitz and Manski (1997). We use data from the Dutch DNB household panel survey, which has the unique feature that it includes questions on both expectations of future income and incomes needed to attain given satisfaction levels, thus enabling us to directly measure both individual expectations and preferences.

Different assumptions can be made regarding the evolution of preferences over time. More precisely, in solving the intertemporal consumption problem, a consumer has to make assumptions about his or her future preferences. A "myopic" consumer may assume that tomorrow's preferences are the same as today's. A (super) rational consumer, on the other hand, may be able to predict tomorrow's preferences perfectly.

Our results suggest that the individually measured welfare functions and expectations have predictive power for the cross-section variation in consumption. We consider several models and find that a model assuming that consumers are myopic fits the data better than a model assuming forward-looking behavior with subjective expectations. Estimates of the intertemporal elasticity of substitution based on the estimated welfare functions are of a similar order of magnitude as those found in the literature.

The remainder of the paper is structured as follows. In the next section, we describe the panel data used in this paper. In Section 3, the welfare function of income (or consumption) and its measurement are described, and in Section 4, we explain how we measure expectations. In Section 5, the Euler equations are derived under the assumptions that intratemporal utility can be described by a lognormal welfare function and the subjective distribution of future consumption follows a lognormal distribution. The empirical strategy is explained in Section 6. We discuss the empirical results in Section 7, where we also compare implied intertemporal elasticities of substitution to existing findings in the literature. Section 8 concludes.

2. Data

The DNB Household Survey (DHS), formerly known as the CentER Savings Survey, is a Dutch panel survey that started in 1993. The survey is conducted by CentERdata and administered over the Internet. If a potential participant has no Internet access, he or she is provided with access through a so-called set top box (also called Web TV or Internet Player) that connects to the Internet via the telephone and a television set. The survey consists of six modules and asks a variety of questions about demographics, health, income, assets, and economic and psychological concepts. For our research, data up to and including the year 2005 were available.

Since the income expectation questions were not asked in 1993 and 1994, we only use data for the years 1995-2005, giving a total of 12,162 observations for 5,214 respondents. See Appendix A for the sample selection (Table A.1) and the number of observations by survey year (Table A.2). We use the unbalanced sample so the number of observations changes across waves, due to both attrition and refreshment.

The consumption measure used in this paper was constructed as the difference between (self-reported) income and savings. For savings, we used the answers to two questions. If the first question, "*Did you put any money aside in the past 12 months?*" was answered affirmatively, the respondent was asked a second question: "*About how much money has your household put aside in the past 12 months?*" The respondent was asked to choose one of seven different brackets, which differed across some of the waves. From these answers and the self-reported (net) income, we constructed lower and upper limits on consumption. Note that, depending on the answers, one of the two limits might be missing. For instance, if a respondent reports that no money was put aside during the last twelve months, this means that savings can have been zero or negative and hence consumption must have been at least equal to income.

Table 1 gives some summary statistics for the first year of data in our sample.

Table 1: Summary Statistics for 1995

Variable	N	Mean
Age	1642	47.9 (12.9)
Female	1642	0.38
Education	1607	
Primary or less		0.03
Lower level		0.24
Intermediate vocational		0.13
Intermediate general		0.13
Higher vocational		0.29
University		0.16
Other		0.02
Income [1]	1642	31,627 (17,129)
Savings		
Lower limit	1301	3,116 (4,948)
Upper limit	1611	6,124 (7,167)
Consumption		
Lower limit	1611	25,471 (16,211)
Upper limit	1301	29,450 (15,572)

[1] *Income is measured after tax. The sample we work with includes observations with at least one non-missing consumption bracket. All income measures are annual and expressed in 2000 Euros. Standard deviations are given in parentheses (except for dummy variables).*

3. Welfare Functions of Income and Their Measurement

A WFI is measured by asking respondents in a survey a so-called Income Evaluation Question (IEQ). The formulation of the IEQ varies somewhat across surveys. The IEQ used in the survey on which the current paper is based reads as follows:

The next question again concerns the net income of the household, that is, the net income of all household members taken together. Consider the current situation of your household when answering this question.

Which NET income of the household would you, IN YOUR SITUATION, find very bad, bad, insufficient, sufficient, good, very good? Please provide annual incomes.

VERY BAD if the annual income would be about: €...........................

BAD if the annual income would be about: €...........................

INSUFFICIENT if the annual income would be about: €............................
SUFFICIENT if the annual income would be about: €............................
GOOD if the annual income would be about: €............................
VERY GOOD if the annual income would be about: €............................

Van Praag's theory assumes an isomorphism between utility theory and probability theory. Utility is assumed to be measurable on a [0,1]-scale. In order to use the answers to the above IEQ to estimate a utility function (or as Van Praag calls it, a welfare function), one needs to assign numerical values between zero and one to the verbal labels "very bad", "bad", "insufficient", "sufficient", "good", and "very good". Based on an information maximization argument, Van Praag (1971) proposes to assign numerical values such that each label represents an equal part of the [0,1]-interval (also see Kapteyn and Wansbeek, 1985, for a detailed explanation). In the formulation used here, this means that "very bad" is assigned the value 1/12, "bad" 3/12, "insufficient" 5/12, "sufficient" 7/12, "good" 9/12, and "very good" 11/12.

The IEQ asks for income levels providing a certain welfare level, and the underlying theory in, e.g., Van Praag (1971) is static. In the standard life-cycle model, utility in a given period depends on consumption in that period. In the intertemporal context of saving and consumption decisions, it therefore seems natural to interpret the (static) welfare function of income as a welfare function of within period consumption. The interpretation of the IEQ is then that respondents' answers reflect the consumption levels that would yield a certain amount of within period utility, assuming that the question refers to a situation without (positive or negative) savings, so that the reported income amounts are actually amounts of total consumption expenditure. From now on, we will interpret the WFI in that way, i.e. as representing the utility of consumption in period *t*. Denote this consumption in period *t* by x_t. Using the isomorphism with probability theory and invoking a Central Limit Theorem, Van Praag shows that under certain conditions the WFI can be approximated by a lognormal distribution function: [1]

---

[1] Throughout, the index representing the individual is suppressed.

$$U\left(x_t,\mu_t,\sigma_t\right)=\int_o^{x_t}\frac{1}{z}\frac{1}{\sigma_t\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\ln z-\mu_t}{\sigma_t}\right)^2\right]dz\ .$$

The parameters μ_t and σ_t are individual preference parameters, which can be estimated from the IEQ above in a straightforward way: Let Φ be the standard normal distribution function, and denote the answers to the IEQ at time *t* by z_{it}, $i=1,...,6$. Furthermore, denote the numerical labels corresponding to the verbal labels "very bad", "bad", etc. by w_i, $i=1,...6$; that is, $w_1=\frac{11}{12}$, $w_2=\frac{9}{12}$, etc. Then we have:

$$w_i=\Phi(\frac{\ln z_{it}-\mu_t}{\sigma_t}),\ i=1,...,6\ .$$

Rewriting and allowing for measurement error ε_{it} in the responses then gives:

$$\ln z_{it}=\mu_t+\sigma_t\Phi^{-1}(w_i)+\varepsilon_{it}\ ,\ \ i=1,...,6\ .$$

Under the assumption that the measurement errors satisfy the classical linear model assumptions, OLS for each individual and each time period (on six observations) gives unbiased estimates for all (individual and year specific) parameters μ_t and σ_t. Note that we estimate parameters μ_t and σ_t for each respondent and for each time period separately.

Table 2 shows the means and standard deviations of the required incomes at each verbal label of the income evaluation question, as well as of the estimated utility function parameters for 1995. To make a comparison with actual income levels possible, the table also includes statistics for actual net household income. The average actual income level is close to what on average is considered a "good" income, in line with the notion that, on average, Dutch people are rather satisfied with their actual income. Table A.3 in the appendix shows how median actual and required incomes (in 2000 prices) evolve over time. Required incomes follow the same pattern as actual incomes – falling in the first few years and then stabilizing. Of course, this may also reflect changes in sample composition, and not necessarily imply similar changes for average individual households. The pattern of the average value of μ_t is similar since this is just a summary measure of the six required income levels. On the other hand, the median value of σ_t first increases towards the middle of the sample period and then falls, showing that the dispersion in income levels needed to attain low and high utility levels first rise and then fall.

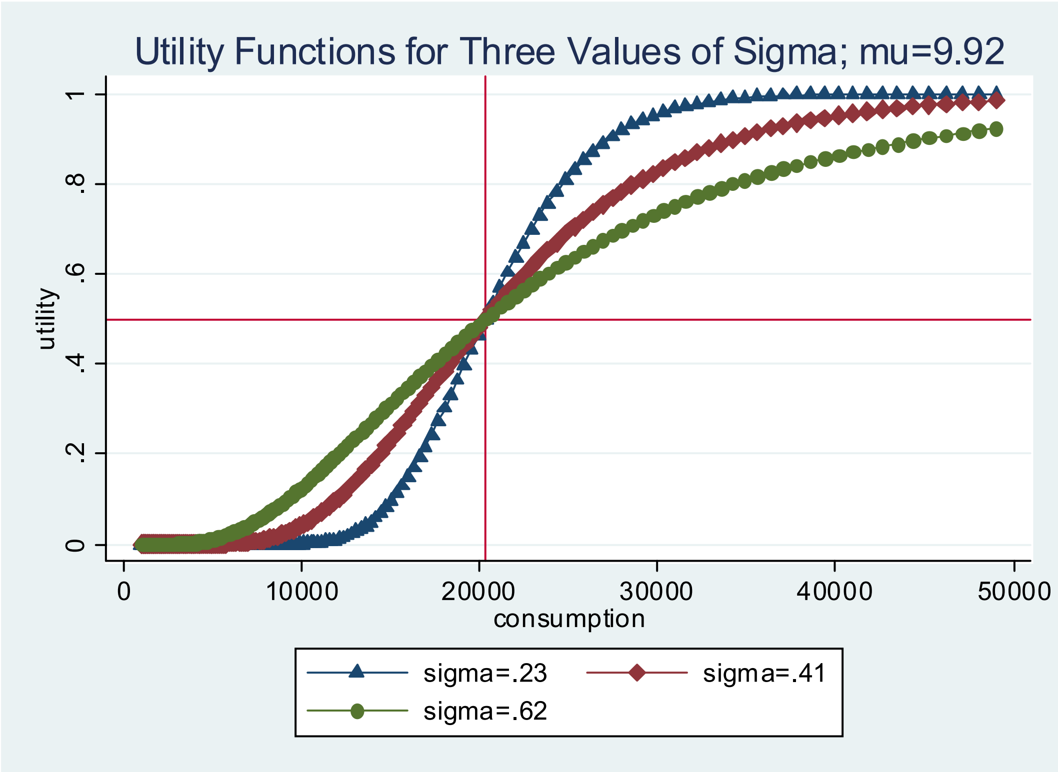
Table 2: Income Evaluation and Estimated Utility Function Parameters: 1995

Variable	N	Mean
Income	1642	31627 (17129)
Log income	1642	10.25 (0.49)
Income Evaluation		
Very bad	1637	13261 (7408)
Bad	1639	16677 (11379)
Insufficient	1638	19843 (13452)
Sufficient	1641	25118 (16665)
Good	1641	30876 (20480)
Very good	1639	40270 (27994)
Utility Function Parameters		
μ_t	1642	9.92 (0.41)
σ_t	1642	0.41 (0.16)

Working sample. All income measures are annual and expressed in 2000 Euros. Standard deviations in parentheses.

Figure 1 shows some welfare functions based on the estimates in Table 2 for individuals with mean $\mu = 9.92$. In the current context, this means that a consumption level of $\exp(9.92) =$ €20,333 is evaluated at satisfaction level 0.5, between "insufficient" and "sufficient." Clearly, a consumer with a larger value of μ needs a higher consumption level to reach a given satisfaction level. The parameter σ determines how steep or flat a welfare function is. The smaller σ, the steeper is the welfare function. We have drawn three welfare functions, all with $\mu = 9.92$, but with different values of σ, namely 0.23 (the 10th percentile in 1995), 0.41 (the sample mean), and 0.62 (the 90th percentile in 1995).

Figure 1: Some Estimated Utility functions, 1995



4. Measurement of Expectations

Most surveys soliciting subjective expectations about future outcomes ask for point estimates. Since future outcomes are intrinsically uncertain, a single point estimate provides incomplete information – it says nothing about the dispersion of the respondent's subjective distribution of the future outcome. Moreover, it is not clear which point estimate respondents give in answer to any such question; this could be, for example, the mode, the median, or the mean (cf., e.g., Manski, 2004).

This problem can be overcome if information is solicited about the subjective probability distribution of the future outcome considered. Dominitz and Manski developed an approach in which respondents first give upper- and lower bounds on their future outcomes, and are then asked for the probabilities that outcomes lie in specific intervals which are subsets of the interval between the upper and lower bound. They applied this methodology to income expectations in the Survey of Economic Expectations (SEE); see, for example, Dominitz and Manski (1997). Das and Donkers (1999) applied this method to income expectations of Dutch households, using the same data source as we use, but the earlier waves only.

Specifically, respondents in the DHS were asked the following questions:

> "*We would like to know a bit more about what you expect will happen to the net income of your household in the next 12 months.*
>
> *What do you expect to be the LOWEST total net income your household may realize in the next 12 months? Please use digits only, no dots or commas*."
>
> "*What do you expect to be the HIGHEST total net income your household may realize in the next 12 months?*"

The wording of the questions was the same in most waves, except that in 1995 and 1996 monthly amounts were asked. We converted these into annual amounts.

Follow up questions were asked about subjective probabilities that income lies within certain intervals. Many respondents did not answer some of these probability questions, even when they did provide the lowest and highest income asked for in the question. We will therefore use two approaches in this study. The first (which we will call: DM, after Dominitz-Manski) only uses the observations with the responses to at least two of the probability questions. In the second approach (to be called LH, for "Low-High") we use a larger data set and only exploit the maximum and minimum incomes provided. The subjective income distributions are imputed by linearly interpolating probabilities between the maximum and minimum, where the probability that income is below the lowest income is set to zero and the probability that the income is below the maximum is set equal to one.

For both procedures we take the observed or imputed probabilities as approximations to a lognormal distribution function. Thus, for each respondent, the lognormal distribution that gives the best fit to the observed or imputed probabilities is determined using non-linear least squares on the six points. For each respondent, this gives estimates of the log median and the log standard deviation of their subjective distribution (the two parameters of the lognormal distribution).

Table 3 shows the summary statistics for actual income, the range of possible future incomes (the lowest and highest possible amounts), and the estimated respondent-specific parameters of the lognormal distribution for 1995 for both approaches.

Table 3: Summary Statistics for subjective income distributions: 1995

Approach	DM	LH
Variable	**Mean**	**Mean**
Income	31,827 (17,204)	31,627 (17,129)
Lowest expected net income	28,691 (14,420)	28,166 (15,029)
Highest expected net income	33,710 (16,648)	33,563 (19,787)
Log-median of subjective distribution in next year	10.23 (0.53)	10.20 (0.53)
Log-standard deviation of subjective distribution in next year	0.04 (0.05)	0.06 (0.07)
Median of subjective distribution in next year	31,432 (15,384)	30,796 (16,446)
Standard deviation of subjective distribution in next year	1.04 (0.06)	1.06 (0.09)
Number of observations	1167	1642

All measures are annual and expressed in 2000 Euros. Standard deviations in parentheses.

The mean of the lowest possible future net incomes is significantly lower than the mean actual income. Similarly, the mean of the highest possible future net incomes is significantly higher than the mean actual income. The average range between highest and lowest possible amount is about €5,400 for the LH dataset and €5,900 for the DM sub-sample. There is substantial variation here, however. For example, for 12.7% of the LH sample and 20.7% of the DM sample, minimum and maximum amount are identical, indicating no subjective uncertainty in future income. In general, the log standard deviations of the subjective income distributions are quite low, with a median of 0.036 for the LH sample and 0.040 for the DM sample.

Table A.3 in the appendix shows how median and dispersion vary over time. The median of the subjective income distribution first falls over time and then remains rather stable. On the other hand, the median standard deviation of the subjective log income distributions increases after 1998, suggesting that respondents' uncertainty increases.

Figures 2 and 3 show the distributions of the respondents' log medians and standard deviations of log incomes of their future income distributions for the LH data in 1995. There is substantial variation across respondents in both graphs. The standard deviations of log incomes are concentrated at rather low values, but there is a substantial right tail with respondents who think that their future income is quite uncertain. The graphs for the DM data look similar (not presented).

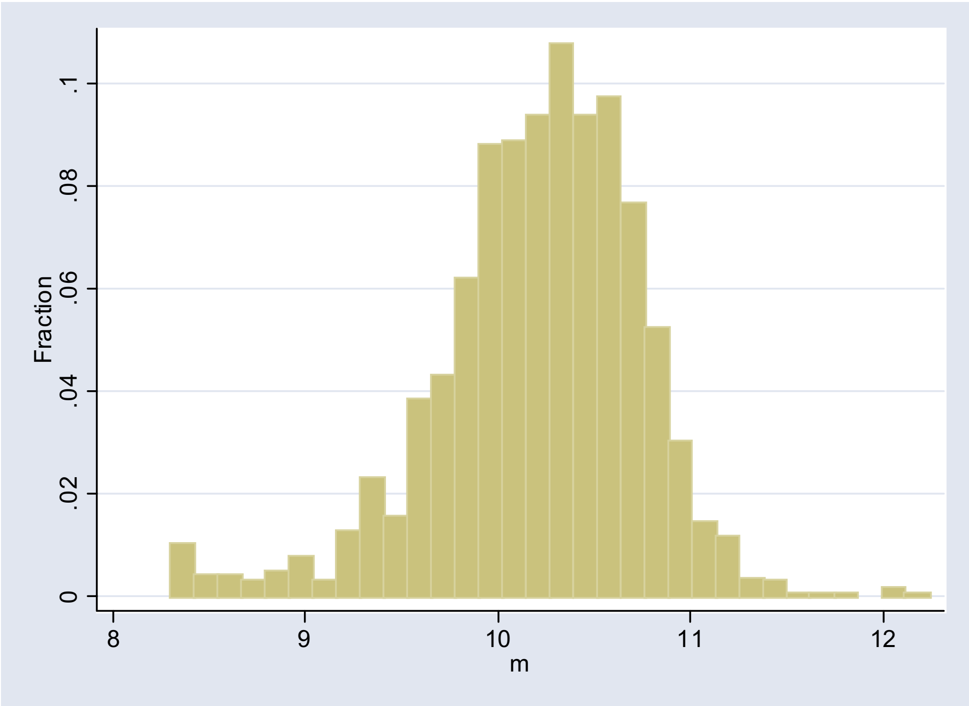


Figure 2. Log medians of subjective income distributions in LH data: 1995 (outliers excluded)

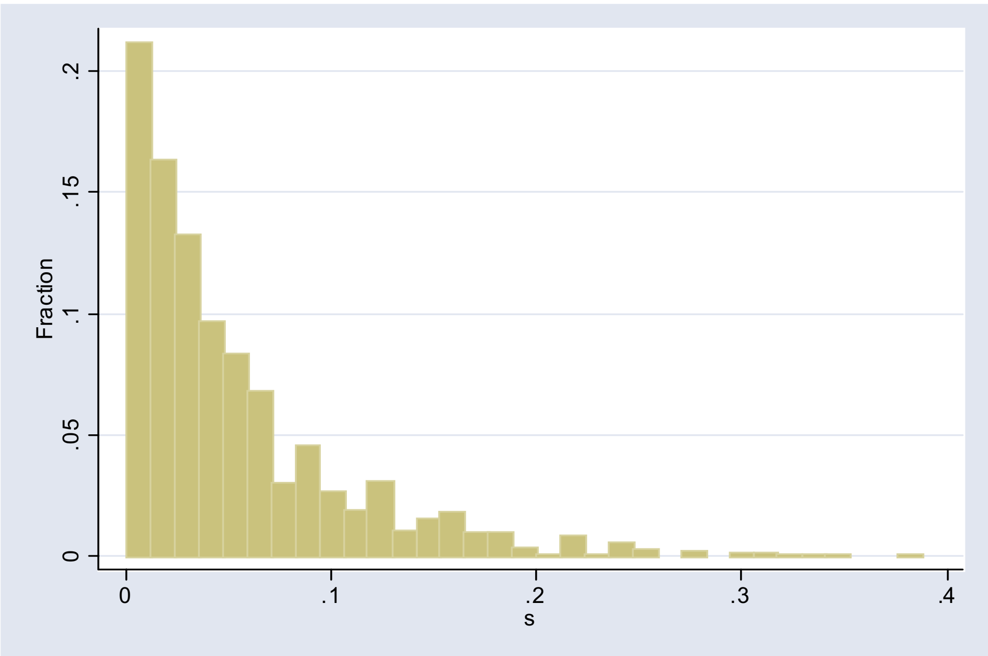


Figure 3. Standard deviations of log incomes distributions LH data: 1995 (outliers excluded)

5. Models of intertemporal consumption

Consider the following standard intertemporal utility maximization problem of a consumer at time t:

$$\text{Max } E_t \sum_{\tau=t}^{T} \left(\frac{1}{1+\delta} \right)^{\tau-t} U\left(x_\tau, \mu_\tau, \sigma_\tau\right) \text{ subject to } \sum_{\tau=t}^{T} \left(\frac{1}{1+r} \right)^{\tau-t} x_\tau = M \, , \tag{1}$$

where:

E_t is the expectation operator conditional on all information available at time t,

r is the real interest rate,

δ is the time preference rate,

M are total lifetime resources,

x_τ is consumption in period τ, and

T is the time horizon.

$U\left(x_\tau, \mu_\tau, \sigma_\tau\right)$ is the utility of consumption x_τ; in line with the discussion in Section 2, this function is assumed to be the distribution function of a lognormal distribution with parameters μ_τ and σ_τ:

$$U\left(x_t, \mu_t, \sigma_t\right) = \int_o^{x_t} \frac{1}{z} \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{\ln z - \mu_t}{\sigma_t} \right)^2 \right] dz \, , \tag{2}$$

The first order conditions of consumption smoothing are:

$$U'\left(x_t, \mu_t, \sigma_t\right) = \frac{1+r}{1+\delta} E_t U'\left(x_{t+1}, \mu_{t+1}, \sigma_{t+1}\right) \; . \tag{3}$$

To be able to characterize the solution to the first order conditions, we make assumptions about the distribution of the random variables that determine the expectation on the right hand side of (3). We will assume that this leads to a distribution of x_{t+1} that is lognormal with parameters m_{t+1} and s_{t+1}^2. Under these assumptions, the following expression for consumption x_t can be derived (see Appendix B for the derivation):

$$\ln x_t = \mu_t - \sigma_t^2 + \sigma_t \sqrt{\left(\sigma_t^2 - 2\mu_t\right) - \left(\sigma_{t+1}^2 - 2\mu_{t+1}\right) + \ln \frac{s_{t+1}^2 + \sigma_{t+1}^2}{\kappa^2 \sigma_t^2} + \frac{\left[m_{t+1} - (\mu_{t+1} - \sigma_{t+1}^2)\right]^2}{s_{t+1}^2 + \sigma_{t+1}^2}} \,. \tag{4}$$

Here we have defined $\kappa \equiv \frac{1+r}{1+\delta}$.

In order to use (4) as the basis for the empirical work, we need to replace the individual specific parameters by estimates. The parameters of the utility function in period *t* will always be estimated using μ_t and σ_t obtained in Section 3. For the future utility function in period *t+1*, we will use the estimates μ_{t+1} and σ_{t+1} in two out of the three models we consider (Models 1 and 3, defined below), assuming that the consumer realizes that preferences change over time and accounts for this when making consumption decisions. In one model (Model 2), we will take $\mu_t = \mu_{t+1} = \mu$ and $\sigma_t = \sigma_{t+1} = \sigma$. That is, we assume that the consumer makes the consumption decision in period *t* under the incorrect assumption that preferences in periods *t* and *t+1* are the same. Thus, the consumer plans as if preferences do not change; this is what we called "myopic" behavior in Section 1. In this model, (4) simplifies to

$$\ln x_t = \mu - \sigma^2 + \sigma \sqrt{\ln\left(\frac{s_{t+1}^2 + \sigma^2}{\kappa^2 \sigma^2}\right) + \frac{\left[m_{t+1} - (\mu - \sigma^2)\right]^2}{s_{t+1}^2 + \sigma^2}} \,. \tag{5}$$

The other choice we have to make is what to do with the parameters of the respondent's subjective future consumption distribution used in making the consumption decision, m_{t+1} and s_{t+1}. In Models 1 and 2, we will replace these by their estimates discussed in Section 4. Thus we interpret the subjective future income distribution as a distribution of future consumption. Clearly, this cannot be exactly right, since after all the whole idea of intertemporal consumption smoothing is to break the contemporaneous link between income and consumption. However, we would expect the future income distribution and future consumption distribution to be related, so that the future income distribution can be seen as a proxy for the future consumption distribution.

An alternative is based upon the theory of preference formation introduced by Kapteyn (1977). This theory states that the respondent's utility of consumption is purely driven by the rank in the consumption distribution in that period. Under this assumption, direct information on future consumption distributions is not needed, and the future consumption distribution is given by the future utility function: $m_{t+1} = \mu_{t+1}$ and $s_{t+1} = \sigma_{t+1}$. This is the assumption we will use in Model 3. In this case, we obtain:

$$\ln x_t = \mu_t - \sigma_t^2 + \sigma_t \sqrt{\left(\sigma_t^2 - 2\mu_t\right) - \left(\sigma_{t+1}^2 - 2\mu_{t+1}\right) + \ln 2\frac{\sigma_{t+1}^2}{\kappa^2\sigma_t^2} + \frac{1}{2}\sigma_{t+1}^2} \; . \tag{6}$$

In the empirical analysis, we will estimate all three non-nested models and test them against each other.

6. Empirical Strategy

Each of the empirical approaches sketched below will be repeated for the two data sets mentioned above, i.e. the dataset that only uses the lowest and highest incomes in the Dominitz-Manski questions (LH) and the subset that also uses information on subjective probabilities of incomes in between the lowest and highest reported incomes (DM). We estimated the models (4) - (6) for a range of values for $\kappa \equiv \frac{1+r}{1+\delta}$.

Since saving is measured in brackets (see Section 2), we use interval regression to estimate the models. Furthermore, as we are using a panel, we account for the fact that individuals may be in the data more than once by using a random effects specification. As described above, we will denote the three different specifications as Models 1 through 3. To be precise:

- The model using (4) with direct estimates of m_{t+1} and s_{t+1} will be called Model 1;
- (5) with $\mu_t = \mu_{t+1} = \mu$ and $\sigma_t = \sigma_{t+1} = \sigma$ will be called Model 2;
- (6) with $m_{t+1} = \mu_{t+1}$ and $s_{t+1} = \sigma_{t+1}$ will be called Model 3.

An unusual feature of these models is that they contain only one unknown parameter: κ. There are different ways to investigate the plausibility of the different specifications. The first approach is to write the models in the form:

$$\ln x_t = \beta_1\mu_t + \beta_2\sigma_t^2 + \beta_3\sigma_t\sqrt{f_t(\kappa)} + u_t, \tag{7}$$

where u_t is an i.i.d. error term, β_1, β_2 and β_3 are parameters, and $f_t(\kappa)$ is the term under the square root in (4), (5) or (6). If one of the models specified here is correct, we would expect: $\beta_1 = 1, \beta_2 = -1$ and $\beta_3 = 1$. Thus, we estimate each of the three specifications and consider the estimates of β_1, β_2 and β_3. Conditional on κ, model (7) is linear. Thus, we perform a grid search over values of κ and estimate (7) conditional on each value of κ. We pick the value of κ that gives the best likelihood value.

In a second approach, we compare the performance of the three different specifications by conducting several non-nested tests. The basic idea of the tests can be summarized by the following simple model:

$$\ln x_t = \alpha m_1(z_t) + (1-\alpha) m_2(z_t) + \varepsilon_t \ , \tag{8}$$

where $m_1(z_t)$ and $m_2(z_t)$ are alternative models explaining $\ln x_t$, and z_t are the explanatory variables. If $\alpha = 1$, $m_1(z_t)$ is the correct model, whereas if $\alpha = 0$, $m_2(z_t)$ is correct. We can rewrite (8) as

$$\ln x_t - m_2(z_t) = \alpha[m_1(z_t) - m_2(z_t)] + \varepsilon_t \ . \tag{9}$$

Applying this to (7), and using superscripts to denote Models 1 and 2, we obtain:

$$\begin{aligned} &\ln x_t - \beta_1^2 \mu_t - \beta_2^2 \sigma_t^2 - \beta_3^2 \sigma_t \sqrt{f_t^2(\kappa^2)} = \\ &\alpha\left([\beta_1^1 - \beta_1^2]\mu_t + [\beta_2^1 - \beta_2^2]\sigma_t^2 + \sigma_t[\beta_3^1 \sqrt{f_t^1(\kappa^1)} - \beta_3^2 \sqrt{f_t^2(\kappa^2)}]\right) + \varepsilon_t \end{aligned} \ , \tag{10}$$

where $f_t^i(\kappa^i), i = 1,2,$ denotes the arguments of the square root in the various models. If we assume that the β s are equal to their theoretical values, then we obtain the following simplification:

$$\ln x_t - \mu_t + \sigma_t^2 - \sigma_t \sqrt{f_t^2(\kappa^2)} = \alpha\left(\sigma_t[\sqrt{f_t^1(\kappa^1)} - \sqrt{f_t^2(\kappa^2)}]\right) + \varepsilon_t \ . \tag{11}$$

We will perform the tests both with the β s set equal to their theoretical values and with estimated β s.

7. Empirical Results

Table 4 provides the results for the first test, showing the estimates of the β s for the best fitting $\kappa \equiv \frac{1+r}{1+\delta}$, which maximizes the log-likelihood and was found by means of a grid search. These estimates of κ are smaller than one for all three models (κ=0.55 and 0.83 for Model 1; κ=0.68 and 0.66 for Model 2; and κ=0.60 and 0.82 for Model 3), but with fairly wide confidence intervals.[2] In view of the definition of κ, its estimates for Models 1, 2 and 3 imply high subjective discount rates. This is a fairly typical finding in the literature.[3]

The estimates for the coefficient β_1 on μ_t have the signs predicted by theory and for all three models (and both datasets), the coefficient of μ_t is very close to the theoretical value of 1. The coefficient of σ_t^2, which should be equal to *–1* according to the theory, is positive for Model 1, and always negative for Models 2 and 3. Similarly, the coefficients for the square root terms, which should be equal to 1, tend to deviate from their theoretical values towards zero, but are closest to 1 for Models 2 and 3. In summary, Models 2 and 3 produce estimates that are closer to the theoretical predictions than Model 1. Model 3 performs better than Model 2 concerning the coefficient on $\sigma\sqrt{\cdot}$. Comparing across the two datasets, the differences are generally fairly small, with model 3 showing the largest variation across the two datasets. This is somewhat surprising, since Model 3 does not use the subjective income expectations at all, so that the only difference between the data sets here is the sample selection. The large differences suggest that dropping those who do not give complete answers to the Dominitz-Manski questions leads to a selection bias. If we

---

[2] As noted, we have estimated κ by means of a grid search. The standard errors presented in the table are therefore conditional on the value of κ listed at the top of each column. A confidence interval for κ can be obtained in a straightforward way, by recognizing that a likelihood ratio test can be used to test any value of κ. If we apply this approach to model 2 for instance, we find that for the LH sample, $\kappa = .885$ would still be in a 95% two-sided confidence interval. For the DM sample $\kappa = .92$ would still be in the 95% confidence interval.

[3] See Frederick et al. (2002) for an overview.

focus on the results using the LH data, model 3 also outperforms model 2 for the coefficient on σ_t^2.[4]

Table 4: Random Effect Interval Regressions

	Model 1 (4)		Model 2 (5)		Model 3 (6)	
	LH	DM	LH	DM	LH	DM
κ	0.55	0.83	0.68	0.66	0.60	0.82
μ	1.013** (0.001)	1.014** (0.002)	1.002** (0.001)	1.004** (0.001)	0.993** (0.002)	1.001** (0.002)
σ^2	0.088* (0.041)	0.217** (0.053)	-0.435** (0.036)	-0.368** (0.048)	-0.511** (0.055)	-0.200** (0.069)
$\sigma\sqrt{\cdot}$	0.131** (0.020)	0.108** (0.024)	0.470** (0.021)	0.417** (0.026)	0.756** (0.043)	0.631** (0.059)
Observations	7,281	4,097	12,162	7131	7,158	3,993
Number of id	3,147	2380	5,214	3926	3,127	2,334
Log Likelihood	-16126.8	-9081.1	-26369.0	-15481.6	-15687.7	-8778.71

*Standard errors in parentheses. * significant at 5%; ** significant at 1%.*

Table 5 provides the results for two variants of the second test using the best fitting values of κ. The first variant uses (10). That is, the three separate models are estimated first and produce estimates of the β s, after which α is estimated. The second variant uses (11). That is, it estimates α while restricting the β s to their theoretical values.

In the columns headed "Model 1", Model 1 represents $m_2(z_t)$ in (8), whereas $m_1(z_t)$ is represented by the other two models. For the comparison of models 1 and 2, we find that the estimated values of α are closer to 1 than to 0 in both variants of the test and for both datasets. As a result, we reject Model 1 against Model 2. The comparison of Models 1 and 3 yields similar results, though the estimates of α are not as close to 1 as for the model 1 versus Model 2 test. In any case, all results suggest that the empirical performance of Model 1 is inferior to that of the other models.

[4] An alternative to using LH or DM is to use the DM approach where possible and the LH approach for all other observations. This gives similar results for LH (identical in case of model 3) but parameter estimates that are somewhat farther from their theoretical values, particularly for model 2.

Table 5: Non-nested Tests of Models 1, 2 and 3 against each other

Test of		Model 1 (4)				Model 2 (5)	
	Against	Model 2 (5)		Model 3 (6)		Model 3 (6)	
		LH	DM	LH	DM	LH	DM
(10)	α	1.198* (0.065)	1.106* (0.092)	0.890* (0.059)	0.812* (0.095)	0.433* (0.048)	0.408* (0.069)
(11)	α	1.051* (0.023)	0.914* (0.030)	0.904* (0.023)	0.759* (0.028)	0.671* (0.025)	0.778* (0.020)

*α =1 means the alternative model (2nd row) is accepted; α =0 means the benchmark model (1st row) is accepted. Standard errors in parentheses. * significant at 1%.*

In the columns headed "Model 2", Model 2 plays the role of $m_2(z_t)$ in (8). Here, the results are different for the two variants of the test. For the variant not imposing the theoretical values of the β s, the comparison of Model 2 against Model 3 yields an estimate of α below 0.5 for both datasets, suggesting a slight preference for Model 2.[5] For the test restricting the parameter values to be equal to the ones predicted by the theoretical model, the estimates of α suggest a slight preference for Model 3 in both datasets.

Thus, as so often in non-nested tests the results are not entirely clear-cut. The empirical performance of Model 1 is clearly inferior to that of either Model 2 or Model 3, but a choice between Models 2 and 3 cannot be made. Table 4 leads to the conclusion that Model 3 is better, Table 5 seems to imply that Model 2 is (somewhat) better.

The results raise the question why the parameter estimates deviate from their theoretical values. The obvious explanation is measurement error. It is known that the measurement of σ is not very precise (Kapteyn, 1977), and hence, both the terms σ^2 on the right hand side of equations (5) and (6) and the square roots, which contain σ,

[5] Strictly speaking, only values of α equal to one or zero are conclusive evidence in favor of one model or the other; a value of α in between zero and one would suggest a mixture of two models outperforms both models.

suffer from measurement error, which will tend to bias their associated coefficients downwards.

Finally, we will compare the implications of our estimates with results found in the literature. An important parameter of interest in intertemporal models of consumption smoothing is the intertemporal elasticity of substitution (IES); see, e.g., Hall (1988) or Kapteyn and Teppa (2003). The IES for the lognormal utility function is equal to:

$$IES = \frac{\sigma_t^2}{\ln x_t - (\mu_t - \sigma_t^2)} \ .$$

We calculate the IES based on Models 2 and 3, using predicted consumption and adding two random draws from a normal distribution to account for the random effect and the estimation error term. Table 6 presents the distribution of the individual estimates of the IES (i.e., the median and the first and third quartiles of the distribution) by education level and income quartile for Model 2. We find that the median IES is 0.209 for the LH dataset and 0.216 for dataset DM. The median is increasing in education and income. Both datasets give quite similar results. The results for Model 3 are very similar (not shown), giving, for example, median IES estimates of 0.210 (LH data) and 0.225 (DM data), and the same increasing patterns with education and income.

The values of our IES estimates are similar to what other authors using microeconomic data have found. For example, Hall (1988), using several different data sets from the US and different estimators, estimated values of the IES for aggregate consumption close to zero. Barsky et al. (1997) report average lower and upper bounds on the absolute value of the IES of 0.007 and 0.36. Yogo (2004) estimates the IES for eleven developed countries using different specifications and finds that in all cases the IES is closer to zero than to one. His estimates of the IES for the Netherlands lie between -0.25 and 0.24.[6] Kapteyn and Teppa, using an approach similar to that of Barsky et al. (1997), find an IES of about 0.5.

[6] Research in macroeconomics typically finds numbers close to zero; see Guvenen (2006) for an attempt to reconcile the different results.

Table 6: Distribution IES by Education and Income: Model 2

	P25		P50		P75	
	LH	DM	LH	DM	LH	DM
All	0.022	0.033	0.209	0.216	0.492	0.490
Education						
Primary or less	- 0.018	0.017	0.183	0.211	0.511	0.441
Lower level	-0.038	-0.056	0.186	0.182	0.454	0.450
Interm. vocational	0.013	-0.029	0.198	0.194	0.512	0.486
Interm. general	0.024	0.028	0.217	0.218	0.498	0.492
Higher vocational	0.035	0.075	0.225	0.231	0.477	0.493
University	0.049	0.062	0.234	0.237	0.520	0.512
Other	- 0.069	0.007	0.192	0.224	0.487	0.414
Missing	0.038	0.066	0.234	0.247	0.522	0.533
Income						
1. Quartile	- 0.049	-0.043	0.178	0.193	0.464	0.442
2. Quartile	- 0.038	-0.017	0.193	0.196	0.474	0.475
3. Quartile	0.026	0.038	0.219	0.220	0.503	0.491
4. Quartile	0.067	0.072	0.246	0.250	0.522	0.534

LH: Income quartiles derived by wave. N=12,162. There are 377 missing observations in education, after imputing missing values from education in other waves. Kappa=0.68.

DM: Income quartiles derived by wave. N=7,131. There are 342 missing values in education. Kappa=0.66.

8. Concluding Remarks

Empirical Models of intertemporal allocation of consumption usually rely heavily on Euler equations. Estimating such models has large data and identification requirements. In this paper, we aim at reducing some of these estimation and identification requirements by exploiting subjective data. Specifically, we investigate if directly measured utility functions and expectations help to explain behavior. It appears that a rather straightforward application of the estimation of Euler equations to individually measured welfare functions and expectations generates results that are in line with the theory. Parameter estimates have the correct sign and are close to their theoretical values when the corresponding variables are estimated with reasonable accuracy. Furthermore, we find that estimates of the Intertemporal Elasticity of Substitution are of similar order of magnitude as found in the literature using very different approaches.

It seems clear that various improvements in the empirical implementation are possible. In particular, the measurement of expectations has been rather crude. Our empirical work was, furthermore, hampered by the fact that saving was measured in brackets, rather than continuously. Continuous measurement would have made the analysis simpler and more powerful.

Having said that, the current results appear sufficiently promising to further pursue this line of research. Both the measurement of expectations and of welfare functions is in principle quite straightforward and should therefore simplify empirical work that tries to improve our understanding of intertemporal decision-making in consumption.

References

Attanasio, Orazio and Hamish Low (2004). Estimating Euler Equations. *Review of Economic Dynamics* 7(2), 406-435.

Barsky, Robert B., Miles S. Kimball, F. Thomas Juster, and Matthew D. Shapiro (1997), Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study, *The Quarterly Journal of Economics*, 112(2), 537-579.

Browning, Martin and Annamaria Lusardi (1996). Household Saving: Micro Theories and Micro Facts. *Journal of Economic Literature* XXXIV (December), 1797-1855.

Carroll, Christopher (2001). Death to the Log-Linearized Consumption Euler equation! (And Very Poor Health to the Second-Order Approximation). *Advances in Macroeconomics* 1(1), Article 6.

Das, Marcel, and Bas Donkers (1999). How Certain Are Dutch Households About Future Income? An Empirical Analysis. *The Review of Income and Wealth* 45(3), 325-338.

Dominitz, Jeff and Charles F. Manski (1997). Using Expectations Data to Study Subjective Income Expectations. *Journal of the American Statistical Association* 92(439), 855-867.

Frederick, Shane, George Loewenstein, and Ted O'Donoghue (2002). Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature* XL (June), 351-401.

Groot, Wim, Henriette Maassen van den Brink, Erik Plug (2004). Money for Health: The Equivalent Variation of Cardiovascular Diseases. *Health Economics* 13(9) 859-872.

Guvenen, Fatih (2006). Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective. *Journal of Monetary Economics* 53(7), 1451-1472.

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

Hall, Robert E. (1988). Intertemporal Substitution in Consumption. *Journal of Political Economy* 96(2), 339-357.

Kapteyn, Arie (1977) *A Theory of Preference Formation*, Ph.D. thesis, Leiden University.

Kapteyn, Arie and Federica Teppa (2003). Hypothetical Intertemporal Consumption Choices. *The Economic Journal* 113 (March), C140-C152.

Kapteyn, Arie and Tom Wansbeek (1985). The Individual Welfare Function. *Journal of Economic Psychology* 6(4), 333-363.

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

Van Praag, Bernard M.S. (1968). *Individual Welfare Functions and Consumer Behavior*. North-Holland Publishing Company, Amsterdam, Contributions to Economic Analysis, nr. 57.

Van Praag, Bernard M. S. (1971). The Welfare Function of Income in Belgium: An Empirical Investigation. *European Economic Review* 2, 337-369.

Van Praag, Bernard M. S. and Arie Kapteyn (1973). Further Evidence on the Individual Welfare Function of Income: an Empirical Investigation in The Netherlands. *European Economic Review* 4, 32-62.

Van Praag, Bernhard M. S. and Ada Ferrer-i-Carbonell (2004), *Happiness Quantified*, Oxford University Press: Oxford.

Yogo, Motohiro (2004). Estimating the Elasticity of Intertemporal Substitution When Instruments are Weak. *Review of Economics and Statistics* 86(3), 797-810.

Appendix A

This appendix explains the sample selection in more detail and shows some summary statistics for the working sample.

Table A.1 shows the sample selection and the number of observations of the working sample. In what follows, we explain why some of the observations were dropped. In particular, two of the sample selection steps might cause some concern. First, a high number of observations (2,562) were dropped because of unusable answers to the income satisfaction questions. Of these, 302 observations were non-monotonic, 1,006 had an implausibly low answer (< 11) to the question of which income the respondent would consider good or very good. The remaining observations (1,254) were dropped because the answers were missing or had coding errors without an obvious correction, or because the answers did not make sense given the wording of the questions (for example when respondents stated very low amounts which could be monthly amounts although the question asked about annual amounts, but a correction would have yielded amounts far outside actual incomes).

Table A.1: Sample Selection

Initial Observations (1995-2005) with income satisfaction questions answered	19,859
Answers to income satisfaction questions appearing to be topcoded	146
Only 1 or 2 answers to the 6 income satisfaction questions	99
Unusable income satisfaction answers	2562
Missing income expectations	138
First question about savings (yes/ no) not answered	3,568
Erroneous income expectations	148
Missing or negative consumption bounds	380
Highest expected income < 5,000 Euros	656
Final Sample	12,162

To reduce the number of observations lost, we recoded obvious errors (such as the wrong number of zero digits). Of the final 12,162 observations used, 665 had some or all of the answers recoded. We also corrected the lowest and/ or highest expected income in 2,779 cases. The vast majority of these cases seemed to have confused monthly and annual amounts because of misleading wording.

In addition, 3,568 observations were dropped because both the lower and upper limits of savings were missing as a result of the nonresponse to the question whether the respondent has put any money aside in the past 12 months. Of these, 257 did not answer this question because the respective part of the questionnaire was only asked of household heads and their partners. All of the other respondents (with the exception of 7) did not answer the first or the subsequent question of the questionnaire.

Table A.2 shows the number of observations per survey year for the working sample used in the analysis. In 1997, the panel was moved from its original institute associated with the University of Amsterdam to CentERdata. In 2000 the technology used for the interviewing of respondents was thoroughly modernized. Each of these changes led to substantial sample loss. Starting 2001, the panel was gradually rebuilt.

Table A.2: Number of Observations by Survey Year (Working Sample)

Year	N
1995	1,642
1996	1,316
1997	1,042
1998	591
1999	596
2000	266
2001	1,215
2002	1,233
2003	1,358
2004	1,458
2005	1,445
Total	12,162

Table A.3 shows the summary statistics for the estimated parameters as well as the lowest and highest expected income by year and for the entire sample.

Table A.3: Summary Statistics: Medians, by Year and for Total

Variable	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	Total
Income	29,477	28,131	25,188	22,900	24,483	24,206	24,931	26,050	26,994	26,868	24,187	26,148
Utility Function Parameters												
μ_t	9.903	9.869	9.866	9.835	9.848	9.772	9.800	9.779	9.939	9.912	9.905	9.868
σ_t	0.388	0.381	0.384	0.367	0.367	0.412	0.410	0.406	0.436	0.409	0.376	0.396
Lowest expected net income	25,731	25,541	25,708	23,758	23,248	20,455	21,769	21,083	21,818	21,486	21,201	22,727
Highest expected net income	30,271	29,699	29,104	26,134	26,536	25,000	26,123	25,299	27,273	26,858	25,760	27,273
Log-median of subjective distribution in next year	10.257	10.247	10.236	10.112	10.120	10.041	10.083	10.051	10.145	10.092	10.061	10.136
Log-standard deviation of subjective distribution in next year	0.034	0.037	0.030	0.029	0.025	0.033	0.037	0.039	0.042	0.042	0.039	0.036
Median of subjective distribution in next year	28,479	28,206	27,885	24,638	24,826	22,939	23,924	23,190	25,453	24,149	23,402	25,237
Standard deviation of subjective distribution in next year	1.035	1.038	1.031	1.029	1.026	1.034	1.038	1.040	1.043	1.043	1.040	1.037
N	1,642	1,316	1,042	591	596	266	1,215	1,233	1,358	1,458	1,445	12,162

All income measures are annual and expressed in 2000 Euros.

Appendix B

This appendix derives equation (4) in Section 5. For convenience of notation, subscripts are suppressed. We have

$$E_t U'\left(x_{t+1},\mu_{t+1},\sigma_{t+1}\right)=\frac{1}{2\pi}\int_o^\infty \frac{1}{x}\frac{1}{\sigma}\exp\left[-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right]\frac{1}{x}\frac{1}{s}\exp\left[-\frac{1}{2}\left(\frac{\ln x-m}{s}\right)^2\right]dz \tag{12}$$

Applying a transformation of variables $z=\ln x$, this can be written as

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{1}{s}\frac{1}{\sigma}\exp\left[-z-\frac{1}{2}\left(\left(\frac{z-\mu}{\sigma}\right)^2+\left(\frac{z-m}{s}\right)^2\right)\right]dz\ . \tag{13}$$

It can be verified directly that this can be further rewritten as:

$$\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{s^2+\sigma^2}}\exp\left\{-\frac{1}{2}\left[\frac{(m-\mu)^2-\sigma^2s^2+2\mu s^2+2m\sigma^2}{s^2+\sigma^2}\right]\right\}.$$
$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{\sqrt{s^2+\sigma^2}}{s\sigma}\exp\left\{-\frac{1}{2}\left[\frac{z-\dfrac{\mu s^2+m\sigma^2-\sigma^2s^2}{s^2+\sigma^2}}{\dfrac{\sigma s}{\sqrt{s^2+\sigma^2}}}\right]^2\right\}dz \tag{14}$$

The second term is the integral of a normal density, and hence equal to one. So we obtain (reintroducing the subscripts):

$$E_t U'\left(x_{t+1},\mu_{t+1},\sigma_{t+1}\right)=\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{s_{t+1}^2+\sigma_{t+1}^2}}\exp\left\{-\frac{1}{2}\left[\frac{(m_{t+1}-\mu_{t+1})^2-\sigma_{t+1}^2s_{t+1}^2+2\mu_{t+1}s_{t+1}^2+2m_{t+1}\sigma_{t+1}^2}{s_{t+1}^2+\sigma_{t+1}^2}\right]\right\}\ . \tag{15}$$

It is of some interest to consider the case $s=0$, i.e. where there is no uncertainty about the future. In that case, the expression should reduce to the marginal utility of consumption in period $t+1$. Indeed we get

$$E_t U'\left(x_{t+1},\mu_{t+1},\sigma_{t+1}\right)=\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_{t+1}}\exp\left\{-\frac{1}{2}\left[\frac{(\ln x_{t+1}-\mu_{t+1})^2+2\ln x_{t+1}\sigma_{t+1}^2}{\sigma_{t+1}^2}\right]\right\}= \\ \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_{t+1}}\frac{1}{x_{t+1}}\exp\left\{-\frac{1}{2}\left[\frac{(\ln x_{t+1}-\mu_{t+1})^2}{\sigma_{t+1}^2}\right]\right\}, \qquad (16)$$

which is the lognormal density function.

Define $\kappa\equiv\frac{1+r}{1+\delta}$. The general solution for $\ln x$ that can now be derived from (2) and (3) is given by

$$\ln x_t=\mu_t-\sigma_t^2+\sigma_t\sqrt{\left(\sigma_t^2-2\mu_t\right)-\left(\sigma_{t+1}^2-2\mu_{t+1}\right)+\ln\frac{s_{t+1}^2+\sigma_{t+1}^2}{\kappa^2\sigma_t^2}+\frac{\left[m_{t+1}-(\mu_{t+1}-\sigma_{t+1}^2)\right]^2}{s_{t+1}^2+\sigma_{t+1}^2}}\;. \qquad (17)$$