

Center



Discussion Paper

No. 2004–26

MODELS FOR ANCHORING AND ACQUIESCENCE BIAS IN CONSUMPTION DATA

By A.H.O. van Soest, M. Hurd

March 2004

ISSN 0924-7815

Models for Anchoring and Acquiescence Bias in Consumption Data¹

Arthur van Soest and Michael Hurd

RAND Corporation

February 2004

Abstract

Item non-response in household survey data on economic variables such as income, assets or consumption is a well-known problem. Follow-up unfolding bracket questions have been used as a tool to collect partial information on respondents that do not answer an open-ended question. It is also known, however, that mistakes are made in answering such unfolding bracket questions. In this paper, we develop several limited dependent variable models to analyze two sources of mistakes, anchoring and acquiescence (or yea-saying), focusing on the first bracket question. We use the experimental module of the AHEAD 1995 data, where the sample is randomly split into respondents who get an open-ended question on the amount of total family consumption - with follow-up unfolding brackets (of the form: is consumption \$X or more?) for those who answer “don’t know” or “refuse” - and respondents who are immediately directed to unfolding brackets. In both cases, the entry point of the unfolding bracket sequence is randomized. We compare models in which the probability of a mistake depends on the deviation between the true consumption amount and the entry point amount \$X and models in which it does not. We find that allowing for acquiescence bias substantially changes the conclusions on the selective nature of non-response to the open-ended question and on the distribution of consumption expenditures in the population. Once acquiescence bias is taken into account, anchoring in the first bracket question plays only a minor role.

Correspondence: mhurd@rand.org, vansoest@rand.org; RAND Corporation, 1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138, U.S.A.

JEL classification: C81, D12

Key words: Unfolding brackets, Non-response, Anchoring, Yea-saying

¹ Financial support was provided by the National Institute of Aging.

1. Introduction

Item non-response in survey data on variables of interest to economists is a well-known problem, particularly if questions relate to sensitive information that respondents are not willing to provide or to information that respondents do not know exactly and find hard to estimate. Examples are the amounts invested in saving accounts or the value of assets such as stocks and bonds, or the value of total family income or consumption. Item non-response becomes particularly problematic if the information is not missing at random, i.e., if the probability of not responding correlates with the amount in question. See, for example, Manski (1989) and Little and Rubin (2002, Chapter 15).

Follow-up bracket questions are often used to reduce the loss of information due to item non-response and to extract at least partial information from initial non-respondents. See, for example, Juster and Smith (1997). Particularly in telephone interviews where it is difficult to show range cards or to let people choose in one step from a larger number of categories, unfolding brackets can be used to collect banded information. For example, a respondent who answers “don’t know” or “refuse” to a question on total family consumption in the past month then gets the question “Can you say whether it was \$2000 or more?” If the answer is affirmative, the next question is “Was it \$5000 or more?” etc.

The majority of initial non-respondents typically appear to be willing to answer one or more unfolding bracket questions, making unfolding brackets a useful tool to collect additional information. Several experiments, however, have shown that the distribution of the categorical answers obtained in an unfolding bracket design depends on the order in which the bracket points are presented. For example, if the first question is “Was it \$ 2000 or more?” i.e., the entry point is \$2000, the distribution gives more weight to the high consumption outcomes than if the entry point were \$500. A psychological explanation is that if people are unsure about the exact amount, the entry point will serve as an anchor that provides information about their own amount. This phenomenon is known as anchoring (see Tversky and Kahneman, 1974).

The literature shows that anchoring exists and becomes more prominent the more uncertain the respondent is about the exact answer (see, for example, Jacowitz and Kahneman, 1995). In order to use the answers to unfolding bracket questions for

meaningful economic analysis, the possibility of an anchoring bias needs to be taken into account. Several models for what the answers to bracket questions look like if they suffer from anchoring have been introduced and estimated, using different sources of experimental data with random entry points. Examples are Hurd et al. (1998), Herriges and Shogren (1996), and Cameron and Quiggin (1994). A comparison between competing models for anchoring, however, has not yet been performed.

Another problem that leads to incorrect answers in an unfolding bracket design is acquiescence or “yea- saying” This is the problem that people have a tendency to answer yes rather than no. See, for example, Holmes and Kramer (1995), Kanninen (1995) and Hurd (1999). It implies that answers may depend on the wording of the question. For example, the number of “yes” answers to the question “Was it 2000 or more?” will be higher than the number of “no” answers to the question “Was it less than 2000?”

In this paper, several models for anchoring and acquiescence are estimated using experimental data on household consumption from the AHEAD 1995 survey. The specific feature of these data is that the sample is randomly split in respondents that start with an open-ended question - and get follow-up brackets if they do not answer the open-ended question - and a sub-sample of respondents who are immediately directed to the unfolding bracket questions without being asked an open-ended question (see Hurd et al., 1998). In a companion paper, we have combined the data from the two sub-samples to construct a non-parametric test for the null hypothesis that the first bracket question is always answered correctly (van Soest and Hurd, 2003). This hypothesis is rejected. Incorporating simple acquiescence and anchoring probabilities that do not depend on the difference between the true amount and the entry point suggests that, for the first bracket question, acquiescence is a larger problem than anchoring.

In the current paper, we develop limited dependent variable models with more flexible forms of anchoring and acquiescence, accounting for the fact that the probabilities of anchoring and yea-saying mistakes fall with the distance between the entry point and the true amount, in line with the notion that uncertainty is a source of anchoring and acquiescence. Our model consists of several equations. First, a regression equation explains the consumption level obtained through open-ended answers (sometimes observed, sometimes not), taking into account focal point answers. Second, a

selection equation explains whether an open-ended answer is given or not. Third, anchoring and acquiescence are modeled as functions of linear combinations of background variables and the difference between the entry point and the unobserved consumption level. The model is used to analyze how anchoring and acquiescence affect conclusions about the consumption function in the AHEAD population.

The remainder of this paper is organized as follows. In the next section, some existing models for anchoring are reviewed. In Section 3, we describe the experimental data. Section 4 introduces the limited dependent variable models. Estimation results and their implications are presented in Section 5. Section 6 concludes.

2. Anchoring Models and Acquiescence Bias

An intuitively appealing anchoring model is the fixed-point model used by Hurd et al. (2001), Herriges and Shogren (1996), and O'Connor et al. (1999) in which the entry point E in the first bracket question serves as an anchor for follow-up questions. The fixed-point model can be interpreted as a Bayesian model, where respondents update their beliefs about an unknown amount if new information (in the form of an entry point or other bracket point) arrives. The intuition is that the respondent is uncertain about the true amount. The entry point is seen as an indication of what the true amount could be, and respondents update their estimates of the true amount towards the entry point.

An important feature of the fixed-point model is that answers to the first bracket question are not affected.² The respondent's estimate of the true amount will be drawn towards the entry point, but will remain on the same side. Thus if anchoring according to the fixed-point model were the only source of incorrect answers, answers to the first bracket question would always be correct. Jacowitz and Kahneman (1995) and Green et al. (1998) conclude that anchoring is already present in the first bracket question. They find that, if a high entry point is used, respondents too often report that the amount exceeds that entry point. Jacowitz and Kahneman ask respondents' estimates of objective quantities such as the length of the Mississippi river or the amount of gas used by an average American. They compare immediate open-ended answers with open-ended

² This is why Hurd et al. (2001) call it a fixed-point model: the transformation from unbiased to biased probability has a fixed point at the entry point.

answers of a group that first got a question on whether the amount was higher or lower than some anchor value. They find that high as well as low anchors pull the amount toward the anchor, which can be explained by the fixed-point anchoring model. They also find, however, that when high anchors are used the amount is sometimes pulled up beyond the anchor. Green et al. (1998) look at willingness to pay for public goods as well as objective quantities, using a similar design. Their main result is in line with the Jacowitz and Kahneman finding: “The proportion of respondents who accept a high value in the referendum question is larger than the proportion of respondents who offer the same value on their own in response to an open-ended question” (Green et al 1998, p. 94). While the authors explain this type of finding with anchoring, the fact that the bias is found for high values only suggests that yea-saying might also play a role.

Hurd et al. (1998) specify a parametric model that captures this phenomenon in a symmetric way, biasing probabilities smaller than 0.5 upward and probabilities larger than 0.5 downward. The idea is that respondents do not compare the true value Y to the entry point E , but instead compare Y to $E+\epsilon$, where ϵ is a mean zero error term, assumed to be normal and independent of all other variables in the model. Hurd et al. (1998) use the same device also at follow-up bracket questions, with question specific errors that are independent across questions and have a variance that can be different at each question in the unfolding bracket design. Their model is called a gating model, since respondents have to pass a number of gates to reach their final bracket answer.

Somewhat similar, though not (yet) applied in this context, are the models of misclassification used in the applied micro-econometrics literature (e.g., Hausman et al., 1998). Suppose that with some “fixed gating probability” $P(\text{Gat})$, people give the wrong answer. This implies $P[\text{“no”}] = P[Y < E](1 - P(\text{Gat})) + P[Y \geq E]P(\text{Gat})$ and $P[\text{“yes”}] = P[Y \geq E](1 - P(\text{Gat})) + P[Y < E]P(\text{Gat})$. If this probability does not depend on the true value and is thus the same whether the true amount is larger or smaller than the entry point, this explains that fractions of reported “yes” or “no” answers are too large if the corresponding population fraction is larger than 0.5 and too small if it is smaller than 0.5.

A direct test of acquiescence bias or “yea-saying” is performed by Hurd (1999). He uses experimental data from the 1996 wave of the Health and Retirement Study. Respondents were asked: “About how much could you sell your home for in today’s

housing market?” They were forced into brackets and were randomly given an entry point from a set of three ($E=\$50,000$, $E=\$100,000$ or $E=\$150,000$). They also got a randomly assigned format of the question: either “Would it be more than E,” or “Would it be E or more?” or “Would it be less than E or more than E?” Hurd found that the first two (unbalanced) formats led to many more “yes” answers than the third (balanced) format for non-financial respondents, while there was hardly any difference for financial respondents (i.e., the person in the household most knowledgeable in financial matters). He interpreted this as evidence of acquiescence related to uncertainty. Acquiescence bias is also a well-known problem in the willingness-to-pay (WTP) literature. For example, Boyle et al. (1998) find evidence of yea-saying for bid levels in the upper tail by comparing open-ended answers and bracket answers on the WTP for a moose hunting site. Holmes and Kramer (1995) compare independent samples of dichotomous choice answers and range card answers on willingness to pay for the protection of a specific forest ecosystem in the southern Appalachian Mountains. They find significant differences in the WTP distributions estimated from range cards and dichotomous choice answers, but they do not find systematic differences in the way in which WTP varies with individual characteristics. Frykblom and Shogren (2000) used experimental data to compare open-ended and discrete choice answers on students’ WTP for a specific book (“The Environment”) and found no differences, leading them to conclude that problems with discrete choice answers can be due to how the survey is framed rather than to the questions themselves. Kanninen (1995) introduces a fixed probability yea-saying model, with $P[\text{“yes”}|Y \geq E]=1$ and $P[\text{“yes”}|Y \leq E]=\gamma$. She estimates a yea-saying probability γ of 0.20 (t-value 7.4) for a wetlands improvement program.

3. The AHEAD Wave 2 Consumption Experiment

We use essentially the same data as Hurd et al. (1998), who describe the data in detail. The AHEAD panel (Assets and Health Dynamics Among the Oldest Old) is roughly representative for the US population over 70 and their spouses. The first wave was collected in 1993. The second wave was done in 1995, mainly by telephone interviews. At the end of the regular survey, participants were asked to complete an

additional experimental module. About 75% of all respondents were willing to do this. This is the sample used by Hurd et al. (1998) and for the current study.

The participants in the experimental module were randomly separated into 7 groups of approximately equal size. All these groups got some questions on total household consumption in the last month, but the question format differed across the groups.³ Three groups (sample I) started with the open question: “How much did your household spend on consumption in the past month?”⁴ Those who answered “don’t know” or “refuse” then got unfolding bracket questions, with different entry points for the three groups. The first unfolding bracket question was formulated as “Would the amount be \$E or more in the past month?” with E=500, E=1000 or E=2000. If a respondent answered “don’t know” or “refuse” to a bracket question, the sequence was stopped. The other four groups (sample II) immediately were given bracket questions, with different entry points for the four groups (\$500, \$1000, \$2000 and \$5000).

The complete sample consists of 4928 observations. About 3.4% of them did not give an answer to either an open-ended question or a follow-up or direct bracket question;⁵ these observations are discarded. This leaves 4760 observations. About 42% of them (sample I; 2097 observations) started with an open-ended question. Almost two thirds of them gave an open-ended answer (1416 observations). The remaining one third gave at least one bracket answer.

Figure 1 presents the empirical distribution function for the open-ended answers and the inferred cumulative distribution function for the follow-up bracket answers, separately for the three entry points. The difference between the distribution functions of open-ended answers and follow-up bracket answers suggests that open-ended answers tend to be lower, on average, than as suggested by the bracket-answers. Some of the differences are statistically significant. An explanation would be that high consumption

³ See Browning et al. (2002) for the usefulness of measuring total consumption expenditures and for an overview of advantages and drawbacks of an overall recall consumption question in household surveys.

⁴ The exact wording of the question was: “About how much did you and your household spend on everything in the past month? Please think about all bills such as rent, mortgage loan payments, utility and other bills, as well as all expenses such as food, clothing, transportation, entertainment and any other expenses you and your household may have.”

⁵ About 0.8% stopped the interview before attaining answering the consumption question and about 2.6% answered “don’t know” or “refuse” to the first bracket question. See the appendix of van Soest and Hurd (2003) for details.

families are more concerned about their privacy and less willing to give a precise number, i.e., open-ended answers could under-represent high amounts (negative selection into open answers; positive selection into follow-up brackets). Alternatively, respondents could make mistakes in their open-ended answers or their bracket answers (or both). In particular, the differences between the distributions of bracket responses for the three entry points suggest that some bracket answers may depend on the (randomly assigned) entry point, which could be explained by fixed-point anchoring.

Figure 2 shows the distributions for the immediate bracket respondents by entry point, together with the distribution of open-ended answers already described in Figure 1. Again, negative selection into open-ended answers might explain why the open-ended answers tend to be small compared to the bracket responses. A salient feature of this figure is that the distribution shifts with the entry point. In this case the numbers of observations are larger than in Figure 1 and most of the differences are significant. The anchoring models in the previous section provide explanations for these differences.

There seems to be common agreement in the literature that fixed-point anchoring biases the answers to the second and later bracket questions. The entry point in the first bracket question acts as an anchor and the respondent's estimate is pulled towards it. In this paper we focus on problems with answers to the first bracket question, which cannot be due to fixed-point anchoring. We do not use the answers to the later bracket questions, thus avoiding the need to deal with fixed-point anchoring in the analysis.

Figure 3 uses the answers to the first bracket question only to estimate the distribution for the sample of immediate bracket respondents as well as the sample of follow-up bracket respondents. It takes the information on entry point bracket answers from Figures 1 and 2 and combines it. For example, for the immediate bracket respondents, the fraction of households with consumption less than \$2000 is based upon the group of immediate respondents who got \$2000 as their entry point. This is the point (20,0.701) on the dash-dotted curve in Figure 2. Similarly, the fraction of households with consumption less than \$1000 is based upon the sub-sample of immediate bracket respondents with entry point \$1000, etc.

The three distributions in Figure 3 will differ if non-response to the open-ended question is non-random. Taking the reports at face value, 81.0% of those in sample I who

gave an open-ended response have consumption less than \$2000, compared to 75.1% of those who gave a follow-up bracket answer. Even if all reports are correct, there is no reason why these two numbers should be estimates of the same thing, since whether or not a respondent gives an open answer may well be correlated with the level of household consumption (a selection effect).

The immediate bracket responses are, because of the random assignment, drawn from the same population as the combined sample of open-ended and follow-up bracket respondents. This implies that, if all answers to open-ended questions and first bracket questions are correct, the distribution of immediate bracket respondents should be a weighted mean of the other two distributions. The figure suggests this is not the case.

In van Soest and Hurd (2003), a formal test is developed based upon this intuition that looks at differences at several entry points. The null hypothesis that all answers are correct is rejected. Van Soest and Hurd (2003) also incorporate a simple form of gating and acquiescence bias, using “fixed” probabilities of acquiescence and gating, i.e., probabilities that do not depend on the difference between the true amount and the amount in the question. First, they assume that there is a fixed probability $P(\text{Acq})$ that a bracket question is automatically answered with “yes” (as in Kanninen 1995). Second, if there is no such automatic “yes,” they assume that there is a fixed probability $P(\text{Gat})$ that people give the wrong answer (“yes” if $Y < E$ or “no” if $Y \geq E$). With these two additions to the nonparametric model, the probability that someone answers “yes” if the true consumption amount is less than E , is given by $P(\text{Acq}) + (1 - P(\text{Acq}))P(\text{Gat})$. The probability that someone answers “no” if the true amount is at least E , is given by $(1 - P(\text{Acq}))P(\text{Gat})$. Hurd and van Soest (2003) find that acquiescence is more important than gating: the estimate of the gating probability is 0.018 with standard error 0.039, while the acquiescence probability is 0.170 with standard error 0.037.

To check whether a fixed acquiescence probability is reasonable, we have repeated the analysis for sub-samples with different household income levels.⁶ If acquiescence and gating depend on the (absolute value of) the difference between the

⁶ We have used the household net income variable in the RAND version of the HRS, which uses a sophisticated imputation procedure to replace missing values of income components, so that an income value is available for almost the complete sample.

entry point and the true amount, we should see a smaller bias for those respondents whose true consumption is unlikely to be close to the entry point. Estimating the unrestricted model for groups with different income levels shows that this is indeed the case. For example, we compared the estimates of $P[Y < 1000]$ based upon samples I and II separately for those with income between the first and third income quartile and for those with income below the first or above the third income quartile. The second and third income quartile households lead to estimates 0.480 and 0.376 for samples I and II; the first and fourth income quartile households give estimates 0.504 and 0.463. The difference is significant in the first case but not in the second. This suggests that acquiescence bias is more prominent in the sample where many consumption values are close to the entry point ($E=1000$), and argues against using a model with a fixed acquiescence probability.

Explanatory Variables

Due to missing values on years of education and self-reported memory skills, two observations had to be discarded. Thus the estimation results will be based upon 4758 observations. The demographic variables included in the analysis are summarized in Table 1.

The financial respondent dummy is included since financial respondents might be less uncertain about the household's total consumption than non-financial respondents. On the other hand, marital status (and household size)⁷ will probably affect the household consumption level, but not response behavior. Age can affect the consumption level but might also act as a proxy for cognitive skills. Gender might influence consumption, but the tendency not to respond to the open question or to give an incorrect answer to a bracket question might also be gender specific. Years of education may serve as a proxy for human capital wealth and for cognitive skills or knowledge of financial household matters in the other equations.

Dummies for whether the respondent smokes or drinks alcohol may proxy individual discount rates and attitudes towards savings and consumption. They may also,

⁷ In this data set of elderly respondents, household size and marital status are strongly correlated. Including them as separate regressors did not lead to sensible results. We therefore do not include household size.

more in general, proxy attitudes towards life related to financial knowledge and/or privacy concerns, and thus may affect response behavior and the sensitivity to anchoring or acquiescence bias. The AHEAD data contain two direct measures of memory skills. The first is based upon a self-report on how good the respondent's memory is these days. The second is the result of a test in which a list of 40 words is read to the respondent, who is then asked to reproduce as many of these words as possible. These variables can have an influence on response behavior and on anchoring or acquiescence errors.

Household income is constructed from income with many missing values. Excluding the observations with missing values would reduce the sample size by more than 50%. Instead, we use the RAND version of the data where missing values have been replaced by imputations. This induces measurement error, however, which will bias the income elasticity towards zero. To account for this, we replace log income by its prediction on the basis of lagged income and all the exogenous variables. This will lead to consistent estimates if the deviation between the predicted and the true (log) income variable is normally distributed and independent of the regressors, and if measurement errors in lagged and current income are independent. Since the imputations of missing values do not make use of lagged information, the latter assumption seems reasonable as far as imputation error is concerned. If respondents persistently over- or underestimate their income in different time periods, then our consumption elasticity will still be biased towards zero, but less than when the original income variable were used at face value.

For wealth, a similar problem arises. In the regressions we include lagged wealth, since current wealth may be affected by consumption decisions. Because the data are from the second wave, this leaves no further lags to be used as instruments. We therefore use the same instruments as for income to predict the log of lagged wealth.⁸

⁸ In principle it would also be possible to use income and wealth from later waves. Due to panel attrition, however, this would substantially reduce the number of observations that can be used.

Table 1. Background Characteristics (4758 observations)

Variable	Description	Mean	Std. Dev.
Dfinr	1 if financial respondent, 0 otherwise	0.745	0.436
Dman	1 if man, 0 if woman	0.346	0.476
Birthyr	Year of birth - 1900	17.905	6.242
Yrs educ	Years of education	11.410	3.398
Nonwhite	0 if white (and not Hispanic); 1 otherwise	0.165	0.371
Catholic	1 if of catholic religion; 0 otherwise	0.244	0.429
Other rel	1 if neither catholic nor Protestant; 0 otherwise	0.102	0.303
Marrlt	1 if married or living together; 0 otherwise	0.552	0.497
Dsmoke	Dummy smoker	0.089	0.284
Ddrink	Dummy drinks alcohol	0.425	0.494
Self_mem	Self-rated memory skills (1: excellent,...5: poor)	2.919	0.979
Words_mem	Immediate word listing memory test (0-10 words from a list of 40) ^a	4.906	1.825
Lhhincpc	Log household income per capita ^b	9.428	0.918
Lhhincpc_p	Prediction of Lhhincpc ^c	9.427	0.535
Lhhwltpc	Log household wealth per capita ^d	10.176	3.635
Lhhwltpc_p	Prediction of Lhhwltpc ^c	10.170	1.745

Notes:

- a. Words_mem has 44 missing observations. These are set to the sample median for the estimations.
- b. Log household income for singles (dmarlt=0), $\log(0.5*[\text{household income}])$ for couples (dmarlt=1).
- c. Prediction based upon regression on all the variables in the table other than income and wealth and on lagged log income per capita.
- d. Log household wealth for singles (dmarlt=0), $\log(0.5*[\text{household wealth}])$ for couples (dmarlt=1). 1.7% has negative wealth and 4.0% has zero wealth. In these cases, \ln wealth is defined as $-\ln(-\text{wealth})$ and 0, respectively.

4. Models with Focal Points, Selection, Acquiescence Bias, and Gating

In this section we will introduce some limited dependent variable models explaining monthly consumption and response behavior, accounting for acquiescence and anchoring. There are several reasons why this is useful. First, the discussion at the end of the previous section and the existing literature suggest that anchoring and acquiescence effects may vary with the deviation between the true amount and the entry point. To incorporate this, we need to model the complete distribution of consumption in the population of interest. Second, we want to know how consumption, response behavior, and anchoring and acquiescence vary with background variables such as age, gender, education level, etc. The sample is much too small to do this in a non-parametric way. Third, simulations on the basis of a completely specified model can be used to analyze how anchoring and acquiescence can bias the conclusions on the distribution of consumption. For given parameter estimates, the conditional distribution of consumption given a set of covariates x is completely known, so that we can compare estimated (conditional) distribution functions, means, standard deviations, etc. for models that do and do not allow for acquiescence and gating.

Fourth, the parametric model will use the information in the open-ended answers efficiently and will lead to tests of anchoring and acquiescence with higher power than the non-parametric tests in van Soest and Hurd (2003). Of course this comes at the cost of parametric assumptions, but we hope to largely overcome this drawback by modeling the distribution of consumption in a flexible way.

Focal Points

Focal points are a salient feature of the open answers, in particular at \$500, \$1000, \$1500 and \$2000 (cf. Hurd et al., 1998, and the jumps in the cumulative distribution of open responses in Figure 1; see also Figure A1 in the appendix for a more detailed picture). Since focal points coincide with entry points, it is essential to take the focal points explicitly into account. A limited dependent variable model generalizing ordered probit and tobit is used to achieve this, modeling the open-ended answers as a mixed continuous discrete distribution.

The observed open-ended answer y (ln consumption) is modeled as a transformation of an underlying latent variable y^* which depends upon background characteristics x (including a constant term) and a normally distributed error term ε in an additive way:⁹

$$y^* = x'\beta + \varepsilon$$

(1)

$$\varepsilon|x \sim N(0, \sigma_\varepsilon^2)$$

The transformation is similar to ordered probit for the discrete focal point answers, but also allows for any answer in the intervals around the focal points. With four focal points $\ln(500)$, $\ln(1000)$, $\ln(1500)$ and $\ln(2000)$, the complete specification is as follows:

	If $y^* \leq m_1$:	$y=y^*$
	If $m_1 < y^* \leq m_2$:	$y=\ln(500)$
	If $m_2 < y^* \leq m_3$:	$y=\ln(500) + (\ln(1000) - \ln(500)) (y^* - m_2) / (m_3 - m_2)$
	If $m_3 < y^* \leq m_4$:	$y=\ln(1000)$
(2)	If $m_4 < y^* \leq m_5$:	$y=\ln(1000) + (\ln(1500) - \ln(1000)) (y^* - m_4) / (m_5 - m_4)$
	If $m_5 < y^* \leq m_6$:	$y=\ln(1500)$
	If $m_6 < y^* \leq m_7$:	$y=\ln(1500) + (\ln(2000) - \ln(1500)) (y^* - m_6) / (m_7 - m_6)$
	If $m_7 < y^* \leq m_8$:	$y=\ln(2000)$
	If $y^* > m_8$:	$y=y^*$

It is easy to see that this model specializes to a standard linear regression model with normally distributed errors and without focal points if $m_1 = m_2 = \ln(500)$, $m_3 = m_4 = \ln(1000)$, $m_5 = m_6 = \ln(1500)$, and $m_7 = m_8 = \ln(2000)$. This also shows that the specification is much more flexible than a linear regression model. The ordered probit model with focal points only is the other extreme special case; it arises if $-\infty = m_1 < m_2 = m_3 < m_4 = m_5 < m_6 = m_7 < m_8 = \infty$. In general, y^* is just a tool to create a flexible mixed

⁹ The index indicating the respondent is suppressed throughout.

discrete-continuous distribution of y (given background variables x) using well-known limited dependent variable techniques. It is not necessary to interpret y^* as a true underlying consumption value and the transformation from y^* to y does not have anything to do with rounding. In the end, we are interested in the distribution of y and not in the distribution of y^* .¹⁰

In the sequel, the category thresholds m_1 and m_8 are normalized at values $\ln(500)$ and $\ln(2000)$, respectively. This excludes the special case with focal points only. Since the reported open answers in the data do not only contain focal points, and in particular also contain values below 500 and above 2000, this loss of generality is harmless for the data at hand.

Selective Response Behavior at Open-Ended Questions

A standard selection equation is used to explain whether respondents in sample I who first get an open-ended question, give an open-ended answer or choose to answer the first bracket question. (As explained before, respondents who neither give an open answer, nor answer the first bracket question, are discarded in the analysis.) The selection equation is specified as follows:

$$s^* = x'\gamma + u$$

(3) $s^* \geq 0$: open-ended answer

$s^* < 0$: bracket answer

The error term u and the error term ε in the main equation (1) are assumed to be jointly normal and independent of background variables x :

$$u, \varepsilon | x \sim N_2((0,0)', \Sigma)$$

¹⁰ A more structural approach is followed by Battistin et al. (2000). They model focal points and the true distribution of consumption using a mixture of several rounding procedures. They identify their model using a second data set with non-rounded (diary based) consumption amounts and do not address non-response or bracket response. Both the nature of our data and our focus on non-response and biases in bracket answers make this approach infeasible in our context.

Here $\Sigma(1,1)$ is normalized to 1, $\Sigma(2,2)=\sigma_\varepsilon^2$ and $\Sigma(1,2) = \rho\sigma_\varepsilon$. If in case of an open-ended answer, y^* were observed rather than y , and the information in the bracket answers would not be used, then this would give the standard Heckman selection model. If $\rho=0$, (1) – (2) could be estimated consistently using the observations on open-ended answers only and ignoring the selection mechanism. In general for unknown ρ , it would still be possible to estimate (1) – (3) without using any information on either the follow-up bracket answers in sample I or the immediate bracket answers in sample II, but identification would then rely on exclusion restrictions in (1) or on functional form assumptions. Using bracket information will help to identify the model non-parametrically, but account has to be taken of gating and/or acquiescence bias.

The bracket question is formulated as: “Is the amount E or more?” Without gating or acquiescence bias, the probability of an answer “yes” for a forced bracket respondent in sample II is given by:

$$\begin{aligned}
 P[\text{“yes”} \mid x] &= P[y \geq \ln E \mid x] = P[y^* \geq m(E) \mid x] = P[\varepsilon \geq m(E) - x'\beta \mid x] = \\
 (4) \quad &= \Phi[(x'\beta - m(E))/\sigma_\varepsilon]
 \end{aligned}$$

where $m(500) = m_1$, $m(1000) = m_3$, $m(1500) = m_5$, $m(2000) = m_7$ and $m(5000) = \ln(5000)$. The discrete nature of the consumption answers is fully taken into account here; if the question would have been “Is the amount more than E?” then $m(500)$ would have been taken equal to m_2 , etc.

For follow-up bracket respondents, similar expressions are obtained, but now conditional probabilities are needed, given selection into the unfolding brackets:

$$\begin{aligned}
 P[\text{“yes”} \mid x, s^* < 0] &= P[y \geq \ln E \mid x, s^* < 0] = P[y^* \geq m(E) \mid x, s^* < 0] = \\
 (5) \quad &= P[\varepsilon \geq m(E) - x'\beta \mid x, u < -x'\gamma] = \\
 &= P[\varepsilon \geq m(E) - x'\beta \text{ and } u < -x'\gamma \mid x] / P[u < -x'\gamma \mid x]
 \end{aligned}$$

Acquiescence

Let us first introduce acquiescence bias, assuming that there is no gating. Acquiescence bias or “yea-saying” implies that some respondents incorrectly answer “yes.” One way of allowing for acquiescence bias is to assume that with probability $P(\text{Acq}|x) = \Phi(x'\delta)$, bracket respondents answer “yes” irrespective of the true value or the entry point. This is the “fixed” acquiescence probability model of Kanninen (1995) and van Soest and Hurd (2003), where the acquiescence probability is allowed to vary with exogenous variables x (but is fixed in the sense that it does not vary with y or E , conditional on x).

As argued above already, this model may be too restrictive. Results of Hurd (1999) imply that acquiescence bias increases with uncertainty. This is also suggested by our comparison of close to median income households with far from median income households at the end of the previous section (although that analysis did not control for x). It implies that, for a given respondent, the probability of an incorrect “yes” answer should be larger if the true value is close to the entry point than if it is far below the entry point. To take this into account, acquiescence bias is introduced as follows, using a (respondent specific) acquiescence parameter A that is added to y^* before the comparison with $m(E)$ is made:

$$\begin{aligned} P[\text{“yes”} | x, A] &= P[y^* + A \geq m(E) | x, A] = \\ (6) \quad &= \Phi[(x'\beta + A - m(E))/\sigma_\varepsilon] \end{aligned}$$

The acquiescence bias (for given x and A) is then given by the difference between (6) and (4). To interpret this, note that for small values of A , this bias can be approximated by

$$A / \sigma_\varepsilon \phi[(x'\beta - m(E))/\sigma_\varepsilon]$$

Thus the acquiescence bias depends on the parameter A as well as on the probability density around the threshold $m(E)$. If E is a focal point, this refers to the number of respondents who are on the borderline between having consumption smaller than E or

consumption equal to E. If E is not a focal point, it refers to borderline respondents with consumption close to E.

Not all answers will be contaminated by acquiescence bias. It seems reasonable to assume that the parameter A varies with observed and unobserved respondent characteristics. Two distributions of A will be considered. The first is a normal distribution, in line with the other distributional assumptions made above:

$$(7a) \quad A|x \sim N(x'\delta, \exp(x'\eta)^2), \text{ independent of } \varepsilon \text{ and } u$$

The second explicitly imposes that the acquiescence bias has to be nonnegative and uses a censored normal, with censoring from below at zero:¹¹

$$(7b) \quad A|x \sim \text{CN}(x'\delta, \exp(x'\eta)^2; 0)$$

The assumption that A is independent of ε and u could be criticized. While it seems reasonable to assume that A is not affected by the true consumption level, it might be the case that the size of the acquiescence bias is related to the tendency not to answer the open-ended question (for sample I). In particular, if not answering the open-ended question is related to uncertainty rather than privacy concerns, uncertainty could drive both response behavior and acquiescence bias. With the data at hand, it seems too ambitious to try and identify a (negative) correlation between u and A that could reflect this. Hopefully, it is picked up by some of the background variables in x (such as the financial respondent dummy or the indicators of memory skills).

Gating

As for acquiescence, two types of models are considered: one with a “fixed” gating probability $P(\text{Gat}|x) = \Phi(x'\theta)$ (fixed in the sense that it does not depend upon y or E) and one in which gating varies with the distance between the entry point and the true amount.

¹¹ $z \sim \text{CN}(\mu, \sigma^2, 0)$ if $z = \max(0, z^*)$ with $z^* \sim N(\mu, \sigma^2)$.

Combining fixed gating with the model incorporating acquiescence bias along the lines discussed above gives the following probabilities for answering “yes” to a bracket question, conditional upon x and the acquiescence parameter A :

$$\begin{aligned}
 & P[\text{“yes”} \mid x, A] = \\
 (8) \quad & = P[y^* + A \geq m(E) \mid x, A] (1-P(\text{Gat})) + P[y^* + A < m(E) \mid x, A] P(\text{Gat}) = \\
 & = \Phi[(x'\beta + A - m(E))/\sigma_\varepsilon](1-\Phi(x'\beta)) + (1-\Phi[(x'\beta + A - m(E))/\sigma_\varepsilon])\Phi(x'\beta).
 \end{aligned}$$

Bivariate probabilities incorporating the decision not to answer the open-ended question can be constructed in the same way.

An alternative is that the probability of gating falls with the distance between E and Y . This can formally be introduced in a similar way as acquiescence bias but has a different interpretation. The “gating error” G will be drawn from a symmetric distribution around 0, leading to incorrect classifications in both directions. Following Hurd et al. (1998), but accounting for focal answers and for acquiescence bias in line with the rest of the model, this leads to the following assumptions, generalizing (6):

$$P[\text{“yes”} \mid x, A] = P[y^* + A + G \geq m(E) \mid x, A]$$

$$G|x \sim N(0, \exp(x'\theta)^2); \quad G \text{ is independent of } A, u \text{ and } \varepsilon$$

The normality assumptions allow for an analytical expression of the probability of a “yes” under gating, conditional on the acquiescence parameter:

$$(9) \quad P[\text{“yes”} \mid x, A] = \Phi[(x'\beta + A - m(E))/\sqrt{(\sigma_\varepsilon^2 + \exp(x'\theta)^2)}]$$

For small $\sigma_G = \exp(x'\theta)$, the gating bias, i.e., the difference between the probabilities in (9) and (6), can be approximated by

$$- 0.5 \sigma_G^2 / \sigma_\varepsilon^2 \phi[(x'\beta + A - m(E))/\sigma_\varepsilon] (x'\beta + A - m(E)) / \sigma_\varepsilon$$

This shows that gating reduces probabilities larger than 0.5 (i.e., cases with $x'\beta + A - m(E) > 0$) and increases probabilities less than 0.5 (i.e., $x'\beta + A - m(E) < 0$), corresponding to the notion that gating increases the probabilities of the tail outcomes. Moreover, the size of the gating effect depends on the relative importance of gating versus genuine unobserved heterogeneity in y^* , on the probability density of borderline consumption values (accounting for acquiescence), and on the distance of the entry point to the mean of the distribution (i.e., the deviation from symmetry). A large negative value of $x'\theta$ implies no gating; the gating probability increases with $x'\theta$. The assumptions about gating and acquiescence bias are similar. In particular, we assume that the gating effect is independent of u , implying that gating effects are similar for follow-up bracket respondents and for forced bracket respondents with the same x and y^* . The fact that, on average, follow-up bracket respondents may be more uncertain than forced bracket respondents, and therefore may be more affected by gating, is taken into account through observed characteristics x only.

Identification and Estimation

Acquiescence and gating enter the model in a similar way; only the distribution of the sum of A and G will be identified. Combining (7b) with (9) gives an asymmetric distribution for $G+A$, which is a mixture of a censored normal and a normal. The model will be identified due to the asymmetry. Combining (7a) with (9) gives a distribution of $G+A$ that is symmetric around $x'\delta$. Although identification is still guaranteed due to the functional form of the variance, this seems infeasible in practice. Instead, we will not try to disentangle G and A for this specification, and replace (7a) and (9) by

$$G+A | x \sim N(x'\delta, \exp(x'\eta)^2)$$

The model can be estimated by maximum likelihood. For those who answer the open-ended question, the likelihood contribution is the product of a univariate normal density and a univariate normal (selection) probability (non-focal points), or a bivariate normal probability (focal points). For follow-up bracket respondents, the likelihood can be written as a function of univariate, bivariate and trivariate normal probabilities. The

latter can be obtained by numerical integration. For the forced bracket respondents in sample II, only univariate and bivariate normal probabilities are needed.

5. Results

Model Selection

Likelihood values of several competing model specifications are presented in Table 2. The table also contains the value of the Akaike Information Criterion of each model (AIC; minus 2 times the log likelihood plus 2 times the number of parameters).

The first model (M1; the benchmark model) has no gating or acquiescence. The second model has fixed gating and acquiescence probabilities that are not allowed to vary with individual characteristics x . This model is already a substantial improvement compared to the benchmark model in terms of likelihood and AIC. This is mainly due to the acquiescence probability which has a point estimate of about 0.05, while the estimated gating probability is only 0.01. Models M3, M4 and M5 allow for “fixed” gating and/or acquiescence probabilities that vary with background variables x (but not with y or E , given x). The model with both acquiescence and gating outperforms the benchmark model and the models with either fixed probability gating or fixed probability acquiescence.¹²

Models M6 – M9 allow for gating and/or acquiescence based upon the normal distribution in (7a). These models perform worse in terms of AIC than the fixed probability models M2-M5 and will be discarded in the sequel. Models M10 and M11 allow for censored normal acquiescence as in (7b) but not for gating. Model M11 outperforms all models with fixed probability gating and acquiescence. In terms of AIC, it also outperforms the more flexible models which allow for a normally distributed gating error in addition to the censored normal acquiescence error, giving a mixture distribution of the total term driving response error, A+G (M12 and M13). Apparently, allowing for a normally distributed gating effect is not necessary once acquiescence is controlled for in a flexible way. Of course this may also reflect an identification problem,

¹² Moreover, models M3 and M4 outperform their simplified versions in which the gating or acquiescence probability does not vary with x , both in terms of AIC and according to a likelihood ratio test.

since identifying the mixture distribution from the bracket answers and their deviations from open-ended answers is difficult in practice.

Table 2. Log likelihood and AIC values of several models

Model	Log likelihood	(# parameters)	AIC
No gating or acquiescence:			
M1. $A=0, G=0$	-5344.34	(22)	10732.7
Fixed gating and/or acquiescence probabilities:			
M2. $P(\text{Gat} x)=\theta_0; P(\text{Acq} x)=\delta_0$	-5281.53	(24)	10611.1
M3. $P(\text{Gat} x)=\Phi(x'\theta), P(\text{Acq} x)=0$	-5277.35	(31)	10616.7
M4. $P(\text{Gat} x)=0, P(\text{Acq} x)=\Phi(x'\delta)$	-5277.38	(31)	10616.8
M5. $P(\text{Gat} x)=\Phi(x'\theta), P(\text{Acq} x)=\Phi(x'\delta)$	-5265.21	(40) ¹³	10610.4
Varying gating and/or acquiescence probabilities:			
<i>Normal distribution:</i>			
M6. $A=0, G\sim N(0, \exp(x'\theta)^2)$	-5294.96	(31)	10651.9
M7. $A+G\sim N(\delta_0, \theta_0)$	-5299.66	(24)	10647.3
M8. $A+G\sim N(x'\delta, \sigma^2)$	-5299.20	(32)	10662.4
M9. $A+G\sim N(x'\delta, \exp(x'\eta)^2)$	-5273.61	(40)	10627.2
<i>Censored Normal:</i>			
M10. $A\sim \text{TN}(x'\delta, \sigma^2; 0), G=0$	-5275.47	(32)	10614.9
M11. $A\sim \text{TN}(x'\delta, \exp(x'\eta)^2; 0), G=0$	-5257.59	(40)	10595.2
<i>Mixtures:</i>			
M12. $A\sim \text{TN}(x'\delta, \sigma^2; 0), G\sim N(0, \exp(x'\theta)^2)$	-5272.24	(41)	10626.5
M13. $A\sim \text{TN}(x'\delta, \exp(x'\eta)^2; 0), G\sim N(0, \exp(x'\theta)^2)$	-5252.87	(49) ¹⁴	10603.7
Fixed gating, varying acquiescence:			
M14. $A\sim \text{TN}(x'\delta, \exp(x'\eta)^2; 0), P(\text{Gat} x)=\theta_0$	-5256.60	(41)	10595.2
M15. $A\sim \text{TN}(x'\delta, \exp(x'\eta)^2; 0), P(\text{Gat} x)=\Phi(x'\theta)$	-5249.80	(49) ¹⁵	10597.6

Finally, since among the models allowing for gating but not for acquiescence bias, the “fixed” gating probability model (model M3 in Table 2) performed better than the model in which the gating probability depends upon the deviation between true value and entry point, while the reverse is found comparing models that allow for acquiescence

¹³ Parameter in θ on dummy smoker converged to $-\infty$.

¹⁴ Parameters in θ on dummies smoker and non-drinker converged to $-\infty$.

¹⁵ Parameter in θ on dummy smoker converged to $-\infty$.

only, we have estimated mixed versions in which gating is incorporated with a fixed probability but acquiescence is incorporated with a varying probability through a censored normal acquiescence term. This is the model given by (6) and (7b) extended with fixed probability gating. That is, in model M15, the probability of a “yes” is given by (8) and in model M14 the gating probability is restricted to be a constant, i.e., $P(\text{Gat}) = \theta_0$ instead of $P(\text{Gat}) = \Phi(x'\theta)$. In model M14, the estimated gating probability is 1.54%. In terms of AIC, models M11 and M14 perform equally well. Both outperform the more general model M15. In this model, the coefficient on the dummy for smokers converges to $-\infty$, implying that for smokers the gating effect is zero. The only other significant coefficient in the gating probability is the dummy for drinking alcohol (positive with t-value 1.98).

Parameter Estimates

In the remainder we will focus on four models: the benchmark model without gating or acquiescence bias (M1 in Table 2, denoted model 1 from now on), the best model with gating only, i.e., the fixed probability gating model (model 2 from now on; M3 in Table 2), the model with censored normal acquiescence bias and no gating (from now on model 3; M11 in Table 2), and the other model with the lowest AIC value which has censored normal acquiescence and fixed probability gating independent of the regressors (model 4 from now on; M14 in Table 2). Table 3 presents the parameter estimates for these four models.

From an economic point of view, the consumption equation is the most interesting, since this gives the underlying distribution of consumption in the population (conditional on x), corrected for selective response behavior and, in models 2, 3 and 4, potential biases in bracket answers. The signs of the significant slope coefficients in the consumption equation are the same for all four models, and significance levels are also similar. There are some changes in the size of the slope coefficients but most of these are small.¹⁶

¹⁶ The coefficients are the estimated marginal effects of changing the regressor on y^* . Due to the nonlinearity of the transformation from y^* to y , the marginal effects on y vary across observations. Because of the chosen normalization, however, they are similar to the marginal effects on y^* , on average. For example, for model 4, additional computations show that, on average over all the observations, the marginal effects on y are 1.023 times as large as the marginal effects on y^* . The marginal effects on y for each individual observation are in between 0.978 and 1.036 times the marginal effect on y^* .

Table 3. Estimation Results Selected Models

	Model 1		Model 2		Model 3		Model 4	
log likelihood	-5344.34		-5277.38		-5257.59		-5238.60	
	Par.	S.e.	Par.	S.e.	Par.	S.e.	Par.	S.e.
y* equation (β)								
constant	1.802*	0.235	1.283*	0.232	1.639*	0.240	1.592*	0.238
0.01*bthyr	0.762*	0.224	0.710*	0.221	0.637*	0.250	0.682*	0.254
marr lt	0.297*	0.029	0.309*	0.028	0.306*	0.029	0.312*	0.030
logincome pc	0.506*	0.038	0.554*	0.037	0.508*	0.039	0.516*	0.038
logwealth pc	0.006	0.015	0.010	0.015	0.008	0.015	0.007	0.015
sig ϵ	0.679*	0.009	0.618*	0.009	0.615*	0.010	0.611*	0.010
Thresholds								
m1 (log 500)	6.215		6.251		6.251		6.251	
m2	6.397*	0.015	6.394*	0.014	6.390*	0.014	6.391*	0.014
m3	6.830*	0.018	6.816*	0.018	6.807*	0.018	6.810*	0.018
m4	7.080*	0.019	7.064*	0.019	7.055*	0.019	7.057*	0.019
m5	7.240*	0.018	7.225*	0.019	7.218*	0.019	7.219*	0.019
m6	7.357*	0.016	7.346*	0.017	7.341*	0.017	7.343*	0.017
m7	7.426*	0.015	7.416*	0.016	7.415*	0.016	7.416*	0.016
m8 (log 2000)	7.601		7.601		7.601		7.601	
Selection (γ)								
constant	0.181	0.173	0.200	0.173	0.240#	0.174	0.234#	0.174
dnonfinr	-0.238*	0.067	-0.240*	0.068	-0.231*	0.068	-0.233*	0.068
d man	0.221*	0.062	0.220*	0.062	0.217*	0.062	0.218*	0.062
0.01*bthyr	0.044	0.495	0.053	0.496	0.106	0.497	0.099	0.497
0.01*yrsted	1.608+	0.926	1.498+	0.929	1.027	0.935	1.106	0.935
d smokes	0.120	0.099	0.120	0.099	0.096	0.101	0.100	0.101
d drinks	0.044	0.062	0.041	0.062	0.045	0.062	0.044	0.062
self_mem	-0.078*	0.030	-0.077*	0.030	-0.076*	0.030	-0.076*	0.030
word_mem	0.052*	0.018	0.052*	0.018	0.053*	0.018	0.053*	0.018
p(u, ϵ)	-0.161*	0.037	-0.139*	0.042	0.065	0.052	0.034	0.060
Gating (θ)								
constant			-1.149*	0.458			1.542+	0.926
dnonfinr			-0.265#	0.206				
d man			-0.165	0.168				
0.01*bthyr			-0.039	1.437				
0.01*yrsted			-2.530	2.054				
d smokes			-0.392	0.320				
d drinks			0.259#	0.166				
self_mem			0.058	0.086				
word_mem			-0.101*	0.047				
Acquiescence (δ)								
constant					-5.091*	1.367	-4.206*	2.202
dnonfinr					0.801#	0.333	0.555+	0.330
d man					0.409#	0.286	0.300	0.236
0.01*bthyr					1.163	2.222	0.107	1.848
0.01*yrsted					20.555*	5.816	18.245*	8.051
d smokes					-0.290	0.559	-0.225	0.464
d drinks					0.372#	0.240	0.251	0.219
self_mem					0.011	0.125	0.034	0.099
word_mem					0.145*	0.077	0.114	0.078
Acquiescence (η)								
constant					2.042*	0.432	1.910*	0.561
dnonfinr					-0.644*	0.226	-0.742*	0.374
d man					-0.195#	0.148	-0.177	0.187
0.01*bthyr					0.146	1.302	0.206	1.520
0.01*yrsted					-8.185*	2.158	-8.351*	2.407
d smokes					-0.061	0.242	-0.062	0.281
d drinks					-0.019	0.149	-0.001	0.178
self_mem					0.024	0.080	0.006	0.088
word_mem					-0.118*	0.042	-0.149*	0.056

Notes: * |t-value|>1.96; +: 1.96>|t-value|>1.64; #: 1.64>|t-value|>1.28
 Model 1: no gating, no acquiescence bias (M1 in Table 2)
 Model 2: fixed probability gating (M3 in Table 2)
 Model 3: Acquiescence term censored normal (M11 in Table 2)
 Model 4: Acquiescence and gating (M15 in Table 2)

Most variables are significant and have the expected sign. Consumption increases with year of birth, i.e., falls with age. It is larger for couples ($d_{\text{mar}}=1$) than for singles, keeping per capita income constant. If instead total household income would be kept constant, then couples would consume slightly less than singles. The estimated income elasticity (keeping other variables constant) varies between 0.506 and 0.554. There is no explicit distinction between transitory and permanent income here, but for this sample of elderly households, transitory income will be of minor importance. For specifications in which we did not take account of measurement error and directly included observed household income, the estimated income elasticity varied between 0.10 and 0.16, clearly illustrating the downward attenuation bias due to measurement error. For specifications that predict household income without using lagged income as one of the instruments, we found still larger income elasticities in the order of magnitude close to one, but with much larger standard errors and with an implausible significantly negative elasticity of consumption with respect to total household wealth. We have therefore chosen to present the models which use lagged household income as one of the predictors of household income, and the estimated income elasticity can be interpreted as a lower bound on the true elasticity if a time persistent measurement error still leads to a downward attenuation bias.¹⁷ This bias will be much smaller than the attenuation bias if income itself is included as a regressor.

The wealth elasticity is essentially zero and not significant. This elasticity appears to vary across specifications in the direction opposite to the income elasticity. It is significantly positive if measurement error is not taken into account.

The slope coefficients in the selection equation indicating whether an open-ended answer is given are also similar for the four specifications. Financial respondents are significantly more likely to give an open-ended answer than others, in line with the notion that financial respondents have less uncertainty about the household's total consumption level. Men are significantly and substantially more likely to give an open-ended answer than women. Since gender will appear to play no role in the acquiescence or gating effects, this will probably not be due to differences in financial knowledge. It

¹⁷ The results concerning selective non-response to the open-ended consumption question and for anchoring and acquiescence bias are qualitatively similar if log income and log wealth are included themselves instead of their predicted values.

might mean that men are less concerned about privacy or confidentiality issues than women.

Education hardly plays a role. Age, smoking, or drinking, do not play any role either. The effects of the two indicators for memory skills are significant and have the expected sign: people with a better memory – either self-reported or according to the word listing test – are more likely to give an open-ended response. The magnitude of the difference is substantial. On average, someone at the first quartile of both memory skills indicators is about 10 percentage-points less likely to give an open-ended answer than someone at the third quartile.

The only selection parameter that changes substantially if acquiescence or gating is allowed for is the correlation coefficient ρ between the error terms ε and u in consumption and response equations. While this parameter is significantly negative in models 1 and 2, it is positive in model 3, although not significant, and negative and not significant in model 4. Allowing for acquiescence bias causes the switch, gating is not important here (models 1 and 2 give a similar ρ , as well as models 3 and 4). The intuition for this result is clear. Taken at face value, bracket reports imply higher consumption levels than open-ended answers (cf. Figure 3). Without acquiescence bias, this has to be attributed to selection. This leads to the negative estimate for ρ in models 1 and 2: those who give the open-ended answers tend to be the people with lower consumption (negative selection into open-ended answers). In models 3 and 4, however, yea-saying can explain why the bracket answers imply so many high consumption levels. Selection effects can then explain differences between forced bracket answers and follow-up bracket answers, if any. The follow up bracket respondents report similar consumption levels as the direct bracket respondents (see Figure 3). Intuitively, this explains why models 3 and 4 do not imply a significant selection effect. This argument is not the complete story, since Figure 3 does not condition on background variables while the models in Table 3 do.

It seems worthwhile to note here that in a selection model that discards the information in the bracket answers completely and uses the same specifications for the selection and consumption equations, the estimate of ρ is -0.75 with t-value -16 . Such a model thus would imply a huge negative selection effect, leading to the conclusion that

people with relatively low consumption levels tend to answer the open question. The model using the bracket information at face value gives a qualitatively similar conclusion but the magnitude of the selection effect is much smaller. The models that allow for acquiescence bias show that if there is a selection effect, it is probably of the opposite sign.

In model 2, the gating probability is larger for non-financial respondents than for financial respondents and is larger for those who drink alcohol than for those who do not. Both of these effects are significant at the 20% level only, however. The latter could reflect something like an easy-going life style for people who drink alcohol, accompanied by sloppier response behavior. For smoking, we find the opposite effect, but this is not significant. Gating falls with years of education but this effect is insignificant also. The only variable that is significant at the 5% level is the objective memory test. It indicates that those with a better memory make fewer gating errors. The negative effects of memory skills, the financial respondent dummy, and years of education are in line with the notion that anchoring becomes more likely if people are less knowledgeable and more uncertain about their actual consumption level.¹⁸

Acquiescence depends on the regressors through δ as well as η . Whether a regressor has a significant effect on acquiescence thus depends on whether the two coefficients on this regressor in δ and η are jointly significant. In model 3, Wald tests show that the dummy for non-financial respondents, years of education, and the words listing memory skills test are significant at the 5% level. The dummy for drinking alcohol is significant at the 14.4% level, other variables are insignificant even at the 20% level. The results for model 4 are similar, but the drinking dummy is insignificant at the 20% level and the non-financial respondent dummy is significant at the 5.3% level only.

One way to interpret the acquiescence parameters in models 3 and 4 is to look at the effect of the regressors on the expected value of the acquiescence parameter A. It is straightforward to show that this is given by

¹⁸ When other variables such as income and wealth are included, these also are insignificant at even the 20% level.

$$\frac{\partial E\{A|x\}}{\partial x} = \Phi(x'\delta / e^{x'\eta})\delta + \varphi(x'\delta / e^{x'\eta})e^{x'\eta}\eta$$

Thus the marginal effects on the acquiescence parameter vary across observations. If coefficients in δ and η have the same sign, the marginal effect always has that sign also. If not, the sign of the marginal effect varies across observations. For all three variables in model 3 that are significant for acquiescence at the 5% level, the sign changes. Averaging over all the observations (or looking at the medians, which are close to the means) for model 3, we find that non-financial respondents are less subject to acquiescence than financial respondents, that years of education has a positive effect, and that respondents with better memory skills – as measured with the word listing test - have a smaller tendency of acquiescence. Only the latter finding is in line with the notion that acquiescence is a sign of uncertainty or lack of financial knowledge.¹⁹ The same results are found for model 4.

Reporting Bias

Table 4 compares the reporting biases of models 2, 3 and 4 for two benchmark respondents. Both are married females of median age in a two persons household with household income \$37,000 and total household wealth \$220,000. Both are the financial respondent in their household and have 12 years of education. The first benchmark respondent (I) has a poor score on memory test and poor self-reported memory skills; the second (II) has a good score on memory test and good self-reported memory skills.²⁰ The first column gives benchmark estimates for the distribution function (which is the same for both respondents, since memory skills are not included in the consumption equation). The other columns give the differences between the reported and the true probability for entry bracket responses, according to each of the models 2, 3 and 4. For example, model 4 predicts that among those with characteristics of respondent I, 60.7% have consumption at most \$1500. The question “is your family consumption more than \$1500?” would be

¹⁹ These signs are in line with the acquiescence parameters in model M10 of Table 2, which is similar to model 3 but does not allow for heteroskedasticity in the underlying normal distribution of the acquiescence parameter. Significance levels of the acquiescence effects in this model are low, however, with only the objective memory test and the drinking dummy significant at the one-sided 10% level.

²⁰ The estimated probabilities of giving an open-ended answer are 0.595 for respondent I and 0.728 for respondent II.

answered with “no” by only 52.2% of such respondents. Only the difference between these numbers, 52.2-60.7=-8.5%-points, is presented in the table. This is the reporting bias in the estimated probability.

In model 2, the reporting bias is due to gating only and thus by assumption symmetric: it is positive for probabilities less than 0.5 and negative for probabilities exceeding 0.5. Large biases are found at very low as well as very high consumption values. In model 3 the bias is always negative (or zero) since it is due to acquiescence, increasing the probabilities of answering “yes” and thus reducing the “no” probabilities that are tabulated here. The bias increases with the consumption level except at very high levels of consumption where the density becomes sparse. In model 4 with gating as well as acquiescence, the bias can be positive or negative for probabilities less than 0.5 and is always negative for probabilities larger than 0.5.

Table 4. Reporting Bias Benchmark Respondents (percentage-points)

	Respondent I				Respondent II		
	True Mod 1	Bias Mod 2	Bias Mod 3	Bias Mod 4	Bias Mod 2	Bias Mod 3	Bias Mod 4
< 250	0.5	6.2	-0.1	1.5	2.0	0.0	1.5
< 500	5.8	5.7	-1.1	0.4	1.8	-0.7	0.8
≤ 500	9.6	5.3	-1.8	-0.4	1.7	-1.2	0.3
< 1000	25.3	3.4	-4.5	-3.6	1.1	-3.0	-1.6
≤ 1000	38.3	1.7	-6.6	-6.0	0.5	-4.2	-3.0
< 1500	47.5	0.4	-7.9	-7.5	0.1	-4.9	-3.8
≤ 1500	54.4	-0.6	-8.8	-8.5	-0.2	-5.3	-4.2
< 2000	58.4	-1.1	-9.2	-9.0	-0.4	-5.5	-4.5
≤ 2000	68.1	-2.5	-10.0	-10.0	-0.8	-5.7	-4.8
< 3000	85.7	-4.8	-10.1	-10.1	-1.5	-5.1	-4.5
< 4000	93.2	-5.6	-9.1	-9.0	-1.8	-4.0	-3.7
< 5000	96.6	-6.0	-8.0	-7.8	-1.9	-3.1	-3.1

Notes:

Respondents I and II: woman, born in 1919, married, not smoking, not drinking, household income \$37,000 and household wealth \$220,000, financial respondent, 12 years of education, born in 1919. Respondent I: self-rated memory fair, word list memory score 3; respondent II: self-rated memory very good, word list score 7.

First column: true (= reported) probability according to model 1 (equal for the two respondents); other columns: probability of reporting minus true probability according to models 2,3,4.

In all models, the bias is typically lower for respondent II than for respondent I, in line with the notion that respondent I is less knowledgeable and more uncertain about the consumption amount, making her more subject to reporting bias. Particularly in models 3 and 4, the estimated reporting bias for respondent I is sometimes substantial, amounting to about 10 percentage-points at consumption levels around \$2000 - \$3000. Respondent II has smaller gating as well as acquiescence bias. In model 4, her total bias never exceeds 5 percentage-points.

Distribution of Consumption

Table 5 shows the estimates of the distribution of consumption for the same benchmark persons that were used in Table 4. Since the consumption equation does not include memory skills, the distribution is the same for the two respondents. Model 2 implies a less dispersed distribution than Model 1, since it attributes part of the observed dispersion in bracket reports to gating. This suggests that estimates that ignore gating would over-estimate the dispersion in the conditional distribution of consumption (given the covariates). The differences do not seem that large, though.

Table 5. Estimated Distributions for Respondents with Benchmark Characteristics (%-points)

	Model 1	Model 2	Model 3	Model 4
< 250	0.5	0.2	0.4	0.3
< 500	5.8	4.3	6.2	5.8
≤ 500	9.6	7.6	10.6	9.9
< 1000	25.3	22.7	28.4	27.4
≤ 1000	38.3	36.5	43.3	42.3
< 1500	47.5	46.6	53.9	52.8
≤ 1500	54.4	54.4	61.7	60.7
< 2000	58.4	58.9	66.2	65.3
≤ 2000	68.1	70.0	76.4	75.6
< 3000	85.7	88.1	91.6	91.3
< 4000	93.2	95.0	96.8	96.6
< 5000	96.6	97.8	98.6	98.6

Notes: See Table 4 for characteristics of benchmark respondents. The table presents the probabilities (in %) that consumption is less than or less than or equal to given value according to the estimates of models 1-4.

Models 3 and 4 almost always give higher probabilities than model 1 for both benchmark respondents, with differences that can amount to 7 or 8 percentage points. Allowing for acquiescence means that some of the “yes” answers to the bracket questions reflect yea-saying rather than high consumption levels. Correcting for this leads to estimates of the (conditional) distribution of consumption that gives more weight to lower consumption levels. This is similar for models 3 and 4. The main difference between these two is that model 4 also corrects for the over-dispersion due to gating, and gives lower probabilities in particularly the left hand tail of the distribution, implying less dispersion than model 3.

In Figure 4, the exercise of Table 4 is repeated for every respondent in the sample, and the means over the whole sample are presented. This gives estimates of the cumulative distribution of consumption according to the four models. Moreover, the estimate based upon a Heckman (1979) selection model extended with focal points that does not use the bracket information is also presented. As already mentioned, this selection model leads to a very large negative estimate of the correlation coefficient driving selection into giving an open-ended answer, implying that in the observed open-ended answers, low consumption levels would be over-represented. Correcting for this leads to the conclusion that consumption levels in the population are much higher than as reflected by the open-ended answers. The mean of log consumption would be 7.0671 (corresponding to a consumption level of \$1173) with standard deviation 0.947. All models that account for the bracket information give smaller selection effects, implying smaller consumption levels for bracket respondents.

Figure 4 shows that there is not much difference between the distributions according to models 1 and 2 and not much difference between models 3 and 4. Differences between models 1 and 2 on the one hand and models 3 and 4 on the other hand are more substantial. Thus accounting for acquiescence (models 3 and 4) makes more difference than accounting for gating (models 2 and 4). It leads to lower consumption values, as already explained for the benchmark respondents. This is also reflected in the estimated overall means of log consumption. Models 1 and 2 give means of 6.908 (consumption level \$1000) and 6.887 (\$979). Models 3 and 4 imply smaller means of log consumption, 6.795 (\$893) and 6.808 (\$905), respectively. All models give

a much lower mean than the selection model, and the models allowing for acquiescence give lower means than those not allowing for acquiescence.

Similar conclusions are obtained concerning the dispersion of the distribution. All four models imply a much lower standard deviation of log consumption than the selection model. Models 1 and 2 give values of 0.823 and 0.771, respectively. The difference is in line with the notion that gating explains part of the dispersion in observed data. Models 3 and 4 imply very similar values of 0.744 and 0.743, respectively. An intuitive explanation why these values are smaller than those for models 1 and 2 is that models 3 and 4 do not need additional unexplained variation to explain the differences between bracket responses and open-ended answers. Apparently, once acquiescence is allowed for (model 3), incorporating gating does not lead to further reduction of the estimate of the dispersion parameter.

6. Conclusions

We have investigated the importance of gating and acquiescence bias at the entry point of an unfolding bracket design. Experimental data on consumption where respondents are randomly either first given an open-ended question on household consumption or immediately directed to bracket questions were used to test whether these phenomena are present without making any model assumptions. The main finding is that acquiescence bias is the main problem. Once acquiescence is taken into account, incorporating anchoring in the form of Hurd et al.'s (1998) gating model or a fixed gating probability version of that, does not further change the conclusion about the underlying distribution of consumption. Ignoring acquiescence, however, leads to misleading conclusions on the selective nature of item non-response and on the inference on the distribution of consumption when selective item non-response is taken into account. It would lead to overestimation of consumption levels and underestimation of poverty rates. Not incorporating bracket information at all, however, would lead to biases in the same direction but of a much larger magnitude, illustrating how misleading the conclusions on the basis of a selection model based upon a loose identification strategy can be.

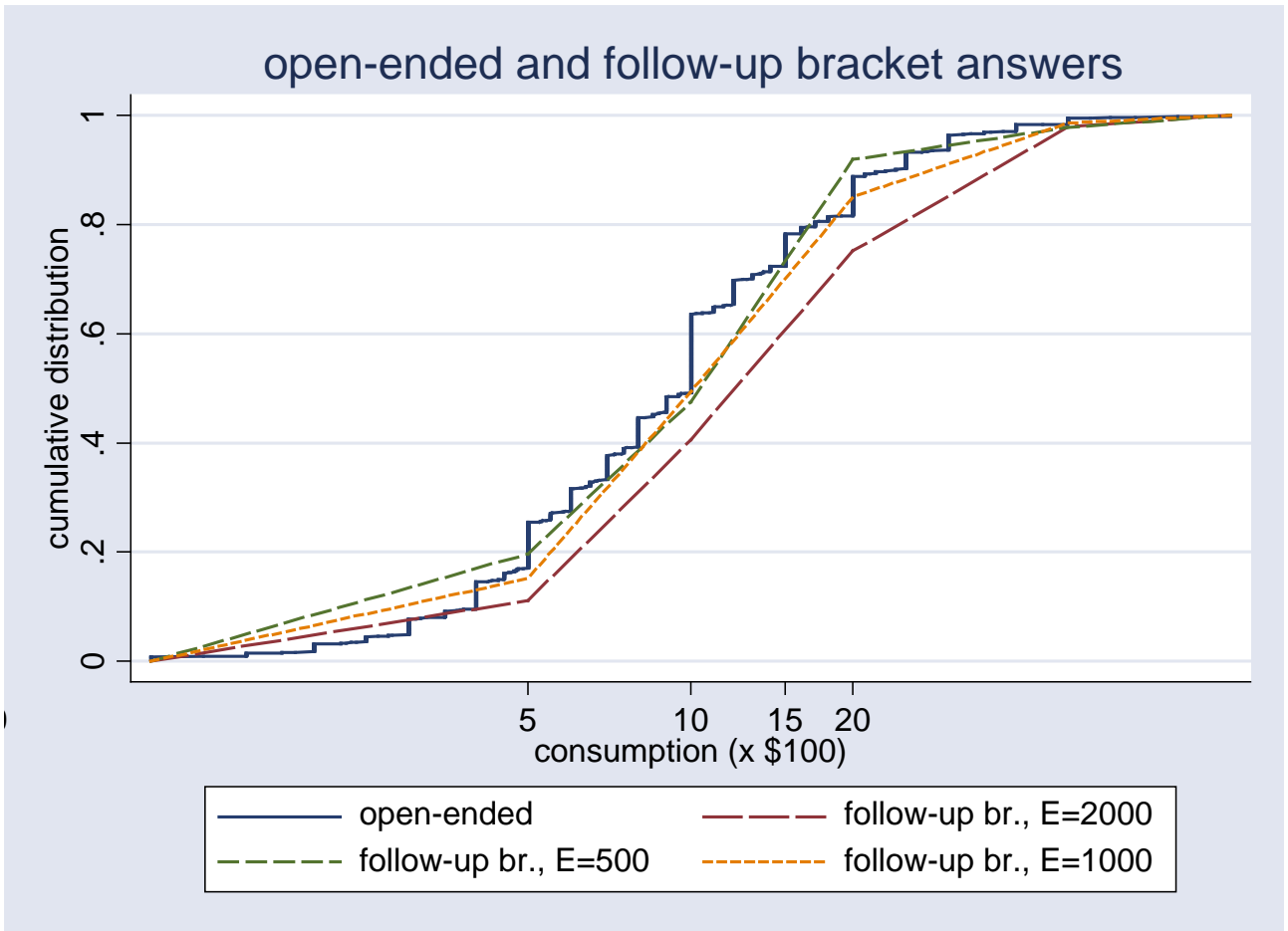
Fortunately, acquiescence bias can at least partially be avoided by adjusting the wording of the questions (cf. Hurd, 1999). The bracket questions can be formulated in a

neutral way, asking people to choose a category rather than answering “yes” or “no”. This has already been implemented in the most recent waves of, for example, the Health and Retirement Study and AHEAD. The results of this study suggest that, with these improved questions, it seems reasonably safe to use the first bracket question in an unfolding bracket design at face value. Given the evidence in the literature that anchoring is a problem also in absence of acquiescence bias, this at the same time means that anchoring at follow up bracket questions must be taken seriously. The fixed point model for anchoring – or another model with Bayesian learning – in which answers to the later bracket questions are affected by the entry point may then be a useful tool.

References

- Battistin, E., R. Miniaci and G. Weber (2000), What do we learn from recall consumption data? IFS discussion paper WP00/10, Institute of Fiscal Studies, London, *Journal of Human Resources*, forthcoming.
- Browning, M., T.F. Crossley and G. Weber (2002), Asking consumption questions in general purpose surveys, SEDAP Research Paper No. 77, McMaster University, Hamilton, *Economic Journal*, forthcoming.
- Boyle, K.J., H.F. MacDonald, H. Cheng and D.W. McCollum (1998), Bid design and yea saying in single-bounded, dichotomous-choice questions, *Land Economics*, 74(1), 49-61.
- Cameron, T. and J. Quiggin (1994), Estimation using contingent valuation data from a “dichotomous choice with follow-up” questionnaire, *Journal of Environmental Economics and Management*, 27, 218-234.
- Frykblom, P. and J. Shogren (2000), An experimental testing of anchoring effects in discrete choice questions, *Environmental and Resource Economics*, 16, 329-341.
- Green, D., K. Jacowitz, D. Kahneman and D. McFadden (1998), Referendum contingent valuation, anchoring, and willingness to pay for public goods, *Resource and Energy Economics* 20, 85-116.
- Hausman, J.A., J. Abrevaya and F.M. Scott-Morton (1998), Misclassification of the dependent variable in a discrete response setting, *Journal of Econometrics*, 87(2), 239-269.
- Heckman, J.J., (1979), Sample selection bias as a specification error, *Econometrica*, 47, 153-161.

- Herriges, J. and J. Shogren (1996), Starting point bias in dichotomous choice valuation with follow-up questioning, *Journal of Environmental Economics and Management*, 30, 112-131.
- Holmes, T. P. and R.A. Kramer (1995), An independent sample test of yea-saying and starting point bias in dichotomous-choice contingent valuation, *Journal of Environmental Economics and Management*, 29, 121-132.
- Hurd, M. (1999), Anchoring and acquiescence bias in measuring assets in household surveys, *Journal of Risk and Uncertainty*, 19(1-3), 111-136.
- Hurd, M., D. McFadden, H. Chand, L. Gan, A. Merrill, and M. Roberts (1998), Consumption and savings balances of the elderly: Experimental evidence on survey response bias, in D. Wise (ed.), *Frontiers in the Economics of Aging*, University of Chicago Press, Chicago, 353-387.
- Hurd, M., A. Kapteyn and J. Zissimopoulos (2001), Recovering the distribution from bracketed responses, paper presented at TMR Savings and Pensions workshop in Evian, France, August 21-23.
- Jacowitz, K. and D. Kahneman (1995), Measures of anchoring in estimation tasks, *Personality and Social Psychology Bulletin*, 21, 1161-1166.
- Juster, T., and J.P. Smith (1997), Improving the quality of Economic data: Lessons from the HRS and AHEAD, *Journal of the American Statistical Association*, 92, 1268-1278.
- Kanninen, B.J. (1995), Bias in discrete response contingent valuation, *Journal of Environmental Economics and Management*, 28, 114-125.
- Little, R.J.A., and D.B. Rubin, *Statistical analysis with missing data*, Wiley, New York.
- Manski, C.F., (1989), Anatomy of the selection problem, *Journal of Human resources*, 24, 343-360.
- O'Connor, R., M. Johansson and P.-O. Johansson (1999), Stated preferences, real behaviour and anchoring: some empirical evidence, *Environmental and Resource Economics*, 13, 235-248.
- Tversky, A. and D. Kahneman (1974), Judgment under uncertainty: heuristics and biases, *Science*, 185(4157), 1124-1131.
- Van Soest, A. and M. Hurd (2003), Anchoring and acquiescence bias in consumption data, mimeo, RAND Corporation, Santa Monica.



**Figure 1. Cumulative Distributions of Monthly Consumption (on logarithmic scale)
Open-ended responses and follow-up bracket responses by entry point; sample I**

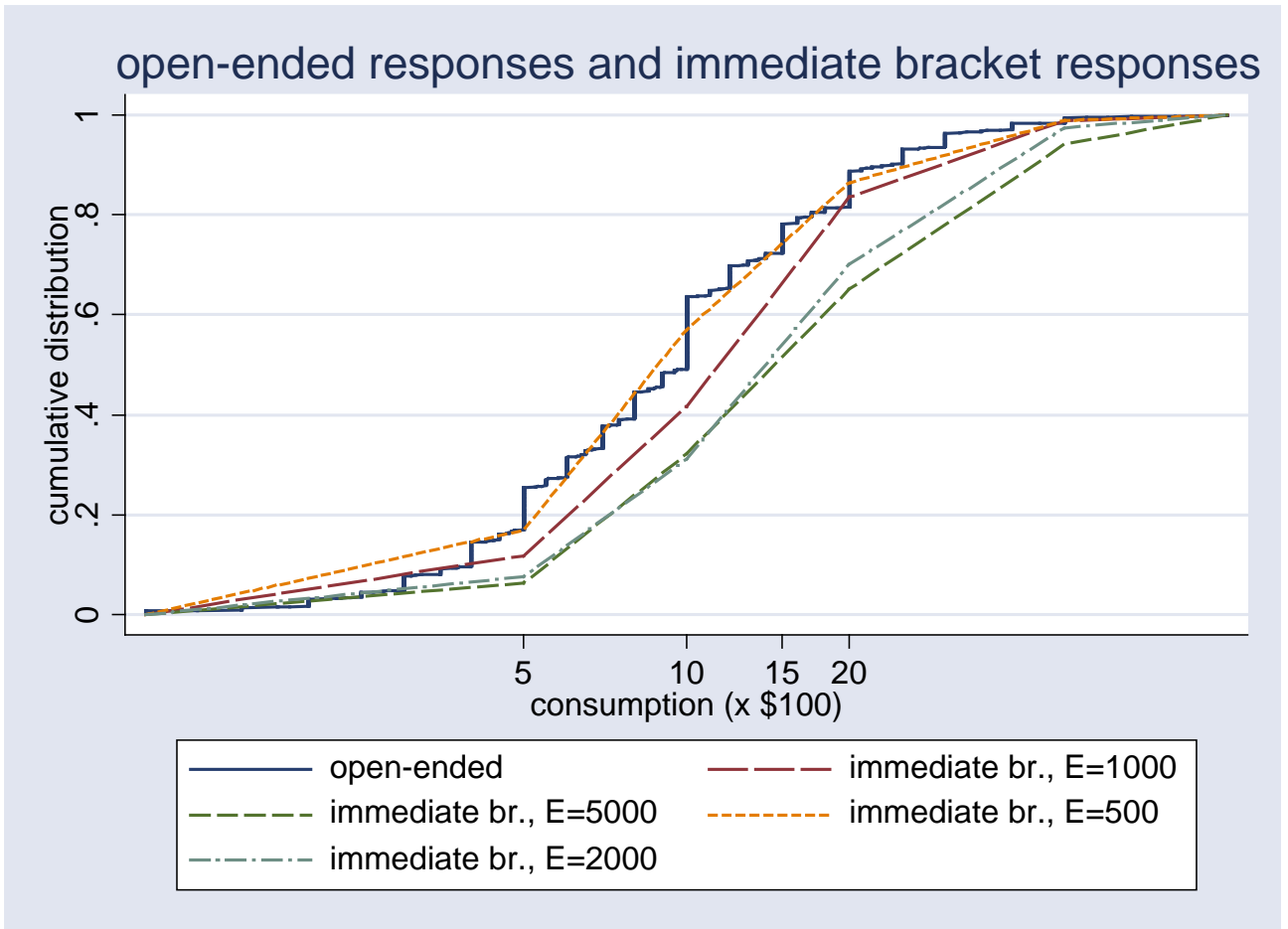


Figure 2. Cumulative Distributions of Monthly Consumption (on logarithmic scale) Open-ended responses (sample I) and immediate bracket responses by entry point (sample II)

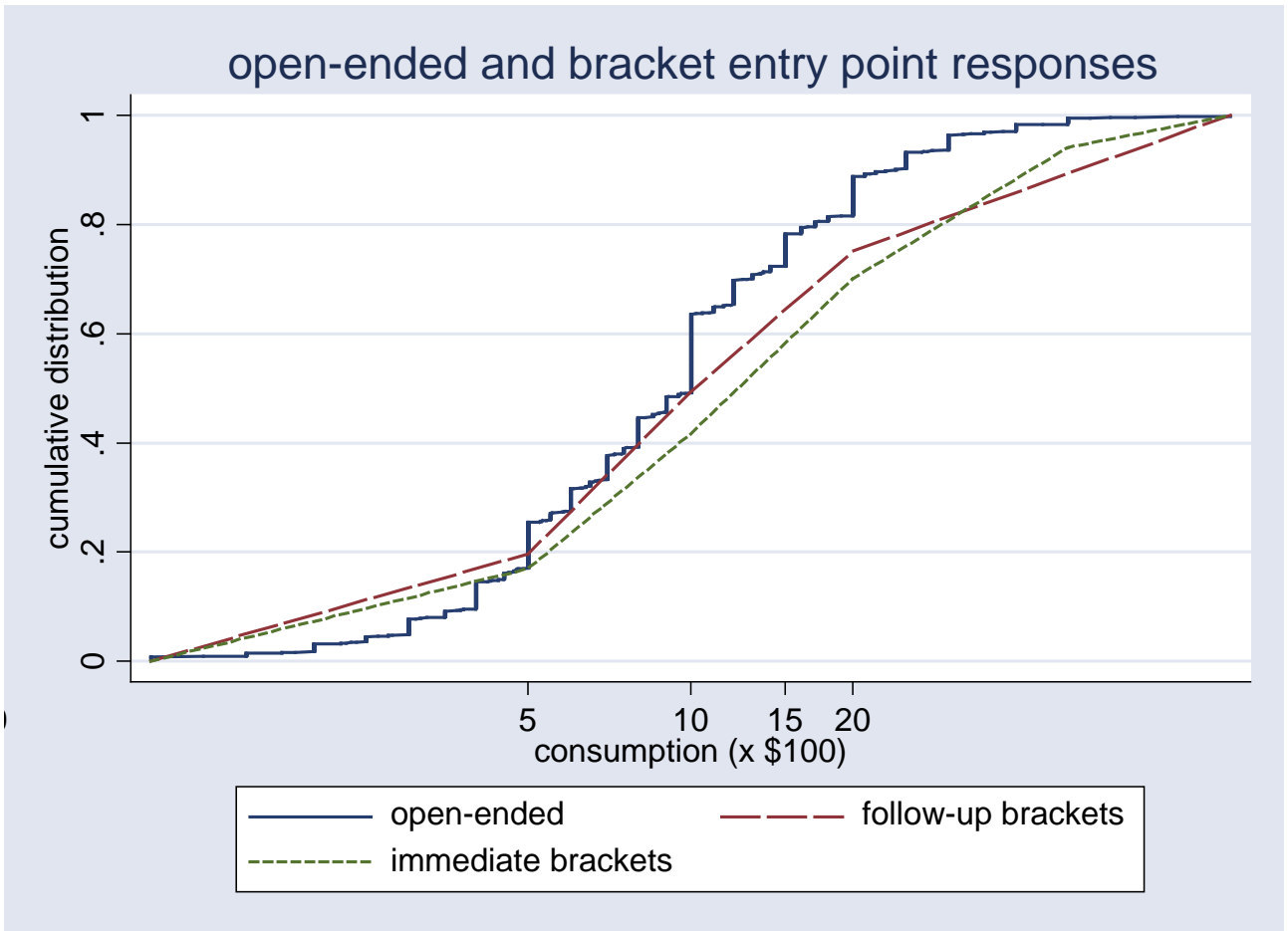


Figure 3. Cumulative Distributions of Monthly Consumption (on logarithmic scale)
Open-ended responses (sample I); Follow-up bracket responses using entry point question only (sample I);
Immediate bracket responses using entry point question only (sample II)

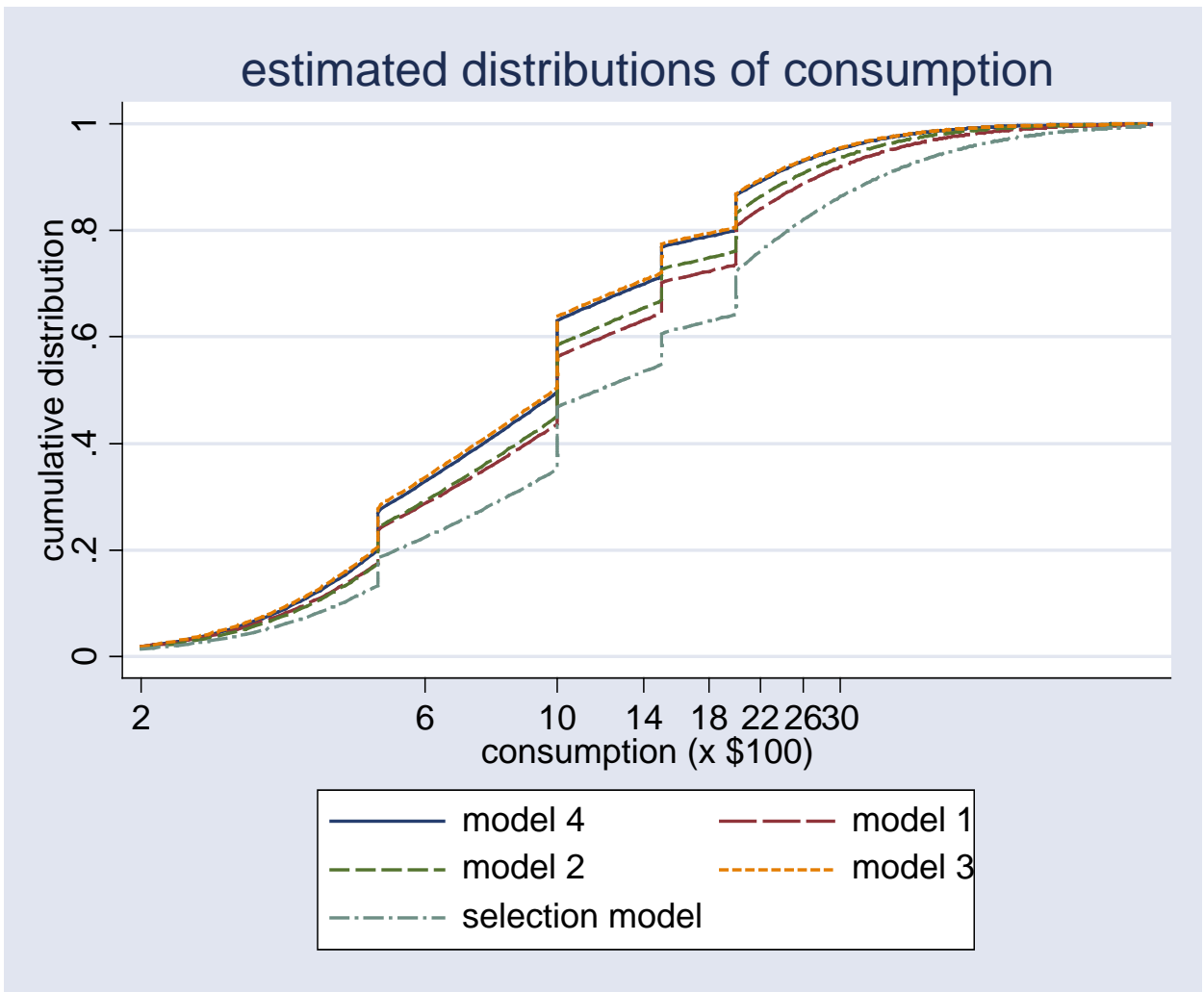


Figure 4. Aggregate Cumulative Distribution of Consumption According to Several Models

Appendix

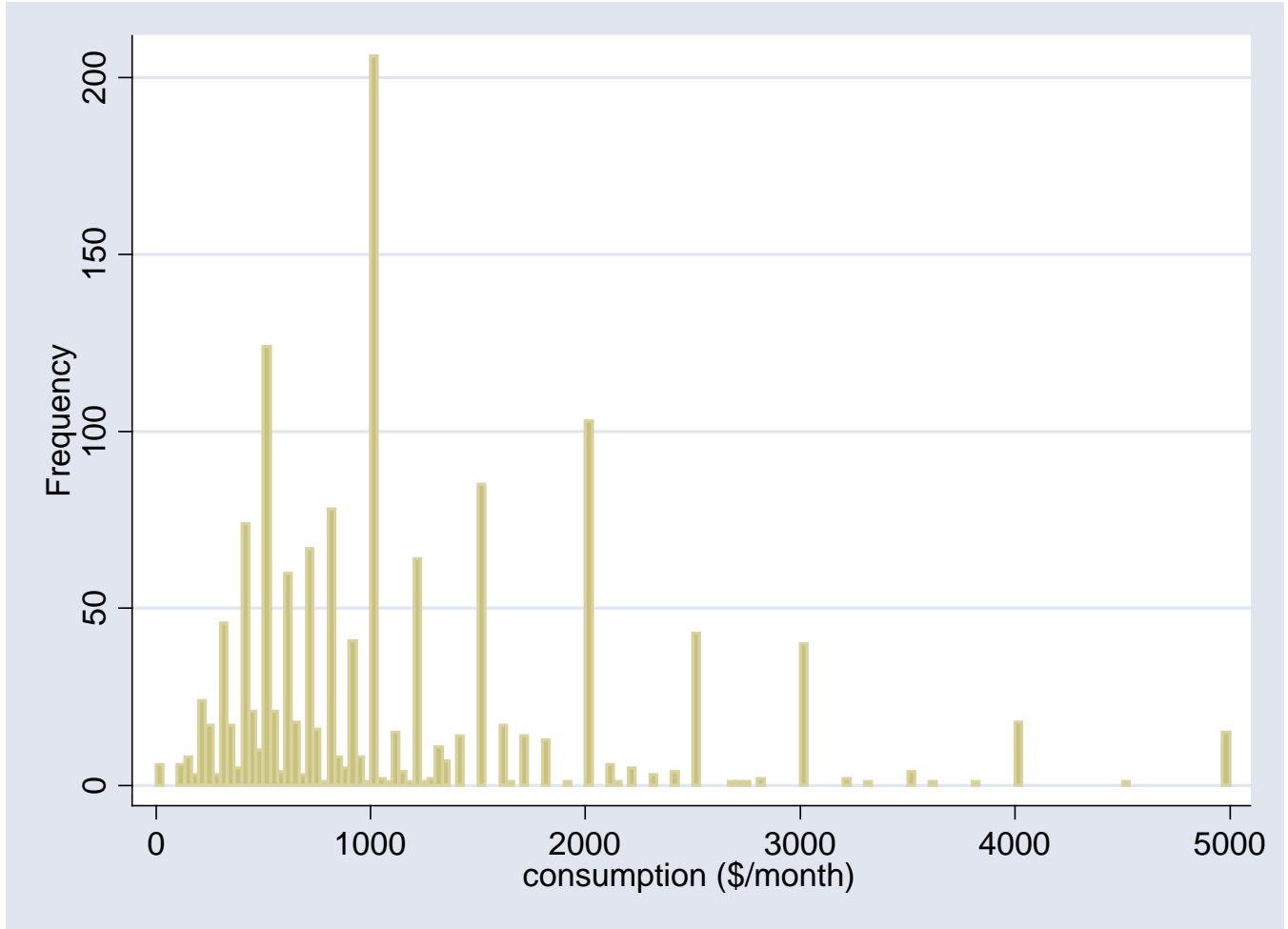


Figure A1. Histogram of Open-ended Answers

The figure shows the histogram of open-ended consumption question answers of at most \$5000 (1407 observations). The 0-5000 range is divided into 150 categories. The figure clearly illustrates the rounded nature of the answers, with many focal points. The most frequent answers are 1000 (205 observations), 500 (120 observations) 2000 (103 observations) and 1500 (85 observations).