Anchoring Bias in Recall Data: Evidence from Central America

Susan Godlonton, Manuel A. Hernandez, and Mike Murphy

Self-reported retrospective survey data is widely used in empirical work but may be subject to cognitive biases, even over relatively short recall periods. This paper examines the role of anchoring bias in self-reports of objective and subjective outcomes under recall. We use a unique panel-survey dataset of smallholder farmers from four countries in Central America collected over a period of three years. We exploit differences between recalled and concurrent responses to quantify the degree of mental anchoring in survey recall data. We assess whether respondents use their reported value for the most recent period as a cognitive heuristic when recalling the value from a previous period, while controlling for the value they reported earlier. The results show strong evidence of sizeable anchoring bias in self-reported retrospective indicators for both objective measures (income, wages, and working hours) and subjective measures (reports of happiness, health, stress, and well-being). We also generally observe a larger bias in response to negative changes for objective indicators and a larger bias in response to positive changes for subjective indicators.

Key words: Anchoring bias, recall data, self-reporting, smallholder farmers, Central America.

JEL codes: C8, O12, Q12.

Collecting accurate data, including retrospective data, is critical for effective economic policy research. However, the potential for recall error in household surveys has been recognized in the literature for several decades (e.g., Neter and Waksberg 1964; Evans and Leighton 1995; Sudman, Bradburn, and Schwarz 1996).[1] Retrospective biases have been shown to exist over a wide range of health, employment, income, and expenditure data obtained from population samples in developed countries. Researchers have also sought to explore measurement biases in low-income settings. For example, in India, Das, Hammer, and Sanchez-Paramo (2012) examine recall error in health outcomes, while de Nicola and Giné (2014) consider recall error among fisherman.[2] In this paper, we contribute to this growing body of literature and evaluate the role of anchoring as a particular source of measurement error by using panel survey data from Central America for a range of outcomes.

We focus on anchoring bias, in which individuals use some easily-observed prior or recently-provided information to guide them in estimating a value under uncertainty (see,

Susan Godlonton is an assistant professor at Williams College and an associated research fellow at the International Food Policy Research Institute (IFPRI). Manuel A. Hernandez is a research fellow and Mike Murphy is a research analyst, both of them at IFPRI. We thank the valuable comments of Eduardo Maruyama, Maria Recalde and seminar participants at the Annual Meeting of the Agricultural and Applied Economics Association and IFPRI. We would also like to thank Francisco Olivet, Luis Arteaga, and the team of Research & Planning for their help in collecting the data. Finally, we would like to thank Travis Lybbert and three anonymous referees for their many useful comments. We gratefully acknowledge financial support from the CGIAR Research Program on Policies, Institutions and Markets. Correspondence to be sent to: m.a.hernandez@cgiar.org.

[1] There is a rich body of interdisciplinary literature in the social sciences examining survey biases resulting from the interaction between respondents' cognitive processes and survey instruments (e.g., see Schwarz and Hippler 1995). Studied topics include optimal recall length (e.g., Clarke, Fiebig, and Gerdtham 2008); the role of telescoping (e.g. Neter and Waksberg 1964); the order of response alternatives (e.g., Krosnick and Alwin 1987); acquiescence bias (e.g., Krosnick 1991); and social desirability bias (see Finkel, Guterbock and Borg 1991). Krosnick (1999a, 1999b) and Sudman, Bradburn, and Schwarz (1996) provide excellent reviews of this extensive literature.

[2] The special edition "Symposium on Measurement and Survey Design" published in the *Journal of Development Economics* provides several other related examples (McKenzie and Rosenzweig 2012).

Amer. J. Agr. Econ. 100(2): 479–501; doi: 10.1093/ajae/aax080
Published online December 8, 2017

e.g., Tversky and Kahneman 1974). Anchoring bias has been well-documented in the economics literature. For example, it has been studied in the context of financial markets (e.g., Campbell and Sharpe 2009), auction prices (e.g., Beggs and Graddy 2009), and the elicitation of willingness-to-pay parameters (e.g., Ariely, Lowenstein, and Prelec 2003).[3] Furnham and Chu Boo (2003) provide a review of this anchoring literature.

Anchoring bias has also been addressed in the survey methodology literature. One strand examines the intersection between questionnaire format and survey biases. For example, Hurd et al. (1998) show that anchoring effects may account for significant differences in responses when using brackets versus open-ended questions on savings and consumption data; Frykblom and Shogren (2000) argue that anchoring effects in discrete-choice questions may result from survey framing rather than from the dichotomous choice per se. Hitczenko (2015) finds strong evidence of sequential anchoring in Likert scale questions, resulting in approximately 13% of responses incorrectly rated for selected questions.[4] We extend this body of literature by examining the role of anchoring bias in retrospective data for objective and subjective measures in a low-income setting. We focus on both objective and subjective indicators, as previous studies find that subjective indicators exhibit lower reliability ratios than objective indicators (e.g., Krueger and Schkade 2008).

We take advantage of a novel panel-survey dataset from IFPRI's Poverty-Sensitive Scorecard pilot program for smallholder farmers in Central America (see Hernandez and Torero 2014). This dataset spans four countries—El Salvador, Guatemala, Honduras, and Nicaragua— and includes three rounds of data collected over three years (2011, 2012, and 2013). The final survey further captured data for both the current and the previous years for a number of objective indicators (monthly total and per capita household income, wages from primary occupation, and hours worked), as well as for subjective measures (reports of happiness, health, stress, and overall well-being).[5] This feature of the survey instrument permits us to exploit potential differences between recalled and concurrent responses to examine the degree of mental anchoring in survey recall data.

For a given outcome, we assess whether respondents in the 2013 survey use the value they report for 2013 as an "anchor" to influence their subsequent recall of the value for 2012, while controlling for the value for 2012 that they reported in the 2012 survey. For clarity, we use the term "concurrent report" when the year of the survey and the reference period are the same, and "recall value" otherwise. Hence, the analysis considers the effect of a respondent's "2013 concurrent report" (the value they reported in 2013 for 2013, i.e., the anchor value) on their "2012 recall value" (the value they reported in 2013 for 2012), while controlling for their "2012 concurrent report" (the value they reported in 2012 for 2012). We account for individual, household, and interview characteristics.

For both objective and subjective indicators, we consistently find that the anchor value strongly predicts the recall value. In all cases, we conclude at conventional statistical levels that the recall measure is largely influenced by the anchor value. Our results provide evidence in favor of recall bias for both objective and subjective indicators. Furthermore, the high degree of mental anchoring observed does not seem to be driven by a specific group of respondents; we generally do not observe systematic differences by location, crop type, or individual characteristics.

Somewhat troublingly, we consistently find an extremely low correlation between recall values and measures reported in the previous period. This holds for both objective and subjective indicators. The 2013 concurrent report is a stronger (in magnitude and statistical precision) predictor for the 2012 recall value compared to the previous period report (2012 concurrent report). In many specifications, the predictive power of the previous period report is indistinguishable from zero, which casts doubts about the signaling value of the recall measure.

[3] The willingness-to-pay literature has led to an ongoing debate about whether preferences are in fact consistent and stable (e.g., Ariely, Lowenstein, and Prelec 2003; Fudenberg, Levine, and Maniadis 2012; Maniadis, Tufano, and List 2014; Alevy, Landry and List 2015). Rather than contributing to this debate, we examine anchoring bias as a source of measurement error.

[4] Other studies also find that respondents' mood (Bodenhausen, Gabriel, and Lineberger 2000), knowledge (Brekke et al. 1996), and cognitive ability (Bergman et al. 2010) can mitigate the magnitude of anchoring bias.

[5] While all measures are self-reported, the distinction between objective and subjective serves to differentiate between measures that can be directly observed by an individual versus measures that are not directly observed.

The estimation results are robust to several different specification checks. We extend the analysis to consider alternative anchors used by respondents, as well as mechanisms that may be driving the anchoring. In so doing, we examine whether positive or negative outcome changes over time play a role in anchoring. We find that positive changes are strongly associated with a positive bias, while negative changes are strongly associated with a negative bias. This pattern of responses further supports the interpretation of the observed measurement error as anchoring bias. We also find that the magnitude of the recall bias for objective indicators is generally larger in response to negative changes, while for subjective indicators, the magnitude of the bias is larger in response to positive changes.

The paper contributes to our understanding of behavioral biases in collecting self-reported data in developing countries. Two studies closely related to ours are de Nicola and Giné (2014) and Maruyama (2007), both of which compare retrospective data obtained from survey respondents with an alternative reference measure to estimate recall error. de Nicola and Giné (2014) use records of reported catches by fishing vessels as a reference measure to assess the reliability of data that fishermen self-report in coastal India. Maruyama (2007) compares a retrospective and a concurrent measure of wage earnings among workers in Indonesia using data from the Indonesian Family Life Survey. We follow a validation approach similar to these studies but focus on the role of anchoring bias in recall error. In our setting, the report value from the previous survey round serves as the reference value that the recall value should approximate. We further add to the literature by also considering subjective measures.

Our results also speak to the generalizability of recall bias as an important feature of retrospective data by extending the evidence base to include smallholder farmers in developing countries. Many original studies on recall bias were conducted in experimental settings on college campuses (Sears 1986) or in the context of large-scale developed country surveys (see Bound, Brown and Mathiowetz 2001; Crossley and Winter 2015). There remains debate, however, regarding whether these findings can be generalized. Recall biases measured primarily among college students may not necessarily extend to the general population, as college students may be easily influenced and readily change their attitudes (Henry 2008; Henrich, Heine, and Norenzayan 2010; Eastwick, Hunt, and Neff 2013). On the other hand, the cognitive process of answering aggregate questions is more demanding and may lead to an increased reliance on simple heuristics among respondents with low levels of education who are asked to recall a series of complex interrelated activities. For example, Arthi et al. (2016) document strong evidence of recall bias among smallholder farmers in Tanzania that is driven not only by failures in memory but also by the mental burdens of reporting on highly-variable agricultural work patterns. Beegle, Carletto, and Himelein (2012) and de Weerdt et al. (2016) also examine the difficulties of collecting agricultural and consumption data in several developing countries. Other empirical evidence consistently finds that respondents of lower socio-economic status (e.g., Gove and Geerken 1977; Das, Hammer, and Sanchez-Paramo 2012) and with lower levels of education (Narayan and Krosnick 1996) exhibit larger survey biases in general. Taken together, it is unclear whether we should anticipate smaller, larger, or similarly sized biases among rural smallholder farmers in developing countries, as these farmers typically have lower education levels and face more volatile economic conditions. Our paper attempts to fill this gap.

In addition, our results offer further support for negativity bias, consistent with a large body of literature in psychology that has established negativity bias across many domains, including disaster, perceptions, and disease (for extensive reviews, see Baumeister et al. 2001; Rozin and Royzman 2001). Our results extend this literature by finding negativity bias in anchoring for objective indicators but positivity bias in anchoring for subjective indicators. This finding has important implications when examining, for example, the evolution of aggregate income or well-being indicators over time and we rely on retrospective survey data.

Data

The data used for the analysis corresponds to a panel survey of households that participated in selected agricultural projects as part of IFPRI's Poverty-Sensitive Scorecard pilot

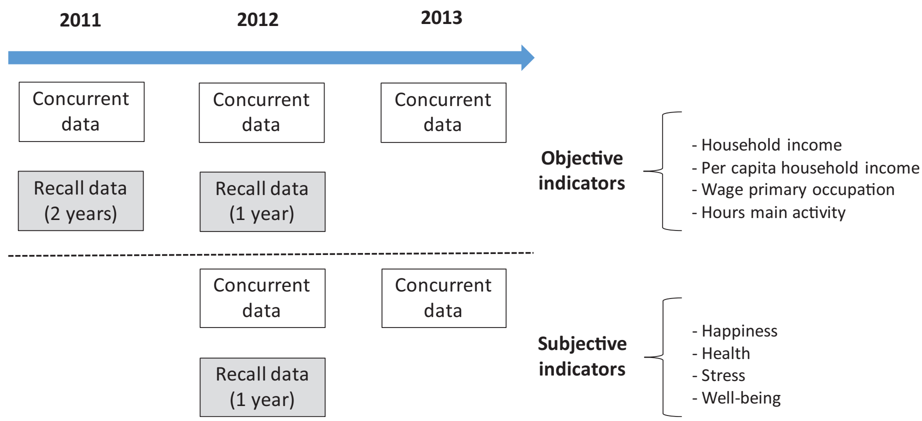

Figure 1. Data availability by reference period for objective and subjective indicators

program in Central America. The objective of the program was to select demand-driven projects using a scorecard tool that can better link poor smallholder farmers to dynamic markets through extension and training activities. The beneficiaries of each project belonged to an existing farmers' association, generally lived nearby, and were involved in the production of the same crop, which varied by project (including basic grains, fruits and vegetables, coffee, and cacao). Hernandez and Torero (2014) provide additional details about the pilot program and selected projects.

A total of three survey rounds were conducted—one baseline survey at the end of 2011 and two follow-up surveys at the end of 2012 and 2013—to evaluate the scorecard tool and the impact of the selected projects.[6] We use the balanced panel of 554 project beneficiaries who were found in each survey year. The sample is representative of the selected projects' beneficiaries located across 24 municipalities (distinct administrative areas) within four countries (El Salvador, Guatemala, Honduras, and Nicaragua; refer to Hernandez and Torero 2016 for further details of the survey sample and attrition over time).[7]

The data collected in the survey include respondents' socioeconomic characteristics, household composition, income and labor activities, household assets and expenditure, and well-being indicators, among others. The dataset also has a special feature that makes it particularly useful for testing anchoring bias in recall data; in the third (2013) survey round, households were asked a series of questions regarding, for example, their household income with different recall periods. That is, in 2013, respondents were asked for their household income in 2013; shortly afterwards, they were asked to recall their income for 2012 and 2011. This was part of the panel survey design in order to make some of the key indicators collected comparable between returning and new households that entered the survey in the last year.[8]

As shown in figure 1, we have repeated values for different objective and subjective outcomes, obtained using different lengths of recall. For the objective indicators (which include monthly total and per capita household income, wages from primary occupation, and working hours in main activity), we have information from each survey round on outcomes for that year (concurrent reports for 2011, 2012, and 2013), as well as information reported in 2013 on those same outcomes for 2011 and 2012 (two-year recall values for 2011 and one-year recall values for 2012). For the subjective indicators (which include measures of happiness, health, stress, and overall well-being on a ten-point scale), we have concurrent reports for 2012 and 2013

[6] The surveys were fielded around the same time of the year to avoid potential seasonal variations in the reported outcomes.

[7] The survey sample maintained the proportion of beneficiaries in each project by gender and location (municipality). There was 10% attrition after each survey year due to internal migration and other factors; however, there were not major differences, at least in observable characteristics, between attrition households and non-attrition households across years.

[8] New households entered the survey in the third year as replacements, and because the sample was expanded to households from nearby areas that did not necessarily participate in the program.

and recall values for 2012, as these measures were not collected in 2011.[9]

It is also important to note that the objective and subjective values used for the analysis result from interviewees' direct responses to a single question rather than from aggregated or constructed values. The only exception is income per capita, which is computed by dividing the response to the household income question by the number of household members reported for the corresponding reference period (i.e., for the 2012 recall measure, the 2012 household size is used). Online supplementary appendix table A.1 details the specific questions asked to obtain the objective and subjective values.

Table 1 presents basic descriptive statistics of the working sample. Panel A reports respondents' demographic and household characteristics in the last period. The sample is evenly split between men and women, and the average age is 43 years. In general, respondents have not completed primary schooling (5.3 years of education, on average). Average household size is close to six members, and average plot size is 2.6 hectares, reflecting the fact that the program targeted poor farmers.

Panel B reports some basic demographic information about the interviewers who collected data from the respondents. While the three survey rounds were conducted by the same firm and under the same supervisors in each of the countries, there was some change in the pool of interviewers across years. In addition, the same subjects were not always assigned to the same interviewer. In the first two rounds of data collection, about 40% of the interviewers were male; in the third round, 55% were male. Similarly, the average age of the interviewers decreased from approximately 35 to 33 years old between the first and the subsequent two survey rounds. We account for these enumerator characteristics in the regression analysis.

Finally, panel C presents summary statistics of the objective and subjective outcomes of interest.[10] Two patterns are worth noting regarding the objective indicators. First, there is a declining trend in the average concurrent measures reported in each survey between 2011 and 2013. For total and per capita monthly household income, the decline is monotonic over time (from \$421 and \$84, respectively, in 2011 to \$214 and \$45 in 2013); primary wages and working hours also show a decrease (from \$207 to \$161 and 39.4 to 38 hours, respectively).[11] Second, and central to this study, the recall values for 2011 and 2012 (captured in 2013) are strikingly similar to the corresponding 2013 concurrent reports both in terms of their mean and in their dispersion (standard deviation). For example, the 2011 and 2012 recall values for per capita income are on average \$46 and \$47, respectively, versus \$45 for the 2013 concurrent report; their standard deviations are \$59, \$66, and \$59. These patterns provide preliminary support of significant anchoring bias in self-reported data measured in the 2013 survey round.

All four subjective indicators are, in turn, elicited using a ten-point scale, on which higher numbers indicate greater magnitude. Higher scores indicate that respondents are better off in terms of happiness, health, and well-being, and the opposite for stress. We observe that respondents are marginally better off when comparing the 2012 concurrent report to the 2013 concurrent report. In most cases, the average 2012 recall values are also closer to the respondent's 2013 concurrent report than to their 2012 concurrent report, which is also indicative of anchoring bias. For instance, the 2012 recall level of happiness is 8.19 versus 8.1 for the 2013 concurrent report and 8.01 for the 2012 concurrent report.

The potential anchoring bias in self-reported objective and subjective data is clearer in table 2. Columns 1 and 2 correspond to the difference in absolute value between the anchor value and the 2012 recall value (2013 concurrent report—2012 recall

---

[9] Focusing on these indicators also helps us to reduce potential framing effects that could influence our results, as these effects are more common among binary survey questions (see, e.g., Goldin and Reck 2015).

[10] Respondents in Guatemala, Honduras, and Nicaragua reported monetary variables in local currency units. For ease of interpretation, we convert these to U.S. dollars using the annual exchange rate published by the local central banks for the corresponding reference period; hence, for a value recalled for 2012, the 2012 exchange rate is used. The estimation results are not sensitive to using the monetary variables in local currency units.

[11] We do not believe that the beneficiaries of the program had an incentive to underreport their income or wages over time, given the design of the program. It was clear from the beginning of the program that the allocated grants were only (one-time) start-up grants for selected projects intended to be sustainable in the medium and long term. The average decline in income could be explained by poor weather (rainfall) conditions in all the municipalities in our study sample between 2011 and 2013. Still, we cannot fully rule out the possibility of underreporting, given that the World Bank Development Indicators in the region show that the national gross domestic product, agricultural value, and crop production index increased in these countries during the study period.

Table 1. Summary Statistics

Variable	Mean	SD	Min.	Max.
Panel A: Individual characteristics (2013)				
Age	43.016	12.661	17	78
If male	0.504	0.500	0	1
Years of education	5.314	3.399	0	17
Household Size	5.745	2.478	1	15
Size of main plot (hectares)	2.616	5.659	0.005	69.888
Proportion missing size of main plot	0.137	0.344	0	1
Panel B: Interview characteristics				
If male interviewer (2011)	0.424	0.495	0	1
If male interviewer (2012)	0.404	0.491	0	1
If male interviewer (2013)	0.549	0.498	0	1
Interviewer age (2011)	34.818	7.266	21	45
Interviewer age (2012)	32.812	7.312	21	46
Interviewer age (2013)	32.948	10.168	20	51
Interview length in minutes (2011)	30.120	13.675	10.633	108.483
Proportion paused (2011)	0.031	0.173	0	1
Interview length in minutes (2012)	38.860	15.190	10.383	117.683
Proportion paused (2012)	0.038	0.191	0	1
Interview length in minutes (2013)	42.667	17.925	11.900	117.967
Proportion paused (2013)	0.049	0.216	0	1
Panel C: Outcomes of interest				
Objective Indicators:				
Income (2011 Concurrent Report)	420.789	746.438	0.313	6515.295
Income (2012 Concurrent Report)	322.530	677.670	0.127	8935.777
Income (2013 Concurrent Report)	213.803	262.150	0.048	3151.955
Income (2011 Recall Value)	232.630	309.122	0.052	3902.599
Income (2012 Recall Value)	234.435	311.808	0.050	3475.243
Income per capita (2011 Concurrent Report)	83.922	160.257	0.078	1628.824
Income per capita (2012 Concurrent Report)	67.118	182.318	0.025	2978.593
Income per capita (2013 Concurrent Report)	44.870	58.926	0.006	488.553
Income per capita (2011 Recall Value)	46.050	59.305	0.007	650.433
Income per capita (2012 Recall Value)	47.409	65.561	0.012	695.049
Primary Wage (2011 Concurrent Report)	206.529	364.018	0.000	3278.184
Primary Wage (2012 Concurrent Report)	155.363	282.484	0.298	3049.053
Primary Wage (2013 Concurrent Report)	161.379	496.677	0.048	8663.994
Primary Wage (2011 Recall Value)	167.772	463.932	0.000	9366.238
Primary Wage (2012 Recall Value)	160.618	441.030	0.050	8936.340
Hours in main activity (2011 Concurrent Report)	39.363	14.619	3	96
Hours in main activity (2012 Concurrent Report)	37.397	15.985	3	96
Hours in main activity (2013 Concurrent Report)	38.032	16.113	0	144
Hours in main activity (2011 Recall Value)	37.818	16.687	1	144
Hours in main activity (2012 Recall Value)	37.426	16.063	1	144
Subjective Indicators:				
Happiness (2012 Concurrent Report)	8.011	1.835	1	10
Happiness (2013 Concurrent Report)	8.101	1.869	1	10
Happiness (2012 Recall Value)	8.195	1.750	1	10
Health (2012 Concurrent Report)	7.724	1.929	1	10
Health (2013 Concurrent Report)	7.749	2.056	1	10
Health (2012 Recall Value)	7.850	1.855	1	10
Stress (2012 Concurrent Report)	5.581	2.642	1	10
Stress (2013 Concurrent Report)	5.446	2.558	1	10
Stress (2012 Recall Value)	5.542	2.460	1	10
Well–being (2012 Concurrent Report)	7.623	1.976	1	10
Well–being (2013 Concurrent Report)	7.731	2.029	1	10
Well–being (2012 Recall Value)	7.825	1.814	1	10
Observations				554

Note: Years of education reported in 2011. Proportion paused indicates interviews that were paused and recommenced. Income and wage variables for Guatemala, Honduras, and Nicaragua were converted to U.S. dollars. Subjective indicators are measured on a 1–10 scale; SD=Standard Deviation.

Table 2. Differences in Absolute Value between Concurrent Report and Recall Measures

Outcome	2013 Concurrent Report – 2012 Recall Value		2012 Concurrent Report – 2012 Recall Value	
	Mean (1)	SD (2)	Mean (3)	SD (4)
Income	79.620	230.626	264.767	659.468
Per capita income	19.700	48.824	55.246	172.579
Primary wage	64.703	280.771	168.051	495.198
Hours in main activity	4.220	11.014	14.170	14.381
Happiness	0.870	1.259	1.758	1.583
Health	0.989	1.391	1.711	1.637
Stress	1.222	1.534	2.671	1.984
Well–being	0.935	1.204	1.783	1.620

Note: The 2013 (2012) Concurrent Report refers to the value for 2013 (2012) reported in 2013 (2012); 2012 Recall Value refers to the value for 2012 reported in 2013. Income and wage variables for Guatemala, Honduras, and Nicaragua were converted to U.S. dollars. Subjective indicators are measured on a 1–10 scale.

value), while columns 3 and 4 pertain to the difference in absolute value between the concurrent report in the 2012 survey round and the 2012 recall value (2012 concurrent report—2012 recall value). The difference between the 2013 concurrent report and 2012 recall value is considerably smaller than the difference between the 2012 concurrent report and 2012 recall value for all objective and subjective indicators. For the objective indicators, the 2012 recall value is 2.6–3.4 times farther from the 2012 concurrent report than from the 2013 concurrent report; for the subjective indicators, it is 1.7–2.2 times farther. Note also the lower level of dispersion in the difference between the 2013 concurrent report and the 2012 recall value than the 2012 concurrent report and the 2012 recall value.

Overall, these summary statistics provide preliminary evidence of substantial anchoring bias in self-reported data across most objective and subjective indicators. Although not presented here, we further observe a large variation at the individual level in the reported concurrent outcomes across years.[12] This large variation emphasizes the need to examine anchoring since anchoring bias will not necessarily introduce meaningful inaccuracies into recalled responses in situations where there is limited year-on-year variation.[13] We turn next to outlining our empirical approach.

[12] Individual variations in income, wages, and working hours' concurrent reports ranged from -99% to more than 900% between 2012 and 2013; in the case of subjective indicators, these ranged from -90% to more than 700%.

[13] We thank an anonymous referee for noting this point.

Empirical Approach

The structure of the survey permits us to compare objective and subjective concurrent reports from each survey year to recall values reported in the final survey year. If recalling information for an earlier period is challenging, the respondent may use her concurrent report for the most recent period as a cognitive heuristic to assist her in the reporting of the subsequent recall value. This mental anchoring may result, for example, in her providing a biased estimate of her income in the earlier period since it has been influenced by her report for the more recent period. The descriptive statistics discussed above suggest that anchoring bias may play an important role in this context.

Formally, we can conceptualize this bias in terms of a simple linear regression framework. Consider $\hat{\boldsymbol{x}}$ to be the recalled value of a given variable $\boldsymbol{x}$, which is reported with measurement error ε,

$$(1) \quad \hat{\boldsymbol{x}} = \boldsymbol{x} + \varepsilon.$$

Following Campbell and Sharpe (2009), we assume an anchoring variable $\boldsymbol{a}$ which is the reported value from a previous question, which influences the process by which the interviewee generates her report of the recalled value $\hat{\boldsymbol{x}}$,

$$(2) \quad \hat{\boldsymbol{x}} = \boldsymbol{p}\boldsymbol{a} + \boldsymbol{\lambda}\boldsymbol{x} + \varepsilon$$

where $\boldsymbol{p}$ and $\boldsymbol{\lambda}$ are the weights that the respondent assigns to the anchoring variable and the true latent variable $\boldsymbol{x}$. In the absence of additional factors affecting recall, we can let $\lambda = 1 - p$; if $p = 1$, the respondent's estimate is solely based on the anchoring value and if $p = 0$, anchoring bias plays no role.

In practice, we can test for anchoring bias in our sample by estimating the following model,

$$(3) \quad y_{ij}^{t-1,t} = \alpha + \beta y_{ij}^{t,t} + \gamma y_{ij}^{t-1,t-1} + \theta Z_{ij} + \kappa_j + \varepsilon_{ij}$$

where $y_{ij}^{t-1,t}$ is the recall value for the variable of interest (e.g., household income), which individual i from project j provides in reference to the prior period t-1 in the current period t (i.e., the income for 2012 reported in 2013); $y_{ij}^{t,t}$ is the concurrent report of the variable for the reference period t reported at t (i.e., income for 2013 reported in 2013); $y_{ij}^{t-1,t-1}$ is the concurrent report of the variable for the prior period t-1 reported at t-1 (i.e., income for 2012 reported in 2012); Z_{ij} is a vector of controls that includes individual, household, and interview characteristics; κ_j is the project-level fixed effect; and ε_{ij} is a white noise error.[14] Hence, the 2013 concurrent report $y_{ij}^{t,t}$ serves as the potential anchor value.

Note that in this setting, the 2012 concurrent report $y_{ij}^{t-1,t-1}$does not necessarily represent the true or actual value in prior period t-1, as there may be other potential sources of error affecting the respondent's report. However, it is still valid to use this measure as the reference value to the extent that this source of error is independent of the anchoring bias.

We estimate the same model for both objective and subjective indicators. The parameters of interest in equation (3) are β and γ. The first parameter measures the partial correlation between the anchor value and the recall value, while the second parameter captures the correlation between the value reported in the previous period and the recall value. If $\beta > \gamma$, then anchoring bias plays a major role when recalling information. The opposite is true if $\beta < \gamma$. For the objective indicators, we further exploit an additional round of data, which permits us to expand the analysis to include the respondents' recall and report values in reference to 2011.

The estimation of equation (3) poses some issues worth discussing. First, there might be certain correlation across time in the reported values, particularly for the objective outcomes, such that the anchoring parameter β could also reflect this autocorrelation. However, the potential correlation across time between the anchor value reported in 2013 ($y_{ij}^{t,t}$) and the recall value for 2012 ($y_{ij}^{t-1,t}$) should be captured by the 2012 concurrent report ($y_{ij}^{t-1,t-1}$), so the anchoring parameter is expected to be net of this correlation (if any).

Second, in the case of the subjective indicators, the elicitation method using a ten-point scale may introduce bias. This bias arises particularly for individuals reporting values in the lower and upper bounds, as their underlying latent values may lie below or above the possible range. We demonstrate that our results are robust to the exclusion of such individuals.

We also extend our base framework to examine whether the direction and magnitude of the bias differs in response to positive or negative changes in the outcomes of interest across time. To do so, we estimate the following regression:

$$(4) \quad \Delta y_{ij} = \alpha + \beta_1 Pos_{ij} + \beta_2 Neg_{ij} + \theta Z_{ij} + \kappa_j + \varepsilon_{ij}$$

where Δy_{ij} is the difference between the 2012 recall value and the 2012 concurrent report, Pos_{ij} is the value of the difference between the 2013 and 2012 concurrent reports if that difference is positive, and zero otherwise, and Neg_{ij} takes the absolute value of the difference if the difference is negative, and zero otherwise.

To be consistent with the hypothesis that anchoring bias drives the measurement error observed, we would expect $\beta_1 > 0$ and $\beta_2 < 0$. In other words, if a household exhibits positive outcome changes between their 2012 and 2013 concurrent reports, we anticipate that the recall bias is positive; we expect the opposite if the household exhibits a negative outcome change. As an additional sensitivity check, we also control for the modal pattern of these differences at the municipality level. Specifically, we control for positive and negative municipality indicators; the

[14] Interview characteristics include two set of variables for the age and gender of the enumerator for the 2012 and 2013 surveys (as households may not have been interviewed by the same person in each round), as well as for the duration of each of the interviews. Similarly, the project-level fixed effects are to account for one of the eight different agricultural projects to which the survey respondents belonged. There were four projects in Honduras, two in Nicaragua, one in El Salvador, and one in Guatemala.

former takes the value of one when the majority of the people in the municipality (excluding the respondent) report a positive change for the given indicator from their 2012 to their 2013 concurrent report, and zero otherwise. The converse is true for the negative dummy.

Results and Discussion

This section presents the estimation results and evaluates the robustness of our findings. We first discuss the base results of the estimated model specified in equation (3). We then evaluate the existence and magnitude of anchoring bias, and assess the validity of the results using alternative model specifications and samples. Finally, we explore alternative anchors and present suggestive evidence of the underlying mechanisms driving the sizeable anchoring bias observed.

Base Results: Anchoring Bias among Smallholder Farmers

Table 3 presents the estimation results for the objective outcomes. Columns (1), (4), (7), and (10) include project-level fixed effects. Columns (2), (5), (8), and (11) add controls for the respondents' demographic and household characteristics, including their age, gender, education (years of schooling), household size, and plot size.[15] Columns (3), (6), (9), and (12) add an additional set of variables for the interview characteristics. These include controls for the ages and genders of the interviewers for the 2012 and 2013 surveys, as well as the duration of each of the interviews.[16] The p-values of the coefficients, reported in parentheses, are clustered at the project level. These values are derived following Cameron, Gelbach, and Miller's (2008) wild-bootstrap procedure to better account for the precision of the standard errors in the presence of only eight clusters.[17]

As shown in table 3, the coefficient of the anchor variable is positive and statistically significant at conventional levels across specifications and for all outcomes. The point estimate of the anchoring coefficient for the full model (with all controls) ranges from 0.66 in the case of working hours to 0.75 in the case of monthly income. Thus, a one-hour increase in the 2013 concurrent report for weekly working hours is associated with a 0.66-hour increase in the working hours recalled for 2012, while a ten-dollar increase in the 2013 concurrent report for monthly income is associated with a \$7.5 increase in the recalled income for 2012. These results provide strong evidence that respondents use the reported value for the current period (2013 outcome reported in 2013) to frame their recall responses in reference to the preceding year (2012 outcome reported in 2013), despite answering a range of other questions between these two responses.

In contrast, the coefficient of the 2012 concurrent report (2012 outcome reported in 2012) is statistically indistinguishable from zero. This result is of concern for studies utilizing retrospective data as it implies that there is very limited signaling value in the recall measure, even over relatively short recall periods. Based on the one-sided Wald tests (p-values) reported at the bottom of the table, we reject the null hypothesis that the 2012 concurrent report coefficient is greater than or equal to the 2013 anchoring coefficient) at conventional levels of significance for all indicators. The ratio of the anchoring coefficient relative to the 2012 concurrent coefficient $\left(\frac{\beta}{\delta}\right)$ is consistently large across indicators. The ratio is typically in the mid-30s and exceeds 100 for income. Despite variation in the difficulty of recalling outcomes (e.g., it might be easier for a respondent to recall her primary wage than her total household income), we observe a significant degree of mental anchoring for all four objective outcomes, and basically a null correlation between the 2012 recall and concurrent values.[18]

Table 4 presents the estimation results for the subjective indicators using the same set of model specifications as those in table 3.

---

[15] We also include an indicator which takes the value one when the size of the main plot is missing (for these cases, we assign the median plot size), and zero otherwise.

[16] For cases in which the interview was paused and re-commenced, we assign the median duration value and include an additional dummy variable that takes the value of one for paused interviews, and zero otherwise.

[17] The statistical significance of the estimated coefficients is not generally sensitive to alternatively deriving the p-values using the six-point weight distribution described by Webb (2014), nor to clustering the standard errors at the municipality level or only allowing for robust standard errors.

[18] The similar findings for own wages and household income could suggest that proxy reporting is not necessarily driving the results. However, formally evaluating proxy reporting is beyond the scope of this study due to a lack of more detailed data.

Table 3. Estimation Results for Objective Outcomes

Coefficient	Dependent variable: 2012 Recall Value											
	Income			Income per capita			Primary Wage			Hours main activity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Anchor Value (2013 Concurrent Report)	0.769***	0.764***	0.750***	0.698***	0.693***	0.684***	0.722**	0.720***	0.721**	0.652***	0.663***	0.660***
	(0.004)	(0.000)	(0.000)	(0.004)	(0.004)	(0.000)	(0.012)	(0.004)	(0.016)	(0.000)	(0.000)	(0.008)
2012 Concurrent Report	0.007	0.003	0.003	0.024	0.020	0.019	–0.071	–0.071	–0.068	0.014	0.018	0.015
	(0.840)	(0.984)	(0.828)	(0.520)	(0.441)	(0.391)	(0.105)	(0.129)	(0.289)	(0.555)	(0.383)	(0.555)
Project Fixed Effects	+	+	+	+	+	+	+	+	+	+	+	+
Individual Characteristics	–	+	+	–	+	+	–	+	+	–	+	+
Interview Characteristics	–	–	+	–	–	+	–	–	+	–	–	+
	Standardized coefficients											
Anchor Value (2013 Concurrent Report)	0.646	0.642	0.630	0.628	0.623	0.615	0.813	0.811	0.812	0.654	0.665	0.662
2012 Concurrent Report	0.014	0.006	0.006	0.066	0.055	0.054	–0.046	–0.045	–0.044	0.014	0.018	0.015
p–value (H_0: Concurrent $\geq$ Anchor)	0.001	0.001	0.001	0.000	0.000	0.000	0.007	0.007	0.007	0.000	0.000	0.000
Observations	554	554	554	554	554	554	554	554	554	554	554	554

Note: Dependent variable is the recall value for 2012 reported in 2013. Anchor value (2013 Concurrent Report) refers to the value reported in 2013 for the outcome in 2013; 2012 Concurrent Report refers to the value reported in 2012 for the outcome in 2012. Individual characteristics include respondent age, gender, years of education, household size, and size of primary agricultural plot. Interview characteristics include the age and gender of the enumerators in both years and the duration of all survey rounds; p–values are estimated clustering over eight projects using the wild–bootstrap cluster procedure described by Cameron, Gelbach and Miller (2008) with bootstrap samples generated using a Rademacher weight distribution with 256 replications. The standardized coefficients result from normalizing the estimated coefficients using the corresponding standard deviations of the variables reported in table 1. The p–value of one–sided Wald test for 2012 Concurrent $\geq$ 2013 Anchor is reported at the bottom using wild–bootstrap. Asterisks *,**, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4. Estimation Results for Subjective Outcomes

Coefficient	Dependent variable: 2012 Recall Value											
	Happy			Healthy			Stress			Well–being		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Anchor Value (2013 Concurrent Report)	0.572***	0.572***	0.581***	0.527***	0.516***	0.518***	0.657***	0.658***	0.646***	0.575***	0.573***	0.577***
	(0.000)	(0.008)	(0.004)	(0.004)	(0.004)	(0.008)	(0.008)	(0.004)	(0.000)	(0.000)	(0.004)	(0.000)
2012 Concurrent Report	0.032	0.034	0.037	0.048	0.036	0.034	0.043	0.041	0.051	0.055***	0.057***	0.057**
	(0.285)	(0.312)	(0.324)	(0.316)	(0.348)	(0.305)	(0.320)	(0.355)	(0.297)	(0.008)	(0.004)	(0.023)
Project Fixed Effects	+	+	+	+	+	+	+	+	+	+	+	+
Individual Characteristics	–	+	+	–	+	+	–	+	+	–	+	+
Interview Characteristics	–	–	+	–	–	+	–	–	+	–	–	+
	Standardized coefficients											
Anchor Value (2013 Concurrent Report)	0.611	0.612	0.621	0.585	0.572	0.574	0.684	0.685	0.671	0.643	0.641	0.646
2012 Concurrent Report	0.033	0.036	0.038	0.050	0.037	0.035	0.047	0.044	0.054	0.060	0.062	0.062
p–value (H_0: Concurrent $\geq$ Anchor)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Observations	554	554	554	554	554	554	554	554	554	554	554	554

Note: Dependent variable is the recall value for 2012 reported in 2013. Anchor value (2013 Concurrent Report) refers to the value reported in 2013 for the outcome in 2013; 2012 Concurrent Report refers to the value reported in 2012 for the outcome in 2012. Individual characteristics include respondent age, gender, years of education, household size and size of primary agricultural plot. Interview characteristics include the age and gender of the enumerators in both years and the duration of all survey rounds; p–values are estimated clustering over eight projects using the wild–bootstrap cluster procedure described by Cameron, Gelbach and Miller (2008) with bootstrap samples generated using a Rademacher weight distribution with 256 replications. The standardized coefficients result from normalizing the estimated coefficients using the corresponding standard deviations of the variables reported in table 1; p–value of one–sided Wald test for 2012 Concurrent $\geq$ 2013 Anchor reported at the bottom using wild–bootstrap. Asterisks *,**, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Similar to our findings for the objective measures, the anchor value has a strong and statistically significant relationship with the recalled measures. The estimated anchoring coefficient for the full model ranges from 0.52 for the health indicator to 0.65 for the stress indicator. The coefficients of the 2012 concurrent reports, in turn, are generally close to zero and are not significant. Only for the well-being indicator is the 2012 concurrent report partially correlated with the 2012 recall value, although this partial correlation is less than one-tenth of the correlation between the anchor and recall measure.[19] As before, the recall value of subjective indicators indicates very limited signaling value. The $\left(\frac{\beta}{\delta}\right)$ ratio is also high in this case (typically in the mid-teens), albeit lower relative to the objective indicators.

For comparison purposes between objective and subjective indicators, we also report the corresponding standardized coefficients for the 2013 concurrent report (anchor value) and the 2012 concurrent report in the bottom panel of tables 3 and 4. These parameters can be interpreted as the change (in standard deviations) in the 2012 recall value associated with a one standard deviation change in the anchor value and the 2012 concurrent report. The standardized anchoring coefficients of the full model for the objective indicators range between 0.62 (per capita income) and 0.81 (primary wages); for the subjective indicators, they range between 0.57 (health) and 0.67 (stress). Hence, the bias is to some extent stronger for objective than for subjective outcomes. This result is not necessarily surprising. On the one hand, there is a natural reference dependence when assessing, for instance, your life satisfaction or well-being across different periods (as opposed to reporting your wage or working hours), and this may result in anchoring. On the other hand, there is more variation in objective than in subjective measures, which could make the former more difficult for respondents to remember and thus make them more prone to use heuristics. Overall, however, the results provide strong evidence that anchoring bias plays an important role when recalling both objective and subjective outcomes.

Relative to the findings of other studies that measure anchoring bias in other contexts, the estimated magnitudes that we find are sizeable. For example, Campbell and Sharpe (2009) find point estimates of the forecast-anchor gap ranging between 0.04 and 0.71, resulting from regressing the forecast error of several macroeconomic indicators on the difference between the consensus forecast and data releases for the previous month (and previous three months) that serve as the anchor. Beggs and Grady (2009), in turn, find point estimates of 0.63 and 0.18 when examining the linear relationship between the final sale price of impressionist and contemporary art and the difference between the previous sale prices and a hedonic price prediction. More generally, Krueger and Schkade (2008) evaluate the reliability of standard life satisfaction and affective experience measures among a group of women in Texas and find a correlation close to 0.6 when the measures are asked two weeks apart.[20]

Sensitivity Analysis of Base Results

Next, we examine the sensitivity of our base results by running three additional sets of estimations. First, we assess whether our base results differ by country or by other observable characteristics of the households studied. In particular, we evaluate whether the observed anchoring bias for both objective and subjective measures is present across individuals in all four countries. Supplementary online appendix table A.2 reports the estimation results in which we interact the country dummies (Honduras is the base category) with the corresponding anchor measure (2013 concurrent report) and 2012 concurrent report. We observe similar anchoring behavior in all four countries. The magnitude of the association between the anchor value and the recalled objective and subjective outcomes is not systematically different across countries. In the case of the association between the 2012 concurrent report and recall measures, we observe differences across countries in three specific cases, but these are small and the overall partial correlations are still close

[19] When excluding the 2012 concurrent report from the estimations, the anchoring coefficient also changes marginally for both objective and subjective indicators (in most cases, by less than 0.01 units). This also confirms the lack of influence of the 2012 concurrent report on the strong linkage between the 2013 concurrent report and the 2012 recall value.

[20] The authors also find that the subjective measures are less reliable than objective measures such as income and education.

to zero. Overall, the high degree of mental anchoring is common among individuals in all surveyed locations.[21]

In a set of unreported results, we also examine whether there are heterogeneous effects by respondent's sex (if male), age (if age above the median), and education level (if secondary education or higher) and by main crop produced (fruits and vegetables, basic grains, and coffee and cacao). We generally do not find systematic differences. As an additional sensitivity check, we estimate the income results separately for households above or below the median household size. We find that larger households exhibit stronger anchoring relative to smaller households, which is consistent with respondents from larger households experiencing greater difficulty recalling certain values (such as total household income) and, hence, rely more heavily on the anchor.[22]

Second, we run an additional set of regressions for our objective variables to assess the effect of the anchoring variable on the recall measure when the recall period is two years.[23] We compare the effect of an individual's 2013 concurrent report on their 2011 recall value. One notable difference in this exercise (relative to the base analysis) is that the respondent now answers approximately twice as many questions between their anchoring value (the concurrent report for 2013) and the two-year recall value (for 2011).

Supplementary online appendix table A.3 presents these results, which provide even stronger evidence of anchoring. The anchoring coefficient ranges between 0.68 and 0.93, and a null or very weak correlation exists between the value recalled for 2011 in 2013 and the concurrent report from the 2011 survey round. One possibility for these stronger anchoring effects in the context of longer recall periods is that respondents are more prone to give more biased answers because it is harder to accurately recall; they are therefore more likely to use heuristics. The stronger effects are also interesting given the increased amount of time between the potential anchoring question and the subsequent report under two-year recall (compared to the report under one-year recall). One could anticipate a lower bias in this context.

Finally, we examine the sensitivity of our subjective indicator results to potential cap effects arising from the fact that these indicators are constructed on a scale of one to ten. Accordingly, supplementary online appendix table A.4 presents the estimations results that exclude the cases in which individuals reported a value of one (or ten) in the 2013 survey—that is, when their 2013 concurrent report is equal to the minimum or maximum value. We find very similar results to those in the original specification, with positive and highly significant anchor values ranging between 0.5 and 0.66. Thus, the anchoring bias among the subjective measures does not seem to be driven by the lower and upper bounds in the scale.[24] Similarly, to test the sensitivity of our objective indicator results to the exclusion of potential extreme values, particularly for income and wages, we examine log specifications and find qualitatively similar results (see supplementary online appendix table A.5).

Choice of Anchor

Our base (and preferred) results discussed above focus on the most recent (2013) reported measure as a potential anchor. This is mainly due to the structure of the survey instrument, in which a set of questions about a given outcome for the most recent period is

21 We are not aware of major weather, crop, or political events in the surveyed areas during the period of analysis that could have differentially affected farmers in the sample. As noted above, there was a general decline in rainfall levels across all locations (municipalities) in the study; this insufficient within-sample variation in rainfall shocks prevents us, for example, from assessing whether people become more reliant on anchoring under particular conditions. Unfortunately, the survey questionnaire was brief and did not capture potential internal or family shocks that could also be exploited.

22 Finally, there are also no differences between individuals experiencing higher (above the median) and lower absolute variation in their 2012 and 2013 concurrent reports. This could be explained, however, by the fact that most individuals in our working sample already exhibit important variations in their reported outcomes, so further distinguishing between relatively higher and lower variations does not result in statistically significant differences. We also examined whether experiencing a good year affected the degree of mental anchoring for subjective measures, but we do not find evidence that individuals with a positive change in their concurrent income exhibit a different pattern of subjective anchoring than the rest.

23 The subjective well-being questions were not included in the 2011 survey, precluding their use for this portion of our sensitivity analysis.

24 Note that this exercise helps to rule out that the caps in the subjective indicators are the source of the anchoring bias. Still, the imposed caps may limit the magnitude of the bias. For example, if an individual reports a value of nine for current well-being and, based on this report (used as an anchor), would like to report a value of 11 for the previous period, she is constrained to reporting a maximum value of ten; a similar (opposite) reasoning applies if her anchor value is close to a value of one.

Table 5. Estimation Results Using Community–level Anchor Value (Municipality Level Median Excluding Respondent's Own Report)

Coefficient	Dependent variable: 2012 Recall Value			
	Panel A: Objective Outcomes			
	Income	Income per capita	Primary wage	Hours main activity
	(1)	(2)	(3)	(4)
Community anchor (2013 Concurrent Report)	–0.427*	–0.640	–0.346	0.274
	(0.063)	(0.105)	(0.527)	(0.398)
2012 Concurrent Report	0.037	0.051***	0.020	0.027
	(0.344)	(0.004)	(0.449)	(0.383)
Project Fixed Effects	+	+	+	+
Individual Characteristics	+	+	+	+
Interview Characteristics	+	+	+	+
p–value (H_0: Concurrent $\geq$ Anchor)	0.010	0.001	0.274	0.118
Observations	553	553	553	553
	Panel B: Subjective Outcomes			
	Happy	Healthy	Stress	Well–being
	(5)	(6)	(7)	(8)
Community anchor (2013 Concurrent Report)	–0.094	0.290	–0.134	0.246
	(0.387)	(0.344)	(0.477)	(0.246)
2012 Concurrent Report	0.115**	0.142***	0.088**	0.117***
	(0.035)	(0.000)	(0.031)	(0.004)
Project Fixed Effects	+	+	+	+
Individual Characteristics	+	+	+	+
Interview Characteristics	+	+	+	+
p–value (H_0: Concurrent $\geq$ Anchor)	0.957	0.254	0.128	0.223
Observations	553	553	553	553

Note: Dependent variable is the recall value for 2012 reported in 2013; Community anchor (2013 Concurrent Report) refers to the median value reported in 2013 for the outcome in 2013 among other respondents in the same municipality (excluding the respondent's own report); 2012 Concurrent Report refers to the value reported in 2012 for the outcome in 2012. Individual characteristics include respondent age, gender, years of education, household size, and size of primary agricultural plot. Interview characteristics include the age and gender of the enumerators in both years and duration of all survey rounds. The *p*–values are estimated clustering over eight projects using the wild–bootstrap cluster procedure described by Cameron, Gelbach and Miller (2008) with bootstrap samples generated using a Rademacher weight distribution with 256 replications; *p*–value of one–sided Wald test for 2012 Concurrent $\geq$ 2013 Community Anchor reported at the bottom using wild–bootstrap. Asterisks *,**, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample has one less observation than the full sample as there is one household that is the only household resident in its municipality.

asked first, followed by questions about that outcome in previous periods. We now explore (to the extent possible) two alternative anchors upon which respondents could also rely.

First, for both objective and subjective indicators we test the community-level concurrent report as a possible anchor. We use the median value of the 2013 concurrent reports of all the other individuals located in the same municipality (i.e., we exclude the report of each particular individual). The results are presented in table 5. We find that this community-level anchor is generally not used by respondents as a reference point for either objective or subjective indicators. There is one notable exception, however; the community-level concurrent report is negatively associated with individuals' recall of their household income. However, we prefer not to over-interpret this particular result, given that the general pattern of results suggests that this anchor is not relevant.[25]

Second, we use the additional data available for objective indicators to test for evidence of recency effects. Specifically, we assess whether the report value of the current period (2013 concurrent report) or the recall value of the previous period (2012 recall value), which is provided after the value for the current period, is used as the reference measure when recalling the value for two

[25] When using the community mean (instead of the median), the anchoring coefficient on income becomes insignificant, while the coefficient on primary wages becomes statistically significant. We believe that the sensitivity of these results provides further support that the community-level concurrent report is not an anchor used by respondents.

Table 6. Estimation Results for Objective Measures, Two–year Recall Controlling for One–year Recall Value

Coefficient	Dependent variable: 2011 Recall Value			
	Income (1)	Income per capita (2)	Primary wage (3)	Hours main activity (4)
Anchor value (2013 Concurrent Report)	0.799**	0.624***	0.107	0.229**
	(0.016)	(0.000)	(0.254)	(0.020)
Anchor value (2012 Recall Value)	0.168	0.179	0.850	0.687**
	(0.508)	(0.164)	(0.223)	(0.012)
2011 Concurrent Report	0.006	0.021*	0.060**	0.003
	(0.609)	(0.070)	(0.035)	(0.871)
Project Fixed Effects	+	+	+	+
Individual Characteristics	+	+	+	+
Interview Characteristics	+	+	+	+
p–value (H_0: Concurrent $\geq$ 2013 Anchor)	0.016	0.001	0.337	0.002
p–value (H_0: Concurrent $\geq$ 2012 Anchor)	0.240	0.117	0.000	0.000
Observations	554	554	554	551

Note: Dependent variable is the recall value for 2011 reported in 2013; Anchor value (2013 Concurrent Report) refers to the value reported in 2013 for the outcome in 2013; Anchor value (2012 Recall Value) refers to the recall value for 2012 reported in 2013; 2011 Concurrent Report refers to the value reported in 2011 for the outcome in 2011. Individual characteristics include respondent age, gender, years of education, household size and size of primary agricultural plot. Interview characteristics include the age and gender of the enumerators in both years and the duration of all survey rounds. The p–values are estimated clustering over eight projects using the wild–bootstrap cluster procedure described by Cameron, Gelbach and Miller (2008), with bootstrap samples generated using a Rademacher weight distribution with 256 replications; p–values of one–sided Wald test for 2011 Concurrent $\geq$ 2013 Anchor and 2011 Concurrent $\geq$ 2012 Anchor are reported at the bottom using wild–bootstrap. Asterisks *,**, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

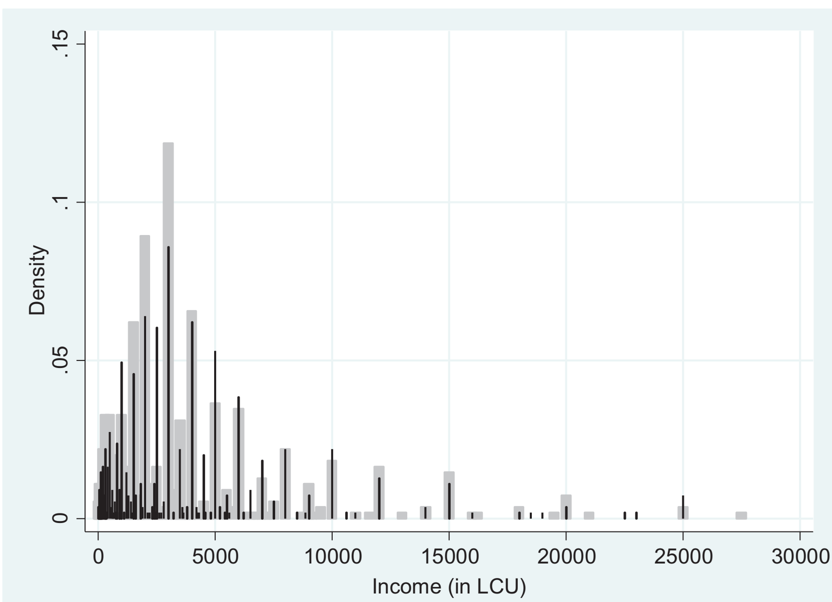

Figure 2. Distribution of selected objective indicator: 2013 concurrent income reported in 2013 versus 2011 recall income reported in 2013

Note: The grey bars correspond to the distribution of the 2013 concurrent income reported in 2013. The black lines correspond to the distribution of the 2011 recall income reported in 2013. Variables are measured in local currency units (LCU).

Table 7. Estimation Results for Objective and Subjective Outcomes, Differential Responses to Positive and Negative Changes in Concurrent Reports

Coefficient	Dependent Variable: Difference between 2012 Recall Value and 2012 Concurrent Report			
	Panel A: Objective Outcomes			
	Income	Income per capita	Primary wage	Hours main activity
	(1)	(2)	(3)	(4)
Positive Difference	0.681**	0.547**	0.732*	0.883***
	(0.039)	(0.016)	(0.066)	(0.000)
Negative Difference	–1.015***	–0.999***	–1.001**	–0.726***
	(0.008)	(0.008)	(0.016)	(0.008)
Project Fixed Effects	+	+	+	+
Individual Characteristics	+	+	+	+
Interview Characteristics	+	+	+	+
p–value (H_0: Positive $\geq$ Negative)	0.081	0.019	0.162	0.978
Observations	554	554	554	554
	Panel B: Subjective Outcomes			
	Happy	Healthy	Stress	Well–being
	(5)	(6)	(7)	(8)
Positive Difference	1.018**	0.922**	0.716**	0.933***
	(0.016)	(0.012)	(0.020)	(0.004)
Negative Difference	–0.527**	–0.547**	–0.875**	–0.583**
	(0.016)	(0.020)	(0.027)	(0.012)
Project Fixed Effects	+	+	+	+
Respondent Characteristics	+	+	+	+
Interview Characteristics	+	+	+	+
p–value (H_0: Positive $\leq$ Negative)	0.000	0.004	0.925	0.000
Observations	554	554	554	554

Note: Dependent variable is the difference between the recall value for 2012 reported in 2013 and the concurrent value for 2012 reported in 2012; Positive difference is the value of the difference between the 2013 concurrent value reported in 2013 and the 2012 concurrent value reported in 2012 if that difference is greater than zero, and takes the value of zero otherwise; Negative difference is the difference of these 2013 and 2012 concurrent reports (in absolute value) if the difference is less than zero, and zero otherwise. Individual characteristics include respondent age, gender, years of education, household size, and size of primary agricultural plot. Interview characteristics include the age and gender of the enumerators in both years and the duration of all survey rounds. The p–values are estimated clustering over eight projects using the wild–bootstrap cluster procedure described by Cameron, Gelbach and Miller (2008) with bootstrap samples generated using a Rademacher weight distribution with 256 replications; p–values of one–sided Wald test for Positive ≥ Negative (Panel A) and for Positive ≤ Negative (Panel B) are reported at the bottom using wild–bootstrap. Asterisks *,**, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

prior periods (2011 recall value).[26] This elicitation method allows us to provide suggestive evidence of whether the observed anchoring is driven by recency effects. Recency effects have been well documented in surveys, particularly in audio presentations of surveys (see, e.g., Schwarz, Hippler, and Noelle-Neumann 1992).

Table 6 presents the corresponding results. We find that for the household income measures, the 2013 concurrent report is the primary anchor when recalling the 2011 value; for wages and working hours, on the other hand, the most recent 2012 recall value plays an important role.[27] This suggests that recency effects drive anchoring for wages and hours, while income and income per capita are less responsive to recency effects.

Why Do Respondents Exhibit Anchoring?

Lastly, we explore potential mechanisms driving the observed anchoring bias. We consider two possible theories that could drive

[26] In particular, individuals were asked a series of questions pertaining to their economic and labor activities for 2013, followed by the same series of questions for 2012 and for 2011.

[27] In the case of wages, while the coefficient of the 2012 anchor value is much larger than the 2013 anchor value, both coefficients are not precisely estimated to conclude that they are statistically different from zero.

respondents' anchoring: rounding and positivity and negativity bias.

Rounding as a driver of anchoring bias. Previous literature has considered rounding as a cause of recall error, particularly for large values (see, e.g., Huttenlocher, Hedges, and Bradburn 1990). In our setting, rounding may drive anchoring bias if respondents first consider their most recent report and then revise it upward or downward to report a rounded number for the recalled period. We would thus expect more frequent rounding of recall values than of concurrent reports. Figure 2 presents the income distribution of both the 2013 concurrent report and the 2011 recall value (reported in 2013).[28] We observe a similar pattern in both distributions, which suggests that the observed anchoring bias is not driven by a systematic rounding process.

Positivity and Negativity bias. Next, we extend the main analysis to examine whether the anchoring effect varies based on changes in the concurrent reports over time. The psychology literature documents that people have a strong tendency to give greater weight to negative events than to positive events (Baumeister et al. 2001; Rozin and Royzman 2001). Such tendencies may also affect recall bias. To explore this possibility, we evaluate how positive and negative changes in the concurrent reports between periods influence the anchoring bias. As shown in equation (4), the dependent variable in this case is the difference between the recall value for 2012 (reported in 2013) and the 2012 concurrent report.[29] The independent variables of interest are two separate variables: one accounting for a positive difference in the 2013 and 2012 concurrent reports (equal to the value of the difference when this difference is positive and zero otherwise) and a second accounting for a negative difference in these reports (equal to the absolute value of the difference between the 2013 and 2012 concurrent reports when this difference is negative, and zero otherwise). Hence, we directly assess the relationship between deviations in the 2012 recall value (relative to the 2012 concurrent report) and positive and negative outcome changes.

Panel A of table 7 presents the estimation results for the objective indicators, while panel B presents the results for the subjective indicators. Two interesting results are worth noting. First, for all objective and subjective measures, we observe a strong and statistically significant association between the magnitude of both positive and negative outcome changes and the magnitude of the anchoring bias on the recall value. The opposite sign of the coefficients for positive and negative outcome changes provides additional support to interpret the observed bias as anchoring bias. A larger positive (negative) change between the 2012 and 2013 concurrent reports results in a larger positive (negative) difference between the 2012 recall value and 2012 concurrent report.

Figures 3 and 4 provide additional insights about the observed outcome changes and the distribution of the measurement bias (the difference between the recall value and concurrent report for 2012). Figure 3 plots the distribution of the measurement bias for income, distinguishing between individuals reporting positive and negative outcome changes, while figure 4 plots the same distributions for the level of happiness. In both cases, we find a leftward distributional shift of the measurement bias for those experiencing a negative outcome change and a corresponding rightward shift for those experiencing a positive outcome change. These patterns further support the opposite direction of the bias when experiencing positive and negative outcome changes and demonstrates that these effects occur throughout the entire distribution of individuals with varying positive (negative) outcome changes and are not simply being driven by outlier values.

The second interesting result is that for the objective outcomes, the magnitude of the bias is frequently stronger for negative than for positive changes between the 2012 and 2013 concurrent reports; for the subjective outcomes, however, we find the opposite relationship. Among the objective indicators, the magnitude of the coefficient is statistically larger for negative differences for the income variables (though not for working hours and only marginally for wages); among the subjective indicators, the magnitude of the coefficient is statistically larger for positive

[28] A similar pattern is observed for wages, as well as when comparing the distributions for the 2013 concurrent report and 2012 recall value.

[29] For individuals reporting no change, the variable takes the value of zero. For the income and wage variables, few individuals report no change; for hours in main activity and the subjective indicators, 12% to -25% of the sample report no change.

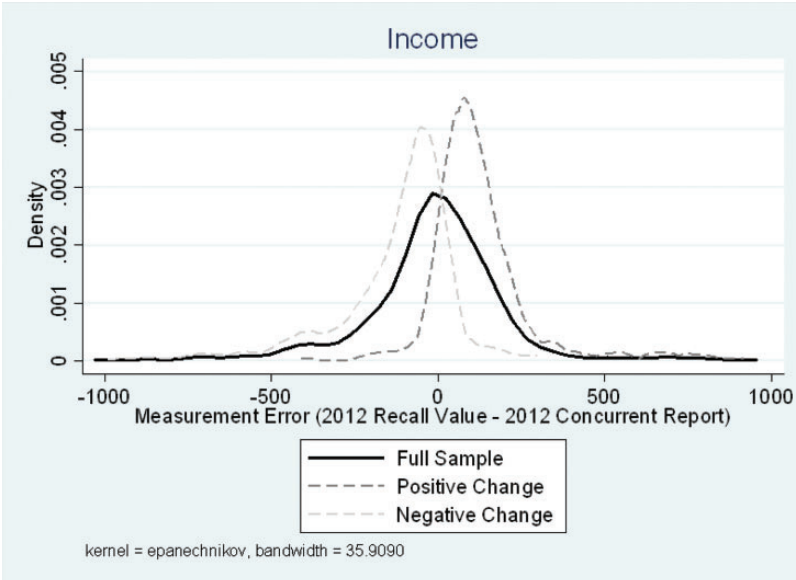

Figure 3. Distribution of measurement bias for selected objective measure

Note: Positive Change corresponds to all individuals with a positive change in their reported concurrent income between 2012 and 2013; Negative Change corresponds to all individuals with a negative change in their reported concurrent income between 2012 and 2013. Density distributions are derived using kernel density estimation. For ease of interpretation, respondents reporting a difference in magnitude greater than 1,000 are excluded (31 observations).

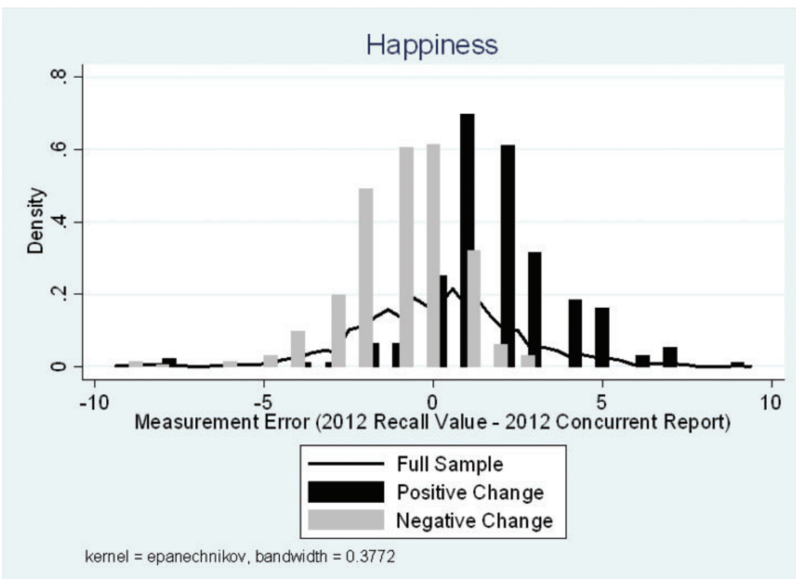

Figure 4. Distribution of measurement bias for selected subjective measure

Note: Positive Change corresponds to all individuals with a positive change in their reported level of happiness between 2012 and 2013; Negative Change corresponds to all individuals with a negative change in their reported level of happiness between 2012 and 2013. Density distributions are derived using kernel density estimation.

differences except for the level of stress (where a negative change actually indicates an improvement for the individual).[30] Hence, the results for objective indicators are

[30] Similar results are obtained for the subjective indicators when excluding the cases in which the value reported in 2013 (i.e., the anchor value) was either the minimum or maximum on the ten-point scale. See supplementary online appendix table A.6.

consistent with a broad body of literature showing that negative information (in our context, a negative reflection of the current state of the respondents' situation) plays a disproportionate role in influencing future impressions (see, e.g., Peeters and Czapinski 1990). However, we also observe positivity bias in the case of subjective indicators. Some research has documented a positivity bias (rather than a negativity bias) in specific populations, such as older adults (see, e.g., Reed and Carstensen 2012).[31]

Supplementary online appendix table A.7 reports the estimation results of equation (4) when further controlling for the modal pattern of outcome changes in the municipality in which the respondent resides. We include both a positive and a negative dummy variable that takes the value of one when the majority of the people in the municipality (excluding the respondent) report a positive or negative change, respectively, in the corresponding variable between their concurrent reports for 2012 and 2013. We find similar results to those shown in table 6. For the objective measures, the magnitude of the bias is larger for negative differences in the reported values over time, while it is larger for positive differences for the subjective measures.[32] In terms of the variables that capture the municipality modal pattern, we find a significant correlation between the modal pattern and the magnitude of the bias in some of the objective and subjective indicators. However, positive (negative) outcome changes in the vicinity do not necessarily translate into a larger positive (negative) bias.

Overall, we observe a strong and consistent anchoring bias across multiple model specifications and for both objective and subjective outcomes. These findings are indicative that relying on retrospective measures without accounting for the effects of anchoring (even over relatively short recall periods) can result in significant mismeasurement of income, labor-related activities, or subjective indicators. This holds especially true in settings in which individuals are subject to considerable variation in their working patterns, such as smallholder farmers in developing countries. The analysis further shows that in situations of economic (income) decline, the recalled income for the previous period may be considerably underestimated. Conversely, in situations of increasing personal satisfaction, recalled satisfaction for the previous period may be considerably overestimated. This can be particularly critical when we are interested in comparing income or well-being measures across periods and we rely on retrospective survey data asking questions about current and past situations in a single survey. A natural solution would be to collect information over multiple rounds, rather than only visiting households one time to ask them questions about their current and past situations. This would result in more reliable measures, but would also require additional resources.[33]

Conclusion

Collecting reliable data, including recall data, is important for policymaking. This paper uses a unique panel-survey dataset from four countries in Central America to examine the role of anchoring as a source of recall error. We find strong evidence in support of anchoring bias for both objective and subjective indicators. Respondents consistently use their reported value for the current period (2013) as a cognitive heuristic to assist them recalling the value for a previous period (2012). The results are robust to alternative model specifications and samples. We further find that positive changes in indicators (measured concurrently) across time are strongly correlated with a positive bias, while negative changes are strongly correlated with a negative bias. Interestingly, objective indicators show larger biases in response to negative changes, while subjective indicators exhibit larger biases in response to positive changes.

We acknowledge some limitations of the results presented. Our dataset permits us to

[31] We also examined if there are heterogeneous effects by age (if age above the median) on the positive and negative indicators but we generally do not find systematic differences.

[32] Alternatively, we included municipality fixed effects (instead of the municipality modal pattern indicators) and find similar results. Similarly, we examined if there are systematic differences between individuals whose outcomes go in the opposite direction than most people in their municipality and individuals whose outcomes go in the same direction, but we do not find major differences.

[33] Implicit Association tests (see Greenwald, McGhee and Schwartz 1998) and multinomial models (e.g., Conrey et al. 2005) offer innovative methods to mitigate measurement biases in the measurement of attitudes and beliefs where issues of self-presentation, self-monitoring, and social desirability biases are of particular concern. Given the indicators studied here, we do not think these methods offer appropriate solutions, but this is an area very much in need of future research.

assess anchoring among a relatively simple set of variables, so our results may not be generalizable to more complex recall variables (although we hypothesize that such biases are likely to be greater for indicators that are harder to recall). Similarly, the survey instrument used to generate the data for this study has a particular structure (i.e., an initial set of questions about a given outcome for the most recent period, with posterior questions about that outcome in previous periods), which both enables us to analyze the extent of anchoring bias more precisely and prevents us from ruling out the possibility that this anchoring would not be present in other survey designs. Similarly, the survey structure prohibits the testing of recommendations to inform survey design to minimize these biases. This matter, while important, is beyond the scope of the paper, but it is the subject of ongoing research (e.g., Heath et al. 2016). Lastly, as the survey was carried out with a very specific population (poor smallholder farmers in Central America), the conclusions may not necessarily be generalizable to different populations, whether in terms of location, wealth, or occupation.

Despite these concerns, it is clear from our data that a naïve analysis using retrospective measures that do not consider potential anchoring effects can result in important mismeasurements. This could lead to incorrect inferences regarding, for example, the evolution of objective and subjective indicators over time. In this sense, it is necessary to take into account respondents' likely cognitive biases when talking about current and past situations, even for relatively short reference periods. This matter may be particularly pertinent in agricultural settings in developing countries, considering the significant variation in the working patterns of smallholders.

While not the primary focus of this study, the weak relationship between outcomes concurrently measured during the reference period and their recall equivalent (i.e., 2012 concurrent report and 2012 recall value) also calls into question the reliability of using retrospective data as an accurate measure. This points to the strong need for additional studies to improve retrospective data collection strategies. The results imply that studies analyzing the time evolution of outcomes relying only on retrospective data could overstate the autocorrelation in outcomes across time. Combined with the anchoring results that objective indicators like income are more heavily biased in the presence of a negative change such studies will document (in the case of objective indicators) trajectories that significantly overstate negative outcomes in previous periods if the current year is not a good year. Hence, we can obtain very different results when relying on retrospective data if the year of data collection is a particularly good or a particularly bad year.

Overall, this study shows the presence of one particular type of cognitive bias in retrospective data from a household survey. Future research should continue investigating the potential biases of self-reported retrospective data, as well as alternative mechanisms that could help reduce the bias in a context in which such data has become widely used in empirical work. For instance, we conceptualize anchoring bias as something directly affecting individual responses (such as household income) that requires additional mental effort and aggregation. It will be interesting to assess the potential influence of anchoring bias on common household measures that typically require the aggregation of multiple responses, such as consumption or expenditure aggregates. Similarly, it would be interesting to assess if the most recent outcome would be as strong an anchor if the question order were randomized—that is, if respondents were asked to first report their outcome for the previous period and then to report their outcome for the most recent period. The differential anchoring bias when comparing individuals who have experienced positive versus negative outcome changes across time also deserves further analysis, probably using an experimental setting.

Supplementary Material

Supplementary material are available at *American Journal of Agricultural Economics* online.

References

Alevy, J.E., C.E. Landry, and J.A. List. 2015. Field Experiments on the Anchoring of Economic Valuations. *Economic Inquiry* 53 (3): 1522–38.

Ariely, D., G. Loewenstein, and D. Prelec. 2003. Coherent Arbitrariness: Stable Demand Curves Without Stable

Preferences. *The Quarterly Journal of Economics* 118: 73–105.

Arthi, V., K. Beegle, J. de Weerdt, and A. Palacios-Lopez. 2016. Not Your Average Job: Measuring Farm Labor in Tanzania. Policy Research Working Paper No. WPS 7773. Washington, D.C: World Bank Group.

Baumeister, R.F., E. Bratslavsky, C. Finkenauer, and K.D. Vohs. 2001. Bad is Stronger than Good. *Review of General Psychology* 5 (4): 323–70.

Beegle, K., C. Carletto, and K. Himelein. 2012. Reliability of Recall in Agricultural Data. *Journal of Development Economics* 98 (1): 34–41.

Beggs, A., and K. Graddy. 2009. Anchoring Effects: Evidence from Art Auctions. *American Economic Review* 99 (3): 1027–39.

Bergman, O., T. Ellingsen, M. Johannesson, and C. Svensson. 2010. Anchoring and Cognitive Ability. *Economics Letters* 107 (1): 66–8.

Bodenhausen, G.V., S. Gabriel, and M. Lineberger. 2000. Sadness and Susceptibility to Judgemental Bias: The Case of Anchoring. *Psychological Science* 11 (4): 320–23.

Bound, J., C. Brown, and N. Mathiowetz. 2001. Measurement Error in Survey Data. In *Handbook of Econometrics*, vol. 5, Chapter 59, ed. J.J. Heckman, and E. Leamer, 3705-843. Amsterdam, The Netherlands: Elsevier.

Brekke, N., K. Etling, C. Houston, and T. Wilson. 1996. A New Look at Anchoring Effects: Basic Anchoring and its Antecedents. *Journal of Experimental Psychology* 125 (4): 387–402.

Cameron, A.C., J.B., Gelbach, and D.L. Miller. 2008. Bootstrap-based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics* 90 (3): 414–27.

Campbell, S.D., and S.A. Sharpe. 2009. Anchoring Bias in Consensus Forecasts and Its Effect on Market Prices. *The Journal of Financial and Quantitative Analysis* 44 (2): 369–90.

Clarke, P.M., D.G. Fiebig, and U.G. Gerdtham. 2008. Optimal Recall Length in Survey Design. *Journal of Health Economics* 27: 1275–84.

Conrey, F.R., B. Gawronski, J.W. Sherman, K. Hugenberg, and C.J. Groom. 2005. Separating Multiple Processes in Implicit Social Cognition: The Quad Model of Implicit Task Performance. *Journal of Personality and Social Psychology* 89(4): 469–87.

Crossley, T.F., and J.K. Winter. 2015. Asking Households about Expenditures: What Have We Learned? In *Improving the Measurement of Consumer Expenditures. Studies in Income and Wealth*, Volume 74, ed. C., Carroll, T.F. Crossley, and J. Sabelhaus, 23–52. Chicago: University of Chicago Press.

Das, J., J. Hammer, and C. Sanchez-Paramo. 2012. The Impact of Recall Periods on Reported Morbidity and Health Seeking Behavior. *Journal of Development Economics* 98: 76–88.

de Nicola, F., and X. Gine. 2014. How Accurate is Recall Data? Evidence from Coastal India. *Journal of Development Economics* 106: 55–65.

de Weerdt, J., K. Beegle, J. Friedman, and J. Gibson. 2016. The Challenge of Measuring Hunger through Survey. *Economic Development and Cultural Change* 64 (4): 727–58.

Eastwick, P.W., L.L. Hunt, and L.A. Neff. 2013. External Validity, Why Art Thou Externally Valid? Recent Studies of Attraction Provide Three Theoretical Answers. *Social and Personality Compass* 7 (5): 275–88.

Evans, D.S., and L.S. Leighton. 1995. Retrospective Bias in the Displaced Worker Surveys. *The Journal of Human Resources* 30 (2): 386–96.

Finkel, S.E., T.M. Guterbock, and M.J. Borg. 1991. Race-of-interviewer Effects in a Preelection Poll: Virginia 1989. *Public Opinion Quarterly* 55: 313–30.

Frykblom, P., and J.F. Shogren. 2000. An Experimental Testing of Anchoring Effects in Discrete Choice Questions. *Environmental and Resource Economics* 16: 329–41.

Fudenberg, D., D.K. Levine, and Z. Maniadis. 2012. On the Robustness of Anchoring Effects in WTP and WTA Experiments. *American Economic Journal: Microeconomics* 4 (2): 131–45.

Furnham, A., and H. Chu Boo. 2011. A Literature Review of the Anchoring Effect. *The Journal of Socio-Economics* 40: 35–42.

Goldin, J., and D. Reck. 2015. *Framing Effects in Survey Research: Consistency-adjusted Estimators*. Mimeo.

Gove, W.R., and M.R. Geerken. 1977. Response Bias in Surveys of Mental Health: An Empirical Investigation. *American Journal of Sociology* 82: 1289–317.

Greenwald, A.G., D.E. McGhee, and J.L.K. Schwartz. 1998. Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *Journal of Personality and Social Psychology* 74(6): 1464–80.

Heath, R., G. Mansuri, D. Sharma, B. Rijkers, and W. Seitz. 2016. Measuring Employment in Developing Countries. Evidence from a Survey Experiment. Working Paper.

Heinrich, J., S.J. Heine, and A. Norenzayan. 2010. The Weirdest People in the World. *The Behavioral and Brain Sciences* 33: 61–83.

Henry, P.J. 2008. College Sophomores in the Laboratory Redux: Influences of a Narrow Data Base on Social Psychology's View of the Nature of Prejudice. *Psychological Inquiry* 19: 49–71.

Hernandez, M.A., and M. Torero. 2014. Poverty-sensitive Scorecards to Prioritize Lending and Grant allocation with an Application in Central America. In *Finance for Food: Towards New Agricultural and Rural Finance*, Chapter 12, ed. D., Köhn, 263–83. New York: Springer.

———. 2016. A Poverty-sensitive Scorecard to Prioritize Lending and Grant Allocation: Evidence from Central America. IFPRI Discussion Paper 1518.

Hitczenko, M. 2013. Modelling Anchoring Effects in Sequential Likert Scale Questions. Federal Reserve of Boston, Research Department Working Series, No. 13–15.

Huttenlocher, J., L.V. Hedges, and N.M. Bradburn. 1990. Reports of Elapsed Time: Bounding and Rounding Processes in Estimation, *Journal of Experimental Psychology. Learning, Memory, and Cognition* 16 (2): 196–213.

Hurd, M.D., 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, chapter 8. In *Frontiers in the Economics of Aging*, ed D.A., Wise. 353–92. Massachusetts: University of Chicago Press.

Krosnick, J.A., and D.F. Alwin. 1987. An Evaluation of Cognitive Theory of Response Order Effects in Survey Measurement. *Public Opinion Quarterly* 51: 201–19.

Krosnick, J.A. 1991. Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology* 5: 213–36.

——— 1999a. Maximizing Questionnaire Quality. In *Measures of Political Attitudes.* ed. J., Robinson, P.R. Shaver, and L.S. Wrightsman, 37–57. San Diego, CA: Academic Press.

——— 1999b. Survey Research. *Annual Review of Psychology* 50: 537–67.

Krueger, A.B., and D.A. Schkade. 2008. The Reliability of Subjective Well-being Measures. *Journal of Public Economics* 92 (8-9): 1833–45.

Maniadis, Z., F. Tufano, and J.A. List. 2014. One Swallow Doesn't Make a Summer: New Evidence on Anchoring Effects. *American Economic Review* 104 (1): 277–90.

Maruyama, E. 2007. *Remembering Earnings During Unstable Times: Evidence from the Indonesian Family Life Surveys*. Washington DC: International Food Policy Research Institute, Mimeo.

McKenzie, D., and M. Rosenzweig. 2012. Symposium on Measurement and Survey Design. *Journal of Development Economics* 98 (1): 1–148.

Narayan, S., and J.A. Krosnick. 1996. Education Moderates Some Response Effects in Attitude Measurement. *Public Opinion Quarterly* 54: 127–45.

Neter, J., and J. Waksberg. 1964. A Study of Response Errors in Expenditures Data from Household Interviews. *Journal of the American Statistical Association* 59 (305): 18–55.

Peeters, G., and J. Czapinski. 1990. Positive-Negative Asymmetry in Evaluations: The Distinction between Affective and Informational Negativity Effects, chapter 2. In *European Review of Social Psychology, Vol. 1*., ed. W. Stroebe and M. Hewstone, 33–60. United Kingdom: John Wiley & Sons Ltd.

Reed A.E., and L.L. Carstensen. 2012. The Theory Behind the Age-related

Positivity Effect. *Frontiers in Psychology* 3:339.

Rozin, P., and E.B. Royzman. 2001. Negativity Bias, Negativity Dominance and Contagion. *Personality and Social Psychology Review* 5 (4): 296–320.

Schwarz, N., and H.J. Hippler. 1995. Subsequent Questions May Influence Answers to Preceding Questions in Mail Surveys. *The Public Opinion Quarterly* 59 (1): 93–7.

Schwarz, N., H.J. Hippler, and E. Noelle-Neumann. 1992. A Cognitive Model of Response-order Effects in Survey Measurement. In *Context Effects in Social and Psychological Research*, ed. N. Schwarz and S. Sudman, 187–201. New York: Springer.

Sears, D.O. 1986. College Sophomores in the Laboratory: Influences of a Narrow Data Base on Social Psychology's View of Human Nature. *Journal of Personality and Social Psychology* 51 (3): 515–30.

Simonsohn, U., and G. Loewenstein. 2006. Mistake #37: The Effect of Previously Encountered Prices on Current Housing Demand. *Economic Journal* 116 (508): 175–99.

Sudman, S., N.M. Bradburn, and N. Schwarz. 1996. *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass Publishers.

Tversky, A., and D. Kahneman. 1974. Judgment Under Uncertainty: Heuristics and Biases. *Science* 185: 1124–31.

Webb, M.D. 2014. Reworking Wild Bootstrap Based Inference for Clustered Errors. Queen's University, Department of Economics Working Paper No. 1315.