UMR 8184

Thema

THéorie Économique, Modélisation et Applications

THEMA Working Paper n°2015-20
Université de Cergy-Pontoise, France

What can we learn from the fifties?

Fabian Gouret

November 2015

UNIVERSITÉ de Cergy-Pontoise

CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

What can we learn from the fifties?*

Fabian Gouret†

First version: May 2011

This version: November 2015

Abstract

Economists have increasingly elicited from survey respondents probabilistic expectations. Subjective probabilistic expectations show great promise to improve the estimation of structural models of decision-making under uncertainty. However, a robust finding in these surveys is an inappropriate heap of responses at "50 percent", suggesting that some of these responses are uninformative. The way these 50s are treated in the subsequent analysis is of major importance. Taking the 50s at face value will bias any aggregate statistics. On the reverse deleting them is not appropriate if some of these answers do convey some information. Furthermore, the attention of researchers is so focused on this heap of 50s that they do not consider the possibility that other answers may be uninformative as well. This paper proposes to take a fresh look at these questions using a new method based on extremely weak assumptions to identify the informativeness of an answer. Applying the method to probabilistic expectations of equity returns in three waves of the Survey of Economic Expectations in 1999-2001, I find that: (i.) at least 65 percent of the 50s convey no information at all; (ii.) it is the answer most often provided among the uninformative answers; (iii.) but even if the 50s are a major contributor to noise, they represent at best 70 percent of the identified uninformative answers. These findings have various implications for survey design.

JEL: C81, D8

Keywords: Subjective probability distribution, survey data, epistemic uncertainty and fifty-fifty.

*I am grateful to Peter Klibanoff for insightful comments on a previous version, as well as François Belot, Russell Davidson, Emmanuel Flachaire, Jérôme Héricourt, Guillaume Hollard, Nicolas Jacquemet, William Kengne, Charles Manski, Christophe Muller, Jean-Christophe Pereau, Jean-Luc Prigent, Eric Renault, Julien Vauday, Jean-Christophe Vergnaud, Murat Yildizoglu and participants at the Xarxa de Referència meeting in Girona, the XREPP D-Day and the BIG Seminar at Universitat de Barcelona, the GREThA Seminar at Bordeaux, the GREQAM Seminar at Marseille, the EQUIPPE Seminar at Lille, the TEMA Seminar at Paris School of Economics, the Théma Seminar at Cergy, the congress of the Société Canadienne de Science Economique at Sherbrooke, the 68th European Meeting of the Econometric Society in Toulouse, the 5th French Econometrics Conference at Toulouse School of Economics, and the 1st annual conference of the International Association for Applied Econometrics at Queen Mary University of London. This research was supported by the Spanish Ministry of Science and Technology through the project CICYT ECO2012-35054, the Generalitat de Catalunya through the project SGR359, and a "Chaire d'Excellence CNRS".

†Théma, Université de Cergy-Pontoise, 33 Bvd du Port, 95011 Cergy-Pontoise Cedex, France. Tel.: +33 1 34 25 22 75; fax: +33 1 34 25 62 33; email: fabiangouret@gmail.com; homepage: https://sites.google.com/site/fabgouret/

1 Introduction

Since the nineties, an important empirical literature has developed to measure probabilistic expectations that individuals hold about future events (Manski, 2004). Subjective probabilistic expectations show great promise to improve the estimation of structural models of decision-making under uncertainty.[1] However, researchers are particularly embarrassed by a seemingly inappropriate high frequency of responses at "50 percent". Table 1 provides some examples of this heap in surveys which have included probabilistic questions. There are at least two interrelated problems with these 50s. The first one is that some of them are probably uninformative, i.e., some respondents provide this answer when they do not know what probability to answer in the interval $[0, 100]$. Fischhoff and Bruine de Bruin (1999), for instance, argue that some respondents use this answer as a proxy for the verbal expression "fifty-fifty", an expression used to say "I really don't know". If so, treating all the 50s as any other answer will bias any aggregate statistics. On the reverse purely suppressing all of them is not appropriate, particularly if some of them do actually convey some information. Given that survey questions usually do not directly reveal if a response provides an uninformative answer, it raises the question: How can one identify if a 50 is uninformative? The second related problem is that the attention of researchers is so focused on this heap of 50s that they do not consider the possibility that other responses may be uninformative as well. This raises the second question: Do the 50s represent a substantial share of all uninformative answers?

The present paper proposes a conservative solution applicable on a specific but widely used format of probabilistic questions used to know the distribution of the future realization of a continuous variable. As argued in this paper, this format, introduced initially in the Survey of Economic Expectations (SEE) by Dominitz and Manski (1997), facilitates ex-post control on the informativeness of a probabilistic answer, i.e., if this answer reflects a sharp expression of belief or a more or less imprecise belief on the event under consideration; and although this method is only applicable in the specific context of this format, its application permits to shed light on the appropriateness of some practices in the related literature. Let me introduce this format first. When the variable is

[1] For work in that direction, see Delavande (2008), Giustinelli (forthcoming), Van der Klaauw and Wolpin (2008), as well as Attanasio (2009, pp.90-91) who provides an overview of some of his papers in progress.

continuous, the SEE asks each respondent i a *sequence of K probabilistic questions* of the type:

What is the percent chance that, one year from now, R_i would be worth over $r_{i,k}$?

Hence, for each respondent i, we observe the probabilistic answers $Q_{i,k} \equiv \mathrm{P}(R_i > r_{i,k})$, $k = 1, 2, \ldots, K$, where $r_{i,1} < r_{i,2} < \ldots < r_{i,K}$ are K thresholds about which the respondent is queried.[2] R_i denotes the variable of interest; it can be the respondent's household income (Dominitz and Manski, 1997), the respondent's personal income (Dominitz, 2001), or the stock market returns according to the respondent (Dominitz and Manski, 2011). The responses to this sequence of K probabilistic questions enable the estimation of each respondent's *subjective distribution*. The sequence of K specific thresholds is chosen among a finite number of possible sequences by an algorithm that uses the average of the the respondent's answers $r_{i,min}$ and $r_{i,max}$ to two preliminary questions asking for the lowest and highest possible values that R_i might take.[3] Table 2 provides an example that I will study in detail in Sections 3 and 4: it considers the questions posed in three waves of the SEE (12, 13 and 14) in July 1999 to March 2001 to measure the probabilistic expectations that Americans hold about equity returns in the year ahead. The SEE first poses a scenario which highlights that the respondent has to think in the performance of an investment of \$1000 in a diversified mutual fund in the year ahead. Then the two preliminary questions are posed, and the responses are used to select via the algorithm a sequence of $K = 4$ threshold values among a set of 5 possible sequences. Finally, the SEE asks the sequence of $K = 4$ probabilistic questions. For instance, if a respondent answers $r_{i,min} = 600$ and $r_{i,max} = 1100$, his midpoint is 850, and he is asked the percent chance that next year's investment in the mutual fund would be worth over $\{500, 900, 1000, 1100\}$.

The objective is to infer the informativeness of a subjective probability $Q_{i,k}$. Note that the response $Q_{i,1}$ is particularly important because it determines the answers for the subsequent thresholds, while the reverse is not true: once a respondent answers $Q_{i,1}$, his answers to the next thresholds

[2]Instead of asking K points on the respondent i's subjective complementary cumulative distribution function of R_i, it is possible to ask points on the respondent i's subjective cumulative distribution of R_i. In that case the probabilistic questions are of the type "*What is the percent chance that, one year from now, the variable R_i would be less than $r_{i,k}$?*". See, e.g., Dominitz (2001).

[3]The two preliminary questions are thus useful to have an idea of the support of the distribution. Following Morgan and Henrion (1990), Dominitz and Manski (1997) give two additional reasons for asking these preliminary questions: they decrease overconfidence with respect to central tendencies and anchoring problems wherein respondents' beliefs are influenced by the questions posed.

have to weakly fall to satisfy the monotonicity of the complementary cumulative distribution function ($Q_{i,1} \geq Q_{i,2} \geq \ldots \geq Q_{i,K}$).[4] Hence, I will mainly focus on the informativeness of $Q_{i,1}$.

To understand the method, consider a respondent who has a precise subjective distribution in mind concerning R_i. If so, he is able to provide informative answers to the preliminary questions and the sequence of probabilistic questions, i.e., answers which reflect sharp expressions of beliefs. If this is the case, he should also provide *coherent* answers between the preliminary questions and the sequence of probabilistic questions, i.e., he should use the same underlying subjective distribution to answer the preliminary questions and the sequence of probabilistic questions. Alternatively, a respondent may have imprecise beliefs concerning R_i, i.e., he does not hold a unique subjective distribution but a set of subjective distributions. This imprecision can occur because he lacks some information; he may also be unable to form precise beliefs in practice, even if he is able to do so in principle, because he lacks time and thinking is costly. Whatever the reason of this imprecision, note that if a respondent provides clearly *incoherent* answers between the preliminary questions and the sequence of probabilistic questions, i.e., he does not use the same underlying subjective distribution, then one can be sure that the respondent has imprecise beliefs; his answers are thus partially informative. And if the *incoherence* is too high, one can be sure that the respondent has extremely imprecise beliefs (a case wherein he lacks any relevant information or he does not want to put any effort in his answers). In that case, his answers are clearly uninformative.

The general idea is thus to exploit the answers $r_{i,min}$ and $r_{i,max}$ to the preliminary questions to make a prediction $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$ of an answer $Q_{i,1}$. The distance between a predicted value and the actual value is a *measure of coherence*. If this measure is too high, $Q_{i,1}$ is highly incoherent with the respondent's answer to the pair of preliminary questions, so the respondent has an extremely imprecise belief on the event $R_i > r_{i,1}$. One can thus infer that $Q_{i,1}$ is uninformative. If this measure is equal to or slightly different from zero, one cannot reject the possibility that the respondent has a relatively precise belief on the event $R_i > r_{i,1}$, rounding only to the nearest 1, 5 or perhaps 10 percent. If so, $Q_{i,1}$ may be broadly informative. In-between, the measure is too far from zero to say

[4] Logically, the interviewer informs the respondent if a probability elicited at threshold $r_{i,2}$, $r_{i,3}$, $\ldots$ or $r_{i,K}$ is higher than one elicited previously to ensure the monotonicity of the subjective complementary cumulative distribution function (Dominitz and Manski, 2004).

that $Q_{i,1}$ may be broadly informative, and too close to zero to be sure that it is uninformative: this answer may be partially informative.

To infer the type of $Q_{i,1}$, we first need to identify the measure of coherence. One can easily understand the basic aspect of such a measure –the distance between a prediction based on $r_{i,min}$ and $r_{i,max}$ and a response $Q_{i,1}$. The crucial elements are however the assumptions to obtain the prediction $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$. There is, in the SEE dataset that we will study, a key feature that permits to learn something about the informativeness of numerous –but not all– Q_1 under some weak assumptions: the algorithm which selects the thresholds does not insure that all of them belong to (r_{min}, r_{max}), and the first threshold r_1 is in practice often outside this interval, in particular lower than r_{min}. More precisely, $r_1 \notin (r_{min}, r_{max})$ for 72 percent of the respondents ($r_1 \leq r_{min}$ for 70 percent, and $r_1 \geq r_{max}$ for 2 percent).

This feature of the dataset is crucial. It is indeed tempting to impose assumptions strong enough to obtain a point-prediction $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$; as such, the measure of coherence will be point-identified. However, to do so, one needs to make a distributional assumption to calibrate a probability distribution based on $r_{i,min}$ and $r_{i,max}$, and interpret perhaps these responses in a too restrictive way. To fully understand, consider that one calibrates a continuous uniform distribution over the interval $[r_{i,min}, r_{i,max}]$. This distributional assumption permits to obtain a point prediction $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$. But it has no foundation and may drive some results of the inferential problem. Why one should choose a uniform distribution? Furthermore, this assumption implies that $[r_{i,min}, r_{i,max}]$ is the support of the subjective distribution. This interpretation seems natural. Some researchers, following Dominitz and Manski (1997, p.858), may think however that *"the phrases "lowest* [...]*" and "highest* [...]*" are too vague to warrant this formal interpretation* [...]", and consider that $r_{i,min}$ and $r_{i,max}$ only *"suggest the general region of the respondent's subjective support of"* the distribution.

With these elements in mind, I propose in Section 2 to identify the measure of coherence under extremely weak assumptions; I then explain how this measure can be used to infer the type of a probabilistic answer. In particular, I make no distributional assumption to obtain a prediction $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$, and capture the idea that the interval $(r_{i,min}, r_{i,max})$ is a suggestive support. These weak assumptions do not point-predict an answer $Q_{i,1}$ to a probabilistic question, but they do partially predict it, yielding a bound rather than a point prediction. As a consequence, these weak assumptions partially identify the measure of coherence, yielding a bound rather than a point estimate; see Manski (2003) for a textbook exposition of partial identification.

It is important to emphasize that because the measure of coherence is partially identified, it is not always possible to infer precisely the type of some answers; in some cases, fundamentally when $r_1 \in (r_{min}, r_{max})$, we cannot even infer if they are broadly informative, partially informative or uninformative because the identification region for the measure of coherence is distressingly wide. Note however that when we are not able to conclude, this is because the additional assumptions necessary to reduce the identification region are stronger. The approach developed in Section 2 is thus deliberately conservative.

Although this method is only applicable in the specific context of the SEE format for continuous variables, its application can shed light on the appropriateness of some practices in the related literature. Contrary to this literature, the measure of coherence does not impose that only some 50s may be uninformative, nor that the other answers, mainly multiples of ten other than $(0, 50, 100)$ –M10– (e.g., 10, 20) and multiples of 5 but not of ten –M5– (e.g., 85, 95), are always broadly informative. Kleinjans and van Soest (2014) propose a parametric model –a model with endogenous regime switching and unobserved regimes– to identify uninformative, partially informative and broadly informative answers. By assumption in their model, only some 50s may be uninformative (i.e., they may imply the interval [0,100]). Concerning the M10, Kleinjans and van Soest (2014) impose that the percent chance of the event is in the range [M10-5,M10+5] in the worst case; and the M5 are assumed to imply the interval [M5-2.5, M5+2.5] in the worst case (except 25 and 75 which can be rounded to a multiple of 25).[5] Manski and Molinari (2010, p.223) use the answers of a respondent to various expectations questions dealing with the same topic (e.g., the set of personal finance questions of the Health and Retirement Study -HRS-) to infer the extent to which he rounds probabilities. If this respondent provides at least one M10 or one M5 to this set of questions, then,

[5]Furthermore, they also impose some distributional assumptions to achieve point identification of the quantities of interest.

by assumption, his answers to the set of questions reflect relatively sharp expressions of beliefs implying maximal intervals of 10 and 5 percentage points. Last but not least, to resolve the unease that researchers have felt with the 50s, and instead of developing an identification strategy, some recent surveys directly reveal if a 50 is uninformative by adding a so-called "epistemic" follow-up question. For instance, if a respondent answers "50 percent" to the stock market probabilistic question of the HRS 08 described in Table 1, then he is asked if he provides this answer to express the belief that "*it is about equally likely that these mutual fund shares will increase in worth as it is that they will decrease in worth by this time next year, or* [if he is] *unsure about the chances*" (see Hudomiet *et al.*, 2011, p.398). The respondents who say they are unsure are then automatically excluded from the analysis.[6] By interpreting reported probabilities as interval data or by considering that only some 50s can be uninformative, all these methods provide more credible results than if these probabilities were taken at face value. But by imposing restrictions on which answers can be uninformative, partially or broadly informative, these methods are potentially subject to misclassification and important biases in the analysis. The empirical findings in Sections 3 and 4 suggest that it is the case.

Sections 3 and 4 focus on the responses to the stock market expectations questions described in Table 2. Previous studies analyzing these data include Dominitz and Manski (2011) and Gouret and Hollard (2011). But they do not really consider the problem of the 50s and other uninformative answers.[7] I choose these data because the topic is similar to the stock market questions of the HRS 08. Thus, it is sometimes a useful point of comparison (although the period studied and the sample of respondents differ). Furthermore, by analyzing these data, this paper also participates to a recent literature which tries to provide an empirical foundation for the study of expectations of equity returns (Dominitz and Manski, 2007, Hudomiet *et al.*, 2011, Hurd *et al.*, 2011, Kandel and Pearson,

[6] The remaining respondents, i.e., those who do not answer "50 percent" or those who answer "50 percent" but say that the shares are equally likely to increase or decrease in value to the epistemic follow-up question, are asked a second probabilistic question for an additional point on their subjective distribution. Assuming that each remaining respondent i has a normal subjective distribution $\mathcal{N}(\mu_i, \sigma_i^2)$, Hudomiet *et al.* (2011, pp.400-401) use the answers to the two probabilistic questions to derive their expected return (μ_i) and their perceived risk (σ_i).

[7] Dominitz and Manski (2011, p.365) note that the pervasiveness of rounding suggests that a researcher should interpret a probabilistic answer as providing an interval rather than a precise probability, and assume a $[40, 60]$ interval when the response is 50. However, they aptly explain that this interval is questionable because "*the responses of 50% that we observe are some mixture of rounded subjective probabilities and expressions of what Bruine de Bruin* et al. *(2000) call epistemic uncertainty*".

1995 and Vissing-Jorgensen, 2004).[8]

In Section 3, I show that the weak assumptions of Section 2 suffice to find that a majority of the 50s are uninformative, and that the 50s represent at most 70 percent of all the answers identified as uninformative. The other answers identified as uninformative are mainly *some* M10 and M5. Hence, by considering that only some 50s can be uninformative, one can miss an important share of uninformative answers. Furthermore, various M10 and M5 (around 10 percent of the sample) cannot be informative, but may be partially informative.

To gauge the relevance of considering as uninformative not only some 50s but *all* the answers identified as such, Section 4 proposes a cross-sectional analysis of the subjective distributions of equity returns in the year ahead conditional on the informativeness of Q_1. The recent literature which studies the expectations of equity returns use the subjective probabilities ($Q_{i,k}$, $k = 1, 2, \ldots, K$) to fit each respondent-specific normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$ (Dominitz and Manski, 2011, p.359, Gouret and Hollard, 2011, p.372, Hudomiet *et al.*, 2011, p.401, and Hurd *et al.*, 2011, p.425). I follow this normality assumption. If a researcher wants to study the cross-sectional distribution of (μ, σ) of the population who has meaningful expectations, and includes in his analysis the (μ, σ) based on uninformative Q_1, his analysis might be biased, given that the (μ, σ) based on uninformative Q_1 are meaningless. I provide some evidence that all the (μ, σ) based on uninformative Q_1, not only those based on uninformative 50s, strongly influence the cross-sectional distribution of (μ, σ).

Finally, Section 5 makes concluding remarks about some implications of my findings for the elicitation of probabilistic questions. I call particularly attention to the HRS 08 and its epistemic follow-up question, as well as a recent approach proposed by Manski and Molinari (2010, p.229) in the case of binary discrete variables which consists in eliciting ranges of probabilities rather than precise probabilities for future events.

[8] Kandel and Pearson (1995) and Vissing-Jorgensen (2004) differ from the the other papers in that they use surveys which seek point predictions of equity returns rather than probabilistic expectations. Kandel and Pearson (1995) analyze the point predictions of financial analysts for the earning of specific firms. Vissing-Jorgensen (2004) considers the UBS and Gallup surveys which ask to some American households possessing more than $10,000 in financial assets their point prediction of the rate of return of the stock market in the year ahead and in the next ten years.

2 Identifying uninformative answers

This Section explains in generality how the specific format of probabilistic questions described in the Introduction and used in the SEE to elicit the distribution of continuous variables permit to identify uninformative answers. Subsection 2.1 shows how this format can be useful to obtain a measure of coherence. I emphasize the importance of making weak instead of strong and untenable assumptions; as a result I obtain a partially identified measure of coherence. Subsections 2.2 and 2.3 explain how this partially identified measure of coherence permits to infer if an answer is uninformative.

2.1 A partially identified measure of coherence

If a respondent i has a precise subjective distribution in mind concerning R_i, he should use the same underlying subjective distribution to answer the preliminary questions and the sequence of probabilistic questions. I thus consider that a respondent has potentially two subjective distributions. The first one is used to answer $r_{i,min}$ and $r_{i,max}$, the lowest and highest possible values of this distribution; I denote by $\widetilde{\mathrm{P}}\left(R_i > r\right)$ the subjective probability of the event $R_i > r$ associated to this first distribution. A second subjective distribution is used by the respondent to answer the sequence of probabilistic questions $Q_{i,k} \equiv \mathrm{P}\left(R_i > r_{i,k}\right)$, $k = 1, 2, \ldots, K$. I focus on $Q_{i,1}$ given that this probabilistic answer determines the probabilistic answers for the subsequent thresholds. Suppose that a researcher has *full knowledge* of the first distribution used by the respondent to answer $r_{i,min}$ and $r_{i,max}$. Then, he can assess the subjective probability of the event $R_i > r_{i,1}$ according to this distribution, i.e., $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$ which is a prediction of $Q_{i,1}$. The distance $d_{i,1}$ between $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$ and $Q_{i,1}$ is then a measure of coherence at the first threshold:

$$d_{i,1} \equiv \left|\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right) - Q_{i,1}\right| \tag{1}$$

All the difficulty to identify $d_{i,1}$ is to assess the prediction $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$. A researcher has only a *partial knowledge* of the first subjective distribution, not a full knowledge: he only knows by definition $r_{i,min}$ and $r_{i,max}$, the lowest and highest possible values of this distribution. Furthermore, Dominitz and Manski (1997, p.860) consider that some respondents may "*associate the phrases "lowest possible" and "highest possible" with low and high probabilities*". I thus consider the conclusions that a researcher can draw concerning the distance $d_{i,1}$ if he only makes the following assumption:

Assumption 1. $0 \leq \widetilde{\mathrm{P}}\left(R_i \leq r_{i,min}\right) \leq \alpha$ and $0 \leq \widetilde{\mathrm{P}}\left(R_i \geq r_{i,max}\right) \leq \alpha$

To understand Assumption 1, consider that the responses $r_{i,min}$ and $r_{i,max}$ of respondent i are the $\{\alpha_i, (1-\alpha_i^{'})\}$-quantiles of his subjective distribution, so $\widetilde{\mathrm{P}}\left(R_i \leq r_{i,min}\right) = \alpha_i$ and $\widetilde{\mathrm{P}}\left(R_i \geq r_{i,max}\right) = \alpha_i^{'}$. However, the researcher does not know if α_i and $\alpha_i^{'}$ are equal to zero (i.e., $r_{i,min}$ and $r_{i,max}$ are the minimum and maximum values of R_i) or are low positive probabilities as highlighted by Dominitz and Manski. Instead of making a point assumption concerning the values of α_i and $\alpha_i^{'}$, he thus assumes that α_i and $\alpha_i^{'}$ can take any value in the interval $[0, \alpha]$. I will carry out the main empirical analysis assuming that $\alpha = 0.10$, which means that α_i and $\alpha_i^{'}$ can be any value in the interval $[0, 0.10]$. Note that all researchers who would have considered a lower value for α must accept $\alpha = 0.10$. In particular, if a researcher considers that $[r_{i,min}, r_{i,max}]$ is the support of the subjective distribution, so $\alpha = \alpha_i = \alpha_i^{'} = 0$, he must agree on $0 \leq \alpha_i \leq 0.10$; the same for $\alpha_i^{'}$. I will also discuss how the results change if one chooses the even weaker assumption $\alpha = 0.20$ (see, in particular, Remark 3 in Subsection 2.3).

By making only Assumption 1, the prediction $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$ in Equation 1 is an interval, not a point. Three cases appear: (i.) $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1} \leq r_{i,min}\right)\right] = [1-\alpha, 1]$, where $\mathrm{H}[\cdot]$ denotes the region for the quantity in brackets (i.e., all possible values that the quantity in brackets can take); (ii.) $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1} \geq r_{i,max}\right)\right] = [0, \alpha]$; (iii.) $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1} \in (r_{i,min}, r_{i,max})\right)\right] = [0, 1]$. As a consequence, Assumption 1 reveals that $d_{i,1}$ lies in the following identification regions in cases (i.), (ii.) and (iii.):

$$\mathrm{H}\left[d_{i,1} | Q_{i,1}, r_{i,1} \leq r_{i,min}\right] = \left[\max\{0, 1 - Q_{i,1} - \alpha\}, \max\{1 - Q_{i,1}, |1 - Q_{i,1} - \alpha|\}\right] \tag{2}$$

$$\mathrm{H}\left[d_{i,1} | Q_{i,1}, r_{i,1} \geq r_{i,max}\right] = \left[\max\{0, Q_{i,1} - \alpha\}, \max\{Q_{i,1}, |Q_{i,1} - \alpha|\}\right] \tag{3}$$

$$\mathrm{H}\left[d_{i,1} | Q_{i,1}, r_{i,1} \in (r_{i,min}, r_{i,max})\right] = \left[0, \max\{Q_{i,1}, 1 - Q_{i,1}\}\right] \tag{4}$$

Although the expressions seem to be cumbersome, the width of the identification region for $d_{i,1}$ cannot be higher than α (and cannot be less than $\frac{\alpha}{2}$) when $r_{i,1} \leq r_{i,min}$ (Equation 2) or when $r_{i,1} \geq r_{i,max}$ (Equation 3). When $r_{i,1} \in (r_{i,min}, r_{i,max})$, the width of the interval is much larger (Equation 4). To see that, remark first that in that case the lower bound is 0 whatever the response $Q_{i,1}$. Second, when $Q_{i,1} = 0.50$, the identification region for $d_{i,1}$ is the interval $[0, 0.50]$. The width of the interval is larger for any other response $Q_{i,1} \neq 0.50$, given that $Q_{i,1} = 0.50$ is the answer with the smallest upper bound (0.50).

Illustration 1. To clearly understand, consider Table 3 which provides the identification region for $d_{i,1}$ conditional on the threshold $r_{i,1}$ and the response $Q_{i,1}$, assuming $\alpha = 0.10$. Column [1] considers the case when $r_{i,1} \leq r_{i,min}$, so when the prediction is [0.90,1]. If the respondent answers, e.g., $Q_{i,1} = 0.50$, the identification region for $d_{i,1}$ is [0.40,0.50]. It means that the distance can take any value in this interval. If he answers, e.g., $Q_{i,1} = 0.95$, the distance $d_{i,1}$ can take any value in the interval [0,0.05]. More generally, one can see that the width of the intervals cannot be higher than 0.10 and cannot be less than 0.05 when $r_{i,1} \leq r_{i,min}$. Column [2] considers the case when $r_{i,1} \geq R_{i,max}$, so when the prediction is [0,0.10]. Again, one can see that the width of the intervals cannot be higher than 0.10 and cannot be less than 0.05. Column [3] considers the case when $r_{i,1} \in (r_{i,min}, r_{i,max})$, so when the prediction is [0,1]. If the respondent answers, e.g., $Q_{i,1} = 0.50$, the distance $d_{i,1}$ can take any value in the interval [0,0.50]. If he answers, e.g., $Q_{i,1} = 0$, the distance $d_{i,1}$ can be any value in the interval [0,1]. More generally, when $r_{i,1} \in (r_{i,min}, r_{i,max})$, the lower bound is always 0 and the interval has a width of at least 0.50.

Remark 1. One should keep two things in mind concerning the bounds obtained in Illustration 1. First, and as often highlighted in partial identification analysis (see, e.g., Manski, 2003, p.3), these bounds establish a domain of consensus among researchers about the value of the distance $d_{i,1}$. For instance, consider a respondent who answers $Q_{i,1} = 0.50$. All researchers who accept Assumption 1 and $\alpha = 0.10$ must agree that the distance is neither less than 0.40 nor greater than 0.50 when $r_{i,1} \notin (r_{i,min}, r_{i,max})$, while they must agree that the distance is in between range 0 to 0.50 when $r_{i,1} \in (r_{i,min}, r_{i,max})$. Second, the bounds clearly show that different assumptions which could permit to point-identify $d_{i,1}$ when $r_{i,1} \in (r_{i,min}, r_{i,max})$ can give very different -so weakly credible- results because the bounds are uncomfortably large in that case. For instance, if $Q_{i,1} = 0.70$, researchers making different point-identifying assumptions may reach very different conclusions about the value of $d_{i,1}$ within the interval [0,0.70].

2.2 The inferential problem under Assumption 1: General considerations

What the bounds of the measure of coherence computed under Assumption 1 can reveal about the informativeness of an answer? I ask whether the answer $Q_{i,1}$ provided by the respondent i is coherent with the prediction $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$. Consider two critical values $\underline{d}$ and $\overline{d}$, such that $0 < \underline{d} < \overline{d} < 1$. If $d_{i,1}$ is high -say $d_{i,1} \in [\overline{d}, 1]$-, the answer $Q_{i,1}$ is clearly uninformative. If $d_{i,1}$ is close to zero -say $d_{i,1} \in [0, \underline{d}]$-, $Q_{i,1}$ may be broadly informative. In-between, $d_{i,1}$ is too far from zero to say that $Q_{i,1}$ may be broadly informative, and too close to zero to claim that it is uninformative -say $d_{i,1} \in (\underline{d}, \overline{d})$-: this answer may be partially informative. There are however two difficulties. One difficulty is to define plausible values for $\underline{d}$ and $\overline{d}$; I will deal with this point in the next Subsection. The other difficulty that I address here is that $d_{i,1}$ is not point-identified, but partially identified. One can infer the type of an answer if the lower and upper bounds of the identification region for $d_{i,1}$ are in the same critical region. But if it is not the case, it is impossible to identify precisely the type of an answer.

Formally, let $d_{i,1,l}$ and $d_{i,1,u}$ be the lower and upper bounds of the identification region for $d_{i,1}$, as defined in Equations 2, 3 and 4; so $\mathrm{H}\left[d_{i,1} | Q_{i,1}, r_{i,1}\right] = [d_{i,1,l}, d_{i,1,u}]$. With the partially identified measure of coherence based on Assumption 1, the data will be classified as follows:

(i.) If $\underline{d} \geq d_{i,1,u} > d_{i,1,l} \geq 0$, $Q_{i,1}$ is identified as broadly informative (the "Broadly informative" category).

(ii.) If $\overline{d} > d_{i,1,u} > d_{i,1,l} > \underline{d}$, $Q_{i,1}$ is identified as partially informative (the "Partially informative" category).

(iii.) If $d_{i,1,l} \geq \overline{d}$, $Q_{i,1}$ is identified as uninformative (the "Uninformative" category).

(iv.) If $d_{i,1,u} \geq \overline{d} > d_{i,1,l} > \underline{d}$, $Q_{i,1}$ cannot be broadly informative, but it is impossible to say if it is uninformative or partially informative (the "Uninformative or partially informative" category).

(v.) If $\overline{d} > d_{i,1,u} > \underline{d} \geq d_{i,1,l}$, $Q_{i,1}$ may be broadly or partially informative (the "Broadly or partially informative" category).

(vi.) If $d_{i,1,u} \geq \overline{d} > \underline{d} \geq d_{i,1,l}$, the type of $Q_{i,1}$ is unknown: it is impossible to infer if $Q_{i,1}$ is broadly informative, partially informative or uninformative (the "Unknown" category).

Illustration 2. Consider again Table 3, and assume furthermore that the critical values are $\underline{d} = 0.10$ and $\overline{d} = 0.40$. First, consider the case when $r_{i,1} \leq r_{i,min}$ (Column [1] in Table 3), so when the prediction is [0.90,1]. If the respondent answers, e.g., $Q_{i,1} = 0.50$, the identification region for $d_{i,1}$ is [0.40,0.50]. This answer is clearly uninformative, given that the lower and upper bounds are in the critical region [0.40,1] which defines uninformative answers. If the respondent answers, e.g., $Q_{i,1} = 0.99$, the identification region for $d_{i,1}$ is [0,0.09]. This answer may be broadly informative because the lower and upper bounds are in the critical region [0,0.10] which defines broadly informative answers. Sometimes it is however difficult to identify precisely the type of the response: if, e.g., $Q_{i,1} = 0.55$, the identification region for $d_{i,1}$ is [0.35,0.45]; the lower bound 0.35 is in the critical region (0.10,0.40) which defines partially informative answers, but the upper bound is in the critical region [0.40,1] which defines uninformative answers. Therefore it is impossible to say if the response is uninformative or partially informative; but at least, we can reject the possibility that it is broadly informative. The situation is worse when $r_{i,1} \in (r_{i,min}, r_{i,max})$ because we cannot infer if an answer is uninformative, partially informative or broadly informative; in other words, the type of the answer is unknown. For instance, if the respondent answers $Q_{i,1} = 0.50$, the identification region for $d_{i,1}$ is [0,0.50] as shown in Column [3]. The lower bound of the identification region for $d_{i,1}$ is 0 so the answer may be broadly informative; but the upper bound is 0.50, so the answer may be uninformative too; and given that $d_{i,1}$ can take any value in this interval, the answer may also be partially informative.

Remark 2. I still have to define plausible critical values $\underline{d}$ and $\overline{d}$, but Illustration 2 suggests that when $r_{i,1} \notin (r_{i,min}, r_{i,max})$, the width of the identification region for $d_{i,1}$ is sufficiently small to learn something about the informativeness of an answer (although it is not always possible to infer precisely the type of an answer because $\alpha > 0$). However, when $r_{i,1} \in (r_{i,min}, r_{i,max})$, one learns nothing because the width of the identification region for $d_{i,1}$ is too wide. By making only Assumption 1, I sacrifice the possibility to make some conclusions (when $r_{i,1} \in (r_{i,min}, r_{i,max})$), but those that I will make (when $r_{i,1} \notin (r_{i,min}, r_{i,max})$) possess credibility. Fortunately, $r_{i,1} \notin (r_{i,min}, r_{i,max})$ for an important share of respondents in the empirical analysis of Section 3, around 70 percent.

2.3 Defining plausible critical values $\underline{d}$ and $\overline{d}$

Remember that $0 < \underline{d} < \overline{d} < 1$. Given that my two main objectives are to learn the share of 50s identified as uninformative and to see if other answers are identified as uninformative, the choice of $\overline{d}$ is the most important and I begin by explaining it.

To identify some 50s as uninformative, it is necessary to make the following additional assumption:

Assumption 2. $\overline{d} \leq 0.50 - \alpha$

Assumption 2 is an identification condition, in the sense that without this assumption no 50s will be identified as uninformative. Indeed, under Assumption 1, $\mathrm{H}\left[d_{i,1}|Q_{i,1} = 0.50, r_{i,1} \notin (r_{i,min}, r_{i,max})\right] = [0.50 - \alpha, 0.50]$ (see Equations 2 and 3), and $\mathrm{H}\left[d_{i,1}|Q_{i,1} = 0.50, r_{i,1} \in (r_{i,min}, r_{i,max})\right] = [0, 0.50]$ (see Equation 4). Hence, to have some 50s identified as uninformative, it is necessary that $\overline{d} \leq 0.50 - \alpha$. Nevertheless, one should also ask if at least $\overline{d} = 0.50 - \alpha$, say $\overline{d} = 0.40$ if we assume $\alpha = 0.10$, appears as sufficiently high as a critical value. If, e.g., $r_{i,1} \leq r_{i,min}$ and $Q_{i,1} = 0.50$, the smallest possible value that $d_{i,1}$ can take is $d_{i,1,l} = 0.40$. Even if $d_{i,1} = 0.40$, this distance between the prediction and the response seems too high to find any information, even partial.

We then obtain the following Proposition:

Proposition 1. *Consider a measure of coherence at the first threshold partially identified via Assumption 1, and Assumption 2 holds. Then:*

(i.) the share of 50s identified as uninformative is similar for all combinations of $\overline{d}$ and α which

satisfy Assumption 2:

$$\frac{\sum_{i=1}^{N} \mathbf{1}\left[r_{i,1} \notin (r_{i,min}, r_{i,max}), Q_{i,1} = 0.50\right]}{\sum_{i=1}^{N} \mathbf{1}\left[Q_{i,1} = 0.50\right]} \tag{5}$$

where $\mathbf{1}[\mathcal{A}]$ is an indicator function which is equal to 1 if $\mathcal{A}$ is true, and equal to 0 otherwise, and N is the number of observations;

(ii.) by choosing the precise value $\overline{d} = 0.50 - \alpha$, the following share provides an upper bound on the share of 50s among the answers identified as uninformative:

$$\frac{\sum_{i=1}^{N} \mathbf{1}\left[r_{i,1} \notin (r_{i,min}, r_{i,max}), Q_{i,1} = 0.50\right]}{\sum_{i=1}^{N} \mathbf{1}\left[d_{i,1,l} \geq \overline{d}\right]} \tag{6}$$

(iii.) by choosing the precise value $\overline{d} = 0.50 - \alpha$, the following share provides a lower bound on the share of uninformative answers other than 50s in the sample:

$$\frac{\sum_{i=1}^{N} \mathbf{1}\left[d_{i,1,l} \geq \overline{d}, Q_{i,1} \neq 0.50\right]}{N} \tag{7}$$

Part (i.) of Proposition 1 is trivial but crucial: it says that the share of 50s identified as uninformative is similar for all $\overline{d}$ and α which satisfy Assumption 2. The 50s identified as uninformative are always those answered to a threshold $r_{i,1} \notin (r_{i,min}, r_{i,max})$. If $r_{i,1} \notin (r_{i,min}, r_{i,max})$, the identification region for $d_{i,1}$ is $[0.50 - \alpha, 0.50]$, so the lower and upper bounds of $d_{i,1}$ are in the critical region $[\overline{d}, 1]$ for all $\overline{d} \leq 0.50 - \alpha$. If $r_{i,1} \in (r_{i,min}, r_{i,max})$, the identification region for $d_{i,1}$ is $[0, 0.50]$, so the type of $Q_{i,1} = 0.50$ is unknown for all $\overline{d} \leq 0.50 - \alpha$. Given that it is impossible to learn the type of a 50 answered to a threshold $r_{i,1} \in (r_{i,min}, r_{i,max})$, note that the share given by Equation 5 is a lower bound on the share of 50s which are uninformative.

Part (ii.) says that as far as one agrees on a critical value which satisfies Assumption 2, choosing $\overline{d} = 0.50 - \alpha$ permits to obtain an upper bound on the share of 50s among the answers identified as uninformative. Indeed the denominator of Equation 6 is lower if $\overline{d} = 0.50 - \alpha$ than if $\overline{d} < 0.50 - \alpha$. Choosing such an upper bound for this share follows the conservative approach of this paper. If we obtain as an upper bound, e.g., 60 percent, we will know that researchers who focus only on the informativeness of the 50s miss an important share -at least 40 percent- of the identified uninformative answers. A potential problem with this share is that the 50s identified as uninformative are a lower bound on the uninformative 50s, as previously mentioned. If so, the share given by Equation 6 may underestimate the share of the uninformative 50s among the uninformative answers. Hence, to have a clearer idea on the importance of the uninformative responses other than 50s, I will also consider the share given by Equation 7 in Part (iii.). This share of uninformative answers other than 50s in the sample is a lower bound for two reasons. First, this share considers the responses other than 50s identified as uninformative. This is a lower bound on the share of responses other than 50s which are uninformative (because it is possible that some of the $Q_1 \neq 0.50$ classified as "Unknown" or classified as "Uninformative or partially informative" are uninformative). Second, the numerator is lower if $\overline{d} = 0.50 - \alpha$ than if $\overline{d} < 0.50 - \alpha$.

Remark 3. As explained previously, it is logical to choose a value not too small for α because $(r_{i,min}, r_{i,max})$ only suggests the support of the distribution. And when $r_{i,1} \notin (r_{i,min}, r_{i,max})$, it is impossible to learn precisely if some answers are uninformative or partially informative (the "Uninformative or partially informative" category) and broadly or partially informative (the "Broadly or partially informative" category) because $\alpha > 0$ (if $\alpha = 0$, the measure of coherence will be point-identified when $r_{i,1} \notin (r_{i,min}, r_{i,max})$, so it will be always possible to learn precisely the type of an answer). Nevertheless remark that as far as $\overline{d} = 0.50 - \alpha$ the precise value of α is not important for the three percentages of Proposition 1 which concern the answers identified as uninformative. Indeed, the set of answers identified as uninformative is the same for all combinations of α and $\overline{d}$ which satisfy $\overline{d} = 0.50 - \alpha$. If $\alpha = 0.10$ and $\overline{d} = 0.40$, on can see in Table 3 that the set of Q_1 identified as uninformative is given by $\{Q_1 \in [0, 0.50] | r_1 \leq r_{min}\} \bigcup \{Q_1 \in [0.50, 1] | r_1 \geq r_{max}\}$. It is easy to see that this set is similar if I choose $\alpha = 0$ and $\overline{d} = 0.50$ or $\alpha = 0.20$ and $\overline{d} = 0.30$.[9]

[9] For a formal proof, remark first that if $\overline{d} = 0.50 - \alpha$, a response is identified as uninformative when $d_{i,1,l} \geq 0.50 - \alpha$. Then consider the three following cases: (i.) $r_{i,1} \leq r_{i,min}$, (ii.) $r_{i,1} \geq r_{i,max}$, and (iii.) $r_{i,1} \in (r_{i,min}, r_{i,max})$. (i.)

For the critical value $\underline{d}$, I choose $\underline{d} = 0.10$, for two reasons. First, and as already noticed in the Introduction, the literature usually assumes that the answers other than $(0, 0.50, 1)$ cannot imply an interval with a width of more than 10 percentage points. If I consider $\underline{d} = 0.10$, then a probability $Q_{i,1}$ will be identified as partially informative if its difference with the answer to the pair of preliminary questions is at least 10 percentage points, i.e., $d_{i,1,l} \geq 0.10$. Second, since Dominitz and Manski (1997), it has been repeatedly observed that respondents tend to report values at 1 percent intervals at the extremes (i.e., 0.01, 0.02, 0.03 and 0.97, 0.98, 0.99). It is often considered that these values are broadly informative.[10] Note that if they are, one can see in Table 3 that the values 0.01, 0.02, and 0.03 will be answered to a threshold $r_{i,1} \geq r_{i,max}$ (and will be identified as broadly informative given that $\underline{d} = 0.10$) or to a threshold $r_{i,1} \in (r_{i,min}, r_{i,max})$ (but in this case we will not be able to infer their type); concerning the values 0.97, 0.98 and 0.99, they will be answered to a threshold $r_{i,1} \leq r_{i,min}$ (and will be identified as broadly informative) or to a threshold $r_{i,1} \in (r_{i,min}, r_{i,max})$ (again, we will not be able to infer their type in that case).

3 Application

This Section revisits SEE data on stock market expectations that has been previously analyzed by Dominitz and Manski (2011) and Gouret and Hollard (2011). Subsection 3.1 describes the data. Subsection 3.2 describes the response patterns by type of answers. We assume that the critical values are $\underline{d} = \alpha$ and $\overline{d} = 0.50 - \alpha$, and choose $\alpha = 0.10$, to classify the respondents in the following categories: "Broadly informative", "Partially informative", "Uninformative", "Partially or broadly informative", "Uninformative or partially informative", "Unknown".

When $r_{i,1} \leq r_{i,min}$, $d_{i,1,l} = \max\{0, 1 - Q_{i,1} - \alpha\}$ (see Equation 2). Hence, when $r_{i,1} \leq r_{i,min}$, $Q_{i,1}$ is identified as uninformative if $\max\{0, 1 - Q_{i,1} - \alpha\} \geq 0.50 - \alpha$, i.e., if $Q_{i,1} \leq 0.50$. (ii.) When $r_{i,1} \geq r_{i,max}$, $d_{i,1,l} = \max\{0, Q_{i,1} - \alpha\}$ (see Equation 3). Hence, when $r_{i,1} \geq r_{i,max}$, $Q_{i,1}$ is identified as uninformative if $\max\{0, Q_{i,1} - \alpha\} \geq 0.50 - \alpha$, i.e., if $Q_{i,1} \geq 0.50$. (iii.) Lastly, when $r_{i,1} \in (r_{i,min}, r_{i,max})$, $d_{i,1,l} = 0$ and $d_{i,1,u} \geq 0.50$ (see Equation 4). Hence, when $r_{i,1} \in (r_{i,min}, r_{i,max})$, the type of $Q_{i,1}$ is always unknown.

[10] For instance, Manski and Molinari (2010, p.219) write that "*when someone states "3%," one might reasonably infer that the person is rounding to the nearest 1%.*"

3.1 The Data

I apply the partially identified measure of coherence to study the patterns of responses of data drawn for the three waves of the SEE where questions on equity returns were asked: wave 12 (where interviews were conducted in the period July 1999-November 1999), wave 13 (February 2000-May 2000) and wave 14 (September 2000-March 2001). The SEE was administered through WISCON, a national computer-assisted telephone interview survey conducted by the University of Wisconsin Survey Center; Dominitz and Manski (1997, 2004) describe the basic features of the survey. 1651 respondents were interviewed in these three cross-sections (547 in wave 12, 465 in wave 13, and 639 in wave 14). Of these 1651 respondents, 1255 answered the two preliminary questions eliciting the lowest and the highest possible values $r_{i,min}$ and $r_{i,max}$ of the investment in the year ahead (to be more precise 1284 answered $r_{i,min}$ and 1255 answered $r_{i,min}$ and $r_{i,max}$). Of these 1255 respondents, 1231 provided an answer $Q_{i,1}$ to the first probabilistic question (415 in wave 12, 342 in wave 13 and 474 in wave 14). The bounds of the measure of coherence at the first threshold can be computed for these 1231 respondents, so for 75 percent of the respondents $\left(\simeq \frac{1231}{1651}\right)$.[11] I refer to this sample as the Preliminary sample. Note also that the Preliminary sample is not the sample that researchers usually use. Researchers usually use the responses to the sequence of probabilistic questions ($Q_{i,k}$, $k = 1, 2, 3, 4$) to fit respondent-specific subjective distributions. I will be more explicit on this point in Section 4, but for various reasons, we can fit these subjective distributions for only 979 respondents (340 in wave 12, 264 in wave 13 and 375 in wave 14). So I will also consider the response pattern at the first threshold of this subsample of 979 respondents called the Final sample. This Final sample represents about 60 percent of the respondents $\left(\simeq \frac{979}{1651}\right)$.

Remember that Assumptions 1 and 2 do not permit to infer the type of $Q_{i,1}$ when $r_{i,1} \in (r_{i,min}, r_{i,max})$, as Remark 2 highlights. It happens that $r_1 \notin (r_{min}, r_{max})$ for 878 of the 1231

[11] Remark that of the 420 respondents ($= 1651 - 1231$) for whom we cannot consider a measure of coherence, 396 did not provide a response $r_{i,min}$ or $r_{i,max}$ to the preliminary questions, so 24 percent of the respondents $\left(\simeq \frac{396}{1651}\right)$ (90 percent of them provided a "don't know" $\left(\simeq \frac{354}{396}\right)$, and the other 10 percent refused to respond). These nonresponses, in particular the "don't know", already indicate that responding to the preliminary questions may be a difficult task, and that many people have little knowledge of the stock market. Note that the HRS 06 and 08 which directly asked the stock market probabilistic question described in Table 1 have similar nonresponse rates. Of the 16754 respondents of the HRS 06, 4017 did not provide a probabilistic response (3865 "don't know" and 152 refusals), so 24 percent. Of the 15727 respondents of the HRS 08, 3099 did not provide a probabilistic response (2998 "don't know" and 101 refusals), so 20 percent.

respondents of the Preliminary sample.[12] So we will learn something about the informativeness of 72 percent of the responses to the first threshold in the Preliminary sample ($\simeq \frac{878}{1231}$), as already mentioned. And the type of the 353 remaining answers ($= 1231 - 878$) will be unknown (111 in wave 12, 91 in wave 13 and 151 in wave 14). Concerning the Final sample, $r_1 \notin (r_{min}, r_{max})$ for 699 of the 979 respondents, so 72 percent ($\simeq \frac{699}{979}$).[13] The type of the 280 remaining answers ($= 979 - 699$) will be unknown (93 in wave 12, 64 in wave 13 and 123 in wave 14).

3.2 Response patterns

Table 4 considers the Preliminary sample, and provides a picture of the response patterns at the first threshold by type and by wave. The column labeled $Q_1 = 0.50$ gives the number of respondents who answered "50 percent". The column labeled $Q_1 = 1$ gives the number of respondents who answered "100 percent". The column labeled $Q_1 = 0$ gives the number of respondents who answered "0 percent". Column M10 gives the number of respondents who answered a multiple of ten other than (0,50,100): e.g., 20, 30 or 90 percent. Column M5 gives the number of respondents who answered a multiple of 5 but not of ten: e.g., 5,15 or 85 percent. The column labeled "Other" gives the number of respondents who did not answer a multiple of 5: e.g., 99, 98 or 1 percent. Three facts appear:

(i.) First, a majority of the 50s are uninformative. More precisely, of the 80 respondents who answered $Q_1 = 0.50$ in wave 12, at least 56 provided an uninformative 50. Hence, computing the share given by Equation 5 in Proposition 1, we find that at least 70 percent of the 50s are uninformative ($\simeq \frac{56}{80}$). This first result is robust across waves: at least 66 percent of the $Q_1 = 0.50$ are uninformative in wave 13 ($\simeq \frac{44}{66}$) and 55 percent in wave 14 ($\simeq \frac{50}{90}$).

(ii.) Second, the 50s represent a majority of -but not all- the answers identified as uninformative. More precisely, in wave 12, of the 86 answers identified as uninformative, 56 are 50s; the second identified uninformative answers most often provided are some M10 with 14 answers. Computing the share given by Equation 6 in Proposition 1, the 50s represent 65 percent ($\simeq \frac{56}{86}$)

[12] More precisely, $r_1 \leq r_{min}$ for 851 respondents, and $r_1 \geq r_{max}$ for 27 respondents.
[13] More precisely, $r_1 \leq r_{min}$ for 689 respondents and $r_1 \geq r_{max}$ for 10 respondents.

of the identified uninformative answers in wave 12. Similar percentages are found for wave 13 ($0.62 \simeq \frac{44}{71}$) and wave 14 ($0.60 \simeq \frac{50}{83}$).

(iii.) Third, computing the share given by Equation 7 in Proposition 1, the uninformative answers other than 50s represent at least 7.2 percent of the answers in wave 12 ($\simeq \frac{3+14+9+4}{415}$). Similar percentages are found for wave 13 ($0.079 \simeq \frac{1+3+18+2+3}{342}$) and wave 14 ($0.07 \simeq \frac{7+17+6+3}{474}$).

The fact (i.) highlights that researchers are right to feel embarrassed with the 50s. A majority of them are clearly uninformative whatever the wave considered. The percentages provided -between 55 and 70 percent depending on the wave- are relatively high despite the fact that they are lower bounds for a researcher who accepts Assumptions 1 and 2. They are however widely believable. As a way of comparison, consider the stock market probabilistic question of the HRS 08 described in Table 1, as well as the answers to the epistemic follow-up question. The wording of the probabilistic question, the period studied (February 2008 through February 2009) and the sample considered (Americans over the age of 50) in the HRS 08 are different, but the topic is similar; so it can give an idea of the plausibility of the percentages we have found. Of the 15727 respondents of the HRS 08, 3399 answered 50 percent. Of these 3399 respondents, 2124 answered that they were "*unsure*" at the epistemic follow-up question, so 62.48 percent of the 50s. This percentage is very similar to those that we have found.

The fact (ii.) suggests however that by focusing only on the informativeness of the 50s, researchers can miss an important share of uninformative answers. Although the 50s always represent a majority of the identified uninformative answers -between 60 and 65 percent depending on the wave-, other answers are identified as uninformative as well, mainly M10 and M5 as shown in Table 4. If one had focused only on the informativeness of the 50s, he would have missed between 35 and 40 percent of the identified uninformative answers. The fact (iii.) confirms that the uninformative answers other than 50s are substantial, given that they represent at least 7 percent of all the answers in the different waves.

Final sample. The same exercise was done with the Final sample. Table 5 provides a picture of the response patterns at the first threshold by type and by wave. Columns [5]-[8] of Table 6 present a summary of the important percentages of Proposition 1. We find again that a majority of the 50s are uninformative (Line [A] in Table 6). The 50s represent 74 percent of the identified uninformative answers (Line [B] in Table 6), so around 26 percent of the identified uninformative answers differ from 50s. This is quite important, although less than in the Preliminary sample. The last percentage of interest, the share of uninformative answers other than 50s, represents at least 4 percent of the 979 respondents of the Final sample. Again, this percentage is lower than the one found for the Preliminary sample, but it still shows that by focusing only on the informativeness of the 50s, one can miss some uninformative answers, mainly some M10 and M5 (see Table 5).

Additional results. There are some additional interesting results in the response patterns of Tables 4 and 5. I have already said that some M10 and M5 are clearly uninformative. This is in sharp contrast with the literature (e.g., Manski and Molinari, 2010, p.223) which usually assumes that these responses cannot imply an interval with a width of more than 10 percentage points, i.e., they are broadly informative. Remark that some M10 and M5 are also classified as "Partially informative" (e.g., 33 in wave 12 in Table 4) and few of them as "Uninformative or partially informative" (e.g., 9 in wave 12 in Table 4), i.e., their difference with the answer to the pair of preliminary questions is at least 10 percentage points, i.e., $d_{i,1,l} \geq 0.10$. These answers represent around 10 percent of the samples (e.g., $(33+9)/415$ in the Preliminary sample and $(27+8)/340$ in the Final sample of wave 12), and again, this result is in contrast with the literature. Lastly, researchers usually assume that the answers which are not a multiple of 5, e.g., 1, 2 or 99 percent, are broadly informative. Most of our results are in line with this assumption, given that of the few respondents who provide these answers to a threshold $r_1 \notin (r_{min}, r_{max})$, most of them are usually identified as informative. Some exceptions appear however in the Preliminary sample: few of these answers are clearly uninformative whatever the wave (e.g., 4 in wave 12 in Table 4).

4 Analyzing the cross-sectional distribution of (μ, σ) conditional on the informativeness of Q_1

This Section gauges the relevance of considering as uninformative not only some 50s but *all* the answers identified as such. Following the literature which studies the expectations of equity returns, I use the sequence of subjective probabilities ($Q_{i,k},\ k = 1, 2, 3, 4$) to fit each respondent-specific normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$. Note that the identification process to infer if a $Q_{i,1}$ is uninformative does not depend on it; it only depends on Assumption 1 and the conservative choice of fixing $\overline{d} = 0.50 - \alpha$.[14] The (μ, σ) based on uninformative Q_1 are meaningless. Suppose that one wants to learn the cross-sectional distribution of (μ, σ) of the population who has meaningful expectations but only excludes the (μ, σ) based on uninformative $Q_1 = 0.50$ (as Hudomiet *et al.* (2011, p.398) do by using the HRS 08 and its epistemic follow-up question; see Footnote 6). Given that he does not consider the fact that other responses are uninformative, is the cross-sectional distribution of (μ, σ) still biased?[15]

Before to step any further, two remarks are in order. The focus is mainly on the uninformative Q_1, and how they can bias the cross-sectional distribution of (μ, σ). But one should keep in mind that the (μ, σ) based on Q_1 classified as "Partially informative" or as "Uninformative or partially informative" are also meaningless: these Q_1 probably do not indicate a lack of knowledge of the stock market as "don't know" and uninformative answers, but they reflect at least a partial knowledge of the stock market. The analysis will also show how the (μ, σ) based on these answers influence the cross-sectional analysis. The second remark has to do with the category "Unknown" and the weakness

[14] Instead of assuming a normal distribution to fit each respondent-specific distribution, it would be possible to use a more flexible (and easily implementable) method proposed recently by Bellemare *et al.* (2012) which maintains much weaker assumptions on the shape of each underlying distribution. Such approach is however beyond the scope of this paper. The objective is to identify uninformative Q_1 under the weak assumptions of Section 2, and this Section shows how the results found in the literature which fits respondent-specific normal distributions may change if one controls for the informativeness of Q_1 (identified under the weak assumptions of Section 2).

[15] If there is no doubt that the uninformative Q_1 are meaningless, one may note however that those who provide these responses (like those who answer "don't know") also take decision to hold or not stocks. Estimating an econometric model of choice behavior is impossible here because the SEE does not ask respondents if they hold stocks, but one may ask how we should have included these respondents in such an analysis. It would have been necessary to consider wide interval to capture the extreme imprecision that these respondents face (some intervals would have been also necessary to capture the imprecision faced by those who provide partially informative answers). Lastly it would have been necessary to know how decision makers behave when they have an incomplete subjective distribution (i.e., which criterion they maximize). If there are some econometric tools to deal with interval data (e.g., Manski and Tamer, 2002), econometric models of choice behavior with interval expectations data are still lacking (Manski and Molinari, 2010, p.231).

of Assumption 1. In Section 3, this assumption, combined with Assumption 2, suffices to show that at least 66 percent of the $Q_1 = 0.50$ are uninformative, and that other Q_1 are uninformative as well. But the category "Unknown" represents an important share of the Final sample (28 percent as mentioned in Section 3). This is a problem to study the cross-sectional distribution of (μ, σ) conditional on the informativeness of Q_1. Given the relative size of this category and the fact that a subset of these Q_1 might be partially informative or uninformative, they can strongly influence the results. It is possible however to study at least how the (μ, σ) based on uninformative Q_1 influence the cross-sectional distribution of (μ, σ), "Unknown" excluded.

Consequently, in a first step (Subsection 4.1), I mainly exclude the "Unknown" from the analysis, and obtain the following interesting results: (i.) respondents who provide Q_1 identified as broadly informative price risk in the same way; that is, regressing expected returns μ on perceived risks σ, I find a robust positive linear relationship and the constant is not significantly different from the risk-free rate. (ii.) Such linear relationship does not appear at all if one considers all the respondents. (iii.) The collapse of the linear relationship is mainly due to the (μ, σ) based on Q_1 identified as uninformative, and in particular the 50s. (iv.) But the linear relationship is not so clear-cut if one only excludes from the analysis the (μ, σ) based on uninformative 50s. In other words, if one had only focused on the 50s as potential uninformative answers, he would not have only failed to capture an important share of uninformative answers, he would have obtained different results than if he had considered all the uninformative answers. Subsection 4.2 discusses some solutions to reduce the size of the category "Unknown" and include these respondents in the analysis.

4.1 A cross-sectional analysis of fitted subjective distributions conditional on the informativeness of Q_1

Following the literature, we use the sequence of subjective probabilities $(Q_{i,k},\ k = 1, 2, 3, 4)$ to fit a respondent-specific normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$. For each respondent i, we find the (μ_i, σ_i)

that solves the following least-squares problem:

$$\inf_{\mu_i,\sigma_i} \sum_{k=1}^{4} [(1 - Q_{i,k}) - F(R_{i,k}; \mu_i, \sigma_i)]^2 \tag{8}$$

where $F(R_i; \mu_i, \sigma_i)$ denotes the cumulative normal distribution function with mean μ_i and standard deviation σ_i, evaluated at any point R_i.

We can solve the least-squares problem in Equation 8 for 979 respondents (340 in wave 12, 264 in wave 13 and 375 in wave 14), i.e., the respondents of the Final sample. Indeed, of the 1231 respondents of the Preliminary sample (those who provided an answer $Q_{i,1}$ to the first probabilistic question), 1125 answered all of the 4 probabilistic questions. For various reasons, we could not fit a subjective distribution to the responses of 146 of these persons. In particular, we cannot find (μ, σ) for 120 respondents because they answered the same probability to all four thresholds.

Consider wave 12. Figure 1 describes the cross-sectional distribution of (μ, σ). Note that, as in Dominitz and Manski (2011), the data are rescaled such that μ_i is expected return (e.g., $\mu_i = 0.06$ means that the expected value of the investment one year from now is 1060). Panel (a) considers the (μ, σ) of the 340 respondents of this wave. Given that we learn strictly nothing about the informativeness of the category "Unknown", Panel (b) proposes the same cross-sectional distribution, excluding the 93 respondents of this category (see Table 5). Remark also that in each panel, some respondents have the parameters (μ, σ) of their fitted subjective distribution depicted by a (red) x mark or a (green) triangle instead of a (blue) dot. The (μ, σ) depicted by a x mark correspond to the 42 respondents who answered $Q_1 = 0.50$ in an uninformative way, and those depicted by a triangle correspond to the 16 respondents (9 M10 and 7 M5) who provided an uninformative Q_1 other than 50 (see Table 5). These two panels show that there is a great deal of heterogeneity. Notes (i.) and (ii.) at the bottom of the Figure confirm this visual impression by providing the (0.25, 0.50, 0.75)-quantiles of μ and σ. Figures A1 and A3 in the Appendix (provided as online supporting information) give similar results for waves 13 and 14.

To evaluate the effect of the informativeness of the answers, Figure 2 presents the cross-sectional distribution of (μ, σ) conditional on the informativeness of Q_1. Note first that most of the respondents classified as “Broadly informative” (Panel (a)), “Partially or broadly informative” (Panel (b)), “Partially informative” (Panel (c)) and “Uninformative or partially informative” (Panel (d)) have a positive expected return, while most of the respondents classified as “Uninformative” have a negative expected return (Panel (e)). It clearly shows that the “Uninformative” strongly bias the cross-sectional distribution of (μ, σ) (“Unknown” excluded). A researcher may be however very intrigued by these empirical regularities. To avoid any doubts, Section B of the online Appendix shows that these regularities mainly follow from the combination of three elements: (i.) the normality assumption made for all the respondents to fit their individual-specific normal distribution (Equation 8), (ii.) the fact that $r_{i,1} \leq r_{i,min}$ for almost all the respondents in Panels (a)-(e) (see Footnote 13), and (iii.) the fact that the first threshold is $r_{i,1} = 1000$ for an important majority of these respondents (given that μ_i is expected return and not the investment one year from now, the first threshold should be rescaled and $r_{i,1} = 1000$ corresponds to a rescaled first threshold of zero).

Furthermore, Panel (a) of Figure 2 shows considerable heterogeneity in the cross-sectional distribution of (μ, σ) for those who provide a broadly informative Q_1, but also shows that the expected return is positively (and *a priori* linearly) related to the risk for these respondents (Spearman's rho$=$ 0.69 with a p-value$=$ 0.000). Conversely, Panel (e) shows that the cross-sectional distribution of (μ, σ) for the uninformative exhibits no particular pattern (Spearman's rho$=$ 0.01 with a p-value$=$ 0.936). This result suggests that the absence of pattern found in Panel (b) of Figure 1 is mainly due to the (μ, σ) based on uninformative Q_1.

Table 7 provides ordinary least-squares (OLS) regressions of μ on σ for the different categories. Column [1] shows that a linear regression fits the monotonic relationship found for the broadly informative remarkably well. It is also possible to conclude that these respondents share a common *ex-ante* price of risk. Indeed, the R^2 is 82 percent, and the estimated intercept term does not differ significantly from the risk-free rate during this period. To see that, note that the estimated intercept term is 0.106, and its estimated standard-error is 0.043. Hence, the 95% confidence interval for this intercept is $[0.02, 0.19]$. Concerning the risk-free rate, we take the one year US LIBOR rate which oscillated in the range [0.056, 0.063] during wave 12.[16] Hence, we cannot reject the null hypothesis

[16] The one year US LIBOR rate is a daily reference rate based on the interest rates at which banks borrow funds

that the intercept term is equal to the risk-free rate, whatever its value in the interval $[0.056, 0.063]$. This linear relationship does not appear at all if we consider the uninformative (Column [5]), or if we consider all the respondents, except the "Unknown" (Column [6]).

Table 7 suggests that the (μ, σ) based on uninformative answers are responsible for the collapse of the linear relationship found in Column [6]. To confirm this result, and to see if excluding all the (μ, σ) based on uninformative Q_1 or only those based on uninformative 50s generate different results, we have used the following procedure. We have preliminary sorted the respondents in ascending order based on the variable X_i: $X_i = 1$ if $Q_{i,1}$ is classified as "Broadly informative", $X_i = 2$ if $Q_{i,1}$ is classified as "Partially or broadly informative", $X_i = 3$ if $Q_{i,1}$ is classified as "Partially informative", $X_i = 4$ if $Q_{i,1}$ is "Uninformative or partially informative", $X_i = 5$ if $Q_{i,1} \neq 0.50$ and is classified as "Uninformative", and $X_i = 6$ if $Q_{i,1} = 0.50$ and is classified as "Uninformative". Then, for respondents with common X, we have sorted them according to $d_{1,l}$. Thereafter, we have considered first a sample of the 80 "Broadly informative" with the lowest $d_{1,l}$, and added individuals one by one according to the final ranking. This has resulted in a series of growing samples of observations, and an OLS regression on each of this sample has been run.[17]

Figure 3 represents the evolution of the R^2, the estimated intercept and slope coefficients. To easily see the impact of the "Uninformative", a (green) triangle is used when the ultimate (μ_i, σ_i) added to the sample is based on an "Uninformative" $Q_{i,1} \neq 0.50$, and a (red) x mark is used when the ultimate (μ_i, σ_i) added in the sample is based on an "uninformative" $Q_{i,1} = 0.50$. At least three results appear:

(i.) First, Panel (a) clearly shows that the (μ, σ) based on uninformative Q_1 are responsible for

from other banks. During wave 12 (July 1999-November 1999), it reached a minimum of 0.056 on 21 July 1999 and a maximum of 0.063 on 27 October 1999.

[17] Two interrelated remarks are in order. First, the final ranking would have been slightly different if we had only sorted the respondents according to $d_{1,l}$. The uninformative 50s have a lower bound $d_{1,l} = 0.40$, while the other uninformative Q_1 have a lower bound $d_{1,l} > 0.40$. Thus the respondents who provided an uninformative 50s would have been ranked before those providing another uninformative response. Given that we want to see the results if we only exclude the (μ, σ) based on uninformative 50s, ranking them at the end of the final ranking is logical. Second, given that we do not know the precise value of d_1 in the interval $[d_{1,l}, d_{1,u}]$, one can argue that there is no justification to rank the respondents with common X according to $d_{1,l}$. The lower bound of the partially identified measure of coherence of respondent i can be less than the one of respondent j, i.e., $d_{j,1,l} > d_{i,1,l}$, despite the fact that their true measures of coherence (that we do not know precisely) can be such that $d_{i,1} > d_{j,1}$. It may occur if the identification regions $\mathrm{H}[d_{i,1}]$ and $\mathrm{H}[d_{j,1}]$ overlap. For instance, if $\mathrm{H}[d_{i,1}] = [0, 0.10]$ and $\mathrm{H}[d_{j,1}] = [0.05, 0.15]$, it is possible that the "true" distances are $d_{i,1} = 0.09$ and $d_{j,1} = 0.05$. Sorting the respondents with common X according to $d_{1,l}$ is only one way to see if there is (or not) a form of stability in the linear relationship in each subcategory X.

an important part of the collapse of the linear relationship between μ and σ.

(ii.) Second, this collapse is particularly due to the (μ, σ) based on uninformative 50s. With the sample of size $N = 205$, i.e., when the (μ, σ) based on uninformative 50s are excluded, the R^2 is 43 percent. When some (μ, σ) based on uninformative 50s are added, the R^2 collapses tremendously to 25 percent (when $N = 209$). Then, it decreases progressively until reaching 11 percent when the 247 respondents are considered for the OLS regression. It is also when these (μ, σ) are added to the sample that the estimated intercept term differs significantly from the risk-free rate during this period (Panel (b)).

(iii.) Third, the other uninformative answers have also a huge impact on the R^2. If only the (μ, σ) based on uninformative 50s are excluded, the R^2 is 43 percent. But if all the (μ, σ) based on uninformative answers are excluded, the R^2 is above 60 percent.

Figures A2 and A4 in the online Appendix provide similar results for waves 13 and 14. The R^2 is always around 60 percent when all the (μ, σ) based on uninformative Q_1 are excluded. If only the (μ, σ) based on uninformative 50s are excluded, the R^2 is always around 40 percent. These results show that the (μ, σ) based on uninformative 50s strongly bias the cross-sectional distribution of (μ, σ) ("Unknown" excluded), but the other uninformative answers also do so.

4.2 Additional results

If the last Subsection suggests that it is important to consider as uninformative not only some 50s but all the answers identified as such, some readers might be however disappointed by the fact that the "Unknown" are excluded from the analysis. Without making additional assumptions to reduce the identification region of $d_{i,1}$ when $r_{i,1} \in (r_{i,min}, r_{i,max})$, it is impossible to include them in a cross-sectional analysis of (μ, σ) conditional on the informativeness of Q_1. This Subsection briefly discusses two solutions to include the "Unknown" in the analysis. For the interested reader, the full descriptions are in Section C of the online Appendix to save space.

4.2.1 Using the responses to the other thresholds

The first solution, instead of adding assumptions, considers a criterion which extracts what we can learn about the distances at the other thresholds: the value of (μ_i, σ_i) is meaningless if $Q_{i,1}$ is uninformative, but it is also meaningless if, e.g., $Q_{i,4}$ is uninformative. We can consider *a priori* four measures of coherence ($d_{i,k}$, $k = 1, 2, 3, 4$). It is obvious that Assumption 1 permits to partially identified these four distances. If $\alpha = 0.10$, the bounds for $d_{i,2}$, $d_{i,3}$ and $d_{i,4}$ are similar to those described in Table 3 for $d_{i,1}$. Assumptions 1 and 2 permit to learn something about the informativeness of $Q_{i,k}$ if and only if the corresponding threshold is outside the suggestive support. Among the subjective probabilities whose thresholds are outside the suggestive support, I have selected the one whose partially identified measure of coherence has the highest lower bound and the highest upper bound, except if $r_{i,1} \leq r_{i,min}$ and $Q_{i,1} = 0.50$ (i.e., if $\mathrm{H}[d_{i,1}] = [0.40, 0.50]$). In this latter case, I have chosen $Q_{i,1} = 0.50$ in order not to overstate the role of the uninformative subjective probabilities other than 50s in the collapse of the linear relationship.[18] I have then studied the cross-sectional distribution of (μ, σ) conditional on the informativeness of the responses selected. A respondent remains classified as "Unknown" if the four threshold values ($r_{i,k}$, $k = 1, 2, 3, 4$) belong to $(r_{i,min}, r_{i,max})$. For instance, in wave 12, 64 respondents are concerned. This is less than the 93 "Unknown" of Section 3 and Subsection 4.1, but it remains imperfect. As shown in Subsection C1 of the online Appendix, we find however similar results: the (μ, σ) based on at least one uninformative Q_k are responsible for the collapse of the linear relationship. When these (μ, σ) are not included, the R^2 is around 60 percent. When we exclude only the respondents whose less informative Q_k is an uninformative 50, the R^2 is around 40 percent.

[18]Indeed, suppose that $r_{i,1} < r_{i,2} < r_{i,min} < r_{i,3} < r_{i,4} < r_{i,max}$, $Q_{i,1} = 0.50$, $Q_{i,2} = 0.40$, and $Q_{i,3} = 0.30$ and $Q_{i,4} = 0.30$. If $\alpha = 0.10$, then $\mathrm{H}[d_{i,1}] = [0.40, 0.50]$, $\mathrm{H}[d_{i,2}] = [0.50, 0.60]$, $\mathrm{H}[d_{i,3}] = \mathrm{H}[d_{i,4}] = [0, 0.70]$. In that case, if I had selected among the subjective probabilities whose thresholds are outside the suggestive support the one whose partially identified measure of coherence has the highest lower bound and the highest upper bound, I will have chosen $Q_{i,2}$. And I will have conclude that (μ_i, σ_i) is meaningless because of an uninformative $Q_{i,2} \neq 0.50$, despite the fact that it is already meaningless because $Q_{i,1} = 0.50$ and is uninformative. This problem occurs because of the monotonicity condition of the complementary cumulative distribution function and because $r_{i,2} < r_{i,min}$.

4.2.2 Adding assumptions to reduce the identification region of $d_{i,1}$

Although choosing as a criterion the less informative probability among the probabilities whose thresholds are outside the suggestive support permits to decrease the size of the "Unknown", it remains imperfect for two reasons. First, some respondents are still classified as "Unknown". Second, this criterion does not focus on the answer for the first threshold. This answer is important because it is not influenced by the answers for the subsequent thresholds, while the reverse is not true. As a consequence, the second solution explored in Subsection C2 of the online Appendix to learn something about the informativeness of the "Unknown" and include this important share of respondents in the analysis is to add assumptions to reduce the prediction region when $r_{i,1} \in (r_{i,min}, r_{i,max})$. Under Assumption 1, the prediction region is $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)\right] = [0,1]$ for all $r_{i,1} \in (r_{i,min}, r_{i,max})$. If we reduce this prediction region, the identification region of the measure of coherence when $r_{i,1} \in (r_{i,min}, r_{i,max})$ will be reduced too. I have not considered point-identifying assumptions, given the difficulty to justify them. I have considered some weaker assumptions that shrink partially the identification region $\mathrm{H}\left[d_{i,1}|Q_{i,1}, r_{i,1} \in (r_{i,min}, r_{i,max})\right]$, such that all (or almost all) the "Unknown" are classified as "Broadly informative", "Partially or broadly informative", "Partially informative", "Uninformative or partially informative", and "Uninformative". These assumptions, presented in Appendix C2, are however stronger than Assumption 1, and clearly illustrates the fact that there is a trade-off between the credibility of the additional assumptions and the possibility to learn something about the informativeness of the "Unknown".[19] Most of the answers classified as "Unknown" in Section 3 are classified as "Partially or broadly informative" or "Uninformative or partially informative" in Appendix C2. For instance, in wave 12, of the 93 respondents classified as "Unknown" in Section 3 (see Table 5), 63 are classified as "Partially or broadly informative", 20 as "Uninformative

[19] To fully understand, consider the prediction region at the midpoint $\frac{r_{i,min}+r_{i,max}}{2}$, i.e., $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1} = \frac{r_{i,min}+r_{i,max}}{2}\right)\right] = [\theta_1, 1-\theta_1]$, where $\theta_1 \in [0, 0.5]$ has to be fixed. If $\theta_1 = 0$, the prediction region $[0,1]$ is similar to the one under Assumption 1: it encompasses all possible point predictions. However, one can argue that the lower bound of the prediction region is strictly greater than 0 because the respondent should have provided initially a much lower $r_{i,max}$; similarly, the upper bound is strictly less than 1 because the respondent should have provided initially a much higher $r_{i,min}$. Hence, it is logical that $\theta_1 > 0$; the difficulty however is to choose a value for θ_1. It is possible that θ_1 is very small. If one chooses a small value, e.g., $\theta_1 = 0.01$, the prediction region is [0.01,0.99]. This assumption is very weak but the width of this interval will be too wide to learn something about the type of an answer at the midpoint. If one chooses however a relatively high θ_1, e.g., $\theta_1 = 0.45$, the prediction region is [0.45,0.55]. The width of this interval is sufficiently small to learn something about the type of an answer at the midpoint, but this assumption is clearly too strong.

or partially informative", 2 as "Partially informative", 1 as "Uninformative", and 7 remains "Unknown" and are included in the "Broadly informative" without influencing the results. Concerning in particular the 17 $Q_1 = 50$ of type "Unknown" in Section 3 (see Table 5), 11 are classified as "Partially or broadly informative" and 6 as "Uninformative or partially informative".

Whatever the wave, the R^2 decreases in an important way when we include the (μ, σ) based on "Partially or broadly informative" Q_1. The R^2 remains however between 50 (in waves 12 and 13) and 60 percent (in wave 14). These percentages are stable when the (μ, σ) based on "Partially informative" Q_1 are included. Then, at least in wave 12, the R^2 strongly decreases when the 6 (μ, σ) based on "Uninformative or partially informative" $Q_1 = 0.50$ are added. This is not a problem per se, given that these Q_1 are clearly not informative. However, to be sure that the (μ, σ) based on "Uninformative" $Q_1 \neq 0.50$ also influence the cross-sectional distribution, we have included them before in the ranking. The results suggest that the (μ, σ) based on uninformative Q_1, not only those based on uninformative 50s, as well as those based on "Uninformative or partially informative" Q_1, strongly influence the cross-sectional distribution of (μ, σ). For instance in wave 14, when the (μ, σ) based on "Uninformative or partially informative" $Q_1 \neq 0.50$ are included, the R^2 decreases from 60 to 50 percent. Then, including the (μ, σ) based on "Uninformative" $Q_1 \neq 0.50$ decreases the R^2 to 40 percent. Lastly, including the (μ, σ) based on "Uninformative or partially informative" and "Uninformative" $Q_1 = 0.50$ decreases the R^2 to 30 percent.

5 Conclusion and implications for the elicitation of probabilistic questions

This paper has proposed a method, a partially identified measure of coherence, to identify uninformative subjective probabilities. The method is based on extremely weak assumptions, and does not impose any restriction on which answers can be uninformative. Hence, although this method is only applicable on a specific format of probabilistic questions used to know the distribution of the future realization of a continuous variable, it adds to our understanding of response behavior to such questions. The empirical application of the paper shows that a majority of the 50s are uninformative, and that these 50s represent a majority of, but far from all, the identified uninformative probabilities. The other identified uninformative responses, mainly some M10 and M5, represent between 24 and 40 percent (depending on the wave and the sample) of these identified uninformative responses. These results highlight that researchers are right to feel embarrassed with the heap of 50s. But by assuming that only some 50s may be uninformative, as does, e.g., the HRS 08 (by asking the epistemic follow-up question only to those who respond "50 percent" to a first probabilistic question), they miss an important share of uninformative answers. Furthermore, our results suggest that the subjective distributions based on uninformative Q_1, not only those based on uninformative 50s, influence the cross-sectional analysis of the subjective distributions.

As a consequence, when the variable of interest is continuous, the format of questions in the SEE appears particularly attractive to identify uninformative responses to the first threshold. This is particularly true if this threshold value $r_{i,1}$ is outside the suggestive support $(r_{i,min}, r_{i,max})$. Ideally, one would like that the algorithm which defines the threshold values satisfies two criteria: (i.) it should insure that some of the thresholds, say $k = 2, \ldots, K$, belong to the suggested support; (ii.) it should insure that at least the first threshold value is such that $r_{i,1} \leq r_{i,min}$ for all i. Criterion (i.) is important because it makes no sense to have all the thresholds outside the suggestive support. Criterion (ii.) permits to learn something about the informativeness of the first probabilistic responses. We would also learn something about the informativeness of a response if $r_{i,1} \geq r_{i,max}$, but in that case, all the other thresholds will be outside the suggestive support. A drawback of the two stages questioning method of the SEE survey is that it does not insure that these two criteria are satisfied. In particular, $r_{i,1} \in (r_{i,min}, r_{i,max})$ for 30 percent of the Preliminary and Final samples. Proposing an algorithm that satisfies these two criteria would be easy. An obvious critic to this refinement is that when the survey medium necessitates an interviewer (face-to-face or telephone interview), the algorithm may make the task of the interviewer too difficult. In the application of this paper, the interviewer had to consider only 5 possible sequences of 4 threshold values. Imposing that $r_{i,1} \leq r_{i,min}$, for all i, may increase the number of different sequences and generate interviewer errors. Note however that most of face-to-face and telephone surveys are, and the SEE was, computer-assisted. It greatly simplifies the task of the interviewer.

One potential criticism of the SEE format that one reader might make is that the preliminary questions attract some nonresponses; it is then impossible to ask the respondents some points on their complementary cumulative distribution function. Note however that the nonresponse rates in the HRS 06 and HRS 08 which directly asked the probability of a gain in the stock market are very similar (see Footnote 11). Furthermore, even it is possible to find a format of question that will generate a lower nonresponse rate, nonresponses, which are mainly "don't know", indicate a lack of knowledge; and it is very difficult to understand how a respondent who is unable to provide the (suggestive) support of his subjective distribution would be able to provide some subjective probabilities which reflect relatively precise beliefs.

Although the paper has mainly focused on uninformative answers, another advantage of the measure of coherence is that it also shows that they are not the only answers that should not be taken at face value: various M10 and M5 are classified as partially informative or as uninformative or partially informative, i.e., their difference with the answer to the pair of preliminary questions is at least 10 percentage points. This result also casts serious doubts on the way the extent of rounding is usually inferred, given that it is usually assumed that these answers cannot imply an interval with a width of more than 10 percentage points. This result suggests that instead of eliciting precise probabilities, it would be better to ask for interval responses in the case of *binary discrete variable* (in the case of a binary variable, there is no suggestive support to elicit in a first stage, so it is impossible to construct a (partially identified) measure of coherence). Before to explain why in this case and not in the case of a continuous variable, note that the format of question I have in mind is of the type: *What do you think is the range of percent chances that event A will occur?* Manski and Molinari (2010, p.229) have recently implemented this kind of question in a wave of the American Life Panel with some encouraging results, in the sense that respondents are willing to report ranges of probabilities.

If the elicitation of a range of probabilities is easy when the variable is a binary discrete variable, it will be more difficult in the case of a continuous variable, given that it will be necessary to ask for a sequence of ranges. Remember that an interviewer asks for a sequence of precise probabilities that should satisfy the monotonicity of the cumulative distribution function. Dominitz and Manski (1997) provide an optimistic view in the sense that only 5 percent (22 of 437 respondents) violate this monotonicity condition. A much more pessimistic view is provided by Van Santen *et al.* (2012): the violation concerns around 30 percent of the respondents. If so, asking for a sequence of ranges can be even worse. Given this potential problem with continuous variables, our results suggest again that asking a first probability for a threshold outside the suggestive support might be useful to detect informativeness of answers to probabilistic questions.

References

ATTANASIO, O. P. (2009), "Expectations and perceptions in developing countries: Their measurement and their use", *American Economic Review*, vol. 99: pp. 87–92.

BELLEMARE, C., BISSONNETTE, L. and KRÖGER, S. (2012), "Flexible approximation of subjective expectations using probability questions", *Journal of Business & Economic Statistics*, vol. 30 n° 1: pp. 125–131.

BRUINE DE BRUIN, W., FISCHHOFF, B. and HALPERN-FELSHER, B. L. (2000), "Verbal and numerical expressions of probability: "It's a fifty-fifty chance"", *Organizational and Human Decision Processes*, vol. 81 n° 1: pp. 115–131.

DELAVANDE, A. (2008), "Pill, patch, or shot? Subjective expectations and birth control choice", *International Economic Review*, vol. 49 n° 3: pp. 999–1042.

DOMINITZ, J. (2001), "Estimation of income expectations models using expectations and realization data", *Journal of Econometrics*, vol. 102 n° 2: pp. 165–195.

DOMINITZ, J. and MANSKI, C. F. (1997), "Using expectations data to study subjective income expectations", *Journal of the American Statistical Association*, vol. 92 n° 439: pp. 855–867.

DOMINITZ, J. and MANSKI, C. F. (2004), "The Survey of Economic Expectations", Technical report, http://faculty.wcas.northwestern.edu/~cfm754/see_introduction.pdf.

DOMINITZ, J. and MANSKI, C. F. (2007), "Expected equity returns and portfolio choice: Evidence from the Health and Retirement Study", *Journal of the European Economic Association*, vol. 5: pp. 369–379.

DOMINITZ, J. and MANSKI, C. F. (2011), "Measuring and interpreting expectations of equity returns", *Journal of Applied Econometrics*, vol. 26: pp. 352–370.

FISCHHOFF, B. and BRUINE DE BRUIN, W. (1999), "Fifty-fifty=50%?", *Journal of Behavioral Decision Making*, vol. 12: pp. 149–163.

Giustinelli, P. (forthcoming), "Group decision making with uncertain outcomes: Unpacking child-parent choices of high school tracks", *International Economic Review*.

Gouret, F. and Hollard, G. (2011), "When Kahneman meets Manski: Using dual systems of reasoning to interpret subjective expectations of equity returns", *Journal of Applied Econometrics*, vol. 26: pp. 371–392.

Hudomiet, P., Kezdi, G. and Willis, R. J. (2011), "Stock market crash and expectations of American households", *Journal of Applied Econometrics*, vol. 26: pp. 393–415.

Hurd, M., Van Rooij, M. and Winter, J. (2011), "Stock market expectations of Dutch households", *Journal of Applied Econometrics*, vol. 26: pp. 416–436.

Hurd, M. D. (2009), "Subjective Probabilities in Household Surveys", *Annual Review of Economics*, vol. 1: pp. 543–562.

Kandel, E. and Pearson, N. (1995), "Differential interpretation of public signals and trade in speculative markets", *Journal of Political Economy*, vol. 103: pp. 831–872.

Kleinjans, K. and van Soest, A. (2014), "Rounding, focal point answers and nonresponse to subjective probability questions", *Journal of Applied Econometrics*, vol. 29: pp. 567–585.

Lillard, L. and Willis, R. J. (2001), "Cognition and wealth: The importance of probabilistic thinking", Working Papers wp007, University of Michigan, Michigan Retirement Research Center.

Manski, C. (2003), *Partial identification of probability distributions*, New-York: Springer-Verlag.

Manski, C. F. (2004), "Measuring expectations", *Econometrica*, vol. 72 n° 5: pp. 1329–1376.

Manski, C. F. and Molinari, F. (2010), "Rounding probabilisitic expectations in surveys", *Journal of Business & Economics Statistics*, vol. 28 n° 2: pp. 219–231.

Manski, C. F. and Tamer, E. (2002), "Inference of Regressions with Interval Data on a Regressor or Outcome", *Econometrica*, vol. 70: pp. 519–546.

Morgan, G. and Henrion, M. (1990), *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*, Cambridge University Press.

Van der Klaauw, W. and Wolpin, K. I. (2008), "Social security and the retirement and savings behavior of low-income households", *Journal of Econometrics*, vol. 145 nº 1-2: pp. 21–42.

Van Santen, P., Alessie, R. and Kalwij, A. (2012), "Probabilistic survey questions and incorret answers: Retirement income replacement rates", *Journal of Economic Behavior & Organization*, vol. 82 nº 1: pp. 267–280.

Vissing-Jorgensen, A. (2004), "Perspectives on behavioral finance: Does 'irrationality' disappear with wealth? Evidence from expectations and actions", in *NBER macroeconomics annual 2003, Volume 18*, National Bureau of Economic Research, Inc, pp. 139–208.

Table 1: Fraction of probability questions heaped at "50 percent": Some examples

Topic	Question	Survey	N	Fraction of 50
Inflation				
	How about the chances that the U.S. economy will experience double-digit inflation sometime during the next 10 years or so?	HRS 98	18884	0.32
Economic depression				
	What do you think are the chances that the U.S. economy will experience a major depression sometime during the next 10 years or so?	HRS 98	4605	0.26
		HRS06	16754	0.24
Stock market$^{(ii.)}$				
	By next year at this time, what is the percent chance that mutual fund shares invested in blue chip stocks like those in the Down Jones Industrial Average will be worth more than they are today?	MSC 02-04	3885	0.22
		HRS 06	16754	0.23
		HRS 08	15727	0.21
Survival probability				
	What is the percent chance that you will live to be 75 or more?	HRS 98	9905	0.25
		HRS 06	6713	0.22

Notes: i. HRS and MSC stand respectively for the Health and Retirement Study and the Michigan Survey of Consumers.
ii. The frame of the stock market question is slightly different in the MSC. The wording is as follows: " *Please think about the type of mutual fund known as a diversified stock fund. This type of mutual fund holds stock in many different companies engaged in a wide variety of business activities. Suppose that tomorrow someone were to invest one thousand dollars in such a mutual fund. Please think about how much money this investment would be worth one year from now. What do you think is the percent chance that this one thousand dollar investment will increase in value in the year ahead, so that it is worth more than one thousand dollars one year from now?*".
iii. The examples drawn from the HRS 98 can be found in Lillard and Willis (2001), and those of the HRS 06 in Manski and Molinari (2010). The answers to the question about investing in the stock market in the MSC 02-04, the HRS 06 and the HRS 08 are analyzed respectively by Dominitz and Manski (2011), Hurd (2009) and Hudomiet *et al.* (2011).

Table 2: An example of the SEE format of probabilistic questions for continuous variables: The stock market expectations questions (waves 12, 13 and 14)

SEE scenario

The next question is about investing in the stock market. Please think about the type of mutual fund known as a diversified stock fund. This type of mutual fund holds stock in many different companies engaged in a wide variety of business activities. Suppose that tomorrow someone were to invest one thousand dollars in such a mutual fund. Please think about how much money this investment would be worth one year from now.

Preliminary questions

What do you think is the LOWEST amount that this investment of $1000 would possibly be worth one year from now? ($r_{i,min}$)

What do you think is the HIGHEST amount that this investment of $1000 would possibly be worth one year from now? ($r_{i,max}$)

Algorithm for selection of investment thresholds

$\left\lceil \frac{r_{i,min}+r_{i,max}}{2} \right\rceil$	$r_{i,1}$	$r_{i,2}$	$r_{i,3}$	$r_{i,4}$
0 to 899	500	900	1000	1100
900 to 999	800	900	1000	1100
1000 to 1099	900	1000	1100	1200
1100 to 1299	1000	1100	1200	1500
1300 or more	1000	1200	1500	2000

Note: The midpoint of the lowest and highest values rounded up to the next integer $\left(\left\lceil \frac{r_{i,min}+r_{i,max}}{2} \right\rceil\right)$ determines the four investment value thresholds $r_{i,k}$, $k = 1, 2, 3, 4$, according to the algorithm presented in this Table.

Sequence of $K = 4$ probabilistic questions

What do you think is the PERCENT CHANCE (or CHANCES OUT OF 100) that, one year from now, this investment would be worth over $r_{i,k}$? ($Q_{i,k}$, $k = 1, 2, 3, 4$)

Table 3: Identification region for $d_{i,1}$ assuming $\alpha = 0.10$

	[1]	[2]	[3]
	$r_{i,1} \leq r_{i,min}$	$r_{i,1} \geq r_{i,max}$	$r_{i,1} \in (r_{i,min}, r_{i,max})$
Prediction	[0.90,1]	[0,0.10]	[0,1]
	$\mathrm{H}[d_{i,1} \vert Q_{i,1}, r_{i,1} \leq r_{i,min}]$	$\mathrm{H}[d_{i,1} \vert Q_{i,1}, r_{i,1} \geq R_{i,max}]$	$\mathrm{H}[d_{i,1} \vert Q_{i,1}, r_{i,1} \in (r_{i,min}, r_{i,max})]$
$Q_{i,1}$			
0	[0.90,1]	[0,0.10]	[0,1]
0.01	[0.89,0.99]	[0,0.09]	[0,0.99]
0.02	[0.88,0.98]	[0,0.08]	[0,0.98]
0.05	[0.90,0.95]	[0,0.05]	[0,0.95]
0.10	[0.80,0.90]	[0,0.10]	[0,0.90]
0.20	[0.70,0.80]	[0.10,0.20]	[0,0.80]
0.30	[0.60,0.70]	[0.20,0.30]	[0,0.70]
0.40	[0.50,0.60]	[0.30,0.40]	[0,0.60]
0.50	[0.40,0.50]	[0.40,0.50]	[0,0.50]
0.55	[0.35,0.45]	[0.45,0.55]	[0,0.55]
0.60	[0.30,0.40]	[0.50,0.60]	[0,0.60]
0.70	[0.20,0.30]	[0.60,0.70]	[0,0.70]
0.80	[0.10,0.20]	[0.70,0.80]	[0,0.80]
0.90	[0,0.10]	[0.80,0.90]	[0,0.90]
0.95	[0,0.05]	[0.90,0.95]	[0,0.95]
0.98	[0,0.08]	[0.88,0.98]	[0,0.98]
0.99	[0,0.09]	[0.89,0.99]	[0,0.99]
1	[0,0.10]	[0.90,1]	[0,1]

Table 4: Number of respondents answering $Q_1 = 0.50$ or another percent according to the partially identified measure of coherence at the first threshold -Preliminary sample-

		[1]	[2]	[3]	[4]	[5]	[6]	[7]
		$Q_1 = 0.50$	$Q_1 = 1$	$Q_1 = 0$	M10	M5	Other	Total
Wave 12								
	Broadly informative		59	3	37	19	14	132
	Partially informative				11	22		33
	Uninformative	56		3	14	9	4	86
	Uninformative or partially informative				8	1		9
	Broadly or partially informative				37	6	1	44
	Unknown	24	4		58	21	4	111
	Total	80	63	6	165	78	23	415
Wave 13								
	Broadly informative		63		28	13	8	112
	Partially informative				8	20		28
	Uninformative	44	1	3	18	2	3	71
	Uninformative or partially informative				7			7
	Broadly or partially informative				27	6		33
	Unknown	22	6	2	40	21		91
	Total	66	70	5	128	62	11	342
Wave 14								
	Broadly informative		67	3	38	17	8	133
	Partially informative				12	27	1	40
	Uninformative	50		7	17	6	3	83
	Uninformative or partially informative				11	5		16
	Broadly or partially informative				47	4		51
	Unknown	40	8	1	64	34	4	151
	Total	90	75	11	189	93	16	474

Table 5: Number of respondents answering $Q_1 = 0.50$ or another percent according to the partially identified measure of coherence at the first threshold -Final sample-

		[1]	[2]	[3]	[4]	[5]	[6]
		$Q_1 = 0.50$	$Q_1 = 1$	M10	M5	Other	Total
Wave 12							
	Broadly informative		54	30	17	12	113
	Partially informative			9	18		27
	Uninformative	42		9	7		58
	Uninformative or partially informative			8			8
	Broadly or partially informative			34	6	1	41
	Unknown	17	4	50	19	3	93
	Total	59	58	140	67	16	340
Wave 13							
	Broadly informative		52	27	12	8	99
	Partially informative			8	14		22
	Uninformative	32	1	8	1		42
	Uninformative or partially informative			7			7
	Broadly or partially informative			24	6		30
	Unknown	13	4	33	14		64
	Total	45	57	107	47	8	264
Wave 14							
	Broadly informative		54	33	12	8	107
	Partially informative			12	22	1	35
	Uninformative	37		11	2		50
	Uninformative or partially informative			11	4		15
	Broadly or partially informative			41	4		45
	Unknown	26	7	54	32	4	123
	Total	63	61	162	76	13	375

Table 6: Summary results

	Preliminary sample				Final sample			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
	Wave 12	Wave 13	Wave 14	Total	Wave 12	Wave 13	Wave 14	Total
[A] **Fraction of 50s identified as uninformative** (Equation 5-Proposition 1)								
	0.70	0.66	0.55	0.63	0.71	0.71	0.58	0.66
[B] **Fraction of 50s among identified uninformative answers** (Equation 6-Proposition 1)								
	0.65	0.62	0.60	0.62	0.72	0.76	0.74	0.74
[C] **Fraction of identified uninformative answers other than 50s in the sample** (Equation 7-Proposition 1)								
	0.072	0.079	0.070	0.073	0.047	0.038	0.035	0.040

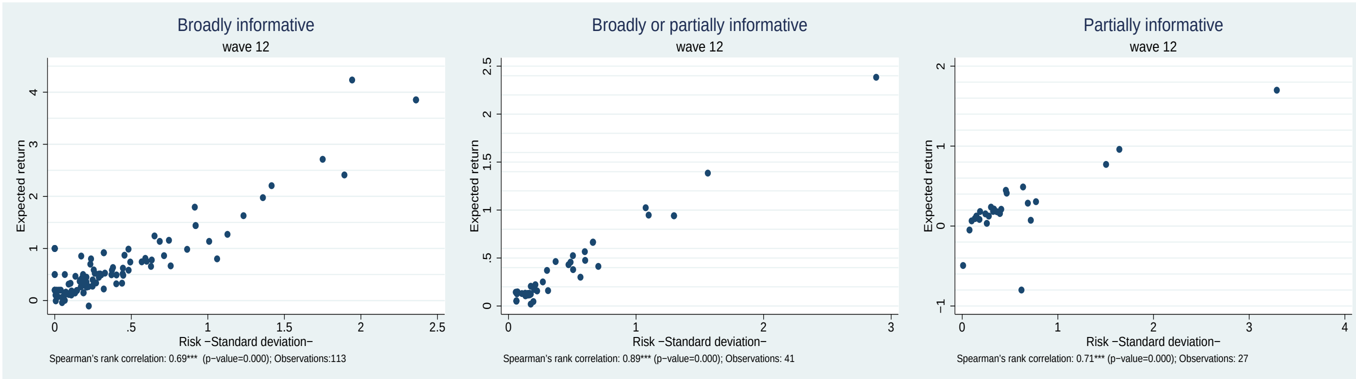
(a) (b)

Figure 1: Plots of the expected return (μ) against risk (σ)-wave 12-

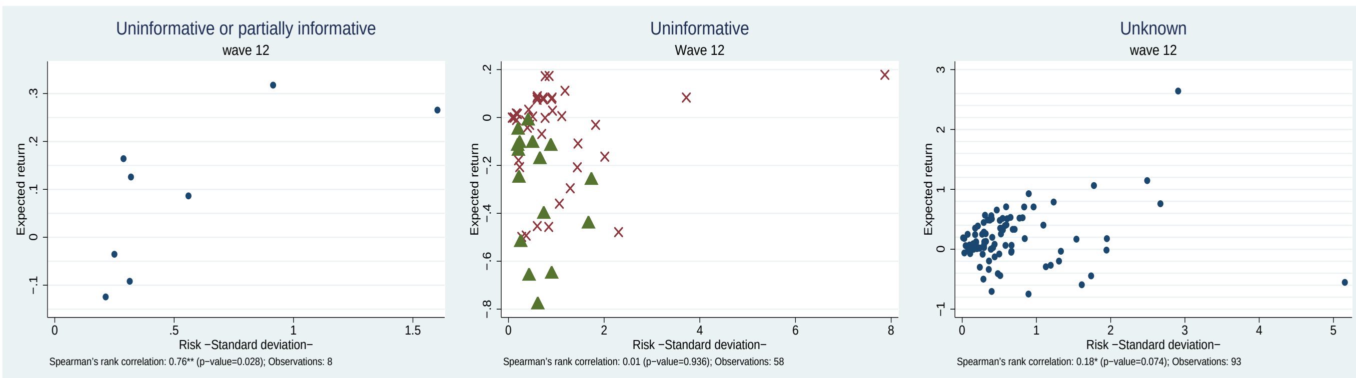
Notes: i. Panel (a)-All categories-: The means of μ and σ are 0.32 and 0.57 respectively. The (0.25,0.50,0.75)-quantiles of μ are (0.03,0.17,0.50). The (0.25,0.50,0.75)-quantiles of σ are (0.18,0.34,0.67).

ii. Panel (b)-All categories except Unknown-: The means of μ and σ are 0.37 and 0.53 respectively. The (0.25,0.50,0.75)-quantiles of μ are (0.05,0.18,0.51). The (0.25,0.50,0.75)-quantiles of σ are (0.15,0.30,0.66).

iii. 58 respondents have the parameters (μ, σ) of their fitted probability distribution depicted by a (red) x mark (42 respondents) or a (green) triangle (16 respondents) instead of a (blue) dot. The 42 (μ, σ) depicted by a x mark correspond to the respondents who answered $Q_1 = 0.50$ in an uninformative way, and the (μ, σ) depicted by a triangle correspond to the 16 respondents (9 M10 and 7 M5) who provided an uninformative Q_1 other than 50 (see Table 5).



(a) (b) (c)



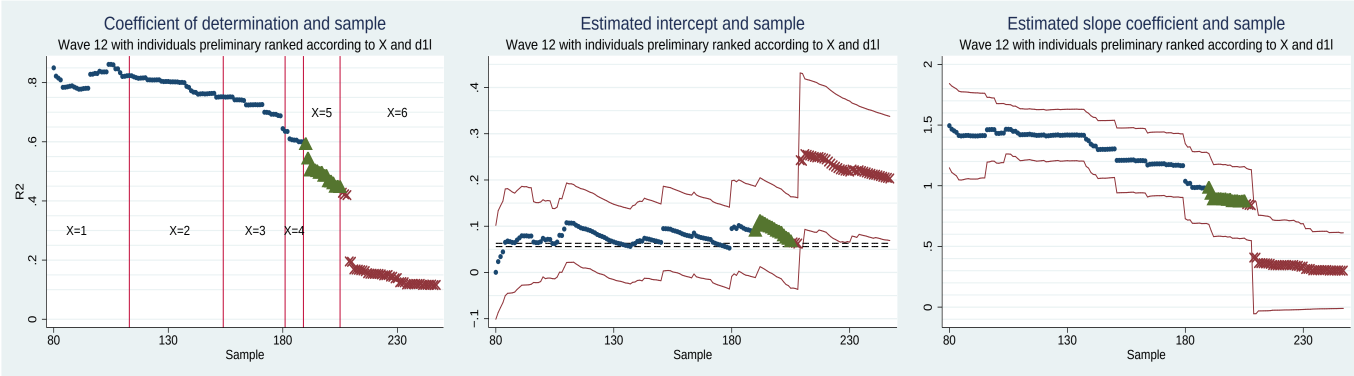
(d) (e) (f)

Figure 2: Plot of expected return (μ) against risk (σ) by type -wave 12-

Table 7: Least-squares regressions of μ on σ -wave 12-

	[1] B.I.	[2] B. or P.I.	[3] P.I.	[4] Un. or P.I.	[5] Un.	[6] All categories except Unknown	[7] Unknown	[8] All categories
Intercept	0.10**	0.015	-0.06	-0.04	-0.15***	0.20***	0.11	0.18***
	(0.04)	(0.01)	(0.05)	(0.06)	(0.03)	(0.06)	(0.08)	(0.05)
σ	1.42***	0.82***	0.53***	0.24**	0.024	0.30*	0.11	0.24**
	(0.10)	(0.02)	(0.02)	(0.08)	(0.018)	(0.16)	(0.15)	(0.11)
R^2	0.82	0.96	0.66	0.52	0.01	0.11	0.03	0.09
Obs.	113	41	27	8	58	247	93	340

Notes: i. Heteroskedasticity-robust standard errors are in parentheses.
ii. "B.I." stands for "Broadly informative" (it corresponds to the (μ, σ) based on broadly informative Q_1). "B. or P.I." stands for "Broadly or partially informative". "P.I." stands for "Partially informative". "Un. or P.I." stands for "Uninformative or partially informative". "Un." stands for "Uninformative".



(a) (b) (c)

Figure 3: Coefficient of determination, estimated intercept and slope coefficients conditional on the sample (respondents preliminary ranked according to X_i and then $d_{i,1,l}$, category "Unknown" excluded) -wave 12-

Notes: i. The two solid red lines in Panels (b) and (c) represent the 95% confidence intervals of the intercept and slope coefficients of the OLS estimates. They are computed using heteroskedasticity-robust standard errors.

ii. The two dashed black lines in Panel (b) correspond to the minimum and maximum values of the daily one-year U.S. LIBOR rate (RF_{12}) during wave 12 (July 1999 to November 1999). It reached a minimum of 0.056 (21 July) and a maximum of 0.063 (27 October), so it oscillated in the range $RF_{12} \in [0.056, 0.063]$. Let γ_{12} be the intercept of an OLS regression of μ on σ. If one wants to test the null hypothesis H_0: $\gamma_{12} = RF_{12}$ versus $\gamma_{12} \neq RF_{12}$, one can look at if RF_{12} lies inside the 95% confidence interval of the intercept -depicted by the two solid red lines-. If RF_{12} lies inside the confidence interval, we cannot reject the hypothesis that the intercept is equal to the risk-free rate, $\forall RF_{12} \in [0.056, 0.063]$. If it is outside, we can reject the null hypothesis with 95% confidence.

Supplement to "What can we learn from the fifties?"

by Fabian Gouret

November 2015

This Appendix presents tables of results for the waves 13 and 14 and supplementary materials. The Appendix is organized as follows. Section A describes some results for the waves 13 and 14, similar to those presented for the wave 12 in Section 4.1. Section B explains why in Figure 2 (Section 4.1 of the main text) most of the respondents classified as "Uninformative" have a negative expected return, while most of the other respondents have a positive expected return. Section C fully describes two solutions to reduce the size of the category "Unknown", and include these respondents in the analysis.

A Cross-sectional analysis of the fitted subjective normal distributions: Waves 13 and 14

This Section presents results for waves 13 and 14, similar to those presented for wave 12 in Section 4.1. Note that, as in Section 4.1, the data are rescaled such that μ_i is expected return (e.g., $\mu_i = 0.06$ means that the expected value of the investment one year from now is 1060).

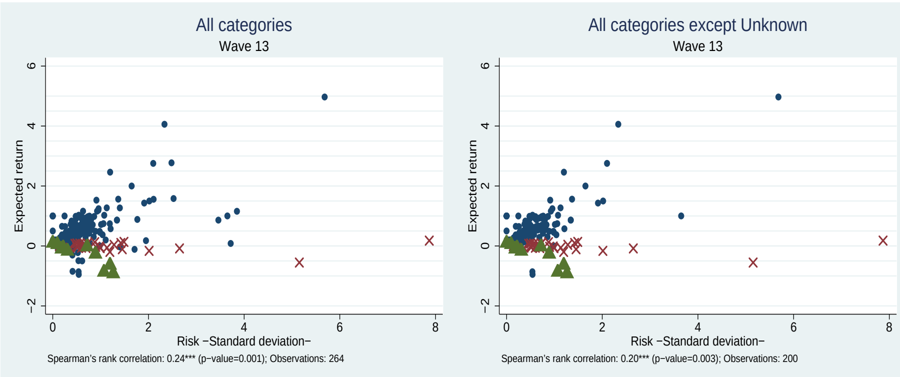
A.1 Wave 13

Figure A1 describes the cross-sectional distribution of (μ, σ) for wave 13. Panel (a) considers the (μ, σ) of the 264 respondents of this wave. Panel (b) proposes the same cross-sectional distribution, excluding the 64 respondents classified as "Unknown" (see Table 5). In each panel, the (μ, σ) depicted by a (red) x mark correspond to the 32 respondents who answered $Q_1 = 0.50$ in an uninformative way, and those depicted by a (green) triangle correspond to the 10 respondents who provided an uninformative Q_1 other than 50 (see Table 5).

Figure A2 provides similar results to those provided for wave 12 in Figure 3. We have preliminary sorted the respondents in ascending order based on X: $X_i = 1$ if $Q_{i,1}$ is classified as "Broadly informative", $X_i = 2$ if $Q_{i,1}$ is classified as "Partially or broadly informative", $X_i = 3$ if $Q_{i,1}$ is classified as "Partially informative", $X_i = 4$ if $Q_{i,1}$ is "Uninformative or partially informative", $X_i = 5$ if $Q_{i,1} \neq 0.50$ and is classified as "Uninformative", and $X_i = 6$ if $Q_{i,1} = 0.50$ and is classified as "Uninformative". Then, for respondents with common X, we have sorted them according to $d_{1,l}$. Thereafter, we have considered first a sample of the 80 "Broadly informative" with the lowest $d_{1,l}$, and added individuals one by one according to the procedure. This has resulted in a series of growing samples of observations, and an OLS regression on each of this sample has been run. Figure A2 represents the evolution of the R^2, the estimated intercept and slope coefficients.

A.2 Wave 14

Figure A3 describes the cross-sectional distribution of (μ, σ) for wave 14. Panel (a) considers the (μ, σ) of the 375 respondents of this wave. Panel (b) proposes the same cross-sectional distribution, excluding the 123 respondents classified as "Unknown" (see Table 5). In each panel, the (μ, σ) depicted by a (red) x mark correspond to the 37 respondents who answered $Q_1 = 0.50$ in an uninformative way, and those depicted by a (green) triangle correspond to the 13 respondents who provided an uninformative Q_1 other than 50 (see Table 5).



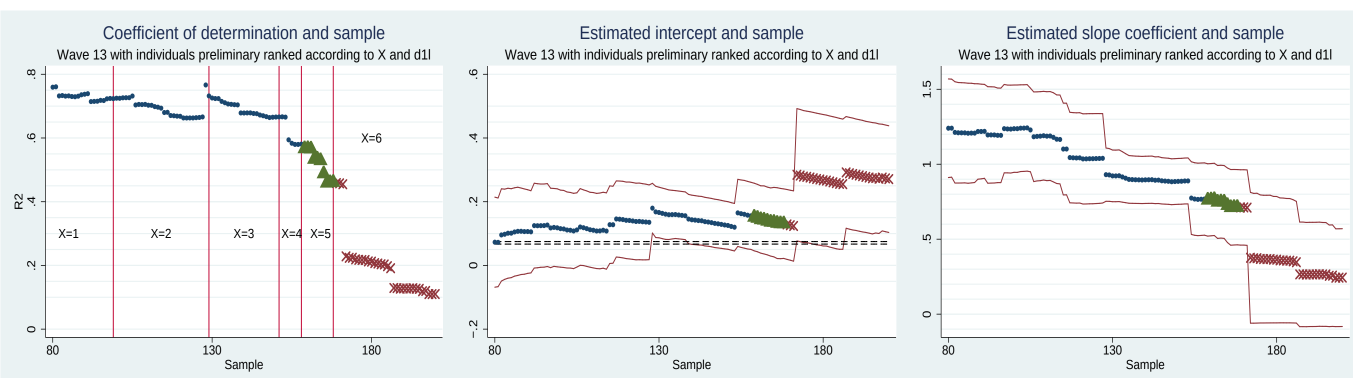
(a) (b)

Figure A1: Plots of the expected return (μ) against risk (σ)-wave 13-

Notes: i. Panel (a)-All categories-: The means of μ and σ are 0.40 and 0.67 respectively. The (0.25,0.50,0.75)-quantiles of μ are (0.08,0.28,0.60). The (0.25,0.50,0.75)-quantiles of σ are (0.22,0.43,0.76).

ii. Panel (b)-All categories except Unknown-: The means of μ and σ are 0.42 and 0.64 respectively. The (0.25,0.50,0.75)-quantiles of μ are (0.10,0.29,0.64). The (0.25,0.50,0.75)-quantiles of σ are (0.20,0.46,0.76).

iii. 42 respondents have the parameters (μ, σ) of their fitted probability distribution depicted by a (red) x mark (32 respondents) or a (green) triangle (10 respondents) instead of a (blue) dot. The 32 (μ, σ) depicted by a x mark correspond to the respondents who answered $Q_1 = 0.50$ in an uninformative way, and the (μ, σ) depicted by a triangle correspond to the 10 respondents (1 $Q_1 = 1$, 8 M10 and 1 M5) who provided an uninformative Q_1 other than 50 (see Table 5).



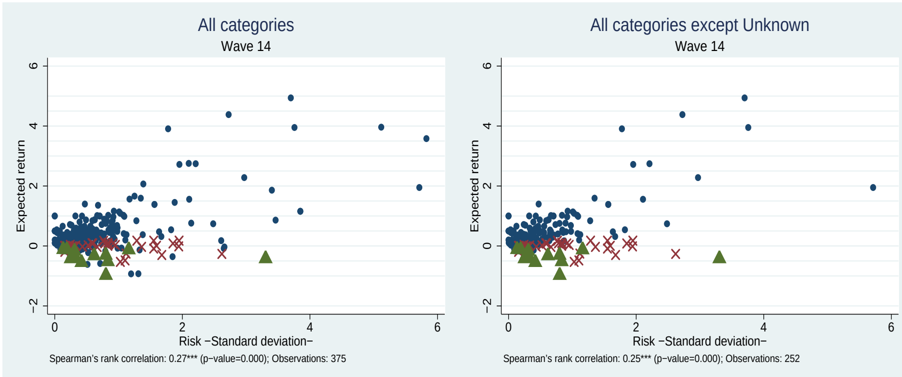
(a) (b) (c)

Figure A2: Coefficient of determination, estimated intercept and slope coefficients conditional on the sample (respondents preliminary ranked according to X_i and then $d_{i,1,l}$, category "Unknown" excluded) -wave 13-

Notes: i. The two solid red lines in Panels (b) and (c) represent the 95% confidence intervals of the intercept and slope coefficients of the OLS estimates. They are computed using heteroskedasticity-robust standard errors.

ii. The two dashed black lines in Panel (b) correspond to the minimum and maximum values of the daily one-year U.S. LIBOR rate (RF_{13}) during wave 13 (February 2000 to May 2000). It reached a minimum of 0.067 (28 February) and a maximum of 0.075 (31 May), so it oscillated in the range $RF_{13} \in [0.067, 0.075]$. Let γ_{13} be the intercept of an OLS regression of μ on σ. If one wants to test the null hypothesis H_0: $\gamma_{13} = RF_{13}$ versus $\gamma_{13} \neq RF_{13}$, one can look at if RF_{13} lies inside the 95% confidence interval of the intercept -depicted by the two solid red lines-. If RF_{13} lies inside the confidence interval, we cannot reject the hypothesis that the intercept is equal to the risk-free rate, $\forall RF_{13} \in [0.067, 0.075]$. If it is outside, we can reject the null hypothesis with 95% confidence.

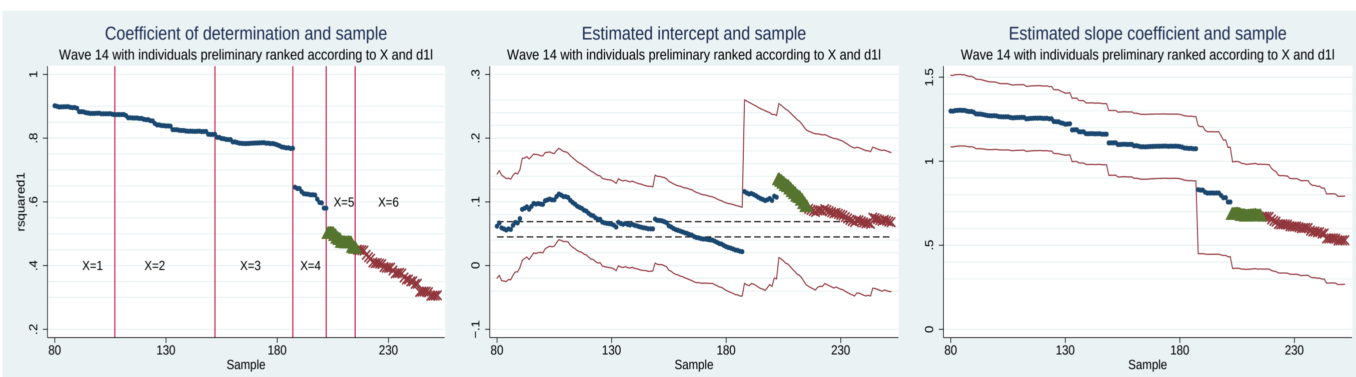
Figure A4 provides similar results to those provided for wave 12 in Figure 3. We have preliminary sorted the respondents in ascending order based on X: $X_i = 1$ if $Q_{i,1}$ is classified as "Broadly informative", $X_i = 2$ if $Q_{i,1}$ is classified as "Partially or broadly informative", $X_i = 3$ if $Q_{i,1}$ is classified as "Partially informative", $X_i = 4$ if $Q_{i,1}$ is "Uninformative or partially informative", $X_i = 5$ if $Q_{i,1} \neq 0.50$ and is classified as "Uninformative", and $X_i = 6$ if $Q_{i,1} = 0.50$ and is classified as "Uninformative". Then, for respondents with common X, we have sorted them according to $d_{1,l}$. Thereafter, we have



(a) (b)

Figure A3: Plots of the expected return (μ) against risk (σ)-wave 14-
Notes: i. Panel (a)-All categories-: The means of μ and σ are 0.36 and 0.65 respectively. The (0.25,0.50,0.75)-quantiles of μ are (0.03,0.20,0.47). The (0.25,0.50,0.75)-quantiles of σ are (0.19,0.39,0.79).
ii. Panel (b)-All categories except Unknown-: The means of μ and σ are 0.37 and 0.58 respectively. The (0.25,0.50,0.75)-quantiles of μ are (0.07,0.20,0.49). The (0.25,0.50,0.75)-quantiles of σ are (0.15,0.33,0.71).
iii. 50 respondents have the parameters (μ, σ) of their fitted probability distribution depicted by a (red) x mark (37 respondents) or a (green) triangle (13 respondents) instead of a (blue) dot. The 37 (μ, σ) depicted by a x mark correspond to the respondents who answered $Q_1 = 0.50$ in an uninformative way, and the (μ, σ) depicted by a triangle correspond to the 13 respondents (11 M10 and 2 M5) who provided an uninformative Q_1 other than 50 (see Table 5).

considered first a sample of the 80 "Broadly informative" with the lowest $d_{1,l}$, and added individuals one by one according to the procedure. This has resulted in a series of growing samples of observations, and an OLS regression on each of this sample has been run. Figure A4 represents the evolution of the R^2, the estimated intercept and slope coefficients.



(a) (b) (c)

Figure A4: Coefficient of determination, estimated intercept and slope coefficients conditional on the sample (respondents preliminary ranked according to X_i and then $d_{i,1,l}$, category "Unknown" excluded) -wave 14-
Notes: i. The two solid red lines in Panels (b) and (c) represent the 95% confidence intervals of the intercept and slope coefficients of the OLS estimates. They are computed using heteroskedasticity-robust standard errors.
ii. The two dashed black lines in Panel (b) correspond to the minimum and maximum values of the daily one-year U.S. LIBOR rate (RF_{14}) during wave 14 (September 2000 to March 2001). It reached a minimum of 0.045 (22 March 2001) and a maximum of 0.069 (1 September 2000), so it oscillated in the range $RF_{14} \in [0.045, 0.069]$. Let γ_{14} be the intercept of an OLS regression of μ on σ. If one wants to test the null hypothesis H_0: $\gamma_{14} = RF_{14}$ versus $\gamma_{14} \neq RF_{14}$, one can look at if RF_{14} lies inside the 95% confidence interval of the intercept -depicted by the two solid red lines-. If RF_{14} lies inside the confidence interval, we cannot reject the hypothesis that the intercept is equal to the risk-free rate, $\forall RF_{14} \in [0.045, 0.069]$. If it is outside, we can reject the null hypothesis with 95% confidence.

B Difference of structure in the cross-sectional distribution of (μ, σ) between the different types

The difference of structure in the data between the different types of respondents in Figure 2 may be puzzling. For instance, it seems that most of the respondents classified as broadly informative have a positive expected return, while most of the respondents classified as uninformative have a negative expected return. Why this is so? This Section shows that these regularities mainly follow from the combination of three elements: (i.) the assumption (made in the literature and in Section 4 of this paper) that all the respondents have a normal distribution (see Equation 8 in the main text), (ii.) the fact that $r_{i,1} \leq r_{i,min}$ for almost all the respondents in Panels (a)-(e) of Figure 2 (see Footnote 13), and (iii.) the fact that the first threshold is $r_{i,1} = 1000$ for an important majority of these respondents. Table B1 presents the relative frequencies of the different first thresholds by type for the wave 12 (the results for the other waves are not presented but are qualitatively similar). Table B1 shows that $r_1 = 1000$ for more than 88 percent of the respondents classified as "Broadly informative". This percentage reaches 92 percent in the case of the respondents classified as "Broadly or partially informative". The lowest percentage, 50 percent, concerns the respondents classified as "Uninformative or partially informative". But remember that this category includes only 8 respondents (see Table 5 in the main text). One can thus conclude that there are a relative intra-type homogeneity in r_1 (i.e., there is a relative homogeneity in r_1 across respondents of the same type) and a relative inter-type homogeneity in r_1 (i.e., the respondents classified in different categories have similar r_1). To understand why this homogeneity matters, I characterize the sets of possible values that (μ, σ) can take in principle conditional on the type of the respondents. Assumptions 1-2 hold, and the critical values are $\underline{d} = 0.10$ and $\overline{d} = 0.40$ as in the rest of the text. Remember that, as in Dominitz and Manski (2011), μ_i is expected return (e.g., $\mu_i = 0.06$ means that the expected value of the investment one year from now is 1060), so R_i, $r_{i,min}$ and $r_{i,1}$ should also be rescaled. I denote $R_i^* = \frac{R_i - 1000}{1000}$, $r_{i,min}^* = \frac{r_{i,min} - 1000}{1000}$ and $r_{i,1}^* = \frac{r_{i,1} - 1000}{1000}$ (so $r_{i,1}^* = 0$ means that the first threshold is $r_{i,1} = 1000$).

Table B1: Relative Frequencies of r_1 (Final sample) by categories -wave 12-

	Wave 12			
	r_1			
	$r_1 = 500$	$r_1 = 800$	$r_1 = 900$	$r_1 = 1000$
Broadly informative	0.0177	0.0973	0	0.885
Broadly or partially informative	0	0	0.732	0.9268
Partially informative	0.1111	0	0.1111	0.7778
Uninformative or partially informative	0	0.25	0.25	0.50
Uninformative	0.1207	0.0517	0.0862	0.7414
Unknown	0.129	0.0538	0.2043	0.6129

Furthermore, let me make the two following additional assumptions.

Assumption B1. $r_{i,1}^* = r_1^*$ *for all* $i = 1, \ldots, N$

Assumption B1 says that the (rescaled) first threshold $r_{i,1}^*$ is the same for all the respondents. As shown in Table B1, the homogeneity in $r_{i,1}$, and so in $r_{i,1}^*$, is strong but not as extreme as in Assumption B1. Making this assumption permits however to simplify the presentation and to understand the implication of the relative homogeneity of $r_{i,1}^*$ around zero.

Assumption B2. $r_{i,1}^* \leq r_{i,min}^*$, *for all* $i = 1, \ldots, N$.

Assumption B2 implies that all the respondents have provided a probability $Q_{i,1} \equiv \text{Prob}\left(R_i^* > r_{i,1}^*\right)$ to a threshold $r_{i,1}^* \leq r_{i,min}^*$. As a consequence, we can learn something about the informativeness of all the $Q_{i,1}$. We already know that Assumption B2 is broadly supported by the data (see Footnote 13 in the main text).

Consider $Q_{i,1} = \text{P}(R_i^* > r_{i,1}^*)$. Suppose for simplicity that the respondent-specific normal distributions perfectly fit the elicited probabilities (see Equation 8 in the main text), and let $\Phi(.)$ denote the standard normal cumulative distribution function, so $Q_{i,1} = 1 - \Phi\left(\frac{r_{i,1}^* - \mu_i}{\sigma_i}\right)$. The expected return of respondent i, μ_i, is thus equal to:

$$\mu_i = r_{i,1}^* - \Phi^{-1}(1 - Q_{i,1}) \times \sigma_i \tag{B1}$$

Equation B1 permits to characterize for each respondent i a set of values that (μ_i, σ_i) can take conditional on the rescaled threshold $r_{i,1}^*$ and $Q_{i,1}$. Assumption B1 holds, so $r_{i,1}^* = r_1^*$ for the N respondents. Suppose furthermore that they have provided the same Q_1. Then, the cross-sectional distribution of (μ, σ) may be heterogenous, but Equation B1 shows that the N values of (μ, σ) are on the same line in a (σ, μ) space. The intercept and the slope coefficients of this line are r_1^* and $\Phi^{-1}(1-Q_1)$, respectively.

The values that $Q_{i,1}$ can take in Equation B1 is closely linked to the type of respondent i. Because of Assumption B2, and given that the critical values are $\underline{d} = 0.10$ and $\overline{d} = 0.40$, then one can see in Table 3 that:

(i.) $0.90 \leq Q_{i,1} \leq 1$ for all the respondents classified as "Broadly informative". If so, $1 - Q_{i,1}$ is bounded in the interval $[0, 0.10]$ for all these respondents. Using Equation B1, and given that $\Phi^{-1}(0) = -\infty$ and $\Phi^{-1}(0.10) \simeq -1.28$, then the set of possible values of μ for the "Broadly informative" is:

$$\text{H}\left[\mu|0 \leq d_{1,l} \leq d_{1,u} \leq 0.10\right] = \{\mu|\mu \geq r_1^* + 1.28\sigma\} \tag{B2}$$

(ii.) $0.80 \leq Q_{i,1} < 0.90$ for all the respondents classified as "Broadly or partially informative". Using Equation B1, and given that $\Phi^{-1}(0.10) \simeq -1.28$ and $\Phi^{-1}(0.20) \simeq -0.84$, then the set of possible values of μ for the "Broadly or partially informative" is:

$$\text{H}\left[\mu|d_{1,l} \leq 0.10 < d_{1,u} < 0.40\right] = \{\mu|r_1^* + 0.84\sigma \leq \mu < r_1^* + 1.28\sigma\} \tag{B3}$$

(iii.) $0.60 < Q_{i,1} < 0.80$ for all the respondents classified as "Partially informative". Using Equation B1, and given that $\Phi^{-1}(0.20) \simeq -0.84$ and $\Phi^{-1}(0.40) \simeq -0.25$, then the set of possible values of μ for the "Partially informative" is:

$$\text{H}\left[\mu|0.10 < d_{1,l} < d_{1,u} < 0.40\right] = \{\mu|r_1^* + 0.25\sigma < \mu < r_1^* + 0.84\sigma\} \tag{B4}$$

(iv.) $0.50 < Q_{i,1} \leq 0.60$ for all the respondents classified as "Uninformative or partially informative". Using Equation B1, and given that $\Phi^{-1}(0.40) \simeq -0.25$ and $\Phi^{-1}(0.50) = 0$, then the set of possible values of μ for the "Uninformative or partially informative" is:

$$\text{H}\left[\mu|0.10 < d_{1,l} < 0.40 \leq d_{1,u}\right] = \{\mu|r_1^* < \mu < r_1^* + 0.25\sigma\} \tag{B5}$$

(v.) $0 \leq Q_{i,1} \leq 0.50$ for all the respondents classified as "Uninformative". Using Equation B1, and given that $\Phi^{-1}(0.50) = 0$ and $\Phi^{-1}(1) = +\infty$, then the set of possible values of μ for the "Uninformative" is:

$$\text{H}\left[\mu|0.40 \leq d_{1,l} < d_{1,u}\right] = \{\mu|\mu \leq r_1^*\} \tag{B6}$$

Using Equations B2-B6, Panels (a)-(e) of Figure B1 describe the sets of possible values of (μ, σ) conditional on the type of respondents in the (σ, μ) space (the shaded sets). A decrease of the possible values of $Q_{i,1}$ rotates the sets of possible values of (μ_i, σ_i) in a clockwise fashion around r_1, and the different sets do not overlap. If $r_1^* = 0$, then all the respondents classified as "Uninformative"have a negative expected return; the other respondents have a positive expected return.

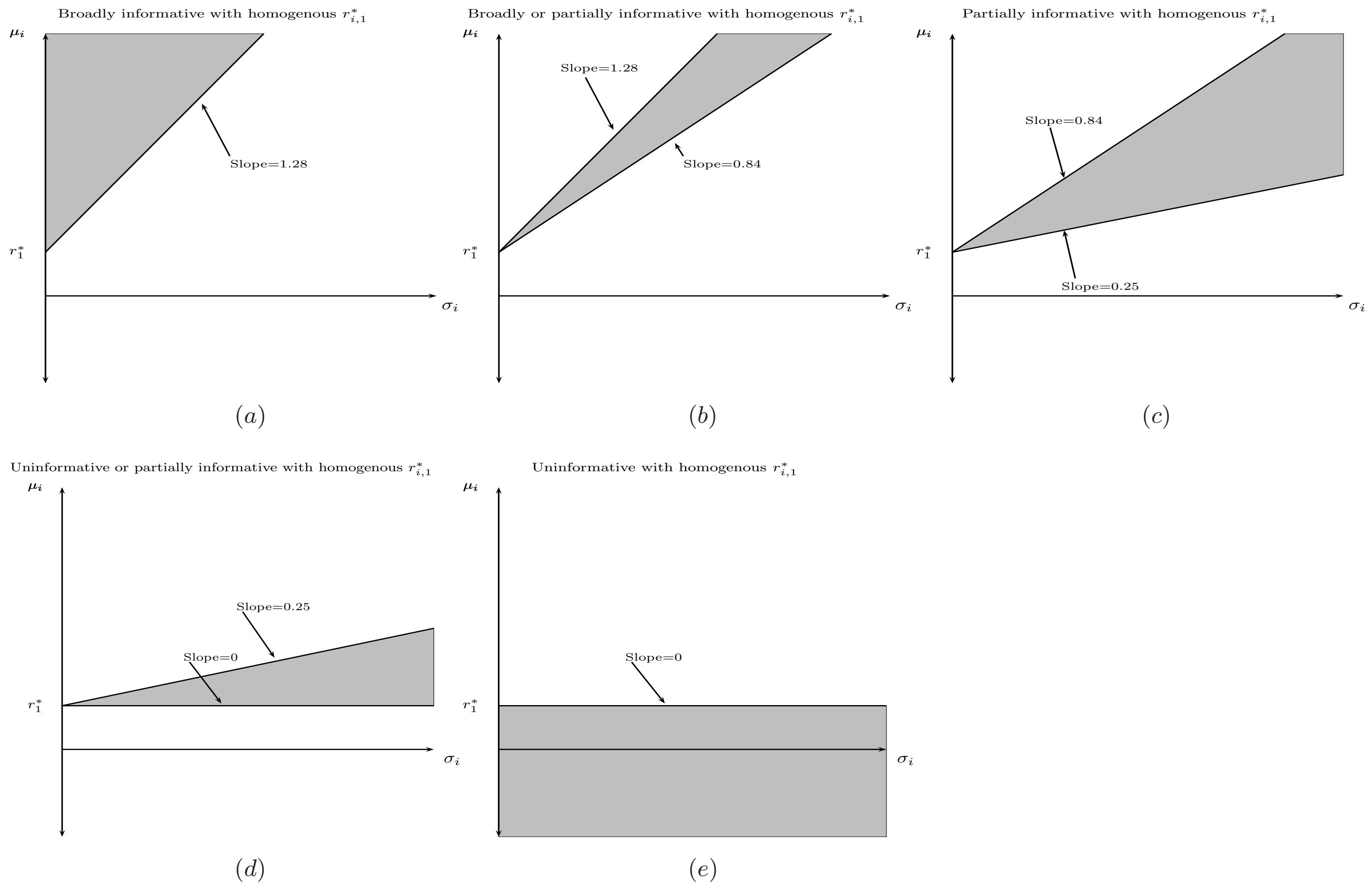


Figure B1: Set of possible values of (μ, σ) by type assuming that $r^*_{i,1} = r^*_1$ and $r^*_{i,1} \leq r^*_{i,min}$ for all $i = 1, ..., N$

C Reducing the size of the "Unknown"

This Section presents in details the two solutions discussed in Subsection 4.2 to reduce the size of the "Unknown", and include them in the cross-sectional analysis of (μ, σ). To keep this Section self-contained, part of it repeats in the main text.

C.1 Using the responses to the other thresholds

I first consider a criterion which extracts what we can learn about the distances at the other thresholds. The general idea is to exploit the fact that the value of (μ_i, σ_i) is meaningless if $Q_{i,1}$ is uninformative, but it is also meaningless if, e.g., $Q_{i,4}$ is uninformative. We can consider *a priori* four measures of coherence ($d_{i,k}$, $k = 1, 2, 3, 4$). It is obvious that Assumption 1 permits to partially identified these four distances. If $\alpha = 0.10$, the bounds for $d_{i,2}$, $d_{i,3}$ and $d_{i,4}$ are similar to those described in Table 3 for $d_{i,1}$. Assumptions 1 and 2 permit to learn something about the informativeness of $Q_{i,k}$ if and only if the corresponding threshold is outside the suggestive support. Among the subjective probabilities whose thresholds are outside the suggestive support, the criterion selects the one whose partially identified measure of coherence has the highest lower bound and the highest upper bound, except if $r_{i,1} \leq r_{i,min}$ and $Q_{i,1} = 0.50$ (i.e., if $\mathrm{H}[d_{i,1}] = [0.40, 0.50]$). In this latter case, $Q_{i,1} = 0.50$ is chosen in order not to overstate the role of the uninformative subjective probabilities other than 50s in the collapse of the linear relationship.[1] A respondent remains classified as "Unknown"

[1] Indeed, suppose that $r_{i,1} < r_{i,2} < r_{i,min} < r_{i,3} < r_{i,4} < r_{i,max}$, $Q_{i,1} = 0.50$, $Q_{i,2} = 0.40$, and $Q_{i,3} = 0.30$ and $Q_{i,4} = 0.30$. If $\alpha = 0.10$, then $\mathrm{H}[d_{i,1}] = [0.40, 0.50]$, $\mathrm{H}[d_{i,2}] = [0.50, 0.60]$, $\mathrm{H}[d_{i,3}] = \mathrm{H}[d_{i,4}] = [0, 0.70]$. In that case, if I had selected among the subjective probabilities whose thresholds are outside the suggestive support the one whose partially identified measure of coherence has the highest lower bound and the highest upper bound, I would have chosen $Q_{i,2}$. And I would have concluded that (μ_i, σ_i) is meaningless because of an uninformative $Q_{i,2} \neq 0.50$, despite the fact

if the four threshold values ($r_{i,k}$, $k = 1, 2, 3, 4$) belong to $(r_{i,min}, r_{i,max})$. By construction, the size of the "Unknown" has to decrease. If $r_{i,1} \in (r_{i,min}, r_{i,max})$, but at least $r_{i,4} \geq r_{i,max}$, then the respondent is no more classified as "Unknown".

More formally, denote by $\mathcal{J}_i$ the set of thresholds whose values are outside the suggestive support; thus, $\mathcal{J}_i = \{k \in \{1, 2, 3, 4\} | r_{i,k} \notin (r_{i,min}, r_{i,max})\}$. If $\mathcal{J}_i \neq \emptyset$, the criterion considered, denoted C_i, is:

$$C_i \equiv \begin{cases} \left| Q_{i,i} - \widetilde{\mathrm{P}}\,(R_i > r_{i,1}) \right| & \text{if } Q_{i,1} = 0.50 \text{ and } r_{i,1} \notin (r_{i,min}, r_{i,max}) \\ \max_{k \in \mathcal{J}_i} \left| Q_{i,k} - \widetilde{\mathrm{P}}\,(R_i > r_{i,k}) \right| & \text{otherwise} \end{cases} \tag{C1}$$

Under Assumption 1, C_i is partially identified. Let $d_{i,k,l}$ and $d_{i,k,u}$ be the lower and upper bounds of the identification region for $d_{i,k}$, and $C_{i,l}$ and $C_{i,u}$ be the lower and upper bounds of the identification region for C_i. If so, the identification region for C_i is given by:

$$\mathrm{H}\,[C_i] = [C_{i,l}, C_{i,u}] = \begin{cases} [0.50 - \alpha, 0.50] & \text{if } Q_{i,1} = 0.50 \text{ and } r_{i,1} \notin (r_{i,min}, r_{i,max}) \\ [\max_{k \in \mathcal{J}_i} d_{i,k,l}, \max_{k \in \mathcal{J}_i} d_{i,k,u}] & \text{otherwise} \end{cases} \tag{C2}$$

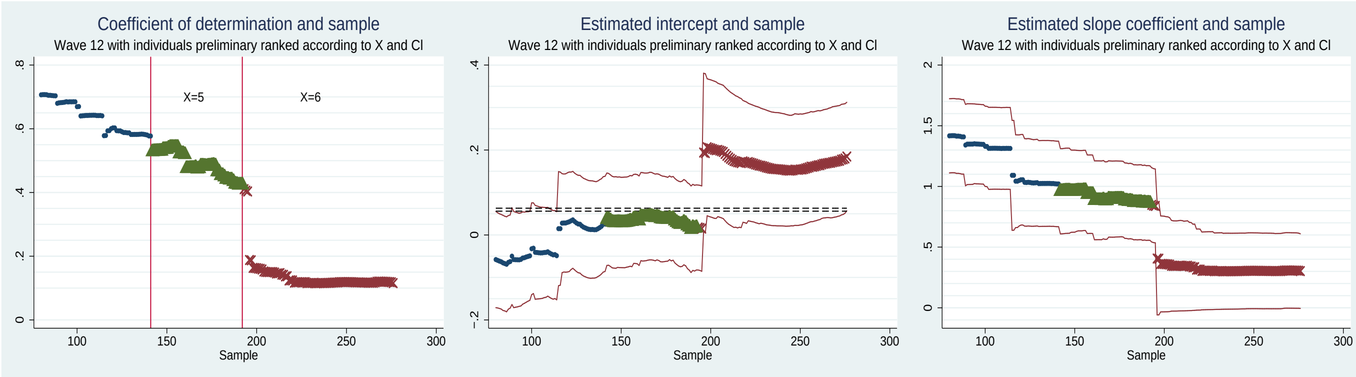
Hence, if $\mathcal{J}_i \neq \emptyset$, the distance C_i can take any value in the interval $[C_{i,l}, C_{i,u}]$. Under Assumption 1, the width of this interval cannot be higher than α (and cannot be less than $\frac{\alpha}{2}$). It is easy to see that Assumption 2 permits to claim that the corresponding subjective probability will be classified in one of the following categories: "Broadly informative", "Broadly or partially informative", "Partially informative", "Uninformative or partially informative", "Uninformative". If $\mathcal{J}_i = \emptyset$, there is no subjective probability at a threshold outside the suggestive support, so we cannot make any conclusion; it corresponds to the category "Unknown".

Except when $r_{i,1} \leq r_{i,min}$ and $Q_{i,1} = 0.50$, note that criterion C_i considered is closely linked to a Kolmogorov-Smirnov (KS) distance. The KS distance, which is often used to compare two continuous one-dimensional probability distributions, is the maximum distance between the two cumulative distribution functions of R_i (the one based on the two preliminary questions and the one based on the sequence of probabilistic questions); or, equivalently, the maximum distance between the complementary cumulative of these two distributions: $KS_i \equiv \max_{k=1,\ldots,4} \left| Q_{i,k} - \widetilde{\mathrm{P}}\,(R_i > r_{i,k}) \right| = \max_{k=1,\ldots,4} d_{i,k}$. Under Assumption 1, the KS distance is partially identified. A problem with the partially identified KS distance is that there is a class of situations where we learn almost nothing about this distance despite the fact that some of the thresholds are outside the suggestive support. Suppose that $r_{i,1} < r_{i,min} < r_{i,2} < r_{i,max} < r_{i,3} < r_{i,4}$, $Q_{i,1} = 1$, $Q_{i,2} = 0.90$ and $Q_{i,3} = 0.15$ and $Q_{i,4} = 0$. If $\alpha = 0.10$, then $\mathrm{H}\,[d_{i,1}] = \mathrm{H}\,[d_{i,4}] = [0, 0.10]$, $\mathrm{H}\,[d_{i,2}] = [0, 0.90]$, $\mathrm{H}\,[d_{i,3}] = [0.05, 0.15]$. As a consequence, we infer that $Q_{i,1}$ and $Q_{i,4}$ may be informative, $Q_{i,3}$ may be broadly or partially informative, but the type of $Q_{i,2}$ is unknown. The bound on the KS distance is $[0.05, 0.90]$. So if one chooses a KS distance as a criterion, he will conclude that the highest distance is between 0.05 and 0.90; hence he learns almost nothing about the KS distance. The difficulty to identify the KS distance is due to the fact that Assumption 1 is too weak to learn something about the distance at a threshold $r_{i,k} \in (r_{i,min}, r_{i,max})$. That is why C_i considers the highest distance among the distances at a threshold outside the suggestive support (except when $r_{i,1} \leq r_{i,min}$ and $Q_{i,1} = 0.50$).

64 respondents are classified as "Unknown" in wave 12 (against 93 in Section 3 and Subsection 4.1), 42 in wave 13 (against 64), and 85 in wave 14 (against 123). This is less than when we only focus on the first threshold, but it remains imperfect. I find however similar results to those found in Subsection 4.1: the (μ, σ) based on at least one uninformative Q_k are responsible for the collapse of the linear relationship. When these (μ, σ) are not included, the R^2 is around 60 percent. When we exclude only the respondents whose less informative Q_k is an uninformative 50, the R^2 is around 40 percent.

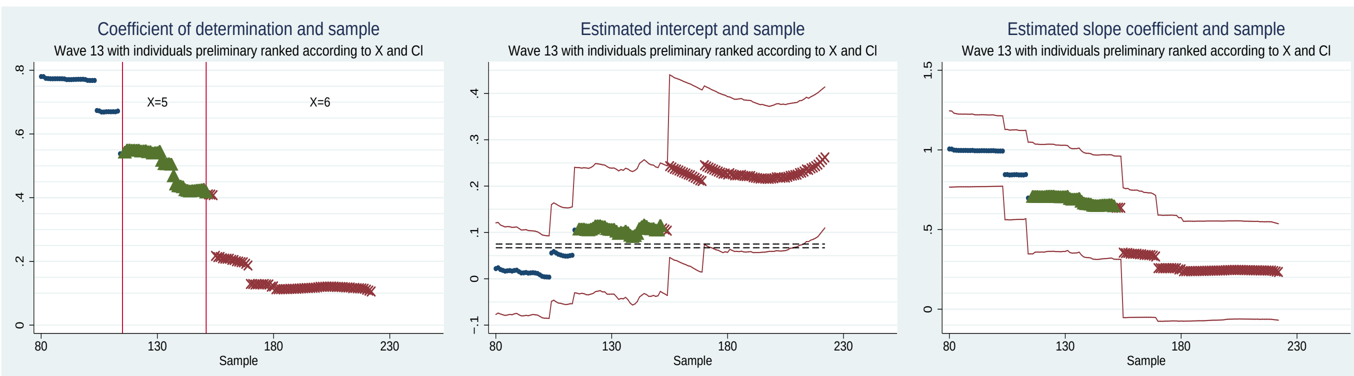
---

that it is already meaningless because $Q_{i,1} = 0.50$ is uninformative. This problem occurs because of the monotonicity condition of the complementary cumulative distribution function and because $r_{i,2} < r_{i,min}$.



(a) (b) (c)

Figure C1: Coefficient of determination, estimated intercept and slope coefficients conditional on the sample (respondents preliminary ranked according to X_i and then $C_{i,l}$, category "Unknown" excluded) -wave 12-
Notes: i. The two solid red lines in Panels (b) and (c) represent the 95% confidence intervals of the intercept and slope coefficients of the OLS estimates. They are computed using heteroskedasticity-robust standard errors.
ii. The two dashed black lines in Panel (b) correspond to the minimum (0.056) and maximum (0.063) values of the daily one-year U.S. LIBOR rate during wave 12 (July 1999 to November 1999).



(a) (b) (c)

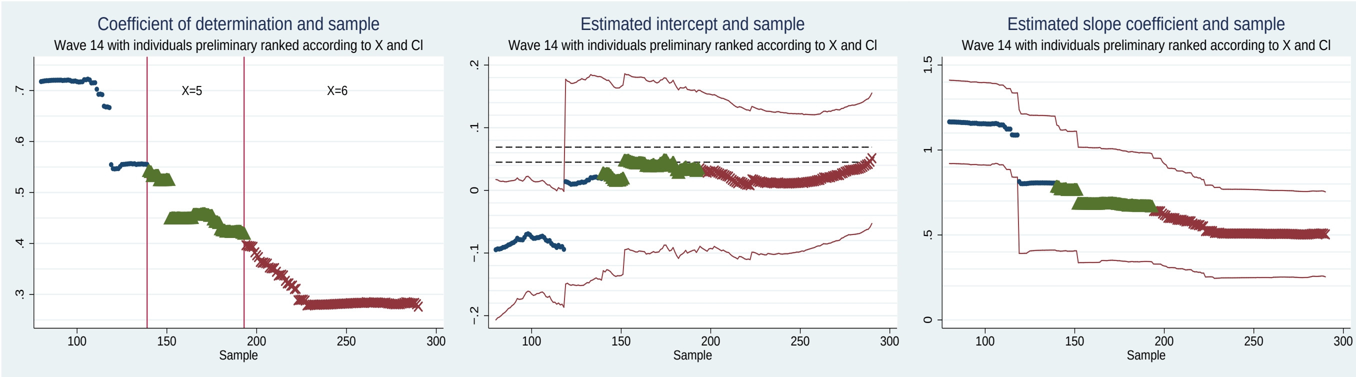
Figure C2: Coefficient of determination, estimated intercept and slope coefficients conditional on the sample (respondents preliminary sorted according to X_i and then $C_{i,l}$, category "Unknown" excluded) -wave 13-
Notes: i. The two solid red lines in Panels (b) and (c) represent the 95% confidence intervals of the intercept and slope coefficients of the OLS estimates. They are computed using heteroskedasticity-robust standard errors.
ii. The two dashed black lines in Panel (b) correspond to the minimum (0.067) and maximum (0.075) values of the daily one-year U.S. LIBOR rate during wave 13 (February 2000 to May 2000).

C.2 Adding assumptions to reduce the identification region of $d_{i,1}$

Although the criterion of Subsection C.1 permits to decrease the size of the "Unknown", it remains imperfect for two reasons. First, some respondents are still classified as "Unknown". Second, this criterion does not focus on the answer for the first threshold. This answer is important because it is not influenced by the answers for the subsequent thresholds, while the reverse is not true. To learn something about the informativeness of the Q_1 classified as "Unknown", and include them in the analysis of the cross-sectional distribution of (μ, σ), this Subsection adds assumptions to reduce the prediction region when $r_{i,1} \in (r_{i,min}, r_{i,max})$. By reducing this prediction region, the identification region of the measure of coherence will be reduced too. This Section considers assumptions which are stronger than Assumption 1 but weaker than point-identifying assumptions. These assumptions clearly highlight that there is a tension between the credibility of these additional assumptions and the possibility to learn something about the informativeness of the "Unknown".

More precisely, I assume that:

Assumption C1. *Assumption 1 holds. Moreover,*



(a) (b) (c)

Figure C3: Coefficient of determination, estimated intercept and slope coefficients conditional on the sample (respondents preliminary sorted according to X_i and then $C_{i,l}$, category "Unknown" excluded) -wave 13-
Notes: i. The two solid red lines in Panels (b) and (c) represent the 95% confidence intervals of the intercept and slope coefficients of the OLS estimates. They are computed using heteroskedasticity-robust standard errors.
ii. The two dashed black lines in Panel (b) correspond to the minimum (0.045) and maximum (0.069) values of the daily one-year U.S. LIBOR rate during wave 14 (September 2000 to March 2001).

(i.) $\theta_1 \leq \widetilde{P}\left(R_i \geq \frac{r_{i,min}+r_{i,max}}{2}\right) \leq 1-\theta_1$, *where* $\theta_1 \in [0, 0.5]$ *is a probability to be fixed.*
(ii.) $\theta_2 \leq \widetilde{P}(R_i \geq r_{i,min} + \delta_i) \leq 1$, *where* $\theta_2 \in [0, 1-\alpha]$ *is a probability and* $\delta_i \in \left[0, \frac{r_{i,max}-r_{i,min}}{2}\right]$ *a positive real number to be fixed.*
(iii.) $0 \leq \widetilde{P}(R_i \geq r_{i,max} - \delta_i) \leq 1-\theta_2$

Parts (i.), (ii.) and (iii.) in Assumption C1 easily provide the prediction region $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)\right]$ when $r_{i,1} = \frac{r_{i,min}+r_{i,max}}{2}$, $r_{i,1} = r_{i,min} + \delta_i$ and $r_{i,1} = r_{i,max} - \delta_i$. When $r_{i,1}$ is between two of these points, linear interpolations are used to find the bound of the prediction region.[2] Panel (a) in Figure C4 depicts the prediction region (the shaded area) under Assumption C1 and the linear interpolations. In order to see how Assumption C1 reduces the prediction region, Panel (b) depicts the prediction region under Assumption 1. Note that there are three new parameters: the two probabilities $\theta_1 \in [0, 0.5]$ and $\theta_2 \in [0, 1-\alpha]$, and a positive real number $\delta_i \in \left[0, \frac{r_{i,max}-r_{i,min}}{2}\right]$ that have to be fixed. Figure C4 shows the trade-off between the credibility of the values chosen for θ_1, θ_2 and δ_i and the possibility to learn something about the informativeness of the "Unknown". It is easy to see in Panel (a) that if $\delta_i = 0$, $\theta_1 = 0$ and $\theta_2 = 0$, the prediction region is similar to the prediction region under Assumption 1, i.e., $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)\right] = [0, 1]$ for all $r_{i,1} \in (r_{i,min}, r_{i,max})$. These values for δ_i, θ_1 and θ_2 are extremely credible because the prediction region encompasses all possible point predictions. However these values will not permit to learn something about the type of the "Unknown" because the width of the prediction region is too wide. On the contrary, if one considers that $\theta_1 = 0.45$, $\theta_2 = 1-\alpha = 0.9$ and $\delta_i = 0$, these values will permit to learn something about the "Unknown" because the width of the prediction region is 0.10 (so sufficiently small) for all $r_{i,1} \in (r_{i,min}, r_{i,max})$. It is however difficult to argue for the credibility of so high values for θ_1 and θ_2. To fully understand, consider, e.g., θ_1 which is associated to the prediction region at the midpoint $\frac{r_{i,min}+r_{i,max}}{2}$, i.e., $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1} = \frac{r_{i,min}+r_{i,max}}{2}\right)\right] = [\theta_1, 1-\theta_1]$.

[2]The prediction regions are thus given by:
(i.) $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | Q_{i,1}, r_{i,1} \leq r_{i,min}\right)\right] = [1-\alpha, 1]$
(ii.) $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1} \in (r_{i,min}, r_{i,min}+\delta_i]\right)\right] = \left[\theta_2 - (1-\alpha-\theta_2)\frac{r_{i,1}-r_{i,min}-\delta_i}{\delta_i}, 1\right]$
(iii.) $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1} \in (r_{i,min}+\delta_i, z_i]\right)\right] = \left[\theta_1 + (\theta_2-\theta_1)\frac{r_{i,1}-z_i}{r_{i,min}+\delta_i-z_i}, 1-\theta_1+\theta_1\frac{r_{i,1}-z_i}{r_{i,min}+\delta_i-z_i}\right]$
(iv.) $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1} \in (z_i, r_{i,max}-\delta_i]\right)\right] = \left[\theta_1\frac{r_{i,1}-r_{i,max}+\delta_i}{z_i-r_{i,max}+\delta_i}, 1-\theta_2+(\theta_2-\theta_1)\frac{r_{i,1}-r_{i,max}+\delta_i}{z_i-r_{i,max}+\delta_i}\right]$
(v.) $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1} \in (r_{i,max}-\delta_i, r_{i,max})\right)\right] = \left[0, \alpha-(1-\theta_2-\alpha)\frac{r_{i,1}-r_{i,max}}{\delta_i}\right]$
(vi.) $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | Q_{i,1}, r_{i,1} \geq r_{i,max}\right)\right] = [0, \alpha]$

It is easy to argue that $\theta_1 > 0$, i.e., this prediction region is not $[0, 1]$: the lower bound cannot be 0 because the respondent should have provided initially a much lower $r_{i,max}$; the upper bound cannot be 1 because the respondent should have provided initially a much higher $r_{i,min}$. If it is logical that $\theta_1 > 0$, the difficulty however is to argue for a too high value of θ_1.

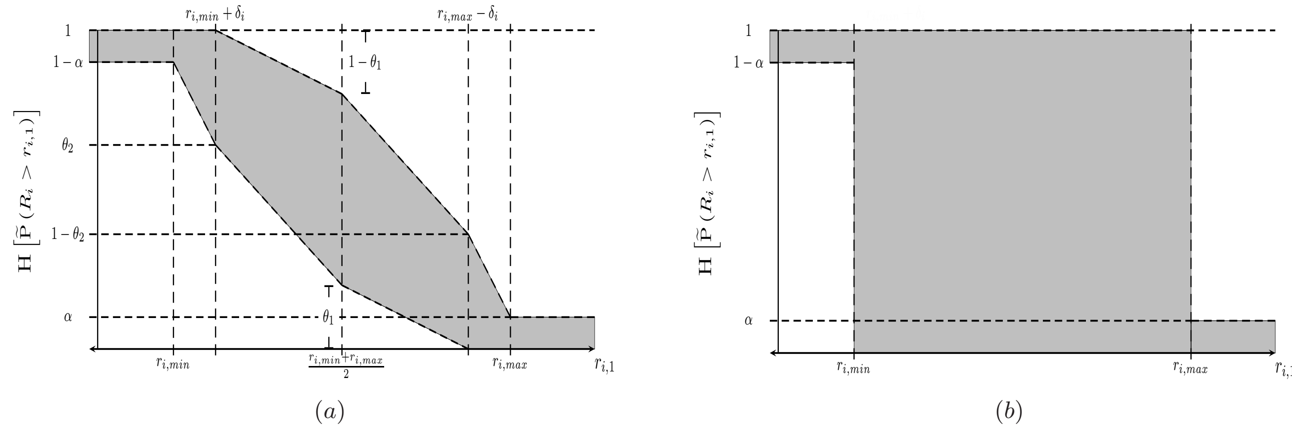


Figure C4: Prediction region $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1} | r_{i,1}\right)\right]$ under Assumption C1 (Panel (a)) and under Assumption 1 (Panel (b))

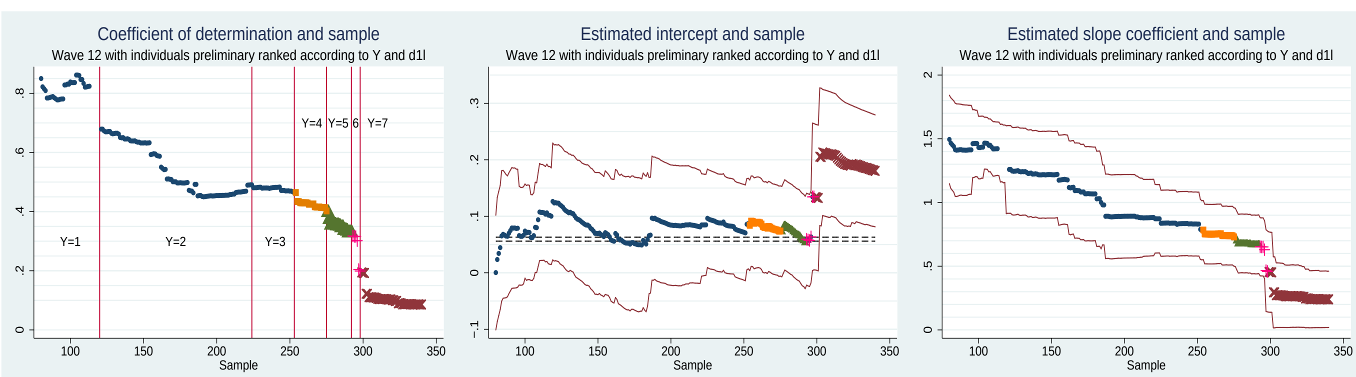
Given that the objective is to classify almost all the "Unknown" of Subsection 4.1, the values for the three new parameters have been chosen to do so: $\theta_1 = 0.30$, $\theta_2 = 0.75$ and $\delta_i = \frac{r_{i,max}-r_{i,min}}{8}$. Given the discussion above, these values may generate some misclassifications, but we believe that they provide a good middle ground. Remember that the prediction region under Assumption 1, i.e., $\mathrm{H}\left[\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)\right] = [0, 1]$ for all $r_{i,1} \in (r_{i,min}, r_{i,max})$, encompasses all possible point predictions. That is all possible point-predicting assumptions generate point predictions which always belong to the prediction region under Assumption 1. Under Assumption C1 and the values mentioned above for the parameters, not all but numerous point-predicting assumptions generate point predictions $\widetilde{\mathrm{P}}\left(R_i > r_{i,1}\right)$ which belong to this prediction region. In particular, using the Final sample, we have verified that it is true when **one assumes that the first subjective distribution used by the respondent to report $r_{i,min}$ and $r_{i,max}$ follows** a continuous uninform distribution over the support $(r_{i,min}, r_{i,max})$. It is also true for a set of point-predicting assumptions based on triangular distributions, with lower limit $r_{i,min}$, upper limit $r_{i,max}$ and different modes $c_i \in [\frac{r_{i,min}+r_{i,max}}{2} - \delta_i, \frac{r_{i,min}+r_{i,max}}{2} + \delta_i])$. Lastly, we have considered a set of point-predicting assumptions based on normal distributions. More precisely, we have considered the following assumptions: (i.) $\widetilde{\mathrm{P}}\left(R_i \leq r_{i,min}\right) = \widetilde{\mathrm{P}}\left(R_i \geq r_{i,max}\right) = \gamma$ and (ii.) $\widetilde{\mathrm{P}}\left(R_i \geq \frac{r_{i,min}+r_{i,max}}{2}\right) = \widetilde{\mathrm{P}}\left(R_i \leq \frac{r_{i,min}+r_{i,max}}{2}\right) = \frac{1}{2}$. Assumption (i.) says that the answers to the preliminary questions give the $(\gamma,(1-\gamma))$-quantiles of a respondent's first subjective distribution, where γ is homogenous across respondents. Assumption (ii.) says that $\frac{r_{i,min}+r_{i,max}}{2}$ is the median of the first subjective distribution. Assumptions (i.) and (ii.) combined with the assumption that the first subjective distribution is normal permit to obtain a point prediction. The point predictions, for all $\gamma \in (0, 0.10]$, always belong to the prediction region under Assumption C1, considering the values mentioned for the parameters θ_1, θ_2 and δ_i.

Using the prediction region under Assumption C1 (see Footnote 2), one can easily find the identification region of the measure of coherence. Considering that $\overline{d} = 0.50 - \alpha$, $\underline{d} = \alpha$ and $\alpha = 0.10$, most of the Q_1 classified as "Unknown" in Section 3 are classified as "Partially or broadly informative" or "Uninformative or partially informative". For instance, in wave 12, of the 93 respondents classified as "Unknown" in Section 3 (see Table 5), 63 are classified as "Partially or broadly informative", 20 as "Uninformative or partially informative", 2 as "Partially informative", 1 as "Uninformative", and

7 remains "Unknown" and are included in the "Broadly informative" without influencing the results. Concerning in particular the 17 $Q_1 = 50$ of type "Unknown" in Section 3 (see Table 5), 11 are classified as "Partially or broadly informative" and 6 as "Uninformative or partially informative".

In some preliminary analyses, at least in wave 12, the 6 (μ, σ) based on "Uninformative or partially informative" $Q_1 = 0.50$ strongly influenced the cross-sectional distribution of (μ, σ). This is not a problem per se, given that these Q_1 are clearly not informative. However, to be sure that the (μ, σ) based on "Uninformative" $Q_1 \neq 0.50$ also influence the cross-sectional distribution, they are included before in the results presented. We have preliminary sorted the respondents in ascending order based on the variable Y_i: $Y_i = 1$ if $Q_{i,1}$ is classified as "Broadly informative", $Y_i = 2$ if $Q_{i,1}$ is classified as "Partially or broadly informative", $Y_i = 3$ if $Q_{i,1}$ is classified as "Partially informative", $Y_i = 4$ if $Q_{i,1} \neq 0.50$ and is classified as "Uninformative or partially informative", $Y_i = 5$ if $Q_{i,1} \neq 0.50$ is classified as "Uninformative", $Y_i = 6$ if $Q_{i,1} = 0.50$ is classified as "Uninformative or partially informative", and $Y_i = 7$ if $Q_{i,1} = 0.50$ is classified as "Uninformative". Then, for respondents with common Y, we have sorted them according to $d_{1,l}$. Figures C5, C6 and C7 present the results for waves 12, 13 and 14. To easily see the impact of the different "Uninformative or partially informative" and "Uninformative", I use an (orange) square (■) when the ultimate (μ_i, σ_i) added in the sample is based on an "Uninformative or partially informative" $Q_{i,1} \neq 0.50$, a (green) triangle (▲) when it is based on an "Uninformative" $Q_{i,1} \neq 0.50$, a (pink) plus (+) when it is based on an "Uninformative or partially informative" $Q_{i,1} = 0.50$, and a (red) x mark (×) when it is based on an "Uninformative" $Q_{i,1} = 0.50$.

Whatever the wave, the R^2 decreases in an important way when we include the (μ, σ) based on "Partially or broadly informative" Q_1, but it remains however around 50 percent (in waves 12 and 13) or 60 percent (in wave 14). These percentages are stable when the (μ, σ) based on "Partially informative" Q_1 are included. The R^2 then quickly decreases when the other (μ, σ) are included. For instance, in wave 14, when the (μ, σ) based on "Uninformative or partially informative" $Q_1 \neq 0.50$ are included, the R^2 decreases from 60 to 50 percent. Then, including the (μ, σ) based on "Uninformative" $Q_1 \neq 0.50$ decreases the R^2 to 40 percent. Lastly, including the (μ, σ) based on "Uninformative" $Q_1 = 0.50$ decreases the R^2 to 30 percent. The (μ, σ) based on "Uninformative or partially informative" $Q_1 = 0.50$ do not have an important impact on the R^2 in this wave (although they do in wave 12).

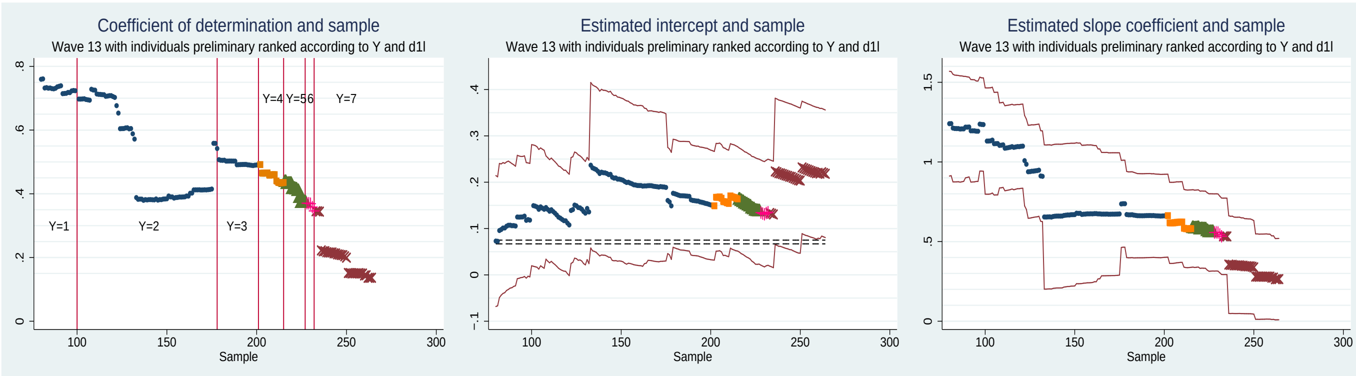


(a) (b) (c)

Figure C5: Coefficient of determination, estimated intercept and slope coefficients conditional on the sample (respondents preliminary sorted according to Y_i and then $d_{i,1,l}$) -wave 12-

Notes: i. The two solid red lines in Panels (b) and (c) represent the 95% confidence intervals of the intercept and slope coefficients of the OLS estimates. They are computed using heteroskedasticity-robust standard errors.

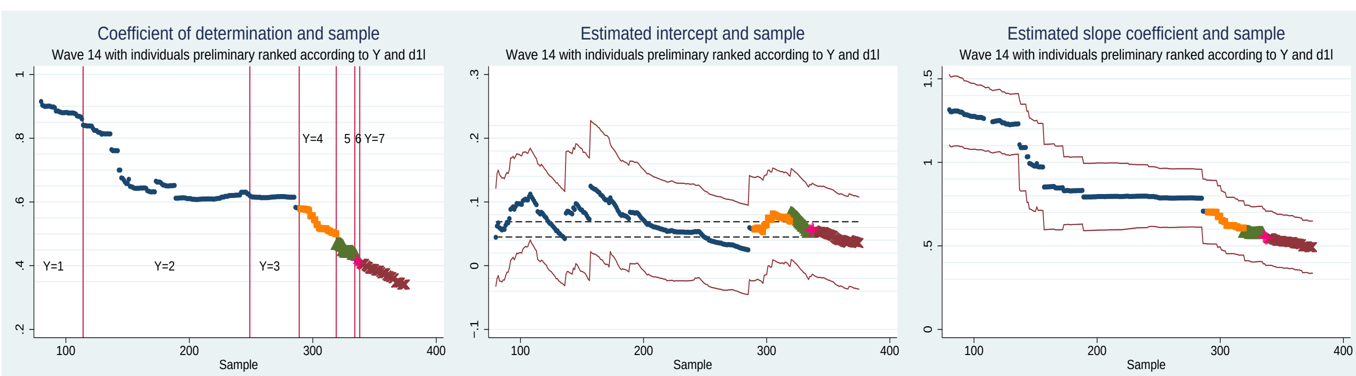
ii. The two dashed black lines in Panel (b) correspond to the minimum (0.056) and maximum (0.063) values of the daily one-year U.S. LIBOR rate during wave 12 (July 1999 to November 1999).



(a) (b) (c)

Figure C6: Coefficient of determination, estimated intercept and slope coefficients conditional on the sample (respondents preliminary sorted according to Y_i and then $d_{i,1,l}$) -wave 13-

Notes: i. The two solid red lines in Panels (b) and (c) represent the 95% confidence intervals of the intercept and slope coefficients of the OLS estimates. They are computed using heteroskedasticity-robust standard errors.
ii. The two dashed black lines in Panel (b) correspond to the minimum (0.067) and maximum (0.075) values of the daily one-year U.S. LIBOR rate during wave 13 (February 2000 to May 2000).



(a) (b) (c)

Figure C7: Coefficient of determination, estimated intercept and slope coefficients conditional on the sample (respondents preliminary ranked according to Y_i and then $d_{i,1,l}$) -wave 14-

Notes: i. The two solid red lines in Panels (b) and (c) represent the 95% confidence intervals of the intercept and slope coefficients of the OLS estimates. They are computed using heteroskedasticity-robust standard errors.
ii. The two dashed black lines in Panel (b) correspond to the minimum (0.045) and maximum (0.069) values of the daily one-year U.S. LIBOR rate during wave 14 (September 2000 to March 2001).