The effect of learning on ambiguity attitudes

Aurélien Baillon[a], Han Bleichrodt[a], Umut Keskin[a], Olivier L'Haridon[b], Chen Li[a]

[a]Erasmus School of Economics, Rotterdam, the Netherlands.

[b]University of Rennes 1-Crem, Rennes, France.

26 October 2016

Abstract

This paper studies the effect of learning information on people's attitudes towards ambiguity. We propose a method to separate ambiguity attitudes from subjective probabilities and to decompose ambiguity attitudes into two components. Under models like prospect theory that represent ambiguity through nonadditive decision weights these components reflect pessimism and likelihood insensitivity. Under multiple priors models, they reflect ambiguity aversion and perceived ambiguity. We apply our method in an experiment where we elicit the ask prices of options with payoffs depending on the returns of initial public offerings (IPOs) on the New York Stock Exchange. IPOs are a natural context to study the effect of learning, as prior information about their returns is unavailable. Subjects perceived substantial ambiguity and they were insensitive to likelihood information. We observed only little pessimism and ambiguity aversion. Subjective probabilities were well-calibrated and close to the true frequencies. Subjects' behavior moved towards expected utility with more information, but substantial deviations remained even in the maximum information condition.

Keywords: ambiguity, learning, updating, neo-additive weighting.

Acknowledgements: Stefan Trautmann, Ferdinand Vieider, Peter P. Wakker, Utz Weitzel, the department editor Manel Baucells, an associate editor, and three reviewers gave helpful comments. The research of Aurélien Baillon and Han Bleichrodt was made possible by grants from the Netherlands Organization for Scientific Research (NWO). Olivier l'Haridon's research was supported by grants from the Institut Universitaire de France and Rennes Metropole district.

1. Introduction

In many real-world decisions, objective probabilities are unknown and decisions have to be made under uncertainty. Decision analysis traditionally analyzes such decisions by assuming that the decision maker assigns subjective probabilities to events, behaves according to expected utility, and updates his subjective probabilities according to Bayes' rule when new information arrives. All these assumptions are open to debate.

First, while people change their beliefs when more information becomes available and these updated beliefs have predictive value (Hamermesh 1985, Smith et al. 2001), empirical evidence suggests that they systematically deviate from Bayes' rule (e.g. Grether 1980, El-Gamal and Grether 1995, Charness and Levin 2005, Hoffman et al. 2011, Poinas et al. 2012, Gallagher 2014). Psychologists have uncovered many updating biases, including under- and overconfidence (Griffin and Tversky 1992), conservatism (Phillips and Edwards 1966), representativeness (Kahneman and Tversky 1972), availability (Tversky and Kahneman 1973), and confirmatory bias (Rabin and Schrag 1999).

Second, and even more fundamental, Ellsberg's (1961) paradox, which shows that people prefer betting on events with known odds over betting on events with unknown odds, undermines not only subjective expected utility, but even the existence of subjective probabilities. To account for Ellsberg's paradox, new models of decision under ambiguity have been proposed (for an overview see Gilboa and Marinacci 2013). While in expected utility decision weights are equal to subjective probabilities, in the ambiguity models they also reflect the confidence people have in their beliefs and their aversion towards ambiguity. The ambiguity models capture an intuition expressed by Keynes (1921):

"The magnitude of the probability of an argument…depends upon a balance between what may be termed the favourable and the unfavourable evidence; a new piece of evidence which leaves this balance unchanged also leaves the probability of the argument unchanged.

But it seems that there may be another respect in which some kind of quantitative comparison between arguments is possible. This comparison turns upon a balance, not between the favourable and the unfavourable evidence, but between the *absolute* amounts of relevant knowledge and relevant ignorance respectively" [p.71].

Keynes conjectured that new information changes both the balance of evidence (people's beliefs) and the total amount of evidence (the amount of ambiguity). In expected utility, ambiguity attitudes play no role and learning only affects beliefs. In the ambiguity models, new information changes both beliefs and ambiguity attitudes. This raises the question of how ambiguity attitudes change when new information becomes available. While several papers have approached this question from a theoretical angle and different rules have been proposed (Gilboa and Schmeidler 1993, Epstein 2006, Eichberger et al. 2007, Epstein and Schneider 2007, Hanany and Klibanoff 2007, Eichberger et al. 2010, Eichberger et al. 2012a), there is a dearth of empirical evidence.[1] This motivated our paper in which we study experimentally how decision makers change their ambiguity attitudes when more information becomes available.

A difficulty in applying ambiguity models is that they involve concepts that are hard to measure empirically. We present a simple method to measure and decompose ambiguity attitudes while controlling for beliefs. We start with the biseparable preferences model of Ghirardato and Marinacci (2001), which for binary acts includes many ambiguity models as special cases. Under biseparable preferences we can observe the effect of ambiguity aversion in the sense of Ghirardato and Marinacci (2002) and explore what happens to it when more information becomes available. The literature usually distinguishes a decision maker's

---

[1] Cohen et al. (2000) and Dominiak et al. (2012) studied updating under ambiguity experimentally but they considered situations in which decision makers receive information that an event cannot occur. In our study decision makers accumulate evidence how often a particular event has been observed in the past.

ambiguity perception from his ambiguity aversion. To study these two components separately, we impose additional structure by assuming the neo-additive preferences of Chateauneuf et al. (2007). Neo-additive preferences have two intuitive interpretations. One interpretation is close to prospect theory and models ambiguity through nonadditive decision weights. We coin this interpretation the *decision weight interpretation*. The other interpretation is close to the multiple priors models and represents ambiguity through a (nonsingular) set of priors. We coin this interpretation the *multiple priors interpretation*.

Both the decision weight interpretation and the multiple priors interpretation permit describing a decision maker's ambiguity attitude by two indices. In the decision weight interpretation, one index measures the decision maker's degree of pessimism, which includes his ambiguity aversion, and the other index measures his insensitivity to likelihood information. Under the multiple priors interpretation, one index measures the decision maker's ambiguity aversion and the other his ambiguity perception.

We applied our method in an experiment. We elicited subjects' ask prices for options with payoffs contingent on the returns of (anonymous) initial public offerings (IPOs). IPOs make it possible to study the effect of more information in a natural decision context for which no prior information about returns is available. We used finance students as subjects because they were familiar with pricing options and we hoped that this would improve the quality of the data.

The results indicated that subjects perceived considerable ambiguity and most subjects were insensitive to likelihood information. We did not observe much pessimism and ambiguity aversion. Aggregate subjective probabilities were close to their true frequencies after correction for ambiguity attitudes. Subjects' behavior moved in the direction of expected utility with more information about the historical performance of the stocks. However, substantial deviations remained even in the maximum information condition.

2. Theoretical framework

Decision model

A decision maker faces uncertainty about the outcome he will receive at time point T. At each time point $t = 0,1,2, \ldots$, one element s_t of a set $\mathcal{S}$ is realized. Thus the complete uncertainty is modeled through the state space $\prod_{t=1}^{T} \mathcal{S}_t$ where $\mathcal{S}_t = \mathcal{S}$ for all $t > 0$. The state space contains all possible *states of the world* s. Only one state occurs, but the decision maker does not know which one. *Events* are subsets of the state space. The decision maker chooses between *binary acts*, denoted by $x_E y$, giving money amount x if event E occurs and money amount $y \leq x$ otherwise.

Information is modeled using an ascending filtration $\{\mathcal{F}_t\}$.[2] The decision maker's information about previous resolutions of uncertainty up to time $t < T$ is formalized by his *history set* $h_t = (s_1, \ldots, s_t)$, where $s_j \in \mathcal{S}$ for all $1 \leq j \leq t$. Complete absence of information is denoted h_0.

The decision maker's preferences are represented through a *history-dependent preference relation* $\succcurlyeq_t$ where the subscript t indicates that preferences depend on the history h_t.[3] The relations $\succ_t$ and $\sim_t$ are defined as usual. A real-valued function V_t *represents* $\succcurlyeq_t$ if for all binary acts $x_E y$ and $z_F w$, $x_E y \succcurlyeq_t z_F w$ iff $V_t(x_E y) \geq V_t(z_F w)$.

The Bayesian approach assumes that preferences $\succcurlyeq_t$ are represented by *expected* utility: $x_E y \mapsto P_t(E)U(x) + (1 - P_t(E))U(y)$, with U a real-valued utility function defined over outcomes and P_t the subjective probability measure given h_t. In expected utility, new information, which expands the history set from h_t to h_v, $v > t$, affects subjective

[2] An ascending filtration is an increasing sequence of σ-algebras on a measurable space.

[3] Strictly speaking we should have written $\succcurlyeq_{h_t}$. For the sake of readability we used $\succcurlyeq_t$. Similarly we only used subscripts t rather than h_t in the rest of the paper.

probabilities but leaves utility unchanged. Updating takes place in the belief (subjective probabilities) part of the representation and tastes (utility) are not influenced by new information about past events. Time-invariant utility is also commonly assumed in the theoretical literature on updating under non-expected utility (e.g. Epstein 2006, Eichberger et al. 2007, Epstein and Schneider 2007) and we will also assume it in this paper.

To account for deviations from expected utility, we assume the biseparable preferences model of Ghirardato and Marinacci 2001), which includes many ambiguity models as special cases.[4] Under biseparable preferences $\succcurlyeq_t$ can be represented by

$$x_E y \mapsto W_t(E)U(x) + \big(1 - W_t(E)\big)U(y), \qquad (1)$$

with U *a real-valued function* unique up to level and unit and W_t a unique *weighing function*,[5] which shares with a probability measure the properties that $W_t\,(\emptyset) = 0$, $W_t(\mathcal{S}) = 1$ and $W_t(A) \leq W_t(B)$ if $A \subseteq B$, but which may be non-additive. The subscript t in W_t expresses that the decision weight W_t depends on the history h_t just like P_t in the Bayesian approach.

Biseparable preferences make it possible to quantify deviations from expected utility and to identify ambiguity averse decision makers. In our experiment, described in Section 4, we partitioned the state space into three events, which were defined by the changes in the prices of initial public offerings (IPOs) on the 21st trading day after their introduction. The events we considered were Up, the price goes up by at least 0.5% on the 21st trading day; $Middle$, the price varies by less than 0.5% on the 21st trading day; $Down$, the price decreases by at least 0.5% on the 21st trading day; and $MiddleUp = Middle \cup Up$. For given payoffs

[4] Examples are maxmin expected utility (Gilboa and Schmeidler 1989), alpha-maxmin expected utility (Ghirardato et al. 2004), contraction expected utility (Gajdos et al. 2008), Choquet expected utility (Schmeidler 1989), and prospect theory (Tversky and Kahneman 1992).

[5] Sometimes the term *capacity* is used instead of weighting function.

$x > y$ and history h_t, we then elicited four certainty equivalents, $CE_t(Up) \sim_t x_{Up}y$, $CE_t(Middle) \sim_t x_{Middle}y$, $CE_t(Down) \sim_t x_{Down}y$, and $CE_t(MiddleUp) \sim_t x_{MiddleUp}y$. With the normalization $U(x) = 1$ and $U(y) = 0$, Eq. (1) implies that $U\big(CE_t(MiddleUp)\big) = W_t(Up)$, $U(CE_t(Middle)) = W_t(Middle)$, $U(CE_t(Down)) = W_t(Down)$, and $U(CE_t(MiddleUp)) = W_t(MiddleUp)$.

Under expected utility, the decision weights W_t are equal to the decision maker's subjective beliefs and thus $W_t(MiddleUp) + W_t(Down) = P_t(MiddleUp) + P_t(Down) = 1$. We will refer to this property as *complementarity*.

Ghirardato and Marinacci (2002) define a decision maker as ambiguity averse if he rejects any ambiguous act that an expected utility maximizer with the same risk attitude would reject. Observation 2 in Appendix A shows that this implies that for an ambiguity averse decision maker $W_t(MiddleUp) + W_t(Down) < 1$. Observation 1 in Appendix A shows that if decision maker A is more ambiguity averse than decision maker B in the sense of Ghirardato and Marinacci (2002), A deviates more from complementarity than B does and, consequently the index $1 - \big(W_t(MiddleUp) + W_t(Down)\big)$ is consistent with relative ambiguity aversion in the sense of Ghirardato and Marinacci (2002).

Most ambiguity models distinguish ambiguity aversion and ambiguity perception. A decision maker who perceives extreme ambiguity might treat all events except the null event and the universal event alike, assigning the same weight to all these events. Such a decision maker may satisfy complementarity (if he assigns the weight 0.5 to all events) and would not be ambiguity averse in the sense of Ghirardato and Marinacci (2002), but he would violate the additivity rule of probabilities and display what Tversky and Wakker (1995) call lower subadditivity. In our experiment this would reveal itself as $W_t(Up) + W_t(Middle) > W_t(MiddleUp)$. To distinguish ambiguity perception and ambiguity aversion we need to impose additional structure. We do so by assuming the neo-additive preferences of

Chateauneuf et al. (2007). As we show next, these preferences have two intuitive interpretations, one in terms of decision weights, the other in terms of multiple priors. Both interpretations permit deriving two indices that jointly capture ambiguity attitudes. In the decision weight interpretation these indices reflect pessimism (which includes ambiguity aversion) and likelihood insensitivity. In the multiple priors interpretation they reflect ambiguity aversion and ambiguity perception.

3. Decomposing ambiguity attitudes

3.1. The decision weight interpretation

For parameters a_t and b_t such that $0 \leq a_t \leq 1$ and $-a_t \leq b_t \leq a_t$, and for a probability measure P_t, *neo-additive decision weighting* W_t, is defined as:

$$W_t(E) = \frac{a_t - b_t}{2} + (1 - a_t)P_t(E) \quad \text{if } 0 < P_t(E) < 1,$$

$$W_t(E) = 0 \quad \text{if } P_t(E) = 0, \text{ and} \tag{2}$$

$$W_t(E) = 1 \quad \text{if } P_t(E) = 1.$$

Equation (2) shows that for each history h_t, the decision maker's preferences are consistent with a probability measure P_t. In other words, neo-additive decision weighting assumes that the decision maker is *probabilistically sophisticated for a given history* h_t. However, a_t and b_t may vary across histories and the decision maker may deviate from probabilistic sophistication when comparing acts involving different histories.

Biseparable preferences with neo-additive weighting can be written as:

$$x_E y \mapsto (1 - a_t)\left[P_t(E)U(x) + \left(1 - P_t(E)\right)U(y)\right] + \frac{a_t - b_t}{2}U(x) + \frac{a_t + b_t}{2}U(y). \tag{3}$$

Equation (3) is a linear combination of the expected utility of the act $x_E y$ ($P_t(E)U(x) + \left(1 - P_t(E)\right)U(y)$), the maximum utility $U(x)$, and the minimum utility $U(y)$. This expression helps to understand the intuition underlying the parameters a_t and b_t, which define two ambiguity indices as we show next.

Pessimism

Figure 1 shows that for a given value of a_t, increases in b_t shift the weighting functions downwards (by $b_t/2$). As can be seen from Eq. (1), the decision weights reflect the weight given to the best outcome and, consequently, increases in b_t (holding a_t constant) imply that the decision maker pays more attention to the worst outcome. We will, therefore, interpret b_t as an index of *pessimism* with higher values indicating more pessimism and negative values reflecting optimism. An expected utility maximizer has $b_t = 0$ (and $a_t = 0$). An extremely pessimistic decision maker, who only considers the worst outcome regardless of its likelihood, has $b_t = 1$ and an extremely optimistic decision maker, who only considers the best outcome, has $b_t = -1$.

Eq. (3) gives:

$$W_t(MiddleUp) + W_t(Down) =$$

$$\frac{a_t - b_t}{2} + (1 - a_t)P_t(MiddleUp) + \frac{a_t - b_t}{2} + (1 - a_t)P_t(Down) = 1 - b_t. \quad (4)$$

Therefore, $b_t = 1 - W_t(MiddleUp) - W_t(Down)$ measures ambiguity aversion in the sense of Ghirardato and Marinacci (2002) (see Observation 3 in Appendix A).[6] This interpretation of b_t as an index of ambiguity aversion is also consistent with the definition of

[6] This follows because the decision maker is an expected utility maximizer for decision under risk in Ghirardato and Marinacci (2002). If the decision maker deviates from expected utility in decision under risk then b_t captures both ambiguity aversion and the lower weight the decision maker gives to the best outcome in decision under risk.

the ambiguity premium in Schmeidler (1989) and with the definition of source preference in Tversky and Wakker (1995) (see Observation 4 in Appendix A).

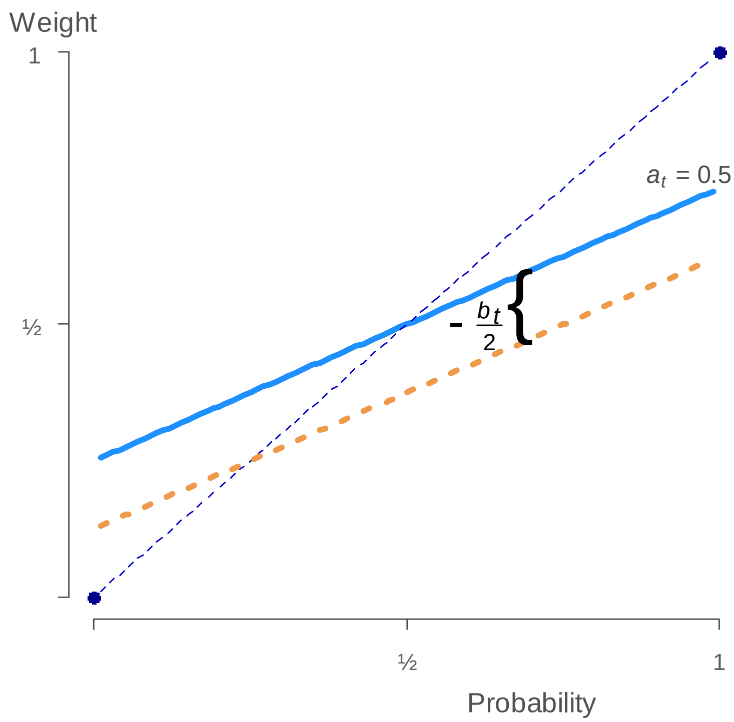


Figure 1. Pessimism and likelihood insensitivity.
The solid line corresponds to $a_t = 0.5$ and $b_t = 0$. The parallel dashed line keeps a_t constant and increases b_t. The figure shows that for constant a_t, increases in b_t shift the neo-additive weighting function downwards leading to an increase in pessimism. The figure also shows that increases in a_t reflect that the decision maker becomes less sensitive to changes in likelihood. The diagonal shows the weighting function when expected utility holds ($a_t = 0$ and $b_t = 0$).

Empirical evidence indicates that pessimism diminishes when the decision maker knows more about a source of uncertainty (Heath and Tversky 1991, Kilka and Weber 2001, Fox and Weber 2002, Di Mauro 2008, and Abdellaoui et al. 2011). This suggests that pessimism will decrease with the size of the decision maker's history set.

As we mentioned above, the range of the parameter b_t depends on a_t. This theoretical dependence, which suggests that the two measures are correlated, has turned out to be of limited empirical relevance. Empirical studies showed that a_t and b_t are independent measures of ambiguity and that they are less correlated than, for instance, the indices obtained using Prelec's (1998) two-parameter weighting function (Abdellaoui et al. 2011, Dimmock et al. 2016a).

Likelihood insensitivity

The parameter a_t in Eq. (3) is an anti-index of the weight that the decision maker gives to expected utility in his evaluation of acts. If a_t is equal to 0 then the decision maker gives maximum weight to expected utility. Larger values of a_t imply that the decision maker gives less weight to expected utility and that he concentrates more on the maximum and minimum utility. In other words, the larger is a_t the more the decision maker ignores the relative likelihoods of x and y. This can also be seen from Eq. (2), where larger values of a_t imply that $P_t(E)$ receives less weight.

Figure 1 also shows the effect of changes in a_t when $b_t = 0$. When $a_t = 0$, the decision maker behaves according to expected utility (dashed line). When a_t increases, the slope of the decision weighting function becomes flatter and the decision maker is less sensitive to intermediate changes in likelihood. As a result, differences between (non-extreme) decision weights are less than the differences between their underlying subjective probabilities. This is called *likelihood insensitivity*. Likelihood insensitivity includes ambiguity perception, but it may be different as the decision maker can also display likelihood insensitivity under risk (as indeed empirical studies show). Likelihood insensitivity is equal to ambiguity perception if the decision maker behaves according to expected utility in decision under risk. We take a_t as an index of likelihood insensitivity with higher values indicating more likelihood insensitivity. Tversky and Wakker (1995) gave behavioral (preference-based) conditions for absolute and relative likelihood insensitivity.

Under neo-additive preferences, we obtain,

$$W_t(Up) + W_t(Middle) - W_t(MiddleUp) = \frac{a_t - b_t}{2}. \qquad (5)$$

Equation (5) shows that the neo-additive model predicts lower subadditivity if $a_t > b_t$. Substituting b_t, which we obtained from Eq.(4), into Eq.(5) gives a_t.

Empirical evidence has shown that likelihood insensitivity is more pronounced for uncertainty than for risk (e.g. Kahneman and Tversky 1979, Kahn and Sarin 1988, Kilka and Weber 2001, Abdellaoui et al. 2005, Wakker 2010, ch. 10) and decreases with the familiarity of the source of uncertainty (Kilka and Weber 2001, Abdellaoui et al. 2011). This suggests that likelihood insensitivity at least partly comprises ambiguity perception and will diminish with the size of the history set (the amount of information).

3.2 The multiple-prior interpretation

The decision weight interpretation is close to Choquet expected utility (Gilboa 1987, Schmeidler 1989) and prospect theory (Tversky and Kahneman 1992) where ambiguity attitudes are modeled through nonadditive decision weights. By contrast, the multiple-prior models capture ambiguity by representing the decision maker's beliefs through a set of priors $\mathcal{C}_t$ rather than a unique probability measure P_t. Chateauneuf et al. (2007) showed that the neo-additive model also has a multiple-prior interpretation and can be written as:

$$x_E y \mapsto alpha_t \, min_{\pi \in \mathcal{C}_t}\big[\pi(E)U(x) + \big(1-\pi(E)\big)U(y)\big] +$$

$$(1-alpha_t)max_{\pi \in \mathcal{C}_t}\big[\pi(E)U(x) + \big(1-\pi(E)\big)U(y)\big]. \quad (6)$$

In Eq. (6), $\mathcal{C}_t = \{\pi: \pi(E) \geq (1-a_t)P_t(E),$ for all events $E\}$. Because this definition holds for both event E and its complement, it follows that $\mathcal{C}_t = \{\pi: (1-a_t)P_t(E) \leq \pi(E) \leq 1 - (1-a_t)(1-P_t(E)),$ for all events $E\}$, which shows that if $a_t = 0$, $\mathcal{C}_t$ consists only of P_t and that the size of the set of priors increases with a_t. The parameter $alpha_t$ in Eq. (6) is defined as $\frac{1}{2} + \frac{b_t}{2a_t}$ when $a_t > 0$ and as $\frac{1}{2}$ when $a_t = 0$.

Under the multiple priors interpretation (Eq. 6) we can also define two indices of ambiguity attitude: one reflecting ambiguity perception, the other (relative) ambiguity aversion.

Ambiguity perception

Ghirardato et al. (2004) argued that the presence of ambiguity in the decision maker's decision problem is revealed by violations of von Neumann and Morgenstern's independence axiom and, consequently, that unambiguous preferences are those for which no violations of independence occur. Under the biseparable invariant model that Ghirardato et al. (2004)

consider, these unambiguous preferences identify a set of priors $\mathcal{C}^*$ that is the smallest among all possible sets of priors. The size of this set of priors then represents the decision maker's ambiguity perception with larger sets revealing more ambiguity (see Proposition 6 in Ghirardato et al. (2004)). Although our $\mathcal{C}_t$ may not coincide with $\mathcal{C}^*$ for the general alpha-maxmin model, within the model of Equation (6) that we consider, our set of priors $\mathcal{C}_t$ is uniquely defined[7], and therefore also the smallest set of priors, given each individual's preferences. Chateauneuf et al. (2007) and Eichberger et al. (2012a) suggest that ambiguity is still reflected by this set of priors, which increases with the value of a_t.

Ambiguity aversion

Ghirardato et al. (2004, Proposition 12) showed that, for a given set of priors, alpha in the general alpha-maxmin is an index of ambiguity aversion which agrees with the relative ambiguity aversion ranking of Ghirardato and Marinacci (2002). It is also the case in Eq.(6); $alpha_t$ is an index of ambiguity aversion (keeping a_t constant; see Observation 3 in Appendix A). However, it is a different measure of ambiguity aversion than b_t. The pessimism index b_t is a global measure of ambiguity aversion and the more ambiguity is perceived the larger the range of ambiguity attitudes that the decision maker can display. On the other hand, $alpha_t$ is a relative measure of ambiguity aversion, which is defined per unit of perceived ambiguity and, therefore, does not depend on the amount of perceived ambiguity. This explains why b_t is bounded by $-a_t$ and a_t, and thus depends to some extent on likelihood insensitivity, and $alpha_t$ by 0 and 1 and does not depend on ambiguity perception.

[7] Eqs. 4 and 5 show that a_t and b_t are uniquely determined. Hence, from Eq. 6, $alpha_t$ and $\mathcal{C}_t$ are also unique.

3.3 Extension

In the decision weight interpretation, the constraints $0 \leq a_t \leq 1$ and $-a_t \leq b_t \leq a_t$ are typically satisfied at the average level, but a sizeable minority of the subjects displays oversensitivity to likelihood information and has a negative a_t (Kilka and Weber 2001, Abdellaoui et al. 2011, Dimmock et al. 2016a, Dimmock et al. 2016b). To accommodate such behavior, we used the generalization of the neo-additive model introduced by Eichberger et al. (2012b). For finite space spaces (as in our experiment), the *generalized neo-additive model* is defined by Equations 1 and 2 for any $a_t \leq 1$, b_t, and P_t such that $\min_{E\in 2^{S}-\{\emptyset,S\}} \frac{a_t-b_t}{2} + (1-a_t)P_t(E) \geq 0$ and $\max_{E\in 2^{S}-\{\emptyset,S\}} \frac{a_t-b_t}{2} + (1-a_t)P_t(E) \leq 1$. These new constraints permit likelihood oversensitivity as long as it does not lead to violations of monotonicity. The interpretations of a_t and b_t as indices of likelihood insensitivity and pessimism remain valid under this generalization (Observations 3 and 4 in the Appendix are still true).

The multiple priors interpretation requires that a_t is positive. As several of our subjects had negative a_t, we could only use the multiple priors interpretation in the aggregate analyses and did not use it in the individual analyses.

4. Experiment

Subjects

Subjects were 64 students from Erasmus University Rotterdam (22 female, average age 24 years, age range 21-33). Subjects were either third-year undergraduate students majoring in finance or graduate finance students. We used finance students because they might better understand and be more motivated to price options (the experimental tasks), which, we hoped, would improve the quality of the data. Each subject received a €5 show-up fee and, in addition, played out one of his choices for real using a procedure described below.

Method

The experiment was computer-run in small group sessions involving 3 subjects at most. We used small groups to further improve the quality of the data as it is easier to monitor for confusion and misunderstanding in small groups. Subjects first received instructions and then answered several questions to check their understanding of the experimental tasks. They could only proceed to the actual experiment after they had correctly answered all these comprehension questions. The experimental instructions and the comprehension questions are in an Online Appendix.

As the source of uncertainty we used the variation in the stock returns of IPOs (Initial Public Offerings) traded at the New York Stock Exchange (NYSE). IPOs are stocks that have just entered the market. We used IPOs for two reasons. First, stock returns are a natural source of uncertainty. Second, because IPOs are new on the market, they have no price history and learning occurs naturally.

We used data on 328 IPOs listed on the NYSE between 1 September 2009 and 25 February 2011. At the start of the experiment, each subject drew four numbers, which determined the underlying stocks. Subjects traded in options which payoffs depended on the four underlying stocks. The identity of these stocks was revealed only after subjects had completed the experiment. Then we also explained where the subjects could verify the stock data that we had presented.

Payoffs were determined by the daily returns of the stocks on the 21st trading day after their introduction. We defined four events: Up = $(0.5, \infty)$, i.e. the stock goes up by more than 0.5% on the 21st trading day, $Middle$ = $[-0.5, 0.5]$, the stock varies by at most 0.5% on the 21st trading day, $Down$ = $(-\infty, -0.5)$, the stock goes down by more than 0.5% on the 21st trading day, and $MiddleUp$ = $[-0.5, \infty)$, the stock returns at least –0.5% on the 21st trading day. We will refer to an option that pays x if event Up obtains as an Up option. $Middle,$

$Down$, and $MiddleUp$ options were defined similarly. We used the variations in the stock prices rather than the absolute prices to ensure that subjects had no information about the stocks and to avoid biases. Stocks with higher prices might, for example, attract more attention leading to biases in the elicitations.

Table 1: The 20 choice questions

Stock	***Condition***	***y***	***x***	***Option type***		***Stock***	***Condition***	***y***	***x***	***Option type***
1	No info	0	10	Up		3	1 week	0	20	Up
1	No info	10	20	Up		3	1 week	0	20	$Middle$
1	No info	5	20	Up		3	1 week	0	20	$Down$
1	No info	10	15	Up		3	1 week	0	20	$MiddleUp$
1	No info	0	5	Up		3	1 week	0	20	$Middle$
1	No info	0	20	Up		4	1 month	0	20	Up
2	No info	0	20	Up		4	1 month	0	20	$Middle$
2	No info	0	20	$Middle$		4	1 month	0	20	$Down$
2	No info	0	20	$Down$		4	1 month	0	20	$MiddleUp$
2	No info	0	20	$MiddleUp$		4	1 month	0	20	$Down$

The columns labeled "Stock" refer to the four different stocks subjects faced. The questions for stock 1 were used to measure utility. The columns labeled "Condition" refer to the amount of information subjects received about the historical performance of the stock. Options were of the type $x_E y$ where the subject received €x if event E occurred and €y otherwise. The columns "Option types" indicate event E.

There were three *information conditions*, each involving a different history set. In the *no information* condition (history set h_0), subjects had no information about the stock returns. In the *one week* condition (history set h_5), subjects knew the daily returns of the stock during the first 5 trading days after its introduction. Finally, in the *one month* condition (history set h_{20}), subjects knew the stock returns during the first 20 trading days after its introduction.

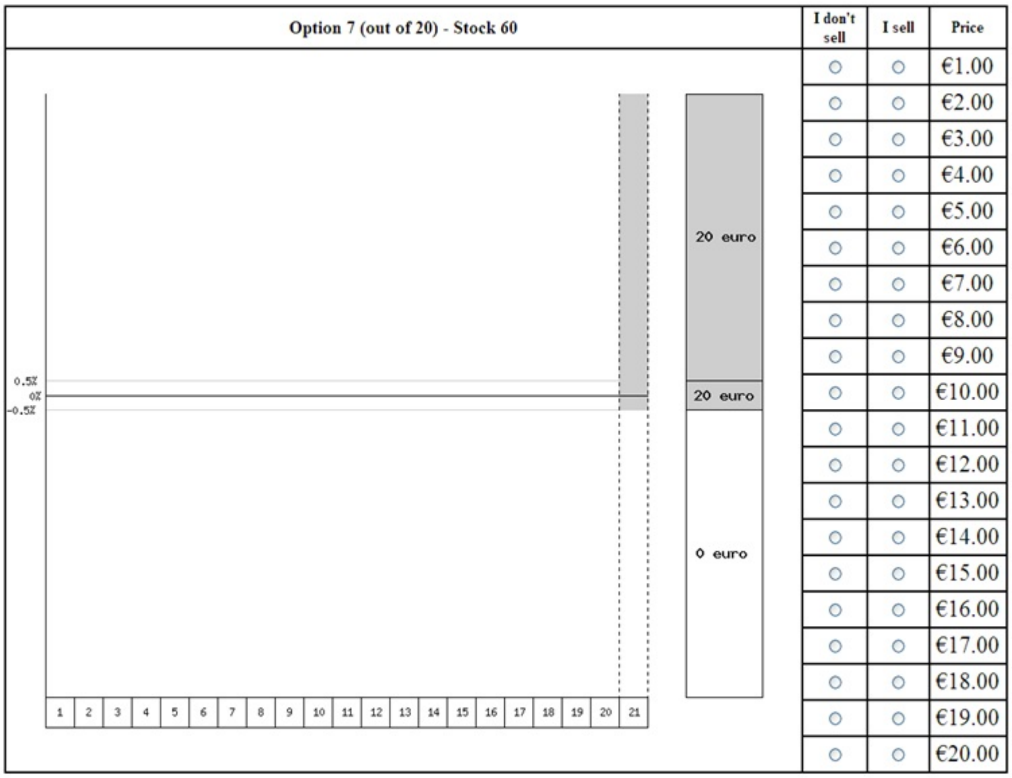


Figure 2. The choice lists used in the experiment. In this example the option pays €20 if event $MiddleUp$ occurs on the 21st trading day after the introduction of the stock and €0 otherwise.

We used choice lists to elicit the ask prices of 20 options, summarized in Table 1. Figure 2 gives an example of a choice list for a $MiddleUp$ option. Subjects were told that they owned the option $x_E y$ and they were asked for each price on the choice list whether or not they wanted to sell the option. The choice lists consisted of 20 prices ranging from € (y + z) to €x in increments of z = €$\frac{x-y}{20}$. The computer program enforced monotonicity. If a subject indicated for some price that he did not want to sell the option then the computer automatically selected "I don't sell" for all prices that were lower. Similarly, if a subject indicated for some price that he wanted to sell the option then the computer automatically selected "I sell" for all prices that were higher. Andersen et al. (2006) found that enforcing monotonicity did not change response patterns, but reduced noise in the data. Two test questions confirmed that subjects understood the principle of monotonicity and agreed with it.

The 20 choices were divided into four groups (see Table 1). Group 1 consisted of six choices to measure utility. The questions in groups 2, 3, and 4 measured the effect of information on ambiguity attitudes. For groups 3 and 4, we repeated one measurement to test the reliability of our measurements. We randomized the order of the options within each group.

The utility questions (group 1) always came first. We counterbalanced the order in which the three information conditions appeared to avoid that a better understanding of the task would confound the effect of giving more information. We had to use different stocks in each group. The main drawback is that the comparison between conditions (especially the one week and the one month conditions) does not show the effect of learning more about the same stock but the effect of learning more about one stock than about another. Yet, if we had used options on the same underlying stock then subjects who had, for instance, received the one month condition first would have used this information in the no information and in the one week conditions.

Incentives

We used a random incentive system. At the end of the experiment, subjects threw a twenty-sided die twice. The first throw selected the option and the corresponding choice list and the second throw selected the line of that choice list that was played out for real. In the selected line, we implemented the choice that the subject had made during the experiment. If the subject had chosen to sell, we paid him the price. If he had chosen not to sell, we played out the option $x_E y$ and he received €x if event E had occurred on the 21st trading day and €y otherwise.

Analysis

To measure utility in the decision weight interpretation, we selected history h_0 and elicited the certainty equivalents CE_k of the binary acts $x_{k_{Up}}y_k, k = 1, \dots, 6$, the first six entries of Table 1. Under biseparable preferences:

$$U(CE_k) = W_0(Up)U(x_k) + \left(1 - W_0(Up)\right)U(y_k). \quad (7)$$

We assumed that utility belonged to the power family: $U(x) = x^{\beta}$ if $\beta > 0$, $U(x) = \ln(x)$ if $\beta = 0$, and $U(x) = -x^{\beta}$ if $\beta < 0$. The power family is widely used in decision theory and generally fits the data well (Stott 2006). To test for robustness, we also analyzed our data under the assumption that utility was exponential and expo-power. This led to the same conclusions except where noted (see also Table B.1 in Appendix B). Moreover, we observed no difference in goodness of fit for power, expo-power, and exponential utility.[8] Dividing all money amounts by the maximum payoff €20 scales the power utility function such that $U(20) = 1$ and $U(0) = 0$ and ensures that higher values of β indicate more convexity of utility. For each subject, we estimated $W_0(Up)$ and β in Eq. (7) by nonlinear least squares. We then substituted β in Eqs. (4) and (5) to derive a_t and b_t and the subjective probabilities.

As an additional robustness check we also estimated the parameters β, a_t, b_t and P_t using a non-linear random coefficient model with individual Fechner errors. Rather than estimating parameters for each individual, the random coefficient model estimates the means and standard deviations of the distributions of individual parameters in the population. This leads to more precise estimates (as more data are used) and limits the effect of outliers. As the mean of a_t was positive, the random coefficient model could also estimate the $alpha$-maxmin model Eq. (6). The parameters a_t, b_t and P_t in the one week and one month

[8] At the aggregate level, we tested this by the Vuong test, at the individual level we compared the sum of the squared residuals.

conditions were defined as the sum of the random coefficients at history h_0 and a history-specific fixed effect. Details of the estimation procedure are in Appendix B.

5. Results

5.1. Consistency

In both questions that we repeated, we observed no significant differences between the original and the repeated ask prices and their correlations were almost perfect (Spearman correlation 0.86 and 0.81, both $p < 0.01$).[9] The mean absolute differences between the ask prices were 1.09 and 1.00 in the two questions, which gives an indication of the average error subjects made.

5.2. Biseparable Preferences

Overall, utility was close to linear both at the aggregate and at the individual level. The median power coefficient β was equal to 1 (interquartile range = $[0.83, 1.28]$) and the number of subjects with concave utility (33) did not differ from the number of subjects with convex utility (31) (Binomial test, $p = 0.90$).[10]

Panel A of Figure 3 shows the kernel density estimates of the distribution of the index $1 - W_t(MiddleUp) - W_t(Down)$ for each information level. As we explained in Section 2, this index reflects ambiguity aversion in the sense of Ghirardato and Marinacci (2002). At the aggregate level there was little ambiguity aversion and we could not reject the null of ambiguity neutrality in each of the three information conditions (Wilcoxon tests, all $p >$

[9] We use the classification scheme of Landis and Koch (1977) to describe the correlations.

[10] With exponential ($p = 0.05$) and expo-power ($p < 0.01$) utility there were more concave than convex subjects.

0.46).[11][12] Moreover, we could not reject the hypothesis that the proportion of subjects for whom the sum of $P(MiddleUp)$ and $P(Down)$ exceeded 1 and the proportion for whom this sum was less than 1 were the same in all three information conditions (Binomial tests, all $p > 0.17$). However, the figure also shows considerable variation at the individual level. In particular, there is a bump in the Figure around the value 1, which indicates that some subjects were extremely ambiguity averse. We did not observe comparable extreme ambiguity seeking.

Panel B of Figure 3 plots the kernel density estimates of the distribution of the lower subadditivity index $W_t(Up) + W_t(Middle) - W_t(MiddleUp)$ for each information condition. For each of the three information levels, we observed significant lower subadditivity (Wilcoxon tests, all $p < 0.01$).[13] The individual analyses confirmed this: the proportion of subjects displaying lower subadditivity was significantly larger than the proportion of subjects displaying lower superadditivity in all three conditions (Binomial tests, all $p < 0.01$). However, the figure also shows a lot of individual heterogeneity. Even though lower subadditivity was clearly the most common pattern, there was a sizeable minority of subjects who behaved according to superadditivity.

[11] The medians of the ambiguity indices were $0.00, -0.05$, and 0.03 for the no information, the one week, and the one month condition, respectively.

[12] With exponential and expo-power utility we found significant ambiguity seeking in the sense of Ghirardato and Marinacci (2002) in all information conditions.

[13] The medians of the lower subadditivity index were $0.15, 0.18$, and 0.11 for the no information, one week, and one month condition, respectively.

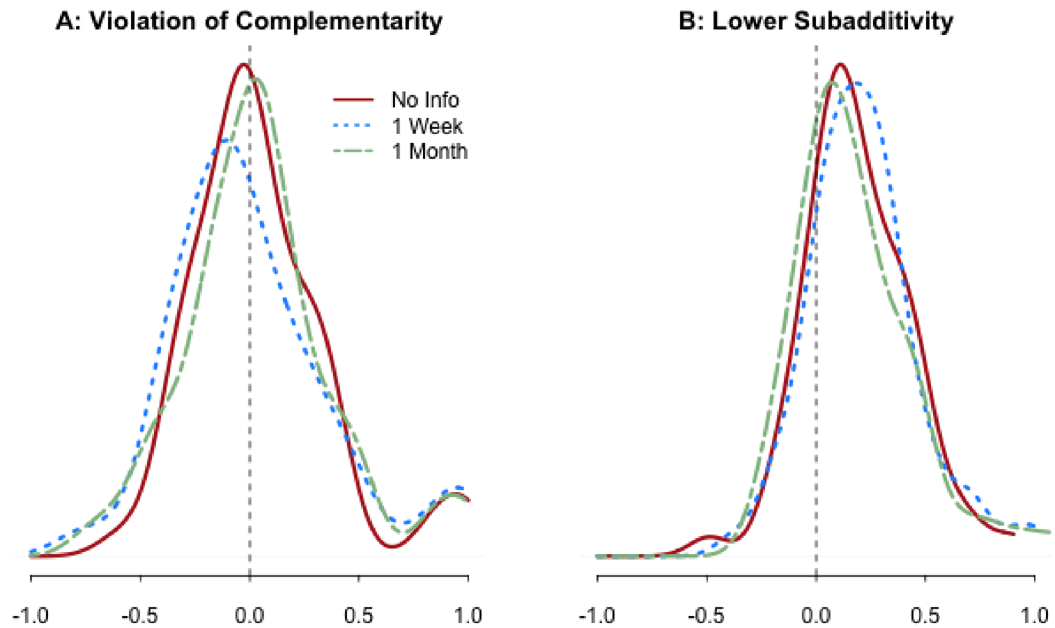


Figure 3. Tests of complementarity and lower subadditivity. The figure shows the kernel density estimates of the ambiguity aversion index in the sense of Ghirardato and Marinacci (2002) (Panel A) and the index of lower subadditivity (Panel B) for the three information conditions. Panel A shows that the average subject was approximately ambiguity neutral, panel B that he displayed lower subadditivity. The figures show substantial individual heterogeneity.

The joint findings of complementarity and lower subadditivity are in line with previous evidence (Tversky and Koehler 1994, Fox and Tversky 1998, Kilka and Weber 2001, Baillon and Bleichrodt 2015). They are consistent with support theory, a psychological theory of belief formation (Tversky and Koehler 1994).

5.3. Individual ambiguity attitudes

Subjects whose (for ambiguity attitudes corrected) subjective probabilities were outside the unit interval deviated from the generalized neo-additive model and had to be excluded from the individual analyses. Because these deviations might just reflect error, we included these subjects in the robustness analysis reported in Section 5.4.

We only excluded subjects for the information condition for which they violated the neo-additive model. This left 56 subjects in the no information condition, 55 subjects in the one week condition, and 52 subjects in the one month condition. To test for robustness and possible selection bias, we also analyzed the data excluding all subjects who violated the generalized neo-additive model at least once. The results were similar and are not reported separately.

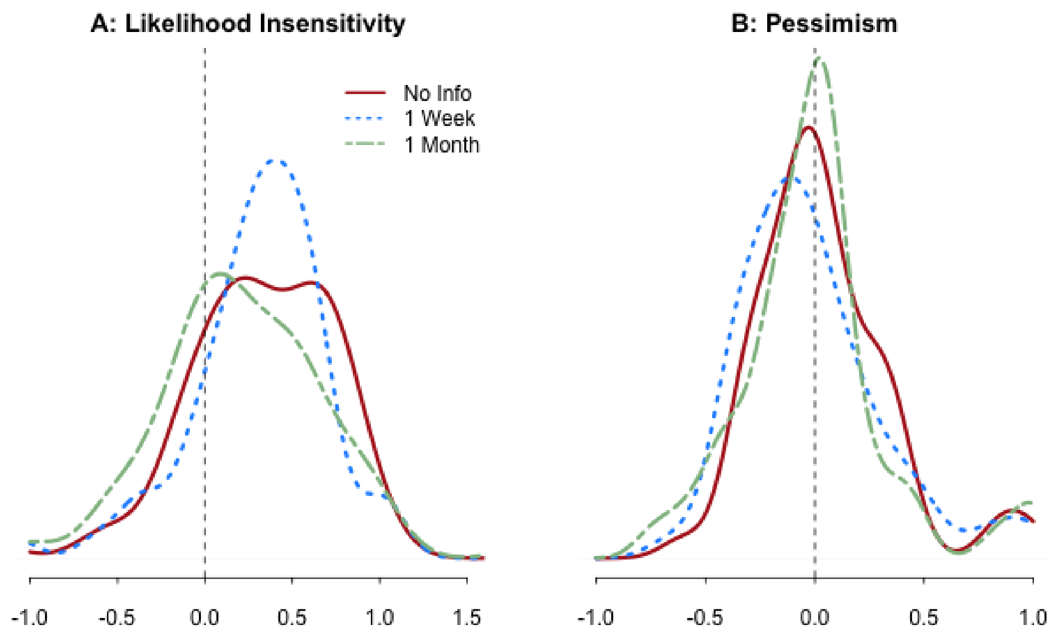


Figure 4. The likelihood insensitivity and pessimism indices. Panels A and B show the kernel density estimates of the distributions of the likelihood insensitivity and pessimism indices for the three information conditions. Likelihood insensitivity falls with more information, but information has no effect on the pessimism indices.

Likelihood insensitivity

Figure 4A shows the kernel density estimates of the distribution of the likelihood insensitivity index (a_t) for the three information conditions. Three things are noteworthy. First, the distributions mainly lie in the positive part suggesting for all three conditions substantial likelihood insensitivity (Wilcoxon tests, all $p < 0.01$).[14] Second, likelihood insensitivity seemed to decrease with more information as the distribution for the one month condition lies to the left of the other two distributions. However, statistical testing gave mixed results. Likelihood insensitivity for one month was smaller than for one week (Wilcoxon test, $p = 0.03$), but the other indices did not differ. Finally, the figure shows that while most subjects were insensitive to likelihood information, some subjects were oversensitive to likelihood information and had negative values of a_t. This illustrates the usefulness of Eichberger et al.'s (2012) generalized neo-additive preferences.

Figure 5 presents pairwise comparisons between the individual values of a_t and gives insight in the effect of information on individual likelihood insensitivity. In each panel, the horizontal axis shows the condition in which less information was available. Points on the diagonal represent subjects with the same likelihood insensitivity for the information conditions depicted. If likelihood insensitivity diminished with the amount of information, then the data points should be located below the diagonal.

The shaded areas of Figure 5 show the subjects who moved in the direction of "correct" sensitivity to likelihood, i.e. to expected utility. The likelihood insensitivity or oversensitivity of these subjects decreased but they did not overshoot and went from insensitivity to even larger oversensitivity or from oversensitivity to even larger insensitivity. In all panels, a majority of points (Binomial tests, $p \leq 0.01$ in all cases) is in the shaded area, which is consistent with convergence towards expected utility with more information.

[14] The medians of a_t were 0.34, 0.36, and 0.22 in the no information condition, the one week condition, and the one month condition, respectively.

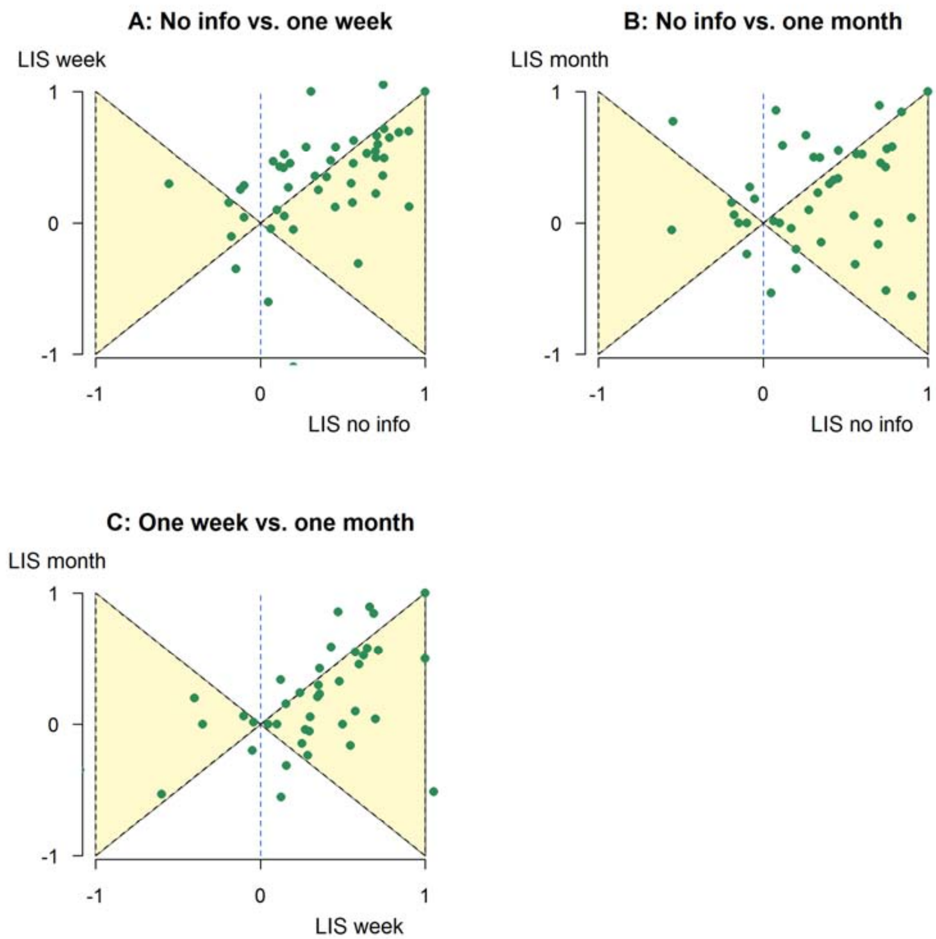


Figure 5. The relations between the individual likelihood insensitivity indices (a_t). If subjects converge to expected utility then the points should lie in the shaded areas. This happens in all panels.

Pessimism

Figure 4B shows that the kernel density estimates of the distributions of the pessimism indices (b_t) were centered around zero for all three information conditions. We could not reject the null of no pessimism in any of the information conditions.[15] There was

[15] For exponential ($p < 0.01$) and expo-power ($p \leq 0.01$) utility we found significant optimism in all information conditions.

more optimism in the one week than in the no information condition (Wilcoxon test, $p = 0.01$), the other differences were insignificant.

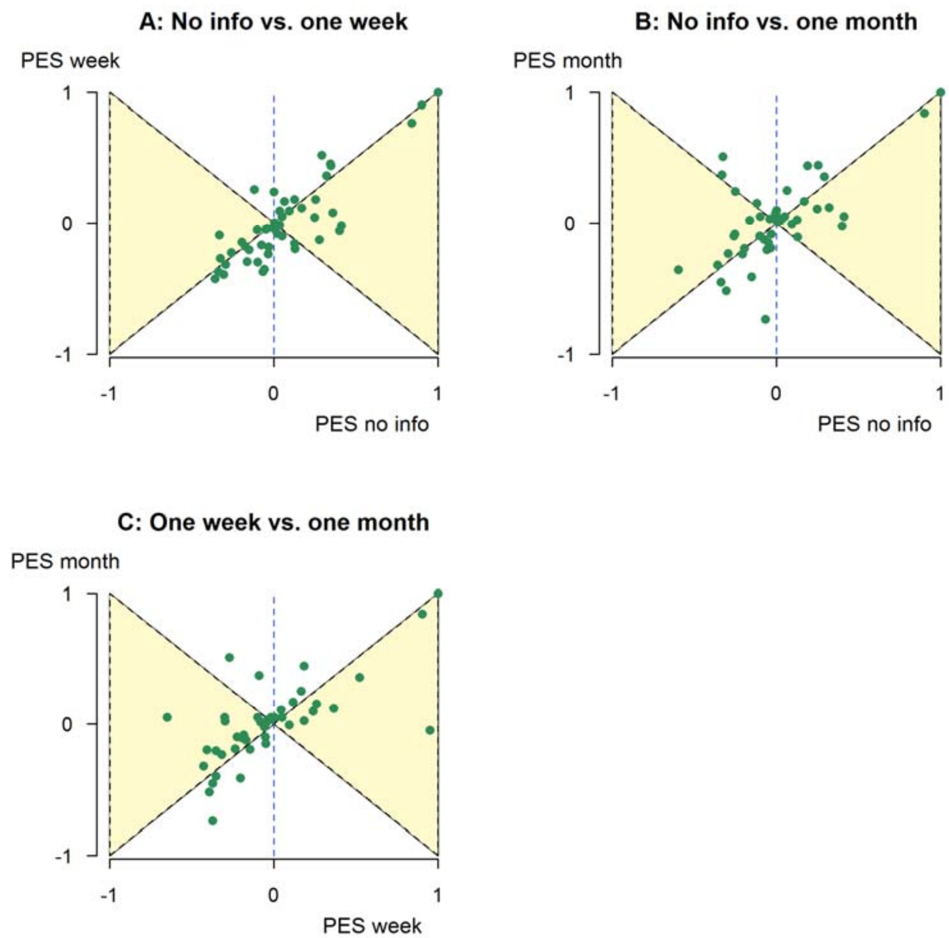


Figure 6. The relations between the individual pessimism indices (b_t). If subjects converge to expected utility then the points should lie in the shaded areas. This happens in Panel C.

Figure 6 plots the individual pessimism indices for the three information conditions with the condition with less information on the horizontal axis. Points on the diagonal represent individuals with the same pessimism in two information conditions. If pessimism decreased with information then individual points should be in the lower halves of the figures. This was the case for a majority of subjects in Panel A (Binomial test, $p = 0.02$), but not in the other two panels. The shaded areas show the subjects who moved in the direction of expected utility as more information became available. We only observed a significant move to expected utility in the comparison between the one month and the one week conditions (Panel C, Binomial test, $p = 0.01$)[16], but it should be kept in mind that most subjects displayed little pessimism anyhow.

5.4. Aggregate ambiguity attitudes

Table 2 summarizes the estimation results of the non-linear random coefficient model. For each parameter, we report the estimate of the mean and of the standard deviation. In this estimation, we could include all elicited certainty equivalents, including repeated measurements and the responses of subjects who violated the generalized neo-additive model. Because at the aggregate level a_t is non-negative, we could estimate the decision weight version (Eq. 3) and infer from it the multiple priors version (Eq. 6) of the neo-additive model.

The decision weight interpretation

With more structural assumptions the random coefficient model led to more significant results.. However, our main conclusions of Section 5.3 remained true. There was significant likelihood insensitivity in all information conditions, but no pessimism (or optimism). Likelihood insensitivity diminished as more information became available, suggesting that subjects moved in the direction of expected utility. Subjects were significantly more optimistic in the one week and one month conditions than in the no information condition. This is consistent with the literature on source-dependent ambiguity attitudes, which indicates that more knowledge/competence reduces ambiguity aversion.

[16] For exponential utility this difference was not significant.

Table 2: Random coefficient models

		Model 1 Decision weight	Model 2 Multiple Priors
Likelihood insensititivity/ Ambiguity perception	a_0	0.52*** [0.04]	
	(one week fixed effect)	−0.16** [0.05]	
	(one month fixed effect)	−0.30*** [0.05]	
	std of random effects	0.16*** [0.02]	
Pessimism	b_0	0.04 [0.04]	
	(one week fixed effect)	−0.07*** [0.02]	
	(one month fixed effect)	−0.05** [0.02]	
	std of random effects	0.07*** [0.01]	
Ambiguity aversion	$alpha_0$		0.55*** [0.05]
	(one week fixed effect)		−0.09** [0.03]
	(one month fixed effect)		0.06 [0.06]
	std of random effects		0.07 [0.00]
$P(Up)$	pup_0	0.41*** [0.02]	
	(one week fixed effect)	−0.08*** [0.02]	
	(one month fixed effect)	−0.07*** [0.02]	
	std of random effects	0.07*** [0.01]	
$P(Middle)$	$pmid_0$	0.30*** [0.02]	
	(one week fixed effect)	0.004 [0.02]	
	(one month fixed effect)	0.05*** [0.02]	
	std of random effects	0.13*** [0.01]	
Utility	β	1.07 [0.06]	
	std of random effects	0.20*** [0.02]	
Noise (Fechner error)	μ_ϵ	−2.33*** [0.05]	
	std of random effects	0.55*** [0.04]	
Log-likelihood		844.32	
N		1280	

Standard errors in square parentheses. ***: significant at 1%, **, significant at 5%, * significant at 10%.

The multiple-priors interpretation

Table 2 also shows the estimates under the multiple priors interpretations. An often-raised objection against the multiple priors models is that they involve a concept (the set of priors) that is unobservable. Figure 7 shows that our method and our assumption of a family of priors characterized by one parameter, allows us to estimate the set of priors. The black dot shows the estimated subjective probabilities P_t. Together with a_t these determine the set of priors (the light grey area). As the Figure shows, the set of priors decreases with more information and is smallest in the one month condition indicating that more knowledge reduced the perception of ambiguity. The estimate of $alpha_t$ was not significantly different from 0.50 in any of the information conditions, which is consistent with $b_t = 0$ and ambiguity neutrality.

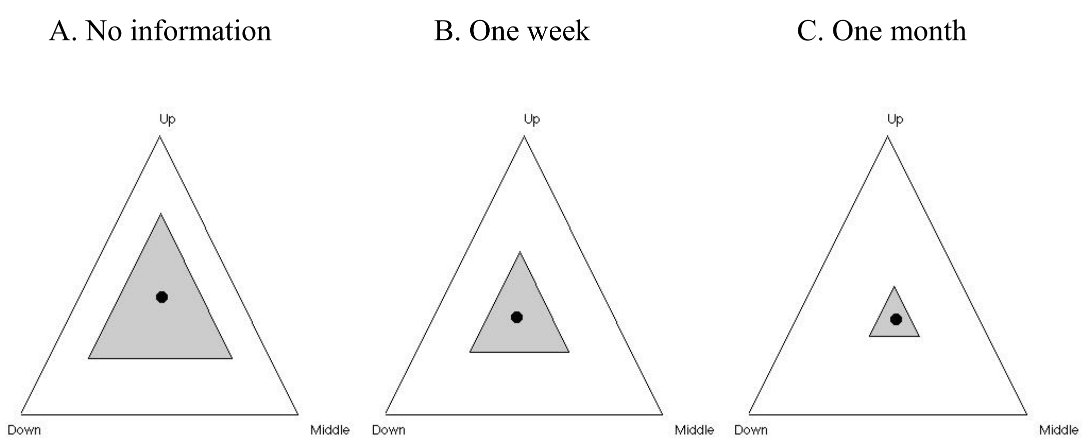


Figure 7: Sets of priors for the three information conditions based on the estimates of Model 2. In each panel, the large triangle is the simplex representing all possible probability measures over the 3 events $Up, Down$, and $Middle$. Each vertex of the simplex denotes an event and corresponds to the measure in which this event is certain. Each opposite side of a vertex represents the probability measures assigning zero probability to the vertex event. The grey triangle is the set of priors and the black dot represents P_t.

Subjective probabilities

Table 2 also shows the estimated subjective probabilities. $P(Up)$, the subjective probabilities that the event Up would occur, tended to decrease as more information became available, whereas $P(Middle)$ increased.[17] $P(Down)$ increased in the one week condition but then decreased in the one month condition to approximately the value in the no information condition.

The elicited probabilities were well-calibrated and close to the true frequencies. For each day from their introduction to the 21st trading day we computed the proportions of the 328 IPOs that went up by more than 0.5% (corresponding with the event Up), the proportions that varied by at most 0.5% (corresponding with the event $Middle$), and those that went down by more than 0.5% (corresponding with the event $Down$). A frequentist may interpret these proportions as the actual probabilities of the events $Up, Middle,$ and $Down$ at each date t in the history.

Figure 8 shows the results of this analysis. Panel A shows the proportions for the event Up, Panel B for the event $Middle$, and Panel C for the event $Down$. The figure also shows the estimated probabilities of $P(Up), P(Middle),$ and $P(Down)$ for the three information conditions (the dots at the end of the line).

All subjective probabilities converged to the actual frequencies in the market. Subjects initially overestimated the probability of the event Up. As more information became available, they adjusted their estimate downwards. On the other hand, subjects underestimated the probability of the event $Middle$. This underestimation decreased with

[17] Most differences are significant ($p < 0.01$) except for the differences between $P(Up)$ in the one week and the one month condition, between $P(Middle)$ in the one week and the no information condition, and between $P(Down)$ in the no information and the one month condition.

information, particularly in the one month condition. Subjects were close to the true frequency of the event $Down$ in the no information condition, but then adjusted their estimate upwards in the one week condition, probably because most stocks did not do well in their first five trading days and their returns were highly volatile. In the one month condition, subjects were, again, close to the true frequency.

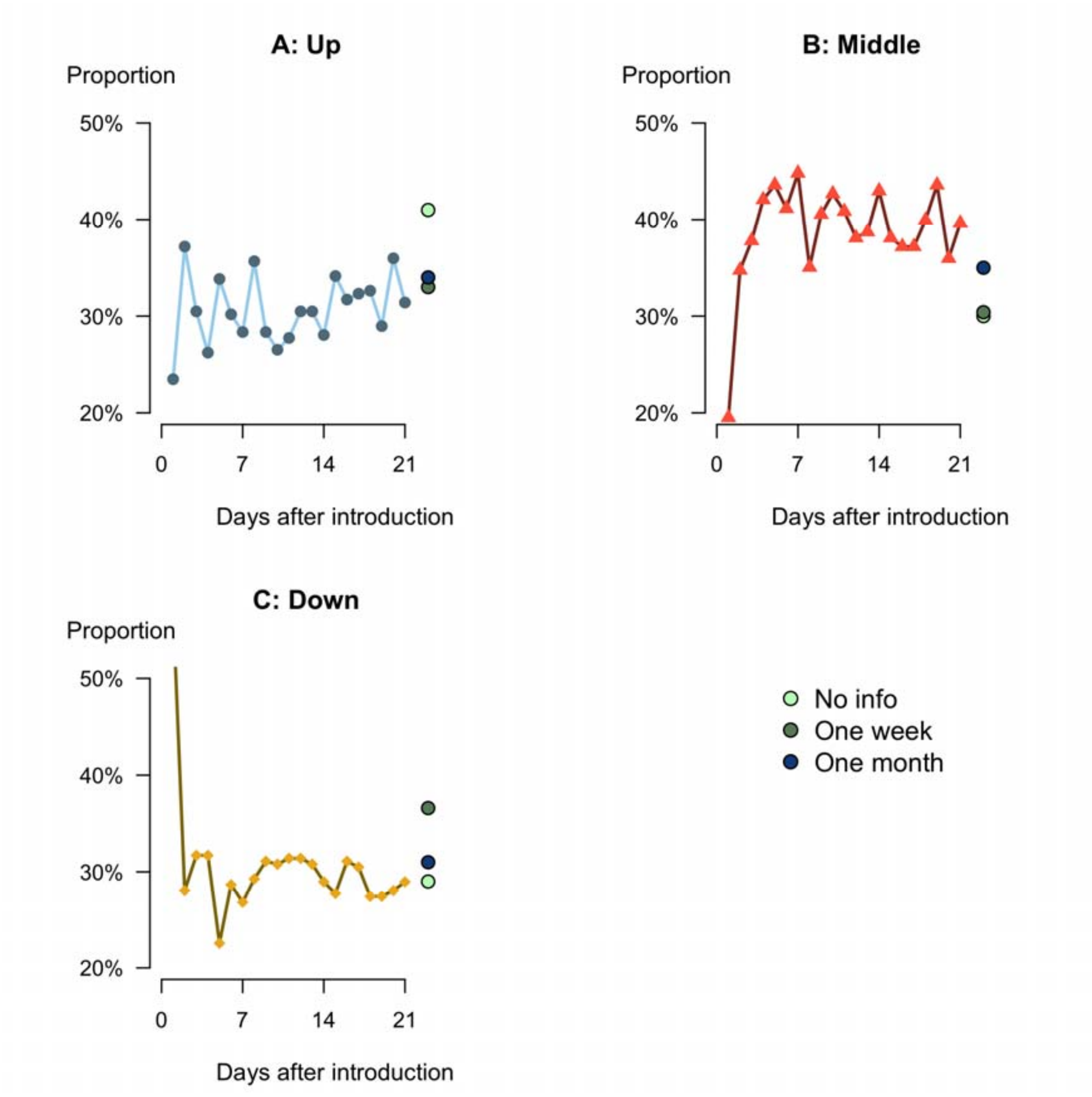


Figure 8. Stock history and subjective probabilities. Panel A shows the proportion of the 328 IPOs that went up by more than 0.5% on each trading day from their introduction to the 21st trading day. Panels B and C show the proportions that varied by at most 0.5% and went down by more than 0.5%, respectively. The dots at the end show the estimated probabilities of $P(Up)$ (Panel A), $P(Middle)$ (Panel B),

and $P(Down)$ (Panel C) under the three information conditions (in Panel A the points for one week and one month overlap).

6. Discussion

This paper has studied the effect of learning new information on ambiguity attitudes. In the decision weight interpretation our method decomposes ambiguity attitudes into likelihood insensitivity and pessimism. At the aggregate level, we found substantial likelihood insensitivity that decreased as more information became available. Subjects moved in the direction of expected utility with more information. We found little evidence of pessimism and information had only a minor effect on it. Likelihood insensitivity is often seen as a cognitive bias, whereas pessimism reflects the motivational component of ambiguity attitudes (Wakker 2010, ch. 7). Our data are consistent with this point of view as one expects that additional information would reduce biases but have little effect on motivation.

In the multiple priors interpretation our method decomposes ambiguity attitudes into ambiguity perception and ambiguity aversion. We found considerable ambiguity perception, which decreased with more information. As in the decision weight interpretation, we could not reject the null of ambiguity neutrality even though the measure of ambiguity aversion in the decision weight interpretation differs from the measure in the multiple priors interpretation. We found no clear relation between ambiguity aversion and more information.

The finding of little pessimism may be surprising given that most empirical studies have found more pessimism than we did (Trautmann and Van de Kuilen 2016). It should be kept in mind though that our subjects were finance students who knew about stocks and options. Empirical evidence suggests that pessimism decreases when subjects feel competent about the source of uncertainty (Heath and Tversky 1991) and this may have explained our findings.

In the decision weight interpretation, subjects tended to move in the direction of expected utility with more information. This agrees with previous findings that learning reduce biases (Myagkov and Plott 1997, List 2004, van de Kuilen and Wakker 2006, Ert and Trautmann 2014). On the other hand, the likelihood insensitivity index differed significantly from zero even in the one month condition. Finally, we found strong evidence of lower subadditivity, which violates expected utility, in all information conditions. We conclude that even though more information may have led to behavior that was more consistent with expected utility, substantial deviations remained.

Our analysis made several assumptions. Our results on ambiguity aversion and lower subadditivity in Section 5.2. only assumed biseparable preferences and that more information about the states of nature did not affect utility. Biseparable preferences are general and include many of the ambiguity models that have been proposed in the literature. The assumption that utility does not change with information about the state space is intuitively plausible as utility reflects preferences over outcomes and information about the state space has no relevance for these. Moreover, the assumption is common in the literature and has, to the best of our knowledge, been assumed by all theories of updating under non-expected utility. Empirical support for the assumption comes from Abdellaoui et al. (2011) and Abdellaoui et al. (2016) who measured utility for different sources of uncertainty and could not reject the null hypothesis that utility was the same across sources.

Our separation of ambiguity attitudes into pessimism [ambiguity aversion] and likelihood insensitivity [ambiguity perception] depended on the assumption that preferences are consistent with the neo-additive model of Chateauneuf et al. (2007). Under the decision weight interpretation, this assumption has two implications, first that probabilistic sophistication holds within histories and, second that the weighting function is neo-additive (Eq. 2). Regarding probabilistic sophistication within histories, we note that different histories can be interpreted as different sources of uncertainty. The notion of sources of uncertainty was first proposed by Amos Tversky in the 1990s (Tversky and Kahneman 1992, Tversky and Fox 1995, Tversky and Wakker 1995). Chew and Sagi (2006, 2008) showed that, if an exchangeability condition holds, subjective probabilities can be defined within sources even when probabilistic sophistication does not hold between sources. Our analysis implicitly assumed this exchangeability condition. Abdellaoui et al. (2011) obtained support for it in all but one of their tests. The only exception was a test involving an unfamiliar source and hypothetical choice. For real incentives, exchangeability always held. Their real incentive system was similar to the one we used. Moreover, because our subjects were finance students, sources were familiar. Abdellaoui et al. (2016) also found support for probabilistic sophistication within sources. Finally, the estimated subjective probabilities were well-calibrated: they were sensitive to more information and they were close to the true frequencies observed in the market. Hence, we are inclined to believe that probabilistic sophistication within histories fitted the preferences of most of our subjects rather well. On the other hand, for a minority of our subjects the estimates did not converge and we found subjective probabilities outside the unit interval, which indicates a poor fit for these subjects.

The assumption of a neo-additive weighting function is not very restrictive as it provides a good approximation to more general weighting functions (Diecidue et al. 2009, Abdellaoui et al. 2010). For most subjects the estimated neo-additive parameters were plausible and within the range allowed by the model.

From a normative perspective, one may wonder whether subjects could really learn anything, beyond perhaps a sense of the volatility in IPO prices, from the information that we gave them as stock prices are largely unpredictable. However, the purpose of our study was descriptive and our results show that the information that we provided led to a decrease in subjects' likelihood insensitivity and ambiguity perception and to an improvement in the calibration of their subjective probabilities. Moreover, it made them behave more in line with subjective expected utility. Apparently, providing more information, even if it only has limited value, moves agents in the direction of more rational behavior. Of course, it would be interesting to repeat our experiment in a decision context where it is more obvious that the information that is provided can help to inform decisions.

In the graphical interface, the area of the event $Middle$ was smaller than that of the other two events. We found no indication that this affected the results in the sense that subjects thought that an event with less space devoted to it would be less common. If his would be the case then the effect would be strongest for the subjects who received the no information condition first. However, we observed no effect of the order in which subjects received the various tasks. We could not reject the null hypotheses that the certainty equivalents, the decision weights, the pessimism indices a_t, and the likelihood insensitivity indices b_t were the same for the one third of subjects who received the no information condition first, the one third who received the one week condition first, and the one third who received the one month condition first.

The finding that we observed no order effects has another implication. Because we wanted to control for order effects in the experiment we had to use different stocks in the three information conditions. Because we found no evidence of order effects, future studies may fix the order and use one single stock for which the amount of information is increased across choices.

In our experiment, we used binary acts for two reasons. First, they increase generalizability by not committing to a particular (class of) model. We can interpret our results both under models that model ambiguity through nonadditive decision weights and under the multiple priors models. Second, binary acts are relatively simple and presenting subjects with relatively simple choices tends to reduce the amount of noise in the data.

Empirical evidence indicates that subjects make more errors in more complex choices and often resort to simple rules of thumb. If the neo-additive model holds, we should be able to predict more complex choices from the parameter estimates that we obtained from the preferences over binary acts. To test whether this indeed happens and whether our findings can be generalized to more complex choices is another topic for future research.

7. Conclusion

New information affects both beliefs and ambiguity attitudes. We have presented a method to measure ambiguity attitude while controlling for beliefs. Our method can be applied both to ambiguity models that capture ambiguity attitudes through nonadditive decision weights, where it decomposes ambiguity attitudes into pessimism and likelihood insensitivity, and to the multiple priors models where our method decomposes ambiguity attitudes into ambiguity aversion and ambiguity perception.

We applied our method in an experiment in which we measured the ask prices of options with payoffs depending on the performance of IPOs. The experiment involved three information conditions about historical performance data.

In terms of decision weights the results indicated that there was significant likelihood insensitivity, but little pessimism. Subjects moved in the direction of expected utility as more information about the historical performance of the stocks became available. In terms of multiple priors, we found evidence for ambiguity perception but not for ambiguity aversion. Ambiguity perception decreased with more information. The estimated subjective probabilities, when corrected for ambiguity attitudes, converged to true frequencies. Subjects moved in the direction of subjective expected utility as more information was provided, but substantial deviations remained even in the maximum information condition.

The effect of more information on subjective probabilities on the one hand, and on likelihood sensitivity and ambiguity perception on the other hand, illustrates that modern ambiguity theories capture Keynes' (1921) intuition about the weight and the balance of evidence. Information changed not only the balance of evidence in favor of an event (e.g. information led to an increase in the probability of the event $Middle$), but it also changed the balance between the "absolute amounts of relevant evidence and relevant ignorance." The information made the subjects rely more on their beliefs, take better account of likelihood information, and perceive less ambiguity.

Appendix A: Proofs

We show that our measure of pessimism b_t is consistent with the definition of (comparative and absolute) ambiguity aversion given by Ghirardato and Marinacci (2002). Consider two decision makers A and B whose utility functions are an affine transformation of each other (i.e., they have the same risk attitude under expected utility). According to Ghirardato and Marinacci (2002), decision maker A is *more ambiguity averse than* decision maker B if A (strictly) rejects all binary acts that B (strictly) rejects. In what follows, we use upper indices to indicate the decision makers.

OBSERVATION 1: If decision maker A is more ambiguity averse than decision maker B in the sense of Ghirardato and Marinacci (2002), then $1 - W_t^B(Down) - W_t^B(MiddleUp) \leq 1 - W_t^A(Down) - W_t^A(MiddleUp)$.

PROOF OF OBSERVATION 1: If A is more averse than B, then $z \sim_t^B x_{MiddleUp} y$ implies $z \succcurlyeq_t^A x_{MiddleUp} y$ and $CE_{Down} \sim_t^B x_{Down} y$ implies $CE_{Down} \succcurlyeq_t^A x_{Down} y$. Consequently, decision maker A has lower certainty equivalents than B. Since he has the same utility, we obtain $W_t^B(Down) + W_t^B(MiddleUp) \geq W_t^A(Down) + W_t^A(MiddleUp)$ and therefore, $1 - W_t^B(Down) - W_t^B(MiddleUp) \leq 1 - W_t^A(Down) - W_t^A(MiddleUp)$.■

Decision maker A is *ambiguity averse* (*seeking)*in the sense of Ghirardato and Marinacci (2002) if there exists an expected utility maximizer B such that A is more (less) ambiguity averse than B.

OBSERVATION 2: If the decision maker is ambiguity averse (seeking) in the sense of Ghirardato and Marinacci (2002), then $1 - W_t(E) - W_t(E^c) \geq 0$ ($1 - W_t(E) - W_t(E^c) \leq 0$) for all E.

PROOF OF OBSERVATION 2: If decision maker A is ambiguity averse then there is an expected utility maximizer B (whose utility is by definition an affine transformation of that of A) such that A is more ambiguity averse than B. From Observation 1 and noting that the decision weights of an expected utility maximizer satisfy complementarity, it follows that $1 - W_t(E) - W_t(E^c) \geq 0$. The proof for ambiguity seeking is similar. ■

OBSERVATION 3: Consider two decision makers A and B whose preferences are represented by the neo-additive model such that $U^A = U^B$, $a_t^A = a_t^B$ and $P_t^A = P_t^B$. Decision maker A is more ambiguity averse than decision maker B if and only if $b_t^A \geq b_t^B$ or equivalently, $alpha_t^A \geq alpha_t^B$.

PROOF OF OBSERVATION 3: That $b_t^A \geq b_t^B$ is necessary follows from applying Observation 1 to any event E. Because $a_t^A = a_t^B$ and $P_t^A = P_t^B$, $b_t^A \geq b_t^B$ and Equation 2 imply that $W_t^A(E) \leq W_t^B(E)$ for all E. This implies ambiguity aversion in the sense of Ghirardato and Marinacci (2002). ■

Finally, we relate the pessimism index b_t to source preference. In our study, histories can be considered sources of uncertainty. Tversky and Wakker (1995) say that a decision maker prefers source (history) h_t to source (history) h_s if for any events E and F, $W_s(E) = W_t(F)$ implies $W_s(E^c) \leq W_t(F^c)$. This is equivalent to the statement that the decision maker is more ambiguity averse for history h_s than for history h_t

OBSERVATION 4: Consider a decision maker whose preferences are represented by the neo-additive model. Assume that there exist two events E and F such that $W_s(E) = W_t(F)$. The decision maker prefers history h_t over history h_s if and only if $b_s \geq b_t$.

PROOF OF OBSERVATION 4: Assume that the decision maker has a preference for history h_s over history h_t. Consider the two events two events E and F such that $W_s(E) = W_t(F)$. By source preference, $W_s(E^c) \leq W_t(F^c)$. Consequently, $1 - W_t(F) - W_t(F^c) \leq 1 - W_s(E) - W_s(E^c)$, which implies $b_s \geq b_t$.

Now assume $b_s \geq b_t$ and consider any events E and F such that $W_s(E) = W_t(F)$. From $W_s(E^c) = 1 - b_s - W_s(E)$ and $W_t(F^c) = 1 - b_t - W_t(F)$ it follows that $W_s(E^c) \leq W_t(F^c)$. Consequently, the decision maker prefers history h_t over history h_s ■

Appendix B: Details of the random coefficient estimation.

Let $CE_{it}(x_{E_j}y)$ denote subject i's certainty equivalent of option j with history t: $x_{E_j}y$, where $E_j = \{Up, Middle, Down, MiddleUp\}$ and $t = \{0, 1\text{ week}, 1\text{ month}\}$. To account for errors in subjects' reported certainty equivalents, we add a stochastic term ϵ_{ijt} to the certainty equivalent predicted by Equation (1) with power utility,

$$CE_{it}\left(x_{E_j}y\right) = \left(W_{it}(E_j)x^{\beta_i} + \left(1 - W_{it}(E_j)\right)y^{\beta_i}\right)^{1/\beta_i} + \epsilon_{ijt}. \quad \text{(A1)}$$

The individual parameter β_i is normally distributed with mean β, and variance σ_β^2. The history dependent individual weighting function W_{it} is defined according to Equation (2) and depends on the parameter vector $\eta_{it} = (a_{it}, b_{it}, pup_{it}, pmid_{it})$ and $pup_{it} = P_{it}(Up)$, $pmid_{it} = P_{it}(Middle)$. The certainty equivalent predicted by Equation (1) with utility parameter β_i and the weighting function parameters η_{it} is denoted $\widehat{CE_{it}}\left(x_{E_j}y\right)$.

We assume that η_{i0} follows a multivariate normal distribution with mean η_0 and diagonal variance-covariance matrix $\sigma_{\eta_0}^2$. For $t \in \{1\text{ week}, 1\text{ month}\}$, let $\Delta\eta_t = \eta_{it} - \eta_{i0}$. We assume that β_i follows a normal distribution with mean β and variance σ_β^2. The error term ϵ_{ijt} is assumed to follow a normal distribution with mean 0 and standard deviation σ_i, where σ_i follows a lognormal distribution with parameters μ_ϵ and σ_ϵ^2. The distributions of β_i, σ_i, and η_{i0} are assumed to be independent.

Let $\xi_i = (\beta_i, a_{i0}, b_{i0}, pup_{i0}, pmid_{i0}, \sigma_i)$ be the vector of individual specific random parameters, which are assumed to be independent of each other. Let $f(\cdot)$ denote the density function of ξ and let $\theta = (\beta, \eta_0, \Delta\eta_{1week}, \Delta\eta_{1month}, \mu_\epsilon, \sigma_\beta, \sigma_{\eta_0}, \sigma_\epsilon)$ denote the vector of model parameters. For a given θ, the contribution to the likelihood for subject i is therefore:

$$l_i(\theta) = \int_{R^6} \left[\prod_{j,t} \frac{1}{\sigma_i} \phi\left(\frac{CE_{it}\left(x_{E_j}y\right) - \widehat{CE_{it}}\left(x_{E_j}y\right)}{\sigma_i}\right)\right] f(\xi|\theta)d\xi, \quad \text{(A2)}$$

where $\phi(\cdot)$ is the standard normal density function. The log-likelihood is given by the sum of the logarithm of l_i for all subjects. To approximate the multiple integral in Eq. (A2), we used simulation techniques, where Halton sequences of length 500 were drawn for each individual (Train 2009, ch. 9). We maximized the log-likelihood function with respect to the vector of model parameters θ using the "fminunc" function in Matlab.

The estimates of the multiple priors interpretation (Eq. 6) were obtained as follows. We defined $alpha_t = \frac{1}{2} + \frac{b_t}{2a_t}$ when $a_t > 0$ and $alpha_t = \frac{1}{2}$ otherwise. The mean of the random effect on $alpha_0$ was inferred from the mean of the random effects on b_0 and a_0 by using a second-order Taylor approximation: $E(alpha_0) \sim \frac{1}{2} + \frac{E(b_0)}{2E(a_0)} + V(a_0)\frac{E(b_0)}{E(a_0)^3}$.

Similarly, the fixed effects $\Delta alpha_t$ for $t = \{1 \text{ week}, 1 \text{ month}\}$ were inferred from $E(alpha_0)$ and $E(alpha_t) \sim \frac{1}{2} + \frac{E(b_t)}{2E(a_t)}$. The variance of the random effect on $alpha_0$ was also obtained using a second-order Taylor approximation: $V(alpha_0) \sim \frac{1}{4}\left(\frac{1}{E(a_0)^2}V(b_0) + \frac{E(b_0)^2}{E(a_0)^4}V(a_0)\right)$.

We obtained the standard errors of $alpha_t$ by the Krinsky and Robb method. This method consists in drawing multiple vectors of estimated coefficients from the multivariate normal distribution of θ that has a mean vector equal to the original estimated coefficient vector and the estimated variance-covariance matrix. The generated series of estimated coefficients drawn from the sampling distribution were used to evaluate the relation $alpha_t = \frac{1}{2} + \frac{b_t}{2a_t}$ for each draw. The estimates of the standard error of the transformed coefficients were obtained by computing the standard deviation of the resulting sample of transformed draws.

Table B.1: Random coefficient model with exponential and expo-power utility functions

		Expo utility Decision weight	Expo-Power utility Decision weight
Likelihood insensititivity	a_0	0.57***	0.51***
		[0.05]	[0.04]
	(one week fixed effect)	−0.19**	−0.12**
		[0.06]	[0.05]
	(one month fixed effect)	−0.32***	−0.29***
		[0.05]	[0.05]
	std of random effects	0.18***	0.21***
		[0.03]	[0.02]
Pessimism	b_0	−0.01	−0.09**
		[0.05]	[0.04]
	(one week fixed effect)	−0.09***	−0.08***
		[0.02]	[0.02]
	(one month fixed effect)	−0.06**	−0.04**
		[0.02]	[0.02]
	std of random effects	0.03***	0.12***
		[0.01]	[0.01]
$P(Up)$	pup_0	0.39***	0.40***
		[0.02]	[0.01]
	(one week fixed effect)	−0.08***	−0.05***
		[0.02]	[0.02]
	(one month fixed effect)	−0.09***	−0.06***
		[0.02]	[0.02]
	std of random effects	0.09***	0.07***
		[0.01]	[0.01]
$P(Middle)$	$pmid_0$	0.27***	0.30***
		[0.02]	[0.02]
	(one week fixed effect)	0.002	−0.012
		[0.02]	[0.02]
	(one month fixed effect)	0.07***	0.04***
		[0.02]	[0.02]
	std of random effects	0.12***	0.12***
		[0.01]	[0.01]
Utility	β	0.31	1.12***
		[0.22]	[0.05]
	std of random effects	0.77***	0.13***
		[0.06]	[0.02]
Noise (Fechner error)	μ_ϵ	−2.22***	−2.40***
		[0.05]	[0.05]
	std of random effects	0.51***	0.57***
		[0.04]	[0.04]
Log-likelihood		835.59	846.18
N		1280	1280

Online Appendix (Experimental instructions)

Instructions

Thank you for participating in our experiment. For your participation, you will receive a show up fee of €5 and an extra payment depending on your choices during the experiment. Please read the instructions carefully. Before starting the experiment, we will ask you several questions to test your understanding of the instructions. If you answer every question correctly, you will proceed to the experiment; otherwise, we will ask you to read the instructions once more and re-answer the questions until all your answers are correct. We want to be sure that you have understood the instructions so that your answers in the experiment reflect your preferences and are not caused by any misunderstandings. If you have any questions, please feel free to ask the experimenter.

During the experiment, you have to answer a series of choice questions. There are **no right or wrong answers** to these questions. We are interested in ***your*** preferences. Your final payment will be determined by the choices you make during the experiment. Hence it is in your own interest to reveal your true preferences in the choices you will face.

***

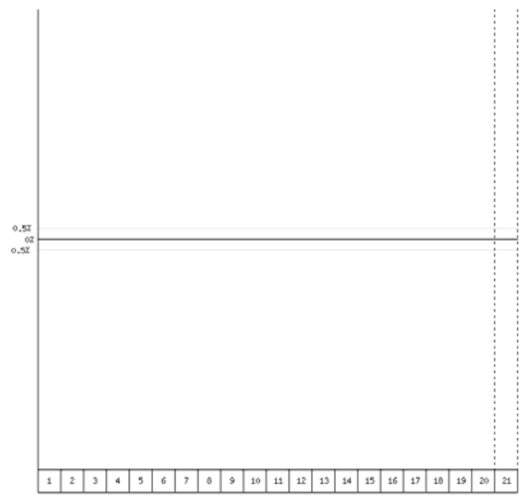
During the experiment, you will be asked to choose between a ***digital option*** for an underlying stock and a sure money amount.

A digital option for an underlying stock pays a pre-specified money amount **H** if a given event occurs and **L** otherwise.

The **underlying stock** is randomly chosen from a database of stocks that were newly-listed on the NYSE between 1 **January** 2009 and 25 February 2011. The stocks in the database are randomly numbered from 1 to 328. At the beginning of the experiment, you will draw 4 numbers from a box, and the 4 corresponding stocks will be used as the underlying stocks of your digital options. At the end of the experiment, the names of the stocks will be revealed, and you can check the historical quotes of the stock prices on Yahoo Finance afterwards. Note that we cannot manipulate the price distribution of the stocks as these are historically given.

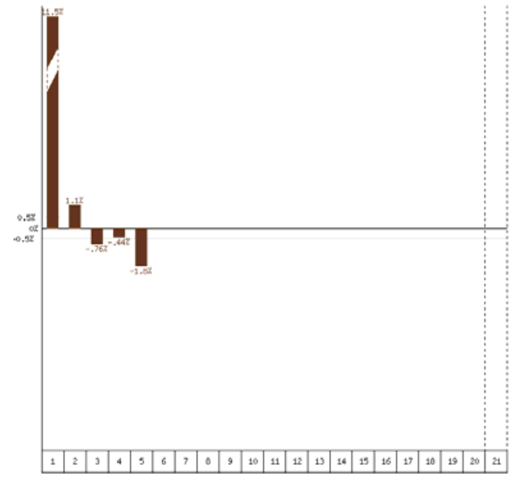
You will face 3 different situations.

Situation 1



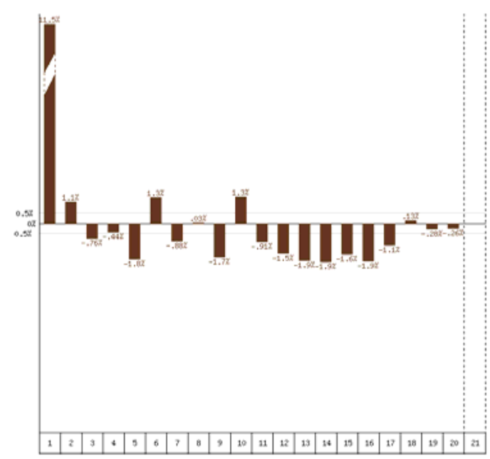
- **Situation 1:** You have an option for an underlying stock, which has just been listed on the Stock Exchange. Consequently, you have no quotes of the historical stock price. You know that the expiration date of the option is the 21^{st} trading day of the stock, and the payoff of the option depends on the daily return of the stock on the 21^{st} trading day. (More explanation about the option payoff will be presented later.)

Situation 2



- **Situation 2:** You have an option for an underlying stock, which has been listed on the Stock Exchange for one week. You have 5 quotes of the historical daily return of the stock, which have been depicted by the brown bars. You know that the expiration date of the option is the (same) 21st trading day of the stock, and the payoff of the option depends on the daily return of the stock on the 21st trading day.

Situation 3

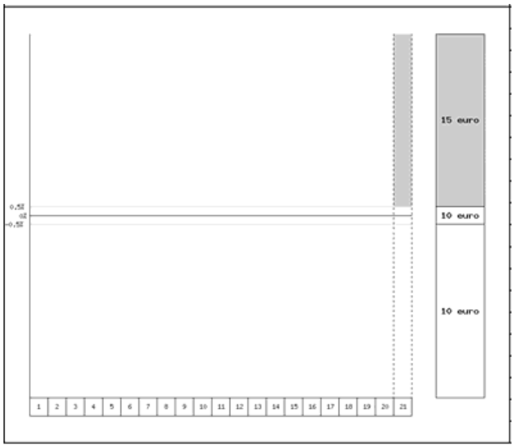


- **Situation 3:** You have an option of an underlying stock, which has been listed on the Stock Exchange for 20 days. You have 20 quotes of the historical daily return of the stock, which have been depicted by the brown bars. You know that the expiration date of the option is the (same) 21st trading day of the stock, and the payoff of the option depends on the daily return of the stock on the 21st trading day.

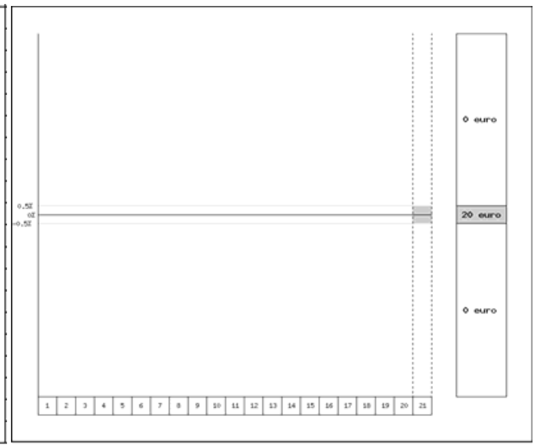
You will face 4 types of *digital option*

For each situation described above, you may face 4 types of digital options. Here, we use the first situation as an example to illustrate the 4 types of digital options.

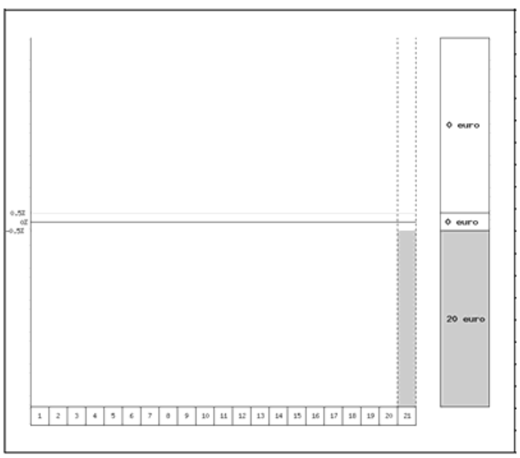
$Up - Option$



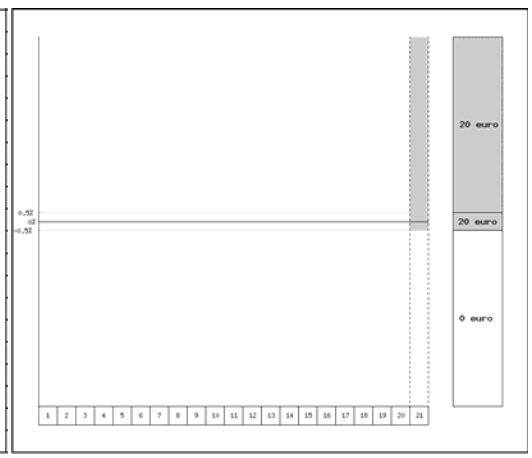
$Middle - Option$



$Down - Option$

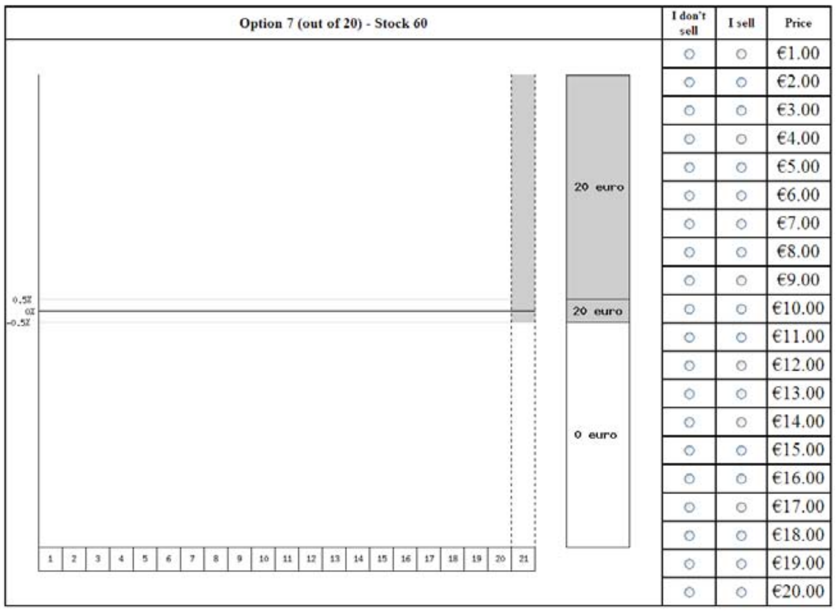


$MiddleUp - Option$



- An **Up-option** pays €H if the daily return (r) of the underlying stock on its expiration day exceeds +0.5% (r > +0.5%) and €L otherwise;
- **A Middle-option** pays €H if the daily return (r) of the underlying stock on its expiration day varied between –0.5% and +0.5% (–0.5% ≤ r ≤ +0.5%) and €L otherwise;
- A **Down-option** pays €H if the daily return (r) of the underlying stock on its expiration day is less than –0.5% (r < –0.5%) and €L otherwise.
- A **MiddleUp-option** pays €H if the daily return (r) of the underlying stock on its expiration day exceeds –0.5% (r ≥ –0.5%) and €L otherwise;

€H and €L are pre-specified money amounts. For instance, the figure above displays an Up-option with H=15 and L=10, and the other three types with H=20 and L=0. You may encounter different H and L in the experiment.



We will determine your selling price of 20 different options through a series of choices between **the option** and **a certain money amount**. An example is given in the above figure. For each of the 20 prices, you are asked to indicate whether you would like to sell the option or not. The money amount where you switch your choice from 'I don't sell' to 'I sell' is taken as your selling price. All sales will be realized on the 21st day.

If you sell at €x, do you agree that you also want to sell at prices higher than €x? Y/N

If you don't sell at €y, do you agree that you don't want to sell at prices lower than €y? Y/N

Payment

After making all the 20 choices, please call the experimenter. The experimenter will let you throw a 20-sided dice twice. The first throw will determine which of your choices will be played for real. The second throw determines the price you are offered.

As an example, imagine that you throw 7 on your first throw and 6 on your second. Hence the 7th choice will be selected and the price you are offered for the option in the 7th choice is €6.

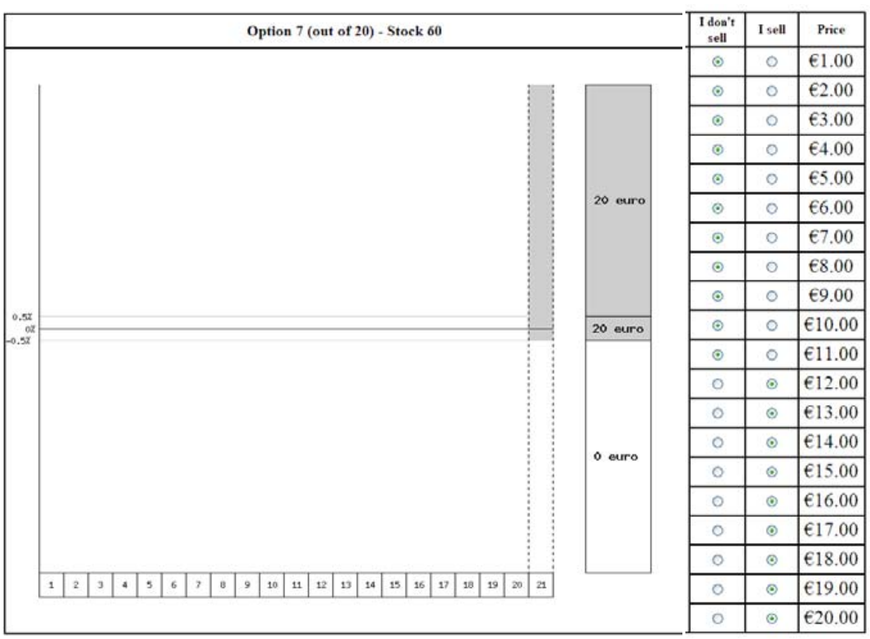
Suppose that option in the 7th choice is a MiddleUp-option with H=20 and L=0, as in the figure above. Suppose further that your selling price for the 7th option was found to be €9. This means that you are not willing to sell the option for a price less than €9 and, hence, you don't accept the offered price of €6 and **thus you keep the option**;

- If the daily return on the 21st trading day of the underlying stock is at least –0.5% (e.g. 0.15%), then we pay you €20 plus the €5 show-up fee. In total you get €25.
- If the daily return on the 21st trading day of the underlying stock is smaller than –0.5% (e.g. –1.49%), then we pay you €0 plus the €5 show-up fee. In total you get €5.

Now imagine that you throw 7 and 10. Then the price offer you are offered is €10. Because you are willing to sell the option if the price is at least €9, you accept the offered price of €10 and thus we pay you €10 plus the €5 show-up fee. In total you get €15.

Note that it is in your best interests to state your selling price truthfully. To see that, suppose your true selling price is €9, but you state a selling price of €11. Then if the price we offer for the option is €10, you keep the option even though it is worth less to you than €10.

Questions:



Suppose you are going to play the choice in the picture above for real.

1. What is the minimum selling price?
2. What is the payoff of the plotted option, if the daily return on the 21st trading day is:
    - 1.4%?
    - −0.45%?
    - −1.4%?
3. Suppose that the daily return on the 21st trading day is 1.4%, what is the total payment you get if the second number you throw is 1?
4. Suppose the daily return on the 21st trading day is 1.4%, what is the total payment you get if the second number you throw is 15?

References

Abdellaoui M, Baillon, A, Placido, L, Wakker, PP (2011) The rich domain of uncertainty: Source functions and their experimental implementation. *American Economic Review* 101: 695-723.

Abdellaoui M, l'Haridon, O, Zank, H (2010) Separating curvature and elevation: A parametric probability weighting function. *Journal of Risk and Uncertainty* 41: 39-65.

Abdellaoui M, Bleichrodt, H, L'Haridon, O, Van Dolder, D (2016) Measuring loss aversion under ambiguity: A method to make prospect theory completely observable. *Journal of Risk and Uncertainty* 52: 1-20.

Abdellaoui M, Vossmann, F, Weber, M (2005) Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Science* 51: 1384-1399.

Andersen S, Harrison, GW, Lau, MI, Rutström, EE (2006) Elicitation using multiple price list formats. *Experimental Economics* 9: 383-405.

Baillon A, Bleichrodt, H (2015) Testing ambiguity models through the measurement of probabilities for gains and losses. *American Economic Journal: Micro* 7: 77-100.

Charness G, Levin, D (2005) When optimal choices feel wrong: A laboratory study of Bayesian updating, complexity, and affect. *American Economic Review* 88: 933-946.

Chew SH, Sagi, JS (2006) Event exchangeability: Probabilistic sophistication without continuity or monotonicity. *Econometrica* 74: 771-786.

Chew SH, Sagi, JS (2008) Small worlds: Modeling attitudes toward sources of uncertainty. *Journal of Economic Theory* 139: 1-24.

Cohen M, Gilboa, I, Jaffray, J, Schmeidler, D (2000) An experimental study of updating ambiguous beliefs. *Risk, Decision, and Policy* 5: 123-133.

Di Mauro C (2008) Uncertainty aversion vs. competence: An experimental market study. *Theory and Decision* 64: 301-331.

Diecidue E, Schmidt, U, Zank, H (2009) Parametric weighting functions. *Journal of Economic Theory* 144: 1102-1118.

Dimmock SG, Kouwenberg, R, Wakker, PP (2016a) Ambiguity attitudes in a large representative sample. *Management Science* 62(5): 1363-1380.

Dimmock SG, Kouwenberg, R, Mitchell, OS, Peijnenburg, K (2016b) Ambiguity aversion and household portfolio choice puzzles: Empirical evidence. *Journal of Financial Economics* 119(3): 559-577.

Dominiak A, Dürsch, P, Lefort, J (2012) A dynamic Ellsberg urn experiment. *Games and Economic Behavior* 75: 625-638.

Eichberger J, Grant, S, Kelsey, D (2007) Updating Choquet beliefs. *Journal of Mathematical Economics* 43: 888-899.

Eichberger J, Grant, S, Kelsey, D (2010) Comparing three ways to update Choquet beliefs. *Economics Letters* 107: 91-94.

Eichberger J, Grant, S, Kelsey, D (2012a) When is ambiguity attitude constant? *Journal of Risk and Uncertainty* 45: 239-263.

Eichberger J, Grant, S, Lefort, J (2012b) Generalized neo-additive capacities and updating. *International Journal of Economic Theory* 8: 237-257.

El-Gamal MA, Grether, DM (1995) Are people Bayesian? uncovering behavioral strategies. *Journal of the American Statistical Association* 90: 1137-1145.

Ellsberg D (1961) Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics* 75: 643-669.

Epstein LG (2006) An axiomatic model of non-Bayesian updating. *The Review of Economic Studies* 73: 413-436.

Epstein LG, Schneider, M (2007) Learning under ambiguity. *The Review of Economic Studies* 74: 1275-1303.

Ert E, Trautmann, ST (2014) Sampling experience reverses preferences for ambiguity. *Journal of Risk and Uncertainty* 49: 31-42.

Fox CR, Weber, M (2002) Ambiguity aversion, comparative ignorance, and decision context. *Organizational Behavior and Human Decision Processes* 88: 476-498.

Fox CR, Tversky, A (1998) A belief-based account of decision under uncertainty. *Management Science* 44: 879-895.

Gajdos T, Hayashi, T, Tallon, JM, Vergnaud, JC (2008) Attitude toward imprecise information. *Journal of Economic Theory* 140: 27-65.

Gallagher J (2014) Learning about an infrequent event: Evidence from flood insurance take-up in the united states. *American Economic Journal: Applied Economics* 6: 206-233.

Ghirardato P, Maccheroni, F, Marinacci, M (2004) Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory* 118: 133-173.

Ghirardato P, Marinacci, M (2002) Ambiguity made precise: A comparative foundation. *Journal of Economic Theory* 102: 251-289.

Ghirardato P, Marinacci, M (2001) Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of Operations Research* 26: 864-890.

Gilboa I, Schmeidler, D (1993) Updating ambiguous beliefs. *Journal of Economic Theory* 59: 33-49.

Gilboa I, Marinacci, M (2013) Ambiguity and the Bayesian paradigm. Acemoglu D., Arellano M. Dekel E., eds. *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society.*(Cambridge University Press, New York).

Gilboa I (1987) Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics* 16: 65-88.

Gilboa I, Schmeidler, D (1989) Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics* 18: 141-153.

Grether DM (1980) Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics* 95: 537-557.

Griffin D, Tversky, A (1992) The weighing of evidence and the determinants of confidence. *Cognitive Psychology* 24: 411-435.

Hamermesh DS (1985) Expectations, life expectancy, and economic behavior. *The Quarterly Journal of Economics* 100: 389-408.

Hanany E, Klibanoff, P (2007) Updating preferences with multiple priors. *Theoretical Economics* 2: 261-298.

Heath C, Tversky, A (1991) Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty* 4: 4-28.

Hoffman RM, Kagel, JH, Levin, D (2011) Simultaneous versus sequential information processing. *Economics Letters* 112: 16-18.

Kahn BE, Sarin, RK (1988) Modeling ambiguity in decisions under uncertainty. *Journal of Consumer Research* 15: 265-272.

Kahneman D, Tversky, A (1972) Subjective probability: A judgment of representativeness. *Cognitive Psychology* 3: 430-454.

Kahneman D, Tversky, A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47: 263-291.

Keynes JM (1921) *A treatise on probability* (McMillan, London).

Kilka M, Weber, M (2001) What determines the shape of the probability weighting function under uncertainty? *Management Science* 47: 1712-1726.

List JA (2004) Neoclassical theory versus prospect theory: Evidence from the marketplace. *Econometrica* 72: 615-625.

Myagkov M, Plott, CR (1997) Exchange economies and loss exposure: Experiments exploring prospect theory and competitive equilibria in market environments. *The American Economic Review* 87: 801-828.

Phillips LD, Edwards, W (1966) Conservatism in a simple probability inference task. *Journal of Experimental Psychology; Journal of Experimental Psychology* 72: 346.

Poinas F, Rosaz, J, Roussillon, B (2012) Updating beliefs with imperfect signals: Experimental evidence. *Journal of Risk and Uncertainty* 44: 219-241.

Rabin M, Schrag, JL (1999) First impressions matter: A model of confirmatory bias. *The Quarterly Journal of Economics* 114: 37-82.

Schmeidler D (1989) Subjective probability and expected utility without additivity. *Econometrica* 57: 571-587.

Smith VK, Taylor Jr, DH, Sloan, FA (2001) Longevity expectations and death: Can people predict their own demise? *American Economic Review* 1126-1134.

Stott HP (2006) Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty* 32: 101-130.

Train KE (2009) *Discrete choice methods with simulation* (Cambridge University Press, Cambidge UK).

Trautmann ST, Van de Kuilen, G (2016) Ambiguity attitudes. Keren G, Wu G, eds. *Blackwell Handbook of Judgment and Decision Making* (Wiley Blackwell), 89-116.

Tversky A, Kahneman, D (1973) Availability: A heuristic for judging frequency and probability. *Cognitive Psychology* 5: 207-232.

Tversky A, Koehler, DK (1994) Support theory: A nonexistential representation of subjective probability. *Psychological Review* 101: 547-567.

Tversky A, Fox, C (1995) Weighing risk and uncertainty. *Psychological Review* 102: 269-283.

Tversky A, Wakker, PP (1995) Risk attitudes and decision weights. *Econometrica* 63: 297-323.

Tversky A, Kahneman, D (1992) Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5: 297-323.

van de Kuilen G, Wakker, PP (2006) Learning in the Allais paradox. *Journal of Risk and Uncertainty* 33: 155-164.

Wakker PP (2010) *Prospect theory: For risk and ambiguity* (Cambridge University Press, Cambridge UK).