

**Learning and Confirmation Bias:
Measuring the Impact of First Impressions and Ambiguous Signals**

Julie Agnew, Hazel Bateman, Christine Eckert, Fedor Iskhakov, Jordan Louviere, and Susan Thorp

August 2018

PRC WP2018-8

Pension Research Council Working Paper

Pension Research Council

The Wharton School, University of Pennsylvania

3620 Locust Walk, 3000 SH-DH

Philadelphia, PA 19104-6302

Tel.: 215.573.3414 Fax: 215.573.3418

Email: prc@wharton.upenn.edu

<http://www.pensionresearchcouncil.org>

**Learning and Confirmation Bias:
Measuring the Impact of First Impressions and Ambiguous Signals**

Julie Agnew, Hazel Bateman, Christine Eckert, Fedor Iskhakov, Jordan Louviere, and Susan Thorp

Abstract

We quantify the widespread and significant economic impact of first impressions and confirmation bias in the financial advice market. We use a theoretical learning model and new experimental data to measure how these biases can evolve over time and change clients' willingness to pay advisers. Our model demonstrates that clients' confirmation bias will reinforce the effect of first impressions. Our results also lend support, in a new financial context, to theoretical models of learning under limited memory where people use unclear signals to confirm and reinforce their current beliefs. We find that almost two thirds of the participants in our experiment make choices that are consistent with a limited memory updating process: they interpret unclear advice to be good advice when it comes from the adviser they prefer. Our results show that models that account for behavioral factors such as confirmation bias may be needed to explain some financial decisions.

Julie Agnew

Mason School of Business
William and Mary
Williamsburg, VA, USA

Hazel Bateman

UNSW Business School
UNSW Sydney
Sydney, NSW, Australia

Christine Eckert

UTS Business School
University of Technology
Sydney, NSW, Australia

Fedor Iskhakov

School of Economics
Australian National University
Canberra, ACT, Australia

Jordan Louviere

University of South Australia
Adelaide, SA, Australia

Susan Thorp

The University of Sydney Business School
The University of Sydney
Sydney, NSW, Australia

1. INTRODUCTION

Around the world, people face increasing responsibility for financial decisions that will directly impact their long term financial wellbeing (Ryan et al. 2011). People with limited financial literacy confront complicated choices over credit, mortgages, investments, and retirement plans (Lusardi and Mitchell 2011). In theory, the expert advice garnered from financial advisers should help consumers make sound decisions (Hackethal et al. 2012). However, empirical research shows that better outcomes are not guaranteed. While some papers show that advice improves portfolios, others find that financial advisers give poor counsel or bolster client's biases or mistakes (Anagol et al. 2017; Bergstresser et al. 2009; Chalmers and Reuter 2015; Hackethal and Inderst 2013; Inderst and Ottaviani 2009, 2012a, b; Mullainathan et al. 2012). As a result, it is not surprising that people find it difficult to distinguish good advice from bad, and consequently, good advisers from bad advisers (Agnew et al. 2018). In this study, we take a deeper look at how and why clients' evaluations of advisers evolve over time, and we also explore the monetary consequences of the behavioral biases that influence their assessments.

New research demonstrates how easily advisers can manipulate consumers into following them by making good first impressions and confirming prior beliefs. However these studies do not address whether these actions have economic consequences nor do they identify the type and proportion of consumers who are most susceptible to these strategies (Agnew et al. 2018). Therefore, the magnitude of this phenomenon and whether, in turn, this is a cause for concern is not clear. Using new experimental data and adapting a new theoretical learning model, this paper addresses these important unanswered questions. It also contributes to the literature by highlighting the widespread and significant economic role first impressions and confirmation bias can play in the financial advice market. To our knowledge, our paper is the first to measure

how these biases can evolve over time and transform a consumer's willingness to pay for advisers.

Our study implements and tests the limited memory learning model in a new, financial context. Until now, the learning model has only been tested using experiments that examine how participants interpret information about two policy issues. In our study, clients must learn about the quality of the advisers through the quality of the advice given. Our results show that the limited memory model can explain features of financial decisions that conventional models cannot, such as the persistent and unwarranted trust that some clients place in financial advisers (ASIC 2012; Mullainathan et al. 2012).

The paper is structured as follows. We begin this paper in Section 2 by discussing learning models and their relationship to confirmation bias. Section 3 follows with a description of the market for financial advice. Section 4 presents a formal model drawn from Fryer et al. (forthcoming) that incorporates confirmation bias and compares it to a rational updating model. Section 5 presents our experimental design. Our empirical results are highlighted in Section 6 followed by a discussion of the findings in Section 7 to conclude the paper.

2. *CONSUMER LEARNING AND CONFIRMATION BIAS*

Consumers frequently make decisions with incomplete information. As a result, learning is best understood not as a simple collection of knowledge, but as a hypothesis-testing process where new information is encoded and integrated with existing beliefs (Hoch and Deighton 1989). A prime example is when clients have to decide whether to follow financial advice in an area where they have little experience, such as how to invest retirement savings. Clients who have incomplete information will usually rely on signals to reach a decision about the quality of

an adviser. For example, clients might rely on an adviser's professional certification, consider advice given on a different topic from the adviser before, or rely on other people's opinions of the quality of the different advisers. All of these signals help clients form an initial belief. New signals and added experience then help them update these beliefs until they can make better-informed decisions. The need to update beliefs formed with incomplete information is common to many situations, such as when consumers purchase for the first time in an unknown product category, when voters choose between two political candidates, when doctors have to choose between two alternative treatments, or, as in the example above, when clients have to decide whether to continue to trust their financial advisers.

People, however, rarely update their beliefs in a rational way. Instead they tend to interpret evidence "in ways that are partial to existing beliefs, expectations, or a hypothesis in hand" (Nickerson 1998) and tend to search harder for information that confirms their beliefs (Muthukrishnan 1995; Snyder and Swann 1978), labelled "confirmation bias". Confirmation bias has proven to be a robust phenomenon in areas as diverse as beliefs about the deterrent effect of the death penalty, nuclear power generation, climate change, brand loyalty and sexual morality (see Fryer et al. forthcoming, Online Appendix C, Table 1, for a summary). Confirmation bias is often founded on a first impression (Beattie and Baron 1988) and also explains how two people can reach opposite opinions after they review common evidence (Darley and Gross 1983). The defining feature of this bias is that additional information leads to the polarization, rather than the moderation, of prior opinions.

Confirmation bias cannot be incorporated into traditional models because it violates a basic assumption of conventional Bayesian learning (Eckstein et al. 1988; Erdem and Keane 1996; Roberts and Urban 1988). Where a rational Bayesian learner would ignore ambiguous new

information, confirmation bias causes the learner to interpret ambiguous new information as a reinforcement of prior beliefs. Thus in contrast to learning models that allow people to give higher weight to new signals from specific sources (Camacho et al. 2011), learning under confirmation bias not only leads to different *weighting* of signals, but actually can reverse the *interpretation* of the signal. Irrespective of the actual signal valence, a person with confirmation bias will treat an ambiguous signal as positive if their prior belief is positive, and will treat it as negative if their prior belief is negative. Such a biased updating in turn leads to overconfidence so that people may come to believe with near certainty in a false hypothesis despite receiving an infinite amount of information (Rabin and Schrag 1999).

In this paper, we investigate whether a learning model that allows for confirmation bias performs better than a rational model that does not. Specifically, we focus on how individuals judge their financial adviser over time based on the advice given to them. We believe that confirmation bias is likely to be more relevant to services such as financial advice, where consumers are not always able to objectively judge quality.

3. *THE MARKET FOR FINANCIAL ADVICE*

The market for financial advice is growing as households face new and difficult choices. “Do-it-yourself” finance is the term Ryan et al. (2011) coined to describe the increased responsibility untrained or inexperienced people have been given for financial decisions. There is ample evidence that people frequently make poor financial decisions when deciding on these new and more complicated products (Campbell et al. 2011). Reasons for these errors include low levels of financial literacy (Lusardi and Mitchell 2011), issues of trust in markets and financial products (Christelis et al. 2010), behavioral biases (Thaler and Benartzi 2004) and limited

cognition (Lusardi and Mitchell 2006). Opportunities to learn from experience in financial contexts are limited because many consequential decisions, such as the choice of a mortgage or retirement account investment, are made infrequently, and feedback from outcomes is often delayed. Financial firms may have incentives to make products more complex to impede consumer learning and preserve profits (Carlin 2009; Carlin and Manso 2010). Financial choices, and the mistakes that often follow, have serious implications for financial welfare and stability, individually and in aggregate (Agarwal et al. 2009; Bar-Gill and Warren 2008; Campbell 2006; Campbell et al. 2011).

Consumers can delegate difficult decisions to financial advisers to compensate for low financial literacy and lack of expertise (Hackethal et al. 2012). However, theory predicts that an adviser's willingness to de-bias and educate clients can be diluted by incentive structures (e.g. Inderst and Ottaviani 2009). Empirical studies likewise show that advisers can exploit the biases of clients (Hackethal et al. 2012; Mullainathan et al. 2012) and that clients credulously continue to trust advisers who deliver poor quality advice (Australian Securities and Investment Commission 2012) (ASIC 2012). In a recent study, Agnew et al. (2018) illustrate how much first impressions in the client-adviser relationship matter, complementing research that shows that clients form opinions of their financial advisers rapidly (Yaniv and Kleinberger 2000). Together, these findings explain how some advisers can successfully use strategies to build and maintain client trust while also providing unhelpful advice (Anagol et al. 2017; Mullainathan et al. 2012). Experimental work also shows that advisers who confirm a client's views on straightforward questions early in an advice relationship are subsequently rated as more trustworthy and competent than advisers who contradict a client's views. Furthermore, clients are more likely to accept their later advice on complicated topics (Agnew et al. 2018). Thus, establishing trust early

on can lead to fruitful business interactions subsequently. More trusted advisers are likely to be able to charge higher fees and thus take a larger share of the benefits of the advice relationship (Gennaioli et al. 2015).

Despite these hazards, citizens of many countries use advice services. For example, Chater et al. (2010) report that 58% of individual investors' stock purchases were influenced by an adviser, in a survey of 6,000 consumers across eight EU countries. Holden et al. (2013) find that people choose to work with advisers because the advisers have expertise in an area that their clients do not. Other studies emphasize that the personal qualities of advisers matter; for example, clients must decide if an adviser is trustworthy and competent before acting on advice. Georgarakos and Inderst (2014) show that clients with limited financial capability are more likely to follow advice if they trust their advisers, but trust depends on many factors, including the client's capability, the accuracy and quality of information provided, and a belief that the adviser and client's incentives are aligned (Sniezek and Van Swol 2001; Yaniv and Kleinberger 2000).

In summary, people are facing challenging new financial choices and may not have the financial knowledge or experience to make sound decisions. Those who turn to advisers for help cannot be certain of getting the best advice, yet they still must form beliefs about advisers' quality using the signals they receive from them. Prior beliefs about advisers' quality will depend on client and adviser characteristics; these beliefs could also influence a client's interpretation of subsequent signals from the advisers. Clients who process information in a limited or biased way are less likely to reach sound judgements about the quality of an adviser.

4. A MODEL OF CONFIRMATION BIAS

Learning under confirmation bias

While confirmation bias and the polarization of opinion that follows is at odds with standard Bayesian updating models, it can be explained by updating models with the form of limited memory introduced in Fryer et al. (forthcoming). Consider a rational Bayesian decision maker who may or may not be able to discern the state of the world, in our case the quality of a financial adviser. Assume that the decision maker, here the “client” forms an expectation over the two states of the world: A (the adviser is good), and B (the adviser is bad). The client holds an initial prior (or starting) belief that $P(A)=\lambda_0$, which they update as they receive a sequence of clear or ambiguous signals from an adviser.¹ We can interpret this starting belief as the client’s initial belief about the adviser. The client receives a clear good (a) or bad (b) signal of the adviser’s quality in the form of a correct or incorrect recommendation from the financial adviser. The signal (recommendation) may either agree with, or contradict, what the client thinks is factual or sound. The client uses clear signals to update their prior belief and form a posterior expectation of the state of the world, i.e., of the quality of the adviser. Let $s>1/2$ denote the probability the client receives a clear, good signal conditional on the adviser being good, $P(a | A)=s$, and assume that the probability of receiving a clear, good signal from a bad adviser is $P(a | B)=1-s$. The parameter s determines how much the beliefs of the client are influenced by the signal, and therefore s can be interpreted as signal strength.

However, the client may also receive an ambiguous signal ab . In our context, an ambiguous signal might be a recommendation on a topic where the client is inexperienced or

¹ We note that Fryer et al (forthcoming) in their general model additionally allow for signals \emptyset that contain no information. Since the financial advice we evaluate here in this study is inherently good or bad we do not allow for empty signals.

uninformed. Ambiguous signals create an opportunity for confirmation bias to operate. Rational Bayesian updaters ignore ambiguous signals and forms a posterior only over the sequence of clear signals. They thus gradually uncover the true state of the world. However, when updaters have a confirmation bias, they will not overlook an ambiguous signal. Rather they will interpret the ambiguous signal in line with their current belief, either as a or b , and thus reinforce their existing view of the state of the world. Fryer et al. (forthcoming) show that limited memory – the need to form a posterior belief on receipt of each signal rather than wait to the end of the sequence - forces an interpretation of ambiguous signals that generate confirmation bias and polarization of opinions.

More formally, in the rational model, beliefs are updated according to

$$(1) \lambda_{t+1} = P(A | \lambda_t, \sigma_{t+1}) = \begin{cases} \frac{s\lambda_t}{s\lambda_t + (1-s)(1-\lambda_t)}, & \text{if } \sigma_{t+1} = a, \\ \frac{(1-s)\lambda_t}{(1-s)\lambda_t + s(1-\lambda_t)}, & \text{if } \sigma_{t+1} = b, \\ \lambda_t, & \text{if } \sigma_{t+1} = ab, \end{cases}$$

where σ_t is the advice received in choice situation t ($t=1, \dots, T$), that is, the signal. When a limited memory updater (Fryer et al., forthcoming, hereafter FHJ) holds a prior belief that the adviser is good quality, he or she interprets an ambiguous signal as a good signal, and as a bad signal when he or she holds a prior that the adviser is poor quality. In the limited memory (FHJ) updating model, the client's beliefs are updated according to:

$$(2) \lambda_{t+1} = P(A | \lambda_t, \sigma_{t+1}) = \begin{cases} \frac{s\lambda_t}{s\lambda_t + (1-s)(1-\lambda_t)}, & \text{if } \sigma_{t+1} = a, \text{ or } \sigma_{t+1} = ab \text{ and } \lambda_t > \frac{1}{2}, \\ \frac{(1-s)\lambda_t}{(1-s)\lambda_t + s(1-\lambda_t)}, & \text{if } \sigma_{t+1} = b, \text{ or } \sigma_{t+1} = ab \text{ and } \lambda_t < \frac{1}{2}, \\ \lambda_t, & \text{if } \lambda_t = \frac{1}{2}. \end{cases}$$

In our experimental application, discussed in detail in Section 5, clients view professionally filmed videos of two advisers giving advice on four different advice topics. (Table 1 shows the scripts of the advice.) The sequence of the topics and the quality of the advice given by each adviser varies by experimental treatment. The videos are adapted from the Agnew et al. (2018) experiment and the signals are the advice given.² Clients perceive each piece of advice given by each adviser as ambiguous or clear, and as good or bad. For each topic, the adviser either provides an unequivocally right or wrong piece of advice based on financial theory but the quality of that advice may not be apparent to all clients, depending on their financial knowledge and experience. For example, for those with strong financial literacy, the advice given by each adviser should provide clear signals of the adviser's quality. However, for clients with limited financial literacy or experience with the topic, some of the advice signals may be ambiguous. The experimental design in this paper includes four advice topics from Agnew et al. (2018) and six previously untested sequences of good/bad advice. The order in which we present topics to experimental participants (clients), combined with the order in which each adviser gives either good or bad advice creates a test of 24 sequences of signals of adviser quality for each of two advisers, or a total of 48 sequences. The new variations in the advice sequences allow us to observe, for the first time, whether participants update their beliefs about an adviser in a conventionally rational way or with limited memory.

INSERT Table 1 ABOUT HERE

Figure 1 illustrates all possible paths of beliefs for the experimental "clients" under different assumptions about initial priors and updating strategies (in rows). Given a client's initial prior beliefs and updating strategy, each path depends on how many advice topics were

² To view an example of the video advice from a treatment in Agnew (2018), please follow this link <https://drive.google.com/file/d/0B-1NMLVfExG1ZzFhZWlrRWlsR2s/preview>.

perceived by the client as clear, and in which choice sets advice on the clear topics was given (in columns, ranging from all topics perceived as clear to all topics perceived as ambiguous). The size of the dots reflects the theoretical proportion of all possible choice patterns that pass through the respective point. We arbitrarily set the probability that a good adviser offers a good signal, s , to 0.75. Rows 1, 3 and 4 reflect FHJ updating with starting priors λ_0 of 0.6, 0.5, and 0.4, respectively. Row 2 is based on rational updating with starting prior λ_0 equal to 0.5.

INSERT Figure 1 ABOUT HERE

For example, consider, the first column of graphs, where we assume that all topics are clear to the client. Differences between paths thus only come from differences in the sequence of good and bad advice that one adviser gives the client over four topics. If all topics are clear, and if all clients have the same starting prior, rational and FHJ updating methods lead to the same posterior beliefs (see column 1, rows 2 and 3). When moving to the next column of graphs, we can see that if one topic is ambiguous, FHJ updating leads to deviations from the rationally updated priors and polarization. The polarization becomes more pronounced as clients perceive more topics to be ambiguous (columns further right), and becomes extreme when they perceive all four topics to be ambiguous: FHJ updating clients reach an almost certain belief that the adviser is good if they start with an initial prior larger than 0.5 (column 5, row 1) and reach an almost certain belief that the adviser is bad if they start with a prior smaller than 0.5 (column 5, row 4). Rational clients do not update their priors if all signals are ambiguous (column 5, row 2). FHJ updaters also do not update when all signals are ambiguous and their starting prior is equal to 0.5 (row 3, column 5).

Model Description

In this section, we describe how confirmation bias can be calibrated and linked to contextual variables using information on a client's sequence of choices of as well as his or her willingness to pay.

Assume a client $k, = 1, \dots, K$, receives a sequence of $t_k = 1, \dots, T_k$ signals $\sigma_{tk} = (\sigma_{Rtk}, \sigma_{Ltk})$ from two different sources (advisers) labeled R and L . We assume that the signals σ_{Rtk} and σ_{Ltk} received in choice set t_k are either both clear or both ambiguous to the client, k . In our case, the sequence represents the choice sets, the signals are advice on a financial topic, and the sources of the signals are financial advisers. We further assume that $(\sigma_{Rtk}, \sigma_{Ltk}) \in \{(a, b), (b, a), (ab, ab)\}$, where from the perspective of the client, a is a clear signal of good quality, b is a clear signal of bad quality, and ab is an ambiguous signal. In each choice set: i) one adviser gives good advice and the other provides bad advice; ii) the client interprets the advice as either ambiguous or clear; iii) the client chooses between the two sources based on their interpretation of the quality of the signal. Thus, in our experiment, the client (k) chooses whether to follow the advice $(\sigma_{Rtk}, \sigma_{Ltk})$ of adviser R or L provided in choice set t . We code the choice data as

$$(3) \quad y_{kt} = \begin{cases} 1, & \text{if } R \text{ was chosen at choice } t_k \text{ by client } k, \\ 0, & \text{if } L \text{ was chosen at choice } t_k \text{ by client } k. \end{cases}$$

Further, assume that after having received T signals from each adviser, the client can choose to purchase an additional unit from each adviser at a certain price p_k . In our example, we ask participants whether they would be willing to pay a particular amount for a one-hour session with the adviser. Let y_k^L (y_k^R) be indicator variables, taking the value 1 if L (R) is chosen by decision maker k in this context³.

³ Note that the model can be extended to include more entities or more attributes to influence the different choices.

Equations (1) and (2) describe how clients update their beliefs depending on whether they are Bayesian rational decision makers, or whether they are prone to confirmation bias, respectively. Which updating scheme a particular client uses is not known by the researcher and needs to be inferred from the observed choices. Similarly, in many cases, the researcher does not know whether a signal is clear or ambiguous to the decision maker and must make inferences about this from the choice data.

In both updating schemes, the posterior belief (or updated prior) depends on the initial prior belief λ_{k0}^r (or starting prior) over adviser $r \in \{R, L\}$ of the decision maker k . The starting prior itself depends on the characteristics of the advisers $r \in \{R, L\}$ and the decision maker, X_0 (see Table 2 for a description of the variables used in our empirical example) and an unknown vector of parameters β_0 . So that we can estimate this parameter, we make the starting prior probability of adviser quality to be a logit function of features of the adviser and the client:

$$(4) \quad \lambda_{k0}^r = \frac{\exp(\beta_0 X_0)}{1 + \exp(\beta_0 X_0)}$$

When combined with a value for signal strength, s , we can calculate λ_{kj}^r , the updated prior about adviser $r \in \{R, L\}$ of decision maker k after choice set j , conditional on the client's updating scheme and signal clarity based on Equations (1) and (2).⁴

INSERT Table 2 ABOUT HERE

We also infer from choice data and survey responses which type of updating scheme the client uses and which topics are ambiguous or clear to him or her. If the client perceives the signal to be clear (i.e., the topic is easy for them), we assume that he or she will choose the

⁴ To enable us to identify parameters, we set s to an arbitrary value greater than 0.5 and check sensitivity of estimation to alternative choices. The results we report below use $s=0.75$.

source that gives a high-quality signal in this choice set (i.e., the adviser who gives correct advice). That is, if the topic is clear, the client chooses the advice based on its quality alone. However, if the client perceives the signal to be ambiguous (i.e., the topic is hard and the client cannot distinguish good from bad advice), we assume that the client will make the choice in this choice set according to his or her posterior belief about the source, in our case, based on posterior beliefs about the advisers' qualities.

We denote $q_{rk}^t \in (a, b)$ as the quality associated with signal σ_{rt} that decision maker k receives. We further define q_k^t equal to 1 if $q_{rk}^t = a$ and -1 otherwise. We acknowledge that decision makers still make some errors when choosing and that this error is extreme value distributed with scale $1/\beta_1$ and $1/\beta_2$, respectively, and thus obtain

$$(5) \quad P(y_{kt} = 1 | \sigma_{ik} = \text{clear}) = \frac{\exp(\beta_1 q_k^t)}{1 + \exp(\beta_1 q_k^t)}$$

and

$$(6) \quad P(y_{kt} = 1 | \sigma_{ik} = \text{ambiguous}) = \frac{\exp(\beta_2 \cdot (\lambda_{kt}^R - \lambda_{kt}^L))}{1 + \exp(\beta_2 \cdot (\lambda_{kt}^R - \lambda_{kt}^L))}.$$

Thus both β_1 and β_2 are scale parameters: as β_1 (β_2) approaches infinity, the expression in the right hand side approaches 1 for $q_k^j = 1$ ($\lambda_{kt}^R > \lambda_{kt}^L$), and 0 otherwise.

Next we turn to the client's willingness to pay. The client can choose to pay for an additional unit from each adviser (i.e. advice from both, one, or none of the advisers). We model this choice as follows:

$$(7) \quad P(\text{willing to pay for } r) = \frac{\exp(\beta_3^0 + \beta_3^1 \lambda_{kt}^r + X_3 \beta_3^2)}{1 + \exp(\beta_3^0 + \beta_3^1 \lambda_{kt}^r + X_3 \beta_3^2)},$$

where X_3 are attributes of the client and r , including the price of another unit of the R or L , β_3^0 is a constant, β_3^1 captures the impact of the posterior on the willingness to pay, and β_3^2 is a vector of unknown parameters.

Summarizing, in our specification, clients' sequences of choices, including their decisions to pay for another round of advice, are functions of posterior beliefs about r . These posterior beliefs are, in turn, a function of the way clients update their beliefs (either rationally or with confirmation bias) and of which signals clients perceive to be clear (i.e., which topics are easy to them). Our model assigns clients to latent classes distinguished by clarity or ambiguity of the signals in choice set t_k and to latent classes distinguished by their updating scheme. In the interest of parsimony, we assume that the probability that a client is a particular updating type and the probability that a client treats any topic as clear are independent, conditioning on the characteristics of the individual client, so that:⁵

$$(8) \quad P_k(\tau) = P_k(\tau_{\text{clarity}}, \tau_{\text{rationality}}) = P_k(\tau_{\text{clarity}})P_k(\tau_{\text{rationality}}),$$

and

$$(9) \quad P_k(\tau_{\text{clarity}}) = \prod_{t_k=1}^{T_k} P_k(\sigma_{t_k} = \text{clear}).$$

Dependence between the latent classes for any client k is captured by allowing class membership probabilities to be influenced by client-specific covariates X_4 and X_5 and associated parameter vectors β_4 , β_5 , and signal specific constants β_5^{tk} , with $\beta_5^{tk} = \beta_5^{t'k'}$ for $\sigma_{t_k} = \sigma_{t'k'}$:

$$(10) \quad P_k(\tau_{\text{rational}}) = \frac{\exp(\beta_4 X_4)}{1 + \exp(\beta_4 X_4)},$$

⁵ Note that in our empirical example without any further assumptions there exist two types of updaters as well as a classification of clear or ambiguous for each of four topics, which results in $2^5=32$ different combinations.

and

$$(11) \quad P_k(\sigma_{kt} = \text{clear}) = \frac{\exp(\beta_5^{tk} + \beta_5 X_5)}{1 + \exp(\beta_5^{tk} + \beta_5 X_5)}.$$

We estimate the parameters $\theta = \{\beta_0, \beta_1, \beta_2, \beta_4, \beta_5\}$ by jointly maximizing the likelihood of choices and willingness to pay decisions. Conditional on the client belonging to one of the C clarity-rationality class combinations τ_c , $c=1, \dots, C$, the likelihood of client k 's sequence of choices is

$$(12) \quad l_k(\theta | \tau_c) = \prod_{t_k=1}^T P(y_{kt} = 1 | \tau_c)^{y_{kt}=1} P(\text{willing to pay for entity } R | \tau_c)^{y_k^R=1} P(\text{willing to pay for entity } L | \tau_c)^{y_k^L=1}.$$

The unconditional likelihood of client k 's sequence of choices is thus:

$$(13) \quad l_k(\theta) = \sum_{c=1}^C P_k(\tau_c) l_k(\theta | \tau_c).$$

Parameter identification

A formal analysis of identification is not feasible for the complex, non-linear learning model discussed above (see also the discussion in Ching et al. 2013). In the following, we sketch our identification strategy for the key model parameters.

First, consider the initial prior belief about r , which is the starting prior $\lambda_{k_0}^r$. The starting prior belief is the basis for the updated posterior belief and thus influences both the choices of r as well as willingness to pay. The starting prior belief itself also influences directly the choices made in choice set 1, as in this set we assume that (up to uncertainty) the adviser (source) with the higher initial prior is chosen if the topic is ambiguous. Since the design of the experiment

(discussed in Section 5 and described in Table 3) ensures that participants (clients) face both easy as well as hard (ambiguous) signals in choice set 1 (refer to panel B in Table 3), we thus obtain sufficient information to estimate the starting prior as well as how it depends on advisers' r and decision maker's characteristics (β_1).

INSERT Table 3 ABOUT HERE

Next we discuss the signal strength $s = \Pr(a | A) = \Pr(b | B)$, which in our empirical application is the probability that a good (bad) signal comes from a good (bad) adviser. We set $s = 0.75$ to allow the probability to be greater than 0.5 but less than one that a good adviser delivers good advice, to ensure that updating can occur. We test for the sensitivity of results at $s = 0.60$ and results remain largely unchanged.

The parameter β_2 is in turn identified via the starting prior $\lambda_{k_0}^r$ and s . These two parameters jointly define the updated beliefs and can thus be considered as pre-determined covariates when participants face an ambiguous topic. Choices made in choice sets with ambiguous signals can thus identify β_2 .

Choices made over the different choice sets allow us to identify the latent "clarity classes." More specifically, our assumption about the choice process can (up to uncertainty in the choice process) be summarized as follows. If we observe that the client chooses the adviser that gives a bad quality signal, we can conclude that the signal was ambiguous for him or her. We cannot make a similar inference if the client chooses the adviser that gives a good quality signal, as this could imply either that the signal was clear for that client or that they chose the adviser because of a higher associated posterior belief. The combined information of updated prior beliefs about the advisers and incorrect choices of advice thus allows us to identify what we call clarity classes.

Since the starting prior belief about the adviser can be inferred from the data without any assumptions about how clients update their beliefs and since signal strength s is fixed, we can calculate the posteriors for both updating schemes. The posterior associated with the higher likelihood then helps to pin down rationality classes.

Discontinuity of the likelihood function and estimation method

Estimation of our model is complicated by the fact that the likelihood function is discontinuous for those cases where participants update their beliefs according to the FHJ updating scheme. The discontinuity in the likelihood appears along the dimensions of the parameters of prior beliefs. Even in the simplest case when the starting prior belief is represented by a single constant (as illustrated in Figure 2), as this constant moves from zero to one and crosses particular thresholds dependent on other parameters, the values and the counts of the possible updated prior beliefs change discontinuously.

INSERT Figure 2 ABOUT HERE

Figure 2 compares the updated prior beliefs λ_4 (after a sequence of four choices has been made) under the rational and FHJ updating for three values of the signal strength parameter s as the starting prior λ_0 changes from zero to one. For each value of the starting prior λ_0 we draw all values of posterior beliefs that are possible in the model (by assuming all possible clarity and quality combinations of signals), with the size of the circles indicating the number of theoretical paths leading to that belief. For example, in row 1 and column 2, at a signal strength of $s=0.75$ and where all participants use FHJ updating, we see that a small change in the prior belief from 0.49 to 0.51 results in large differences in the posterior: for a prior of 0.49 most posteriors are smaller than 0.1, while at a prior of 0.51 most posteriors are larger than 0.8. Since both the prior

as well as the posterior beliefs influence the likelihood function, a small change in the prior can thus lead to a huge - and discontinuous - change in the likelihood function. We use Sequential Adaptive Bayesian Learning (SABL) proposed by Durham and Geweke (2014) to overcome this challenge and estimate the model. Appendix A provides a further discussion of the discontinuity problem and outlines the estimation procedure.

5. *EXPERIMENTAL DESIGN*

We designed our experiment to achieve two goals. First, we aimed to collect data to estimate and compare a model of client confirmation bias (FHJ) with a model of rational (standard Bayesian) choice. Second, we aimed to measure the participants' willingness to pay for advice and how it is affected by confirmation bias. Since clients who lack experience or financial literacy are probably more susceptible to manipulation, we also collect an array of demographics, preferences, financial capability measures, and psychological inventories to help identify these clients.

We fielded a four-part online survey that included an incentivized choice experiment in December 2014.⁶ Members of a nationally representative online panel were invited to complete the survey. Those who responded to the invitation had to pass two screening questions to meet age and gender quotas. This resulted in 2,003 "clients". To ensure incentive compatibility, we compensated participants who completed the experiment for their time and rewarded them if they chose correct advice in each choice set and in a post experiment quiz. Participants first answered a set of questions that measured their general financial literacy and numeracy (Lusardi and Mitchell 2011; Lipkus et al. 2001), and that evaluated their understanding and experience of

⁶ The survey offered participants who completed all questions a small compensation for their time (around \$4) and one entry in a prize draw for \$A50 for each correct choice of advice.

the participants of the four advice topics covered by advisers in the subsequent discrete choice experiment (DCE).

Our DCE offers “clients” a sequence of videos of advisers who give financial advice on four common and important consumer finance topics.⁷ The topics include credit card debt repayment, retirement savings account consolidation, diversification in equity investment, and index fund fees. Table 1 records the scripts for the good and bad advice for each topic.

To identify the effect of confirmation bias in the DCE, we ensured advice topics, advisers, the environment and the mode of advice delivery to be uniform. Figure 3 shows the advisers from the videos and their “names”. The videos allow us to control the two advisers shown to each participant, the order of advice topics, the quality of advice given by each adviser for each topic and the attributes of advisers giving advice.

INSERT Figure 3 ABOUT HERE

In the DCE videos, two advisers give a recommendation on each of the four topics: each participant received four pairs of advice; the two advisers were the same across the four advice topics for each participant; and in each case, one adviser provided a correct recommendation, while the other provided an incorrect recommendation. Correct and incorrect advice hereafter is termed “good” and “bad” in line with the description of the method in the previous section. The videos systematically varied adviser factors - the adviser’s gender (2 options: male or female) and age (2 options: young or old), professional certification (2 options: certification presented or not) - the order of the advice topics (4 options: first, second, third, or fourth) and the quality of the advice (2 options: correct or incorrect).

⁷ The topics were previously used by Agnew et al. (2018) based on their relevance for people around the world, that they had unequivocally right and wrong answers and were based on the mistakes often made in these areas.

The experiment used a between-subjects design. As noted, the advice viewed by any one participant is provided by the same two advisers; hence, variation in adviser factors (age, gender, certification) is a between-subjects manipulation. To minimize the between-subjects treatment groups, we used a fold-over design in which we created the 2^3 complete factorial of possible advisers and paired each of them with their “mirror image” (that is, the exact opposite level, so that a younger woman adviser was matched with an older male adviser). This produced pairs of advisers who were orthogonal in the differences in factor levels. The resulting design is optimally efficient under the assumption that a conditional multinomial logit choice model underlies the participant choices (Street et al. 2005; Street and Burgess 2007). This design approach produced four between-subject treatment groups and is shown in panel A of Table 3.

Further variation in the DCE relates to between-subject manipulation of a) topic sequence and b) order in which good and bad advice is given by each adviser. Variation in these orders is essential to test hypotheses about formation of persistent participant preferences for advisers. The fold-over design used to create the between-subjects manipulations ensures variation in quality of advice. We also maximized variation in adviser attributes by ensuring that both financial advisers gave advice on the same topic in each pair. Thus, we combined the between-subjects treatment groups (4) with a design to vary the orders of topics (4 levels) and good and bad advice (2 levels). A full factorial design would have required a very complex survey program and a very large sample, since it implies 16 possible sequences of good (G) and bad (B) advice and 24 possible sequences of topics. To maximize variation and to enable a test of confirmation bias we used six sequences of good and bad advice orders – those where each

adviser gives two good and two bad recommendations (see panel C of Table 3) - and topic sequences with an equal number of hard and easy financial topics (see panel B of Table 3).⁸

When we combined the four possible pairs of advisers with the six possible sequences of topics and the six possible sequences of advice quality, we obtained a design with $6*6*4 = 144$ conditions. We randomly assigned at least 10 and up to 14 participants to each condition.

After the DCE task, participants rated the trustworthiness, competence, attractiveness, understanding, professionalism, financial expertise, genuineness and persuasiveness of the advisers they saw, and stated their willingness to pay \$X for a one-hour session with both, one or none of the advisers. We assigned to participants fixed fee values

$X \in \{50, 100, 150, 250, 500, 750\}$ so as to minimize their predictability from the other manipulated characteristics of the experiment condition. After they had answered questions about demographics (e.g., marital status, household size and number of dependents, education, labor market status, income, gross assets, and debts/liabilities) and personal characteristics, including personality traits and risk attitudes, participants read debriefing information that explained correct advice. The survey closed with an incentivized quiz on the debriefing material.⁹

6. EMPIRICAL RESULTS

In aggregate, participants chose correct recommendations 79% of the time. The percentages of correct choices by topic are 86% for retirement account consolidation, 88% for

⁸ We rely on the results by Agnew et al (2018) who find that two topics (debt repayment and retirement account consolidation) are relatively easy (E) and that the other two topics (diversification and index fund fees) are relatively hard (H).

⁹ Appendix A compares the characteristics of the sample with the Australian Census data from 2011. Our sample reports slightly higher educational attainments and a higher probability of being married than the census shows, but otherwise is representative of the population.

credit card debt repayment, 79% for stock diversification and 64% for index fund fees.

Participants chose the advice of the young female adviser two percentage points more often than the advice of the older male who appears alongside her in the experiment. Participants chose the younger male and older female equally often. On average, participants were also more likely to choose the certified adviser (52%) than the uncertified adviser (48%), illustrating the importance of an initial good impression. The fact that this difference was smaller for the “easy topics” (retirement account consolidation and credit card debt repayment; 2.6%), than for the “hard topics” (stock diversification and index fund fees; 5%), indicates that beliefs about the adviser are more important if the signal quality is ambiguous.

We estimated our model using data from 1,903 of the 2,003 participants and held back the remaining responses to assess hold-out fit. In-sample fit was satisfactory and hold-out fit did not deviate very much from in-sample fit, which shows that our model does not over-fit the data: The model predicted an average (over all choice sets) probability of 0.69 for the estimation sample that the adviser who was in fact chosen would be chosen, and it predicted a related probability of 0.69 for the hold out sample. When the adviser was not chosen, the predicted choice probability decreased to 0.29 for the estimation sample, whereas it decreased to 0.28 for the hold-out sample. The predicted probabilities were less discriminating in the willingness to pay choice probabilities. When a participant chose to pay the adviser, the model’s average predicted probability was 0.48 for the estimation sample data and 0.44 for the hold-out data. When a participant chose not to pay the adviser, the average predicted probability of being paid was 0.28 for the estimation sample data and 0.34 for the hold-out sample data. Thus, the model slightly underestimates the probability that a participant is willing to pay the proposed fee for the adviser.

We also compared our model to a model that assumes pure rational updating (where $P_k(\tau_{rational})=1$ for all k), in line with traditional learning models. The logmarginal density for this restricted model is -5887.17 compared to a logmarginal density of -5827.47 for our proposed model for in-sample fit, and -305.62 versus -298.35 for hold-out sample fit, thus demonstrating that accounting for confirmation bias significantly improves model fit. Indeed, as also discussed further below, 63% of participants update their beliefs in a way that is not consistent with rational Bayesian updating, which leads to both polarized choice probabilities as well as willingness to pay estimate.

Table 4 reports the model estimates. For each parameter, we report the mode of its posterior distribution as well as the 2.5 and 97.5 percentiles of this distribution, that is, the corresponding equi-tailed credible interval (CI). There is a 95% probability that the parameter is not zero if zero does not fall in the CI. Next, we discuss each of the model components.

INSERT Table 4 ABOUT HERE

Prior belief about adviser

We allow the starting prior belief about an adviser's quality to depend on the trust that a participant has in financial advisers (Guiso et al. 2008) and also whether the adviser displayed a professional certification in the experiment (Agnew et al. 2018). Both factors have been shown by earlier studies to influence whether people will take financial advice. The mode of the distribution for the trust parameter equals 0.520, and the 95% credible interval does not contain zero. This shows that participants who rate financial advisers as trustworthy hold a higher prior belief that the adviser is good, as we would expect. The mode of the distribution for the non-certification parameter equals -0.085 and the 95% CI again does not contain zero. Based on the

posterior draws of the parameters (not reported here), we can infer the distribution of the difference in prior beliefs for certified versus uncertified advisers. From the perspective of participants who trust financial advisers already (“Trust in advisers” variable =1), this posterior distribution has a mean of 0.016 with an associated 95% CI of [0.001, 0.036]. For participants who generally distrust financial advisers (“Trust in advisers” variable =-1), the impact of certification is double: this posterior distribution has a mean of 0.032 with a 95% CI of [0.003, 0.073]. The mean of the posterior for participants who are neutral about financial advisers (“Trust in advisers” variable =0) is 0.023 [0.002, 0.054]. We infer that if an adviser displays a professional certification, participants form a significantly higher prior belief that he or she will give good advice and this higher expectation will be subsequently reflected in higher choice probabilities in the case of ambiguous topics and higher willingness to pay for additional advice from this adviser. Certification has a stronger influence on participants who are generally skeptical of adviser quality. This is in line with the findings by Agnew et al. (2018) who show that displaying a certification significantly increases an adviser’s likelihood of having this or her advice accepted by a client.

Choice of advice

Participants’ ability to choose good advice depends on the clarity of the topic. In the case of easy topics, the participant will choose the good advice (up to some error, equation (5)). In the case of hard topics, equation (6) posits that (up to some error) participants will choose the adviser they rate as better, according to participants’ updated (posterior) beliefs. The parameters associated with both the quality of the advice and the belief about the adviser are positive, equaling 4.138 and 2.663, respectively, and the 95% credible intervals based on their posterior distributions do

not include zero. This translates into the following choice probabilities: the probability that the participant chooses good advice if the topic is easy equals $\exp(4.296) / (1 + \exp(4.296)) = 0.99$; if the topic is ambiguous, the probability of an adviser R with associated belief $\lambda^R = 1$ being chosen when being evaluated against an adviser L with associated belief $\lambda^L = 0$ is $\exp(2.510 \cdot (1 - 0)) / (1 + \exp(2.510 \cdot (1 - 0))) = 0.92$. These results confirm that ambiguous signals are related to more uncertainty and variability in participants' choices.

Willingness to pay for advice

Our model assumes that the willingness to pay a particular price for an additional hour with the adviser depends on the actual price charged, several characteristics of the participant, as well as the participant's posterior belief about this adviser.

Table 4 shows that parameters here have the expected signs but some have credible intervals that include zero. The impact of price is negative with a mode of -0.085 and a 95% CI interval that does not include zero. On the other hand, the impact of the posterior belief about the adviser is positive (18.309) with the associated 95% CI also not including zero. Of the remaining participant characteristics, the only parameter with a CI that does not include zero is the indicator for whether the participant has paid for financial advice in the past. The mode of this parameter is positive at 0.466 with a 95% CI of [0.348, 0.570] and we conclude that participants who have paid for advice in the past are more willing to pay than those who have not.

Based on these parameters and Equation (7) it is possible to calculate the associated price difference $\Delta \text{price} = \text{price}_{new} - \text{price}_{old}$ that a participant is willing to pay for a specific difference in posterior beliefs $\Delta \text{belief} = \text{belief}_{new} - \text{belief}_{old}$, namely

$$(9) \quad \Delta\text{price} = -\frac{\beta_2^{\text{posterior}}}{\beta_2^{\text{price}}}\Delta\text{belief}\cdot 100,$$

where the multiplication with 100 is necessary because the price was divided by 100 before entering the estimation. We can use this formula to calculate the additional dollar amount that participants are willing to pay for their preferred adviser. Based on the posterior distribution of the estimates, we obtain additional willingness to pay estimates that have a mean of \$1722 with the lower bound of the 95% CI equal to \$189 and the upper bound equal to \$4639.

Rational versus FHJ updating

Our model shows that a 62.9% of participants exhibit confirmation bias. In our setup, we use participants' conscientiousness and impulsiveness to explain which participants are more likely to display confirmation bias. Table 4 shows that participants with high impulsiveness are less likely to be rational updaters (mode of -0.3445, the 95% CI does not include zero). This parameter implies that more impulsive participants are more likely to interpret ambiguous signals as a confirmation of their prior belief. In contrast, high conscientiousness has a positive mode (0.243) but the 95% CI does include zero.

Clarity of topics

In our model, we assume that whether a topic is perceived as clear or ambiguous by a participant depends on the participant's characteristics as well as on the topic itself. More specifically, we find that participants with more expertise are more discerning. Results show that participants with high knowledge of the financial products related to the advice, high financial literacy and high numeracy are more likely to perceive a topic as clear. For all these variables, the modes of the posterior distributions are positive and the 95% CIs do not contain zero. Gender

and participants' age also affect whether a topic is perceived as clear versus ambiguous: female participants are significantly more likely to perceive a topic as clear and so are participants who are 40 years or older. In addition, the size and sign of the topic-specific constants is in line with the share of correct answers for these topics. The advice related to index fund fees is perceived as significantly more difficult than all other topics since the associated 95% CI does not overlap with the CI of any other topic.

Table 5 reports the percentage of participants who belong to each of the 16 possible clarity classes. For example, 18.2% of participants perceive all topics to be clear; 3.8% of participants perceive all topics to be ambiguous; and 21.9% of participants struggle to understand advice on index fund fees even though all other topics are clear to them.

INSERT Table 5 ABOUT HERE

Illustration of Model Implications

Our model allows us to compare the impact of the participants' two different updating strategies on their choices. It additionally allows us to measure the impact of first impressions on subsequent choices. To illustrate, consider two participants A and B who update their beliefs according to the rational and biased updating scheme, respectively. Let the right adviser (R) display a certification and the left adviser (L) not display a certification. For the sake of simplicity, we assume that the participants distrust financial advisers, we set other characteristics at the medians of the survey sample distributions, and we fix estimated parameters at the mode of the posterior distributions. Both participants will thus have the same prior belief about the

right (R) and the left (L) adviser of $\lambda_{A0}^R = \lambda_{B0}^R = \frac{\exp(1.728 + 0.085 - 520)}{1 + \exp(1.728 + 0.085 - 520)} = 0.785$ and

$\lambda_{A0}^L = \lambda_{B0}^L = \frac{\exp(1.728 - 0.085 - 0.520)}{1 + \exp(1.728 - 0.085 - 0.520)} = 0.755$. Assume that Adviser R gives good advice on

a clear topic in the first choice set and that both advisers give (from the client's perspective) ambiguous advice in the remaining three choice sets.

Table 6 shows how the updated prior beliefs, choice probabilities for the advisers in each choice set evolve in this scenario. Both clients update their beliefs in the same way at the first choice because they get clear information about adviser quality. Participant A's beliefs about the advisers, as well as the associated choice probabilities, remain the same throughout the later three choice sets as this participant simply ignores the ambiguous information and ends the experiment still favoring Adviser *R*. In contrast, Participant B interprets all new information in line with current beliefs, so this participant will treat all ambiguous information as evidence that Adviser *R* is good and that Adviser *L* is bad. Thus, Participant B's updated beliefs about Adviser *R* rise steadily and so does his or her probability of choosing Adviser *R*.

INSERT Table 6 ABOUT HERE

The table thus shows that FHJ updating leads to a choice probability that is very close to one for Adviser *R* and close to zero for Adviser *L*, while the same probabilities are 0.9 (Adviser *R*) and 0.1 (Adviser *L*) for the rational updater. It also shows the difference a first impression makes. An early clear signal has a stronger influence on the FHJ updater, whose opinion approaches certainty over few choices. Combined with Equation (13), the results in Table allow us to calculate the monetary value of a first impression. While both consumers are willing to pay \$176 more for Adviser *R* after the first piece of advice has been given, this amount rises by 18% to \$214 after all four pieces for the FHJ updater whereas it stays constant for the rational updater.

7. CONCLUSION

Using a unique discrete choice experiment, this paper contributes to both the financial advice and behavioral finance literature by measuring the economic consequences of two common biases on clients' willingness to pay for financial advisers. To our knowledge, this is the first time this has been done. Our model puts a dollar value on differences in first impressions, as well as on differences in updating schemes, and demonstrates that ignoring confirmation bias will underestimate the impact of first impressions. More specifically, we show that, although immediately after the first impression the willingness to pay is the same for the rational and the biased learner, three additional pieces of ambiguous information can lead the biased learner to be willing to pay 18% more than their rational counterpart.

Furthermore, our results show that it is too strong to assume that all learners apply rational Bayesian methods when they receive and process signals of attribute quality. However, to our knowledge, we are the first to show, in a financial context, how many consumers are actually failing to update in a rational way. Utilizing the learning model developed by Fryer et al. (forthcoming), we find that almost two thirds of participants in our experiment did not use the commonly assumed rational method, instead they made choices consistent with the FHJ model, a limited memory updating process where people use unclear signals to confirm and reinforce their current beliefs. This significant percentage is consistent with Fryer et al.'s (forthcoming) experimental test examining how opinions related to public policy issues are formed. Our experiment also presents evidence that people who are unsure of how to interpret the signals they receive and who do not ignore them, not only end up with strongly biased beliefs, but will spend accordingly. We also demonstrate how polarizing opinions about financial advisers can result, even when consumers are given identical signals but start with different priors.

Our research provides a glimpse into the type of decision maker that is more prone to become a biased updater. By using participant characteristics to predict the probability of using rational versus FHJ updating, we show that particularly impulsive people are more likely to suffer financially from the information processing bias. Also, our model segments decision makers by their updating strategies. This information has many potential applications for companies. For example, companies can gain valuable insight into how customer segment(s) are likely to respond to new information, based on their prior beliefs, either towards more moderate or more polarized preferences.

In the context of financial advice, our model provides useful insights for financial advisers and public policy. For the former, we show that displaying recognizable professional certifications has a significant positive impact on first impressions which then filters through to a higher chance that clients will accept advice and a higher willingness to pay for additional advice. Thus, advisers should assess the costs of gaining a qualification in the light of these possible future gains. The implications for public policy are even more interesting: regulators should consider how advisers are able to use credentials to increase their pay. If credentials are signals of superior service and recommendations, then credentials can provide helpful information to the consumer. However, many different certifications of varying quality may be available to advisers. Future research could study more extensively whether clients can distinguish among the many credentials available and whether the quality of the designation is properly incorporated into clients' willingness to pay for advisers. If not, regulators should consider whether they should limit the designations available to those meeting certain criteria. In addition, while current research emphasizes improvements to financial literacy to encourage

sensible financial decisions, our model shows that less impulsivity can also increase client welfare. Thus, our model proposes another potentially important lever.

Finally, our results could stimulate research into learning models that account for behavioral tendencies, so as to capture how people with limited financial knowledge approach decision making. One possible modification of our model involves the updating strategies. That is, we assume that participants are either purely FHJ or purely rational updaters so they interpret ambiguous signals as exactly confirming their prior belief (FHJ updating) or as not being informative at all (rational updating). It is possible that participants interpret ambiguous signals as only partly confirming their priors, meaning that there is a continuum between extreme updating processes.

References:

Agarwal, Sumit, John C Driscoll, Xavier Gabaix, and David Laibson (2009), "The age of reason: Financial decisions over the life cycle and implications for regulation," *Brookings Papers on Economic Activity*, 2009 (2), 51-117.

Agnew, Julie R, Hazel Bateman, Christine Eckert, Fedor Iskhakov, Jordan Louviere, and Susan Thorp (2018), "First Impressions Matter: An Experimental Investigation of Online Financial Advice," *Management Science*, 64(1), 288-307.

Anagol, Santosh, Shawn Allen Cole, and Shayak Sarkar (2017), "Understanding the advice of commissions-motivated agents: Evidence from the Indian life insurance market," *Review of Economics and Statistics*, 99(1), 1-15.

Australian Securities and Investment Commission, ASIC (2012), "Shadow shopping study of retirement advice, Report 279, Sydney.," (accessed September 21, 2017, [available at <http://download.asic.gov.au/media/1343876/rep279-published-27-March-2012.pdf>].

Bar-Gill, Oren and Elizabeth Warren (2008), "Making credit safer," *University of Pennsylvania Law Review*, 1-101.

Barchard, Kimberly Anne (2001), "Emotional and social intelligence: Examining its place in the nomological network," *University of British Columbia*.

Beattie, Jane, and Jonathan Baron (1988), "Confirmation and matching biases in hypothesis testing," *The Quarterly Journal of Experimental Psychology*, 40(2), 269-97.

Bergstresser, Daniel, John MR Chalmers, and Peter Tufano (2009) "Assessing the costs and benefits of brokers in the mutual fund industry." *The Review of Financial Studies* 22(10), 4129-4156.

Camacho, Nuno, Bas Donkers, and Stefan Stremersch (2011), "Predictably non-Bayesian: quantifying salience effects in physician learning about drug quality," *Marketing Science*, 30(2), 305-20.

Campbell, John Y (2006), "Household finance," *The Journal of Finance*, 61(4), 1553-604.

Campbell, John Y, Howell E Jackson, Brigitte C Madrian, and Peter Tufano (2011), "Consumer financial protection," *The Journal of Economic Perspectives*, 25(1), 91-113.

Carlin, Bruce I (2009) "Strategic price complexity in retail financial markets," *Journal of financial Economics*, 91(3), 278-287.

Carlin, Bruce I, and Gustavo Manso (2010) "Obfuscation, learning, and the evolution of investor sophistication," *Review of Financial Studies*, 24(3), 754-785.

Chalmers, John, and Jonathan Reuter (2015) "Is Conflicted Investment Advice Better than No Advice?." Working Paper, University of Oregon.

Chater, Nick, Steffen Huck, and Roman Inderst (2010), "Consumer decision-making in retail investment services: A behavioural economics perspective," Report to the European Commission/SANCO.

Chernozhukov, Victor and Han Hong (2004), "Likelihood estimation and inference in a class of nonregular econometric models," *Econometrica*, 72(5), 1445-80.

Ching, Andrew T, Tülin Erdem, and Michael P Keane (2013), "Invited Paper—Learning models: An assessment of progress, challenges, and new developments," *Marketing Science*, 32(6), 913-38.

Christelis, Dimitris, Tullio Jappelli, and Mario Padula (2010), "Cognitive abilities and portfolio choice," *European Economic Review*, 54(1), 18-38.

Darley, John M and Paget H Gross (1983), "A hypothesis-confirming bias in labeling effects," *Journal of Personality and Social Psychology*, 44(1), 20.

Douc, Randal and Olivier Cappé (2005), "Comparison of resampling schemes for particle filtering," in *Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium on: IEEE*.

Durham, Garland and John Geweke (2014), "Adaptive sequential posterior simulators for massively parallel computing environments," in *Bayesian Model Comparison: Emerald Group Publishing Limited*.

Eckstein, Zvi, Dan Horsky, and Yoel Raban (1988), "An Empirical Dynamic Model of Brand Choice," Working Paper No. 88, University of Rochester.

Erdem, Tülin and Michael P Keane (1996), "Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets," *Marketing Science*, 15(1), 1-20.

Fryer, Roland G, Philipp Harms, and Matthew O Jackson (forthcoming), "Updating Beliefs when Evidence is Open to Interpretation: Implications for Bias and Polarization," *Journal of the European Economic Association*.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny (2015), "Money doctors," *The Journal of Finance*, 70(1), 91-114.

Georgarakos, Dimitris and Roman Inderst (2014), "Financial advice and stock market participation." ECB Working Paper, No. 1296, European Central Bank (ECB), Frankfurt.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2008), "Trusting the stock market," *The Journal of Finance*, 63(6), 557-600.

Hackethal, Andreas, and Roman Inderst (2013) *How to make the market for financial advice work*. Oxford: Oxford University Press.

Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli (2012), "Financial advisers: A case of babysitters?," *Journal of Banking and Finance*, 36(2), 509-24.

Hoch, Stephen J and John Deighton (1989), "Managing what consumers learn from experience," *The Journal of Marketing*, 1-20.

Holden, Sarah A (2013), "When, why, and how do mutual fund investors use financial advisers?," In O.S. Mitchell and K. Smetters (Eds.) *The Market for Retirement Financial Advice*, Oxford: University Press Oxford, 249-274.

Hsee, Christopher K, Yang Yang, Xingshan Zheng, and Hanwei Wang (2015), "Lay rationalism: Individual differences in using reason versus feelings to guide decisions," *Journal of Marketing Research*, 52(1), 134-46.

Inderst, Roman and Marco Ottaviani (2009), "Misselling through agents," *The American Economic Review*, 99(3), 883-908.

Inderst, Roman, and Marco Ottaviani (2012a), "Competition through commissions and kickbacks," *American Economic Review*, 102(2), 780-809.

Inderst, Roman, and Marco Ottaviani (2012b) "Financial advice," *Journal of Economic Literature*, 50(2), 494-512.

Jasra, Ajay, David A Stephens, and Christopher C Holmes (2007), "On population-based simulation for static inference," *Statistics and Computing*, 17(3), 263-79.

Lipkus, Isaac, Greg Samsa and Barbara Rimer (2001), *General Performance on a Numeracy Scale among Highly Educated Samples*, 21(1), 37-44.

Lusardi, Annamaria and Olivia S Mitchell (2006), "Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth," *National Bureau of Economic Research*.

Lusardi, Annamaria and Olivia S Mitchell (2011), "Financial literacy around the world: an overview," *Journal of Pension Economics and Finance*, 10(4), 497-508.

Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar (2012), "The market for financial advice: An audit study," *National Bureau of Economic Research Working Paper 17929*.

Muthukrishnan, Anaimalai V (1995), "Decision ambiguity and incumbent brand advantage," *Journal of Consumer Research*, 22(1), 98-109.

Nickerson, Raymond S (1998), "Confirmation bias: A ubiquitous phenomenon in many guises," *Review of General Psychology*, 2(2), 175.

Rabin, Matthew and Joel L Schrag (1999), "First impressions matter: A model of confirmatory bias," *The Quarterly Journal of Economics*, 114(1), 37-82.

Roberts, John H and Glen L Urban (1988), "Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice," *Management Science*, (2), 167-85.

Ryan, Andrea, Gunnar Trumbull, and Peter Tufano (2011), "A brief postwar history of US consumer finance," *Business History Review*, 85(03), 461-98.

Sniezek, Janet A and Lyn M Van Swol (2001), "Trust, confidence, and expertise in a judge-advisor system," *Organizational Behavior and Human Decision Processes*, 84(2), 288-307.

Snyder, Mark and William B Swann (1978), "Hypothesis-testing processes in social interaction," *Journal of Personality and Social Psychology*, 36(11), 1202.

Street, Deborah J, Leonie Burgess and Jordan Louviere (2005), "Quick and Easy Choice Sets: Constructing Optimal and Nearly Optimal Stated Choice Experiments," *International Journal of Research in Marketing* 22(4), 459-470.

Street, Deborah J and Leonie Burgess (2007), *The Construction of Optimal Stated Choice Experiments*, Wiley.

Thaler, Richard H and Shlomo Benartzi (2004), "Save more tomorrow™: Using behavioral economics to increase employee saving," *Journal of Political Economy*, 112(S1), S164-S87.

Yaniv, Ilan and Eli Kleinberger (2000), "Advice taking in decision making: Ego-centric discounting and reputation formation," *Organizational Behavior and Human Decision Processes*, 83(2), 260-81.

Table 1: Financial Advice Script

This table reports the scripts for the four advice topics used in the choice experiment. Participants make four choices in total, one for each advice topic, with topic orders following the experimental design shown in Table 3. Advice is delivered to participants in videos. Each choice set begins with a narrator's introduction, then two advisers provide identical advice (the underlined advice) at the beginning of their talk and then divergent advice at the end (the italicized part). After participants have viewed both advisers' videos, they choose which adviser they would follow, and proceed to the next topic.

| Narrator Introduction | Advice | Narrator Introduction | Advice |
|--|--|--|--|
| <p>Paying Down Debt In this scenario, you have accumulated some large outstanding credit card debt with a high associated interest rate. Recently, you have inherited some money unexpectedly and would like to know what to do with it. The next 2 financial advisers will recommend what you should do.</p> | <p>Good Advice: <u>I understand that you have some large credit card debt but recently inherited money. It is important to think about your overall financial position when making a decision about what to do. It is easy to simply save this big sum of money in a savings account to achieve a savings goal, but the interest gained is far smaller than the high interest expense of not paying down your credit card debt. Therefore, I recommend you pay off your credit card debt to eliminate the high interest charges.</u> Bad Advice: [Insert underlined above] <i>It is hard to save big sums of money so it is important to think about your special savings goals when making this decision. Therefore, I recommend you ignore your credit card debt for now and put your inheritance in a separate savings account.</i></p> | <p>Choosing an Index Fund In this scenario, you are thinking about investing in a managed share index fund. The next 2 financial advisers will recommend what you should do about it.</p> | <p>Good Advice: <u>I understand you need help regarding your choice of share index fund. Did you know that all share index funds invest with the aim of matching the overall share market return? These various share index funds provide an almost identical product so why pay a fund manager more than the others for the same thing.</u> <i>Therefore, I recommend that you choose the share index fund with the lowest management fees.</i> Bad Advice [Insert underlined above] <i>but some fund managers have better reputations than others and you get what you pay for. Therefore, I recommend that you avoid the share index funds with low management fees.</i></p> |
| <p>Consolidating Retirement Accounts In this scenario, suppose you have just changed jobs and started a new superannuation account. Currently, you already have two other superannuation accounts from past jobs. The next 2 financial advisers will recommend what you should do about it.</p> | <p>Good Advice: <u>I see that you have three superannuation accounts with different super funds. Did you know that people are typically charged regular fixed administration fees on all of these superannuation accounts? As a result, I recommend that you roll all of these accounts together so you are not paying extra fees.</u> Bad Advice: [Insert underlined above] <i>Despite that, I recommend that you not roll all of these accounts together so you are diversified across different superannuation funds.</i></p> | <p>Diversifying a Stock Portfolio In this scenario, you are thinking about investing in the share market. The next 2 financial advisers will recommend what you should do about it.</p> | <p>Good Advice: <u>I understand you need help regarding how to invest your superannuation money. Did you know money invested in shares can go up and down? It is good to try to balance out the shares that go up with the shares that go down. Therefore, I recommend that you spread your money across a variety of shares in different types of companies and industries.</u> Bad Advice: [Insert underlined above] <i>That is why it is good to invest in something you know and can easily monitor. Therefore, I recommend that you invest your money in one blue chip company.</i></p> |

Table 2: Variable description

This table reports definitions of variables used in the estimation of choice model (eqn 13) where X_i are vectors of explanatory variables for the components of the model (eqns , consisting of elements shown by “x” in the corresponding column). Variables are computed from responses to an online survey of a representative sample of 2003 Australian adults conducted in December 2014.

| Variable Name | X ₀ | X ₃ | X ₄ | X ₅ | Description |
|------------------------------------|----------------|----------------|----------------|----------------|--|
| Constant | x | x | x | x | Constant; topic specific for X ₅ |
| <i>Adviser characteristics</i> | | | | | |
| Displays NO credential | x | | | | Indicator variable that equals 1 if only adviser’s name was displayed and -1 when “Certified Financial Planner” and adviser’s name was displayed. |
| Price | | x | | | Price in \$ (divided by 100) for one additional hour with this adviser |
| Posterior | | x | | | Posterior belief about adviser after advice on all four topics has been provided – estimated within the model |
| <i>Advice</i> | | | | | |
| Good advice | | | | | Indicator variable that equals 1 if the wrong advice was given in the particular choice set, -1 otherwise. Enters the model via the choice specification in Equation (5) |
| Topic: Account consolidation | | | | x | Indicator variable that equals 1 if the topic was account consolidation, 0 otherwise. |
| Topic: Stock diversification | | | | x | Indicator variable that equals 1 if the topic was stock diversification, 0 otherwise. |
| Topic: Index fund fee | | | | x | Indicator variable that equals 1 if the topic was index fund management fees, 0 otherwise. |
| Topic: Debt repayment | | | | x | Indicator variable that equals 1 if the topic was debt repayment, 0 otherwise. |
| <i>Participant characteristics</i> | | | | | |
| Participant female | | | | x | An indicator variable that equals 2 if the participant is a female, 1 otherwise. |
| Participant older than 39 years | | | | x | An indicator variable that equals 1 if the participant is a older than 39 years, 0 otherwise. |
| Trust in advisers | x | | | | An indicator variable that equals 1 if the participant reported general trust in financial advisers, -1 if distrust, 0 otherwise |
| Paid for advice | | x | | | Indicator variable that equals 1 if the participant has ever paid for financial advice, -1 if they have not |
| Household income | | x | | | Household income (\$’000, mean centered) |
| Confidence in financial decisions | | x | | | Indicator variable that equals 1 if participant has high confidence in their own ability to make financial decision, -1 if low |
| Financial risk tolerance | | x | | | Indicator variable that equals 1 if participant’s risk tolerance is high and -1 if low |
| Decision maker | | x | | | Indicator variable that equals 1 when the participant is most responsible for financial decisions, 0 when jointly responsible and -1 when someone else is responsible. |
| Financial literacy | | | | x | An indicator variable that equals 1 if the participant’s correct percentage on four financial literacy questions is above the sample median, 0 otherwise. Questions test simple interest, inflation, diversification, and compound interest. |
| Numeracy | | | | x | An indicator variable that equals 1 if the participant’s correct percentage on three numeracy questions is above the sample median, 0 otherwise. Questions test fractions, percentages and probabilities. |
| Product knowledge | | | | x | An indicator variable that equals 1 if the participant’s correct percentage on four financial product questions is above the sample median, 0 otherwise. Questions test topics used in |

| | | |
|-------------------|---|---|
| | | advice experiment: debt, index funds, account consolidation, diversification. |
| Conscientiousness | x | An indicator variable that equals 1 if the participant's conscientiousness is above the sample median, 0 otherwise. Participants rated themselves as organized, responsible, hardworking and careless (reverse coded) on a four-point scale. Ratings are averaged. |
| Impulsiveness | x | An indicator variable that equals 1 if the participant's impulsiveness is above the sample median, 0 otherwise. Participants rated themselves as buying too much, buying impulsively, buying without planning, and/or buying unnecessarily on a five point scale. Ratings are averaged. |
| Market experience | x | An indicator variable that equals 1 if the participant's percentage on owning four financial securities is above the sample median, 0 otherwise. Participants reported whether they owned a credit card (debt), units in an index fund (fees), a superannuation account (consolidation) and stocks (diversification). |

Table 3: Experimental design

This table shows the structure of the experiment. Each participant in the experiment makes four choices of financial advice where the design of the four choice sets consists of: one row from Panel A (adviser characteristics); one row from Panel B (sequence of advice topics); and one row from Panel C (sequence of delivery of good or bad advice from Adviser 1 and Adviser 2). Panel A shows the combinations of adviser characteristics: each pair of advisers consisted of an adviser with a triple (gender, age, certification) and an adviser with its reverse. Adviser 1 appeared on the left-hand side of the choice set screen and Adviser 2 appeared on the right-hand side. Each participant saw the same two advisers for the entire experiment and each adviser stayed on the same side of the screen throughout the experiment. Panel B shows the sequence of advice topics for each condition in the experiment where “E” stands for one of the easy topics (debt and account consolidation) and “H” stands for one of the hard topics (fees and diversification). Panel C shows the eight sequences of advice quality for each condition in the experiment where “G” stands for good advice, while “B” stands for bad advice.

Panel A. Design of adviser pairs

| Pair | Adviser 1 | | | Adviser 2 | | |
|------|-----------|-------|---------------|-----------|-------|---------------|
| | Gender | Age | Certification | Gender | Age | Certification |
| 1 | Female | Young | Yes | Male | Old | No |
| 2 | Female | Old | No | Male | Young | Yes |
| 3 | Male | Young | No | Female | Old | Yes |
| 4 | Male | Old | Yes | Female | Young | No |

Panel B. Sequence of advice topics

| Sequence | Choice 1 | Choice 2 | Choice 3 | Choice 4 | Clarity |
|----------|-----------------|-----------------|-----------------|----------|---------|
| 1 | Diversification | Fees | Consolidation | Debt | HHEE |
| 2 | Consolidation | Debt | Diversification | Fees | EEHH |
| 3 | Diversification | Consolidation | Fees | Debt | HEHE |
| 4 | Consolidation | Diversification | Debt | Fees | EHEH |
| 5 | Diversification | Consolidation | Debt | Fees | HEEH |
| 6 | Consolidation | Diversification | Fees | Debt | EHHE |

Panel C. Design of the sequence of advice quality

| Quality Sequence | Advice from adviser 1 | | | | Advice from adviser 2 | | | |
|------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | 1 st topic | 2 nd topic | 3 rd topic | 4 th topic | 1 st topic | 2 nd topic | 3 rd topic | 4 th topic |
| 1 | G | G | B | B | B | B | G | G |
| 2 | G | B | G | B | B | G | B | G |
| 3 | G | B | B | G | B | G | G | B |
| 4 | B | G | G | B | G | B | B | G |
| 5 | B | G | B | G | G | B | G | B |
| 6 | B | B | G | G | G | G | B | B |

Table 4: Empirical Results

This table reports statistics from the posterior distributions of estimated parameters of the choice model (eqn 13). Data are survey responses of 2003 participants collected in December 2014. Variables are defined in Table 2. For each parameter, we report the mode of its posterior distribution as well as the 2.5 and 97.5 percentiles of this distribution, i.e., the equi-tailed credible interval (CI). There is a 95% probability that the parameter is not zero if zero does not fall in the CI. Estimation was conducted using SABL; see Appendix B for details.

| | Mode | 2.5 Percentile | 97.5 Percentile |
|--|--------|----------------|-----------------|
| <i>Prior belief about adviser, Eq. (4)</i> | | | |
| Trust in financial advisers | 0.520 | 0.421 | 0.610 |
| Displays NO credential | -0.085 | -0.199 | -0.009 |
| Constant | 1.728 | 1.511 | 1.903 |
| <i>Choice of Advice, Eqs. (5) & (6)</i> | | | |
| Quality (β_1) | 4.296 | 3.599 | 5.060 |
| Posterior belief (β_2) | 2.510 | 1.494 | 3.622 |
| <i>Willingness to pay, Eq. (7)</i> | | | |
| Constant | -7.782 | -9.687 | -6.228 |
| Price | -0.085 | -0.124 | -0.043 |
| Posterior | 18.309 | 14.808 | 22.230 |
| Paid for advice | 0.466 | 0.348 | 0.570 |
| Household income | 0.094 | -0.021 | 0.163 |
| Confidence in financial decisions | -0.088 | -0.186 | 0.051 |
| Financial risk tolerance | 0.055 | -0.047 | 0.156 |
| Decision maker | 0.034 | -0.125 | 0.186 |
| <i>Rational vs FHJ updating, Eq. (9)</i> | | | |
| Constant | -0.454 | -0.994 | -0.185 |
| High Conscientiousness | 0.243 | -0.046 | 0.485 |
| High Impulsiveness | -0.344 | -0.724 | -0.154 |
| <i>Clarity of Topics, Eq. (10)</i> | | | |
| High Market Experience | 0.073 | -0.046 | 0.151 |
| High Product Knowledge | 0.267 | 0.171 | 0.358 |
| Participant older than 39 | 0.554 | 0.441 | 0.646 |
| Participant female | 0.138 | 0.053 | 0.237 |
| High Financial Literacy | 0.372 | 0.244 | 0.458 |
| High Numeracy | 0.357 | 0.278 | 0.482 |
| Consolidation | 1.405 | 1.148 | 1.632 |
| Diversification | 0.615 | 0.395 | 0.814 |
| Fees | -0.545 | -0.794 | -0.358 |
| Debt | 1.768 | 1.511 | 1.995 |
| <i>Estimated % of participants using rational learning: 37.11; limited memory learning: 62.89.</i> | | | |

Table 5: Proportion of participants in latent classes of clear and ambiguous advice topics.

This table shows estimated posterior proportion of participants assigned to 16 latent classes differentiated by clarity or ambiguity of the four advice topics: account consolidation (column 1), diversification (column 2), index fund fees (column 3), and debt (column 4). A “1” indicates that participants in that class treated the topic as clear and “0” indicates that they treat the topic as ambiguous. For example, the model assigns 18.2% of participants to latent class 1 (row 1) that treats all topics as clear, and assigns 3.8% of participants to latent class 16 (row 16) that treats all topics as ambiguous. We infer latent classes from estimation of the choice model (eqn 13) – see Table 4 for estimation results.

| Latent class | Consolidation | Diversification | Fees | Debt | Segment Size |
|--------------|---------------|-----------------|------|------|--------------|
| 1 | 1 | 1 | 1 | 1 | 18.2 |
| 2 | 1 | 1 | 1 | 0 | 2.2 |
| 3 | 1 | 1 | 0 | 1 | 21.9 |
| 4 | 1 | 1 | 0 | 0 | 4.5 |
| 5 | 1 | 0 | 1 | 1 | 6.9 |
| 6 | 1 | 0 | 1 | 0 | 1.4 |
| 7 | 1 | 0 | 0 | 1 | 14.0 |
| 8 | 1 | 0 | 0 | 0 | 4.8 |
| 9 | 0 | 1 | 1 | 1 | 3.1 |
| 10 | 0 | 1 | 1 | 0 | 0.6 |
| 11 | 0 | 1 | 0 | 1 | 6.4 |
| 12 | 0 | 1 | 0 | 0 | 2.3 |
| 13 | 0 | 0 | 1 | 1 | 2.0 |
| 14 | 0 | 0 | 1 | 0 | 0.7 |
| 15 | 0 | 0 | 0 | 1 | 7.2 |
| 16 | 0 | 0 | 0 | 0 | 3.8 |

Table 6: Evolution of beliefs with one clear and three ambiguous topics.

This table compares the effects of first impressions on subsequent choices when clients use either rational (standard Bayesian) or FHJ (limited memory) to update beliefs about adviser quality. The example assumes client A uses rational updating and client B uses FHJ updating, that both clients are initially distrusting of financial advisers, and that otherwise both clients have characteristics at the medians of the sample distributions. Parameters are set to the modes of the posterior distributions. Adviser R shows a professional certification and Adviser L does not. Both participants thus have the same prior beliefs that the right (R) and the left (L) adviser are of good quality, λ_0 . Adviser R delivers good advice on a clear topic at choice 1 but topics 2-4 are ambiguous to both clients. Both clients update their beliefs in the same way at the first choice because they get clear information about adviser quality λ_1 . Client A's beliefs about the advisers ($\lambda_2 - \lambda_4$), and choice probabilities, $\Pr(y_2 = 1)$ to $\Pr(y_4 = 1)$, remain constant because the rational client does not update using ambiguous signals. Client B treats ambiguous information as evidence in favor of his or her priors and continues to update in favor of Adviser R.

| | λ_0 | $\Pr(y_1 = 1)$ | λ_1 | $\Pr(y_2 = 1)$ | $\Pr(y_3 = 1)$ | λ_3 | $\Pr(y_4 = 1)$ | λ_4 | |
|------------------------|-------------|----------------|-------------|----------------|----------------|-------------|----------------|-------------|-------|
| Adviser R, client A | 0.785 | 0.987 | 0.916 | 0.886 | 0.916 | 0.886 | 0.916 | 0.886 | 0.916 |
| Adviser R, client B | 0.785 | 0.987 | 0.916 | 0.886 | 0.970 | 0.913 | 0.990 | 0.921 | 0.997 |
| Adviser L, client A | 0.755 | 0.013 | 0.098 | 0.114 | 0.098 | 0.114 | 0.098 | 0.114 | 0.098 |
| Adviser L, client B | 0.755 | 0.013 | 0.098 | 0.114 | 0.035 | 0.087 | 0.012 | 0.079 | 0.004 |

$\lambda_i =$ prior belief about adviser quality at choice set i ;

$\Pr(y_i = 1) =$ probability of choosing to follow advice of adviser at choice set i

Figure 1: Simulation of belief updating

This figure illustrates all possible paths of beliefs about the probability that an adviser is of good quality for experimental “clients” under different assumptions about initial prior probabilities of good quality and updating strategies (in rows). Larger dots indicate a higher proportion of all possible updating paths that pass through each point. We set the probability that a good adviser offers a good signal, s , to 0.75. Rows 1, 3 and 4 reflect limited memory (FHJ or “Fryer”) updating with starting priors λ_0 of 0.6, 0.5, and 0.4, respectively. Row 2 reflects rational updating where prior $\lambda_0 = 0.5$. Columns 1-5 show belief paths where 4, 3, 2, 1, and 0 advice topics are clear to clients. At column 1 (all topics clear), difference between paths come from different sequences of good and bad advice. As more topics become ambiguous difference between paths additionally come from different sequences of clear and ambiguous signals. FHJ updating leads to increasing polarization. Rational clients do not update their priors if all signals are ambiguous (column 5, row 2). FHJ updaters also do not update when all signals are ambiguous and their starting prior is equal to 0.5 (row 3, column 5).

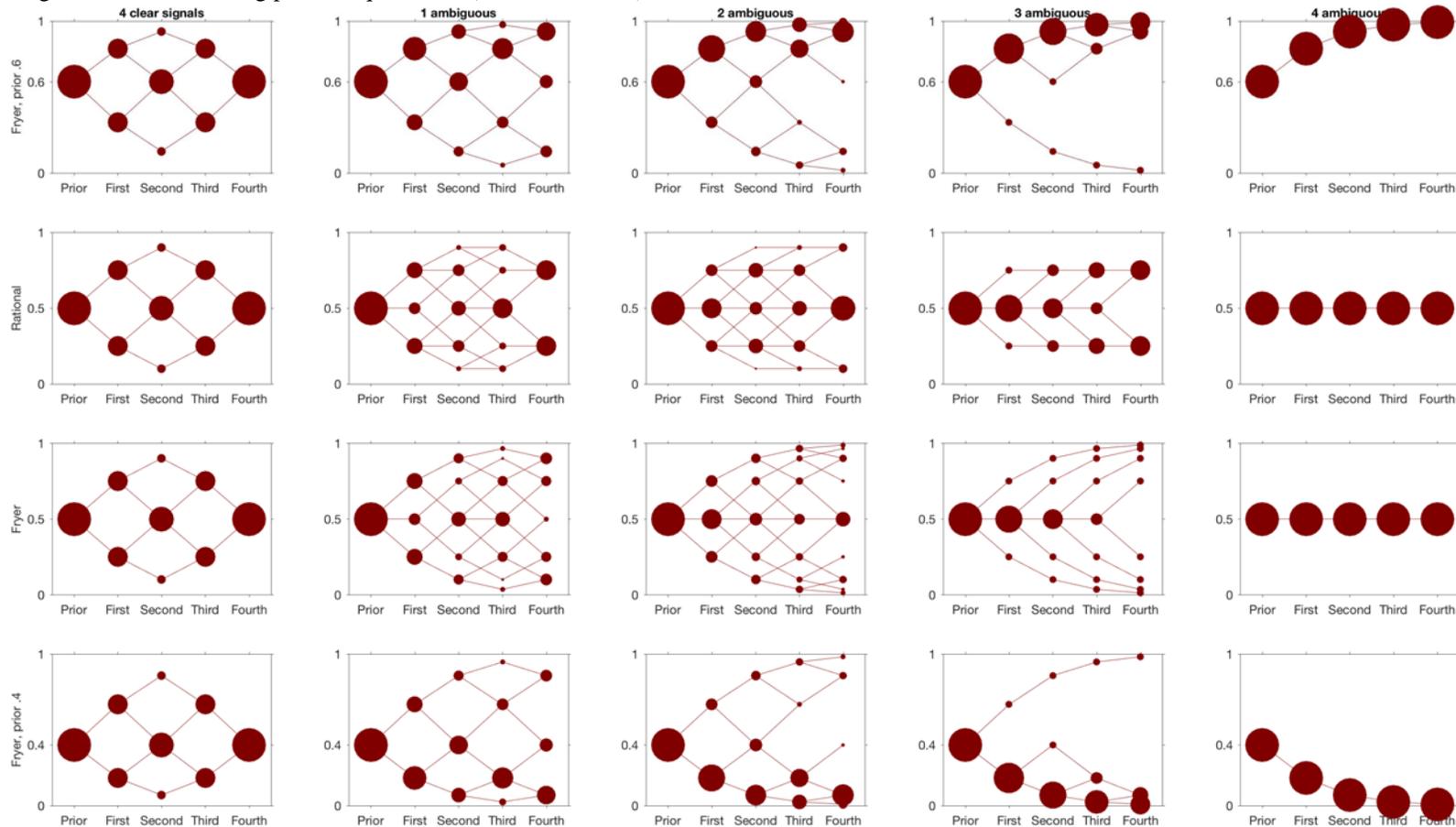


Figure 2: Discontinuities in likelihood function under FHJ updating

This figure shows proportion of theoretical belief paths reaching updated beliefs, λ_4 , after four signals from advisers to experimental clients who use either rational or FHJ (“Fryer”) updating. Rows show simulations for three arbitrary values of the signal strength parameter s , that is, the probability that a good adviser delivers good advice. For each value of the initial prior belief that an adviser is good, λ_0 (horizontal axes), we draw all possible values of posterior beliefs that an adviser is good after four signals, λ_4 (vertical axes), supported by the model. Paths vary by the pattern of clear or ambiguous signals delivered by the adviser. Larger circles indicate that more possible paths lead to any specific value of λ_4 . For example, in row 1 and column 2, at a signal strength of $s=0.75$ and where all participants use FHJ updating, as λ_0 moves from 0.49 to 0.51, λ_4 jumps from mostly less than 0.1, to mostly larger than 0.8, creating a discontinuous likelihood function.

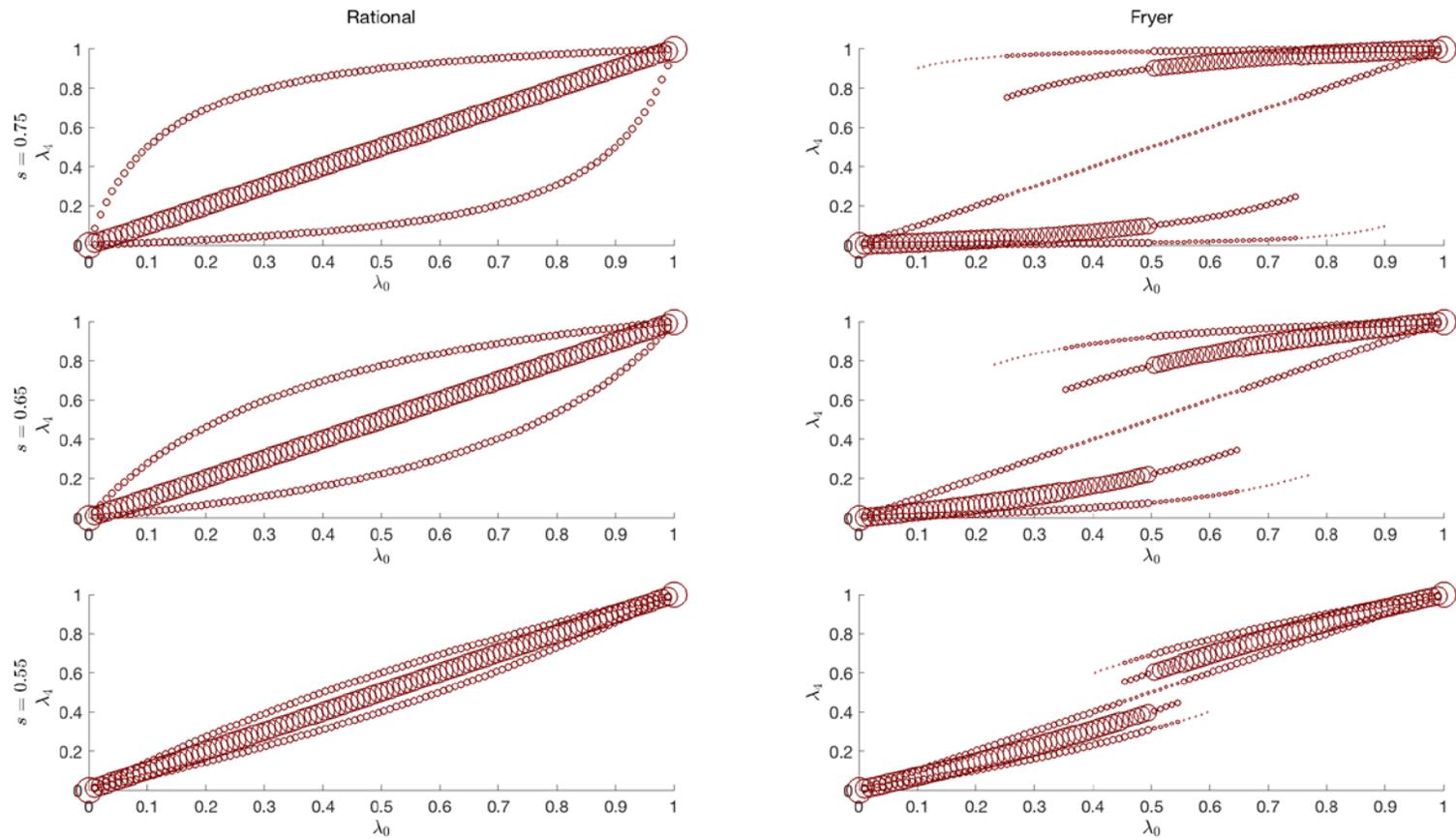
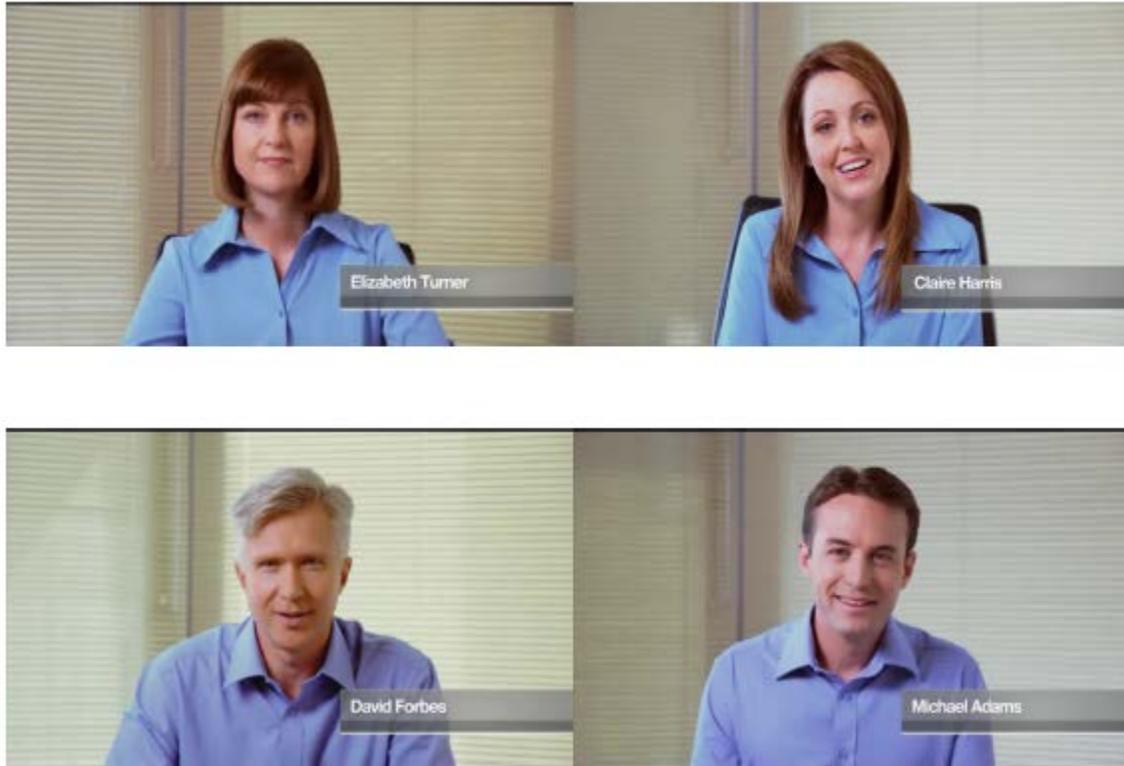


Figure 3: Advisers

This figure shows screen shots of four “advisers” who delivered video advice in the experiment. Each participant in the experiment viewed advice delivered by two of the four advisers as matched pairs of gender, age and certification opposites (e.g., young, male, certified v. older, female, not certified).



Appendix A: Demographics – comparisons between survey sample and Australian population (18-79 years), 2011 Census.

This table compares demographics of sample of 2003 participants drawn from a nationally representative online panel by email invitation in 2014 with 2011 (most recent) Australian census.

| | Survey Participant Sample | 18-79 yrs Australian Population | | Survey Participant Sample | 18-79 yrs Australian Population |
|-------------------------|---------------------------------|---------------------------------------|--|---------------------------------|---------------------------------------|
| Gender | | | Marital Status | | |
| Male | 50% | 49% | Never Married | 26% | 30% |
| Female | 50% | 51% | Divorced/ Separated | 10% | 13% |
| Age | | | Widowed | 2% | 3% |
| 18-24 years | 8% | 10% | Married or long term relationship | 62% | 54% |
| 25-29 years | 8% | 10% | Personal Income | | |
| 30-34 years | 12% | 10% | \$1-\$20,799 (i.e. less than \$399 a week) | 24% | 25% |
| 35-39 years | 12% | 10% | \$20,800-\$51,999 (i.e. \$400-\$999 a week) | 35% | 32% |
| 40-44 years | 12% | 10% | \$52,000-\$103,999 (i.e. \$1,000-\$1,999 a week) | 25% | 23% |
| 45-49 years | 9% | 10% | \$101,000 (i.e. \$2,000 a week) or more | 7% | 7% |
| 50-54 years | 12% | 10% | Negative or Nil Income | 9% | 6% |
| 55-59 years | 12% | 9% | Not Started | 0% | 7% |
| 60-64 years | 13% | 8% | Highest Level of Education | | |
| 65-69 years | 2% | 6% | High School or Less | 26% | 40% |
| 70-79 years | 0% | 8% | Vocational/Technical certificate | 21% | 20% |
| Work Status | | | Tertiary diploma | 11% | 9% |
| Employed | 62% | 63% | Bachelor degree | 23% | 15% |
| Unemployed | 8% | 3% | Graduate certificate, diploma or degree | 19% | 6% |
| Not in the labour force | 18% | 29% | Not stated | 0% | 10% |
| Retired | 12% | not broken out | | | |
| Not stated | 0% | 5% | | | |

Appendix B: Sequential Adaptive Bayesian Learning Estimation

The discontinuities in the likelihood function of our model cause implementation problems due to the inherent computational difficulty for maximum likelihood estimators (see also Chernozhukov and Hong 2004). We overcome these difficulties with Bayesian estimation methods. More specifically, we use Sequential Adaptive Bayesian Learning (SABL) proposed by Durham and Geweke (2014). SABL is an extension of sequential Monte Carlo methods that additionally exploits the benefits of parallel computing environments. SABL does not require the modeler to specify conjugate priors and it is also robust to multimodal posteriors which can arise in high dimensional problems (Jasra et al. 2007) such as ours. When used for Bayesian inference, SABL is a posterior simulator. Our interest only lies in the latter; thus we focus the following basic description of SABL on this while also ignoring the aspects that make SABL an efficient tool to address very complex problems.

As with any Bayesian estimation approach SABL requires the user to specify the likelihood function $l(\theta)$ as well as prior distributions $p^{(0)}(\theta)$ for the parameters to be estimated. SABL then produces draws from the posterior $p^*(\theta)$ as follows:

- Draw parameters from the prior distributions. To do this SABL represents initial information by $\theta_{gh}^{(0)} \sim_{iid} p^{(0)}(\theta)$, organized into H groups¹⁰ of G draws each (SABL defaults to $H=16$ and $G=192$). Let $p^{(0)}(\theta)$ be a very flat version of the likelihood function, i.e. $p^{(0)}(\theta) = l(\theta)^{r_0}$ with r_0 very small.
- For a sequence of cycles $n=1, 2, \dots$

¹⁰ SABL organizes the draws into groups to exploit the parallel processing possibilities of the algorithm.

- a. Correction (C) phase: Determine $p^{(n)}(\theta)$ by raising the likelihood function to a higher power, i.e. $p^{(n)}(\theta) = l(\theta)^{r_n}$ with $r_n > r_{n-1}$. Calculate for each draw a weight $w^{(n)}(\theta_{gh}^{(n-1)}) = p^{(n)}(\theta_{gh}^{(n-1)}) / p^{(n-1)}(\theta_{gh}^{(n-1)})$, $h=1, \dots, H, g=1, \dots, G$;
- b. Selection (S) phase, applied independently to each group $h=1, \dots, H$: Use multinomial residual resampling (e.g. Douc and Cappé 2005) based on $\{w^{(n)}(\theta_{gh}), g=1, \dots, G\}$ to select $\{\theta_{gh}^{(n,0)}, g=1, \dots, G\}$ out of $\{\theta_{gh}^{(n-1)}, i=1, \dots, I\}$.
- c. Mutation (M) phase, applied independently to each group $h=1, \dots, H$: The M phase is a Metropolis random walk. In each step o ($o > 0$) of the random walk obtain for each $g=1, \dots, G$ a proposal $\theta_{gh}^{(n,o)*}$ is drawn from $N(\theta_{gh}^{(n,o-1)}, \Sigma^{(n,o-1)})$, where $\Sigma^{(n,o-1)}$ is proportional to the sample variance of the particles $\{\theta_{gh}^{(n,o)}, g=1, \dots, G\}$. Accept $\theta_{gh}^{(n,o)*}$ with probability α where α is defined as
- $$\alpha = \min\{1, p^{(0)}(\theta_{gh}^{(n,o)*})l(\theta_{gh}^{(n,o)*}) / p^{(0)}(\theta_{gh}^{(n,o-1)})l(\theta_{gh}^{(n,o-1)})\}$$
- and set $\theta_{gh}^{(n,o)} = \theta_{gh}^{(n,o)*}$, otherwise set $\theta_{gh}^{(n,o)} = \theta_{gh}^{(n,o-1)}$. The proportionality factor is thereby increased when the rate of accepting the proposal draws is higher than a particular threshold (the default in SABL is 0.25), and decreases otherwise. The random walk terminates once the dependence among particles has been sufficiently broken, that is when the particles are sufficiently independent (note that the S-phase introduces dependence via

repeated sampling of the same $\theta_{gh}^{(n-1)}$). SABL assumes sufficient independence of the particles when the variance (calculated across the H group means) falls below a certain threshold. The last set of $\theta_{gh}^{(n,o)}$ is then denoted as $\theta_{gh}^{(n)}$.

- If $p^{(n)}(\theta) = l(\theta)$ then $N=n$ and the algorithm terminates with draws $\theta_{gh}^{(N)}$ from the posterior distribution $p^*(\theta)$.

We chose uninformative priors. We assumed that the prior for each parameter of interest is independent normal with a mean of zero and a standard deviation of five. We evaluated the sensitivity of prior influence by a careful visual examination of the posterior distribution against the prior distribution.

The advantage of using SABL (or a Bayesian approach in general) is that the posterior distribution of draws can help in assessing the identification of the model parameters (see also discussion in the previous section). More specifically, a high correlation between the posterior draws of two parameters may suggest that these are not separately identified by the choice data. In addition to including different covariates in the different model parts (see also discussion in section 4) and specifying different uninformative priors, we used this correlation matrix check to further assess the identification of our model.¹¹

¹¹ We run SABL using its MATLAB interface. SABL itself can be downloaded from http://www.quantosanalytics.org/garland/mp-sps_1.1.zip. The time to estimate our model using SABL is approximately 60 minutes.