DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Coarse Thinking and Persuasion

Citation
Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer. 2008. Coarse thinking and persuasion. Quarterly Journal of Economics 123, no. 2: 577-619

Published Version
http://dx.doi.org/10.1162/qjec.2008.123.2.577

Permanent link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:11022284

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

COARSE THINKING AND PERSUASION[1]

Sendhil Mullainathan

Joshua Schwartzstein

Andrei Shleifer

Revised October 2007

Abstract

We present a model of uninformative persuasion in which individuals "think coarsely": they group situations into categories, and apply the same model of inference to all situations within a category. Coarse thinking exhibits two features that persuaders take advantage of: (i) transference, whereby individuals transfer the informational content of a given message from situations in a category where it is useful to those where it is not, and (ii) framing, whereby objectively useless information influences individuals' choice of category. The model sheds light on uninformative advertising and product branding, as well as on some otherwise anomalous evidence on mutual fund advertising.

[1] This paper replaces an earlier draft with the same title, as well as "Persuasion in Finance" by Mullainathan and Shleifer. We are grateful to Nicholas Barberis, Gary Becker, Dan Benjamin, Daniel Bergstresser, Olivier Blanchard, Lauren Cohen, Stefano DellaVigna, Daniel Gilbert, Edward Glaeser, Xavier Gabaix, Matthew Gentzkow, Simon Gervais, Robin Greenwood, Richard Holden, Emir Kamenica, Lawrence Katz, Elizabeth Kensinger, David Laibson, Owen Lamont, Steven Levitt, Ulrike Malmendier, Richard Posner, Andrew Postlewaite, Matthew Rabin, Christina Romer, Jesse Shapiro, Jeremy Stein, Rene Stulz, Richard Thaler, Robert Waldmann, Glen Weyl, Gerald Zaltman, Eric Zitzewitz, three anonymous referees, and especially Nicola Gennaioli, Giacomo Ponzetto, and the fourth anonymous referee for helpful comments. We also thank Michael Gottfried, Tim Ganser, and Georgy Egorov for excellent research assistance. Schwartzstein acknowledges financial support from an NSF graduate fellowship. All errors remain our own.

1. Introduction

Most societies devote huge resources to persuasion (McCloskey and Klamer 1995). Selling, advertising, political campaigns, organized religion, law, much of the media, and some education are devoted to changing beliefs in a way advantageous to the persuader. Persuasion is not simply an expenditure of resources: the content of the message crucially shapes its effectiveness[2]. But what constitutes persuasive content?

Economists usually assume that only one type of persuasive content matters: objectively useful information. Stigler (1987, p. 243) defines advertising as "the provision of information about the availability and quality of a commodity." Economists typically model persuasion, including advertising (Stigler 1961), political campaigns (Downs 1957), or legal argument (Milgrom and Roberts 1986b; Dewatripont and Tirole 1999) as provision of information. In some models, such as Crawford and Sobel (1982), Glazer and Rubinstein (2001, 2004), Grossman (1981), Grossman and Hart (1980), Milgrom (1981), Okuno-Fujiwara, Postlewaite, and Suzumura (1990), and Holden and Kamenica (2006), the persuader uses information strategically, but conveys information nonetheless.

Psychologists and marketers understand persuasion quite differently. They argue that people evaluate various propositions or objects using representativeness, metaphors, analogies, and more generally associative strategies (Gilovich 1981; Edelman 1982; Kahneman and Tversky 1982; Lakoff 1987; Zaltman 1997). The strategy of persuasion is to take advantage of these mental strategies, which we refer to as coarse thinking, to improve the audience's assessment of the persuader's issue or product.

[2] A vast advertising literature makes this point (see, e.g., Zaltman 1997; Sutherland and Sylvester 2000). Bertrand et al. (2006) present some clear evidence that persuasive content matters in a field experiment using loan advertisements by a South African consumer lending institution.

In this paper, we present a model of coarse thinking and persuasion. We distinguish two ways in which persuaders, such as product advertisers, can take advantage of coarse thinking. First, the audience might already have some analogy or model of this product in mind; it already thinks of this product in terms of something else. In this case, one way to persuade is to advertise attributes of the product that are positively related to quality in the analogous situation. The coarse thinker transfers the informational content of these attributes across analogous situations, and so improves his view of the product. We call this form of coarse thinking transference.

Second, and more fundamentally, persuaders may themselves try to shape or create the relevant analogy by advertising attributes associated with that desired analogy. This method persuades successfully when it changes the lens through which a thinker views all features of the product. Following Goffman (1974), we call this form of coarse thinking "framing." In many instances, successful persuasion takes advantage of both transference and framing, but these forms of coarse thinking are conceptually distinct[3].

To illustrate these ideas, consider several examples:

Alberto Culver Natural Silk Shampoo was advertised with a slogan "We put silk in the bottle." The shampoo actually contained some silk. During the campaign, the company spokesman conceded that "silk doesn't really do anything for hair" (Carpenter, Glazer and Nakamoto 1994).[4] This example is a relatively pure case of the persuader relying on transference. Culver takes advantage of the co-categorization of shampoo with hair, which leads consumers to value "silk" in shampoo (not very sensible) because

[3] Recently, the term framing has been used much more narrowly to describe the coding of gains and losses in prospect theory (Kahneman and Tversky 1979). Goffman's original and broader meaning of framing as a lens or "model" through which the audience interprets data is what we try to capture below.

[4] A recent reincarnation of this marketing idea, discovered by inspecting products in a drugstore, is *Pure Cashmere Softsoap*, which contains "cashmere extract." Cashmere adds quality to sweaters not soap.

they value "silky" in hair. By adding silk to the bottle, Culver effectively transfers a positive trait from hair to shampoo to make its product more attractive.

For about forty years, Avis Car Rental advertised itself with a slogan "We are number two. We try harder." Besides this key message, most ads contain little data. We think of this campaign as a relatively pure case of framing. When Avis was getting started, many of its potential customers knew that it was smaller than Hertz. How the attribute "being second" is interpreted depends on categorization: either negatively as a loser or positively as an underdog. Since the underdog image favors Avis, the campaign primed this frame by stressing the attributes of underdogs (they try harder). The fact that Avis lagged behind Hertz in sales became a sign of higher rather than lower quality.

Most persuasive messages take advantage of both transference and framing. Take two examples, one from the economic and one from the political sphere.

Over the course of the internet stock market bubble (1994-2003), the brokerage firm Merrill Lynch ran six advertising campaigns, respectively called "a tradition of trust", "the difference is Merrill Lynch," "human achievement," "be bullish," "ask Merrill," and "total Merrill." The motto of each campaign always appeared in the ad. Roughly speaking, the first two campaigns preceded the bubble, the third and the fourth appeared during it, and the last two ran after the sharp market decline.

One way to compare these campaigns is to look at a representative ad from each. "A tradition of trust" ads often portray a grandfather and a grandson fishing together. The ads talk about slow accumulation of wealth and Merrill's expertise. The activities of fishing and, even more so, teaching to fish suggest slowness, tradition, skill, consistency, and patience. The ads advise on how to protect oneself and one's family financially. Ads

from “the difference is Merrill Lynch” campaign likewise show grandfathers and grandsons with fishing rods, and recommend saving.

This message changes in 1999. A “human achievement” ad from 1999 shows a 12 year old girl wearing a helmet and carrying a skate board. The image is much hipper than those from the previous campaigns. The next campaign, from 2000-2001, simply intones: “be bullish”. One ad shows a Merrill Lynch bull wired as a semiconductor board (the word “wired” itself has two meanings, connected and hyperactive). The theme of protection is gone; growth and opportunity emerge, unfettered by data.

After the market declines, Merrill switches to the “ask Merrill” campaign, with its emphasis on uncertainty in the world, and the company’s expertise in protecting and advising its customers. A representative ad is dominated by a page-size question mark, invoking insecurity, uncertainty, and the need for answers. Finally, by the end of the decade, the firm moves to the “total Merrill” campaign, with its familiar emphasis on expertise and intergenerational fishing.

Merrill’s campaigns take advantage of both transference and framing. In quieter markets, Merrill seeks to frame or position itself as an expert and adviser, much like a doctor or a grandfather teaching his grandson to fish. It then wants the audience to transfer the positive attributes of this expert to enhance the perceived value of its own advice. At the peak of the bubble, Merrill recognizes that many investors are seeking to get rich quickly, to grab the opportunities created by the technology boom. It then frames itself as the agent of such opportunities, and wants the audience to transfer the positive attributes of “technology” to its own services. After the crash of the bubble, the patient advisor again becomes an attractive frame.

As a final example – from the area of political persuasion -- consider Arnold Schwarzenegger's memorable speech at the 2004 Republican National Convention. In the best remembered part of his speech, Schwarzenegger defended free trade: "To those critics who are so pessimistic about our economy, I say: Don't be economic girlie men!... Now they say India and China are overtaking us. Don't you believe it. We may hit a few bumps – but America always moves ahead. That's what Americans do."

Schwarzenegger's speech takes advantage of both framing and transference. He frames international trade as war, with winners or losers, not as an economic model in which everyone gains. With this frame in mind, he transfers America's as well as his own (a former champion as well as The Terminator) propensity to win onto trade. This message of victory, while helpful for assessing US military engagements and Schwarzenegger's own accomplishments, is of limited value in evaluating globalization.

We will come back to these examples as we discuss our analytical results. In our model, individuals deviate from Bayesian rationality in two crucial ways: first, they group situations into categories based on the data they receive and, second, they fail to differentiate between co-categorized situations and use one model of inference for all situations in the same category. Such coarse thinking allows the persuader to create uninformative messages that frame the interpretation of public information (laggard in car rentals) through category choice (underdog). Coarse thinking also allows the persuader to take advantage of transference by creating messages (we will win) that are uninformative in the relevant situation (free trade) but still induce a reaction because they are informative in the co-categorized one (war).

The model presented in this paper shares many elements with Crawford and Sobel's (1982) model of strategic information transmission. Indeed, our model of transference is mathematically similar to an extension of Crawford and Sobel in which the audience does not know the exact situation (the underlying game being played), and so may react to messages that do not contain decision-relevant information in the relevant situation.[5] On the other hand, framing through category choice does not have a natural counterpart in the cheap talk literature. Even if we focus on transference, our interpretation of the underlying mathematics is very different. We suppose that the audience knows the situation it is in, but uses one model of inference to interpret information in multiple situations. We do not think that the audience reacts to the message "contains silk" because it is uncertain whether adding silk improves the shampoo. Rather, we suggest that the audience thinks coarsely and reacts to the message "contains silk" because it interprets messages about shampoo and hair similarly.

Associational or analogical thinking is both extremely common, and extremely useful in everyday life, because it reduces the evaluation of new situations to comparison with familiar ones. Edelman (1992) thinks that, for this reason, our brains have evolved so as to make metaphor and analogy standard hard-wired forms of reasoning. Of course, the patterns of thought that are *usually* extremely helpful are not *always* so. Persuaders

[5] Crawford (2003) presents a cheap talk model in which uncertainty surrounding the Sender's type (e.g., whether or not he is strategic) may enable a strategic Sender to "fool" a sophisticated Receiver into taking a suboptimal action in equilibrium. Kartik, Ottaviani, and Squintani (2007) present a related model that incorporates the possibility of equilibrium deception or misinterpretation of information. They extend the basic cheap talk model by assuming that a fraction of the Sender's audience misinterprets equilibrium messages with some non-equilibrium based rule (e.g., they always blindly believe the Sender's recommendation). In the equilibria they identify, the Sender always sends an inflated message that deceives the "naïve" agents.

take advantage of people utilizing a strategy which, though generally useful, may not be useful in the situation of interest to the persuader.[6]

A large literature looking at persuasion deals with advertising. Nelson (1974) broadens the range of what might be seen as informative advertising.[7] Stigler and Becker (1977) and Becker and Murphy (1993) put advertising into the utility function. Gabaix and Laibson (2006) and Shapiro (2006) offer behavioral models of advertising. Recent research on persuasion has gone beyond advertising, and includes studies of hatred (Glaeser 2005), media (Mullainathan and Shleifer 2002, 2005; Gentzkow and Shapiro 2006; DellaVigna and Kaplan 2007), and political persuasion (Becker 2001; Murphy and Shleifer 2004; Glaeser, Ponzetto, and Shapiro 2005). As far as we know, our paper is the first to study persuasion in a model of associative thinking so central to psychological work.

The next section outlines our model of coarse thinking, and compares Bayesian and coarse decision makers. Section 3 presents the results on persuasion, showing how the persuader takes advantage of transference and framing. Section 4 uses the model to understand a crucial aspect of marketing, namely product branding. Section 5 applies the ideas of the model to the case of mutual fund advertising, and presents some evidence on such advertising during the internet bubble. Section 6 concludes.

[6] Economic theory has also considered analogical reasoning in thinking both about how individuals forecast the payoffs to different actions under uncertainty (Gilboa and Schmeidler 1995) and about how they forecast opponents' strategies in game theoretic environments (Eyster and Rabin 2005; Jehiel 2005). Ettinger and Jehiel (2007) apply Jehiel's equilibrium concept to a model similar to Crawford's (2003) and show how a strategic Sender can exploit the fact that a Receiver who thinks coarsely about strategies may misinterpret the Sender's actions in equilibrium.

[7] Research following Nelson (1974) (e.g., Kihlstrom and Riordan 1984; Milgrom and Roberts 1986a) focuses on how, in equilibrium, the amount of advertising may signal quality. This literature does not make predictions regarding the equilibrium content of advertisements.

2. Model

Basic setup

An individual must assess the quality of a given object, such as a shampoo, a mutual fund, or a political candidate. We denote this underlying quality of the object by $q \in Q$, where Q is some subset of $\mathbb{R}$. In assessing the quality of the object, the individual faces one of three similar, but not identical, observable situations $s \in \{0,1,2\}$. For instance, $s = 0$ could be "selecting a mutual fund", $s = 1$ could be "selecting a professional service", and $s = 2$ could be "grabbing an opportunity". Our analysis pertains to the assessment of quality in situation $s = 0$.

The individual observes a piece of public information $r \in \{u,d\}$ sent by nature that is potentially informative about q. The public information is meant to capture any piece of data available to the audience that cannot be controlled by a persuader. For instance, r could stand for past stock market performance and u and d could stand for up and down respectively. Later, the individual receives a potentially informative message $m \in \{a,b\}$ about q. In situation $s = 0$, this message is sent by a persuader who privately observes signal $x \in \{a,b\}$ prior to sending m, but in situations $s = 1$ and $s = 2$ nature sends signal $m = x$ directly. The individual then uses (r,m) to form expectations about underlying quality.

The timeline of the individual's decision problem in $s = 0$ is given in Figure 1.

Bayesian Thinking

The underlying joint distribution function over quality, public information, private signals, and situations is $p(q,r,x,s)$ and is common knowledge. We assume that the induced joint probability mass function of public information and private signals

conditional on situations, $p(r,x \mid s)$, satisfies $p(r,x \mid s) > 0$ for each *r, x*, and *s*. We also assume that the marginal probability mass of situations, $p(s)$, satisfies $p(s=0)+p(s=1)+p(s=2)=1$ and $p(s)>0$ for each *s*.

The individual has beliefs about the probability mass of messages conditional on public information, private signals, and situations (i.e., beliefs about the strategy of the persuader and of nature), $\hat{\sigma}(m \mid r,x,s)$. These beliefs, combined with *p*, generate a joint probability distribution over quality, public information, private signals, messages and situations, $\hat{p}(q,r,x,m,s)$.

Prior to making some decision in *s = 0*, the individual uses (r,m) to form an expectation of the underlying quality of the object. Since the situation is observable, the updated Bayesian expectation of quality is

$$E[q \mid r,m,s=0] \quad = \quad \sum_{x' \in \{a,b\}} E[q \mid r,x',s=0]\hat{p}(x' \mid r,m,s=0), \tag{2.1}$$

where $E[q \mid r,x,s=0]$ is calculated using *p* and $\hat{p}(x \mid r,m,s=0)$ is uniquely derived from *p* and $\hat{\sigma}$ using Bayes' rule whenever possible (otherwise it is an arbitrary distribution).

To study the efficacy of persuasion, it is useful to examine the marginal effect of the message sent by the persuader on the individual's assessment of quality conditional on public information *r*. We define the Bayesian's *reaction* to *m* to be the difference between the Bayesian's expectation of quality prior to receiving a message from the persuader and his expectation of quality after receiving the message. Hence, the Bayesian's reaction, $E[q \mid r,m,s=0]-E[q \mid r,s=0]$, is given by

$$\left(\hat{p}(a \mid r,m,s=0)-p(a \mid r,s=0)\right)\left(E[q \mid r,a,s=0]-E[q \mid r,b,s=0]\right) \tag{2.2}$$

As illustrated in equation (2.2), the Bayesian's reaction to *m* is the product of two terms: (i) the revision in the probability placed on the persuader's private signal being $x = a$ as a result of the message and (ii) the extent to which the conditional expectation of quality is different under private signal $x = a$ than under $x = b$. Thus, the Bayesian only reacts to a message when he believes both that it is informative about the persuader's private signal and that the private signal is predictive of quality.

When $E[q \mid r, m, s = 0] \neq E[q \mid r, s = 0]$, we say that *message m* is *informative in s = 0 given r*; when $E[q \mid r, m, s = 0] = E[q \mid r, s = 0]$, we say that *message m is uninformative in s = 0 given r*. The message *m* = "looks silky" is informative in situation *s* = "evaluating hair" when it comes from a trustworthy source but *m* = "contains silk" is always uninformative in situation *s* = "evaluating a shampoo" because silk does not affect the quality of shampoo. A Bayesian would react to *m* = "silk" in the former situation but not in the latter.

Coarse Thinking

The first essential assumption in our model is that different pieces of data may prime different mental representations or categorizations. For example, the advertisement "be bullish" may lead individuals to think of investing with Merrill Lynch as grabbing an opportunity to get rich. They then interpret decision-relevant information about Merrill accordingly.

The second essential assumption is that individuals do not have separate mental representations for every situation. Instead, they have only one representation for all the situations in a category and are effectively unable to differentiate between these

situations. If individuals perceive foreign trade as a kind of war, they interpret the message "Americans are likely to win" in the same way when assessing globalization and military conflict. Categorical thinking has been modeled by Mullainathan (2000), Fryer and Jackson (2004), and Peski (2006)[8]; we rely most closely on Mullainathan (2000).

Specifically, we assume that coarse thinkers either group situation *s = 0* together with similar situation *s* = 1, denoted by categorization $C_1 \equiv \{0,1\}$, or group situation *s = 0* together with similar situation *s = 2*, denoted by categorization $C_2 \equiv \{0,2\}$. They can think of investing with Merrill Lynch as grabbing an opportunity, or as hiring professional advice. Crucially, we assume that coarse thinkers do not have a separate mental representation for *s = 0* and that they cannot group *s = 0* together with both *s = 1* and *s = 2* at the same time. The latter assumption could arise out of a richer model of associations where *s = 1* and *s = 2* are sufficiently dissonant.

The assumption that coarse thinkers have the same mental representation for distinct situations is motivated by evidence from psychology. Krueger and Clement (1994) asked experimental participants to estimate the average temperatures of 48 different days of the year. Participants tended to under-estimate the difference between the average temperature of two days belonging to the same month (e.g., August 12th and August 20th) and to over-estimate the difference between the average temperature of two days belonging to neighboring months (e.g., August 25th and September 2nd). Because months are mental categories, coarse thinkers may estimate temperatures to be more similar within months than across months.

[8] There are some related finance models. Barberis and Shleifer (2003) present a model in which some investors group risky assets into categories and do not distinguish among assets within a category when formulating their demand. Hong, Stein, and Yu (2007) present a model in which investors use simplified univariate theories to forecast dividends. The theories investors use change over time in response to data.

Carpenter, Glazer, and Nakamoto (1994) asked experimental participants to rate the quality of hypothetical products described by a list of attributes. Participants preferred products with irrelevant differentiating attributes ("alpine class fill" for down jackets) to products without such attributes even when told that such attributes are irrelevant. One interpretation is that participants responded positively because alpine-ness contains decision-relevant information in a similar situation, e.g., buying skis.

The coarse thinker relies on categorization C_1 or C_2 to form beliefs about quality. The specific categorization depends on the data received. Denote the map between the data received and the chosen categorization by $C : \{u,d\} \times \{\varnothing, a, b\} \to \{C_1, C_2\}$. With a slight abuse of notation, $(r,m) = (r,\varnothing)$ denotes the information available to the individual prior to receiving the message from the persuader. We assume that individuals choose the most likely category given the data received:[9] [10]

$$C(r,m) = \arg\max_{C \in \{C_1, C_2\}} \hat{p}(s \in C \mid r,m) \text{ for all } (r,m) \in \{u,d\} \times \{\varnothing, a, b\} \qquad (2.3)$$

and ignore the alternative category (Mullainathan 2000). This map is meant to capture the idea that different mental representations or schema may be primed by different

[9] Maximizing the expression in (2.3) is equivalent to solving $\max_{s' \in \{1,2\}} \hat{p}(s' \mid r,m)$ for all $(r,m) \in \{u,d\} \times \{\varnothing, a, b\}$ since *s = 0* is in both C_1 and C_2, where

$$\hat{p}(s \mid r,m) = \frac{\left(\sum_{x' \in \{a,b\}} \hat{\sigma}(m \mid r, x', s) p(r, x' \mid s) \right) p(s)}{\sum_{s' \in \{0,1,2\}} \left(\sum_{x' \in \{a,b\}} \hat{\sigma}(m \mid r, x', s') p(r, x' \mid s') \right) p(s')}.$$

[10] With multiple solutions to $\max_{C \in \{C_1, C_2\}} \hat{p}(s \in C \mid r,m)$ for some $(r,m) \in \{u,d\} \times \{\varnothing, a, b\}$, pick C_1.

stimuli.[11] For instance, an individual might associate – and therefore co-categorize - investing with Merrill Lynch with grabbing another opportunity when looking at the ad "be bullish" or with seeking professional advice when looking at "a tradition of trust."

The coarse thinker applies the same model of inference to all situations in the category $C(r,m)$. One natural way to model coarse thinking is as Bayesian thinking with a coarser information set; i.e., the coarse thinker forms his beliefs about quality as a Bayesian who observes the category, $C(r,m)$, he is in but not the specific situation:

$$\hat{p}(q \mid r,m,C_i) = \hat{p}(q \mid r,m,s=0)\hat{p}(s=0 \mid r,m,C_i) + \hat{p}(q \mid r,m,s=i)\hat{p}(s=i \mid r,m,C_i) \quad (2.4)$$

where $C_i = C(r,m)$ for some $i \in \{1,2\}$.

This update rule (2.4) implies that, upon receiving (r,m), the coarse thinker's expectation of quality is:

$$E[q \mid r,m,C_i] \quad = \quad E[q \mid r,m,s=0]\hat{p}(s=0 \mid r,m,C_i) + E[q \mid r,m,s=i]\hat{p}(s=i \mid r,m,C_i) \quad (2.5)$$

An alternative way to model coarse thinking, which we actually follow, is to assume that the coarse thinker does not condition on the information received in weighing each situation, but instead uses constant weights $p(s \mid C_i)$. Such a coarse thinker forms beliefs about quality according to the rule

$$\hat{p}^{C_i}(q \mid r,m,s=0) = \hat{p}(q \mid r,m,s=0)p(s=0 \mid C_i) + \hat{p}(q \mid r,m,s=i)p(s=i \mid C_i) \quad (2.6)$$

where $C_i = C(r,m)$ for some $i \in \{1,2\}$.

Update rule (2.6) implies that, upon receiving (r,m), a coarse thinker's expectation of quality is:

$$E^{C_i}[q \mid r,m,s=0] \quad = \quad E[q \mid r,m,s=0]p(s=0 \mid C_i) + E[q \mid r,m,s=i]p(s=i \mid C_i) \quad (2.7)$$

[11] Smith (1998) reviews mental representation and psychological models of association. The idea that the coarse thinker chooses the most likely category given the observed data and ignores alternative categories is motivated by experimental evidence (Murphy and Ross 1994; Malt, Ross, and Murphy 1995).

These two models share many features. Both capture the idea that the coarse thinker updates as if he cannot distinguish the situations within a category: Under each specification, the coarse thinker's expectation of quality in $s = 0$ is a weighted average of the Bayesian's expectation across all situations co-categorized with $s = 0$ given (r, m). In the first model, a coarse thinker weighs the expectation of quality in situation s by the likelihood of s in the category given the information received: $\hat{p}(s \mid r, m, C_i)$. In the second, a coarse thinker weighs the expectation of quality in situation s by the likelihood of that situation in the category: $p(s \mid C_i)$. In both models, a coarse thinker's inference approaches the Bayesian's as $p(s = 0 \mid C_i)$ tends to 1.

We do not favor one model over the other based solely on psychological evidence. Since the model of coarseness with weights $p(s \mid C_i)$ greatly simplifies the formulas and proofs, we focus on it in what follows. We demonstrate in Appendix 2 that our main results hold with inessential modifications under the alternative assumption that coarse thinkers update using more "Bayesian" weights $\hat{p}(s \mid r, m, C_i)$.

We are interested in how the coarse thinker reacts to the message sent by the persuader conditional on the public information r. Fixing public information r, assume (without loss of generality) that the coarse thinker co-categorizes situation $s = 0$ with situation $s = 1$ before receiving the persuader's message. That is, let $C(r) \equiv C(r, \varnothing)$ and assume that $C(r) = C_1$. Then, the coarse thinker's expectation of quality is $E^{C_1}[q \mid r, s = 0]$ (as given by (2.7)) before receiving a message from the persuader.

To evaluate the coarse thinker's reaction to m, we distinguish the cases where the coarse thinker either does or does not re-categorize situation $s = 0$ after receiving the persuader's message. We say that message m *is pivotal given* r if

$$C(r,m) \neq C(r) \tag{2.8}$$

Message *m* is pivotal given *r* if it leads the coarse thinker to re-categorize situation *s = 0*.

First, consider the case where *m* is not pivotal given *r*. In this case, the model produces empirically plausible patterns of under- and over-reaction to data relative to the Bayesian benchmark. Comparing the magnitude of the coarse thinker's reaction to *m*, $\left|E^{C_1}[q \mid r,m,s=0] - E^{C_1}[q \mid r,s=0]\right|$, to the Bayesian's, $\left|E[q \mid r,m,s=0] - E[q \mid r,s=0]\right|$, suggests two distortions. First, since $p(s=0 \mid C_1) < 1$, the reaction of the coarse thinker only mutedly depends on his reaction to message *m* in situation *s=0* itself, as measured by $E[q \mid r,m,s=0] - E[q \mid r,s=0]$. This effect can lead to an under-reaction to data. Take, for example, the case where $E[q \mid r,m,s=0] \neq E[q \mid r,s=0]$, so *m* is informative given *r* in *s = 0*, and $E[q \mid r,m,s=1] = E[q \mid r,s=1]$, so *m* is uninformative given *r* in $s=1$. We see here that

$$\begin{aligned}\left|E^{C_1}[q \mid r,m,s=0] - E^{C_1}[q \mid r,s=0]\right| &= \left|E[q \mid r,m,s=0] - E[q \mid r,s=0]\right| p(s=0 \mid C_1) \\ &< \left|E[q \mid r,m,s=0] - E[q \mid r,s=0]\right|\end{aligned}$$

Coarse thinking leads to an under-reaction to news relative to Bayesian thinking. The uninformativeness of the data in the co-categorized situation dilutes its impact in situation *s = 0* precisely because the current situation is under-weighted in the update rule.

However, continuing to consider the case where *m* is not pivotal given *r*, the coarse thinker's response also depends on a term that the Bayesian's does not depend on: the informativeness of the message in the other situation $s=1$ in the same category. This implies that the coarse thinker could react to non-information or over-react to information. Take, for example, the case where $E[q \mid r,m,s=0] = E[q \mid r,s=0]$, but

$E[q \mid r, m, s=1] \neq E[q \mid r, s=1]$: the message *m* is uninformative in situation *s = 0* given *r*, but is informative in co-categorized situation $s=1$. Then,

$$\begin{aligned} \left|E^{C_1}[q \mid r, m, s=0] - E^{C_1}[q \mid r, s=0]\right| &= \left|E[q \mid r, m, s=1] - E[q \mid r, s=1]\right| p(s=1 \mid C_1) \\ &> \left|E[q \mid r, m, s=0] - E[q \mid r, s=0]\right| \\ &= 0 \end{aligned} \tag{2.9}$$

The coarse thinker now reacts to an uninformative message in situation *s = 0* because it is informative in the co-categorized situation. His use of the same model to interpret messages in all situations in the category may lead him to over-react to non-informative messages. This, we suggest, is part of what a person responding to "We put silk in the bottle" is doing. We call this process transference. This transference of the informational content of messages across situations within a category drives the first set of our results below. Indeed, if the persuader cannot affect how individuals categorize a situation, the strategy of uninformative persuasion is to trigger such transference: successful persuasion in that case takes advantage of over-reaction.

Now consider the case where message *m* is pivotal given public information *r*, so *m* leads to the re-categorization of *s = 0* from C_1 to C_2. Comparing the magnitude of the coarse thinker's reaction to *m* when categorization depends on the message received, $\left|E^{C_2}[q \mid r, m, s=0] - E^{C_1}[q \mid r, s=0]\right|$, to that when categorization is fixed, $\left|E^{C_1}[q \mid r, m, s=0] - E^{C_1}[q \mid r, s=0]\right|$, highlights additional effects from re-categorization.

Since message *m* leads to the re-categorization of *s = 0* from C_1 to C_2, the coarse thinker's reaction now depends on more than the informativeness of that message about quality in some situation. To take an extreme example, the coarse thinker may react to *m*

even when it is uninformative about quality in *all* situations because it affects how he categorizes *s = 0*. To see this, take the case where *m* is uninformative about quality in all situations, but where the expectation of quality conditional on *r* in category C_2 is different from that in category C_1: $E[q \mid r, m, s] = E[q \mid r, s]$ for all *s*, but $E^{C_2}[q \mid r, s = 0] \neq E^{C_1}[q \mid r, s = 0]$ [12]. When message *m* prompts the coarse thinker to re-categorize *s = 0* from C_1 to C_2, the magnitude of his reaction to *m* is:

$$\begin{aligned} \left|E^{C_2}[q \mid r, m, s = 0] - E^{C_1}[q \mid r, s = 0]\right| &= \left|E^{C_2}[q \mid r, s = 0] - E^{C_1}[q \mid r, s = 0]\right| \\ &> \left|E^{C_1}[q \mid r, m, s = 0] - E^{C_1}[q \mid r, s = 0]\right| \quad (2.10) \\ &= 0 \end{aligned}$$

The coarse thinker reacts to the message because it affects how he categorizes situation *s = 0*, not because it is informative about quality in any situation. We call this phenomenon framing. "We try harder" frames Avis as an underdog. The message itself is uninformative about quality in all situations, but it encourages the re-categorization of Avis from loser to underdog, so public information (Avis lags behind Hertz in sales) becomes an indicator of higher quality. When the persuader can influence categorization, uninformative persuasion optimally frames that situation in the mind of the audience.

In summary, we have presented a simple model that naturally describes two manifestations of coarse thinking: framing and transference. Framing refers to *how* the

---

[12] Applying the definition of $E^{C_i}[q \mid r, s = 0]$ and rearranging terms,
$E^{C_2}[q \mid r, s = 0] \neq E^{C_1}[q \mid r, s = 0]$ if and only if
$E[q \mid r, s = 0]\big(p(s = 0 \mid C_2) - p(s = 0 \mid C_1)\big) \neq E[q \mid r, s = 1]p(s = 1 \mid C_1) - E[q \mid r, s = 2]p(s = 2 \mid C_2)$.
It is clear that this inequality holds for a wide range of parameter values. For instance, if
$E[q \mid r, s = 0] = 0$ and $p(s = 1 \mid C_1) = p(s = 2 \mid C_2)$ then the requirement is just that the public signal is interpreted differently in *s = 2* than in *s = 1*: $E[q \mid r, s = 2] \neq E[q \mid r, s = 1]$.

audience thinks about the data, transference refers to *what* it thinks about it by analogy. We show next how a rational persuader takes advantage of these mental strategies.

3. Persuasion

Setup

Fixing public information $r \in \{u, d\}$, the persuader in situation $s = 0$ maximizes the individual's expectation of quality, net of persuasion costs. Presumably, the higher the expected quality, the greater the support the persuader receives, whether through sales, votes, or membership in his organization.

The persuader can alter messages in situation $s = 0$. Specifically, we assume that the persuader observes the signal x that the individual would see absent intervention (for example, absent an advertising campaign). He can then either intervene prior to the individual's observing x and send altered message $m \in \{a, b\}$ not equal to x at cost $c \geq 0$, or simply report $m = x$ at zero cost. The individual then observes m and never sees the original x.

As mentioned above, we make the simplifying assumption that nature directly sends the signal $m = x$ in situations $s = 1$ and $s = 2$. This should be a reasonable assumption if $s = 1$ and $s = 2$ are situations in which persuaders are restricted (e.g., by law) to truthfully reveal all private information. Alternatively, $s = 1$ and $s = 2$ may be common, everyday situations in which there is no persuader to send an altered message (e.g., underdogs in many real-life situations observably try harder).

Denote the persuader's strategy in $s = 0$ by $\bar{m}_0 : \{u,d\} \times \{a,b\} \to \{a,b\}$, where $\bar{m}_0(r,x) = m$ represents the strategy of reporting message m whenever the public information is r and the private signal is x.[14] Denote the corresponding "strategy" of nature in situations $s = 1$ and $s = 2$ by $\bar{m}_1(r,x) = \bar{m}_2(r,x) = x$ for all r and x since nature always reveals the private signal in those situations.

An optimal strategy for the persuader in $s = 0$ selects an m to maximize his payoff conditional on (r,x) for each (r,x), where this conditional payoff is given by:

$$\begin{array}{ll} E[q \mid r,m,s=0] & \text{if } m = x \\ E[q \mid r,m,s=0] - c & \text{if } m \neq x \end{array} \tag{3.1}$$

if the audience consists of Bayesians and by

$$\begin{array}{ll} E^{C(r,m)}[q \mid r,m,s=0] & \text{if } m = x \\ E^{C(r,m)}[q \mid r,m,s=0] - c & \text{if } m \neq x \end{array} \tag{3.2}$$

if the audience consists of coarse thinkers.

Optimal strategies depend on how individuals respond to messages. The standard assumption is that individuals are *sophisticated*, meaning that their beliefs are consistent with Bayes' rule applied to the reporting strategies of the persuader and nature. More formally, this implies the following restrictions on beliefs:

(S) $\hat{p}$ is derived from p and $\hat{\sigma}(m \mid r,x,s) = \begin{cases} 1 \text{ if } \bar{m}_s(r,x) = m \\ 0 \text{ if } \bar{m}_s(r,x) \neq m \end{cases}$ using Bayes' rule whenever possible.[15]

[14] We restrict attention to pure strategies except in Appendix 2.

[15] When there exists a message m that is sent with zero probability under the persuader's strategy in situation $s = 0$, relevant conditional beliefs about quality are derived from p, $\hat{\sigma}$, and an arbitrary distribution $\mu(x \mid r,m,s=0)$ over $\{a,b\}$.

An alternative to sophistication is that individuals take messages at face value: individuals, including both Bayesians and coarse thinkers, take the message *m* they see and update as if the persuader always non-strategically reveals the private signal *x*. Formally, this implies the following restriction on beliefs[17]:

(F) $\hat{p}$ is derived from p and $\hat{\sigma}(m \mid r,x,s) = \begin{cases} 1 \text{ if } m = x \\ 0 \text{ if } m \neq x \end{cases}$ using Bayes' rule.

We will show that, given an additional assumption described below, it does not matter whether we assume that individuals are sophisticated or that they take messages at face value in considering the persuader's optimal strategies.

Definition: An equilibrium $(\bar{m}_0^{\,*}, \hat{p}^*)$ satisfies:

(1) For each (r,x) the persuader chooses $\bar{m}_0^{\,*}(r,x)$ to maximize the audience's expectation of quality minus costs of persuasion, as given by (3.1) if the audience consists of Bayesians and by (3.2) if the audience consists of coarse thinkers. The persuader takes the audience's beliefs $\hat{p}^*$ as given.

(2) $\hat{p}^*$ satisfies (S) given the reporting strategy of the persuader, $\bar{m}_0 = \bar{m}_0^{\,*}$, if the audience is sophisticated and (F) if the audience takes messages at face value.

To simplify the remaining analysis as well as to focus on truly uninformative persuasion, we make the *uninformative persuasion assumption*: the signal the persuader privately observes is always uninformative about quality in *s = 0*:

[17] Shin (1994) also refers to updating on *ex ante* probabilities as taking messages at face value. Other authors (e.g., Kartik, Ottaviani, and Squintani 2007) refer to such updating as "naïve".

$$E[q \mid r, x, s = 0] = E[q \mid r, s = 0] \tag{3.3}$$

for all *r, x*. The uninformative persuasion assumption implies that all messages are uninformative in situation *s = 0* by (2.2). Hence, Bayesians never react to messages sent by the persuader in *s = 0*.

Before characterizing the equilibria of this model, we establish the following useful result.

Lemma: Under our assumptions:

(i) Categorization rule (2.3) is independent of the persuader's strategy.

(ii) Sophisticated and face value audiences hold the same expectations of quality conditional on messages whether they consist of Bayesians (with expectations given by (2.1)) or of coarse thinkers (with expectations given by (2.7)).

(iii) From (i) and (ii) it follows that $\bar{m}_0^*$ is an equilibrium strategy for the persuader given some sophisticated beliefs if and only if it is an equilibrium strategy when the audience takes messages at face value. This is true whether the audience consists of Bayesians or of coarse thinkers.

Proof: In Appendix 1. ■

This lemma establishes that, in considering the equilibrium strategies of the persuader, we can assume without loss of generality that the audience takes messages at face value.[18] The intuition is that face value and sophisticated individuals differ only in their beliefs regarding the equilibrium strategy of the persuader in *s = 0*. Since the persuader's underlying private signal is uninformative about quality and categorization rule (2.3) is independent of the persuader's strategy, these differences do not generate

[18] This lemma does not hold when coarse thinkers update according to the "more Bayesian" rule explored in Appendix 2, where we focus on the sophisticated case.

disagreement between face value and sophisticated thinkers in what messages imply about quality and hence do not affect the persuader's best response.

This lemma is useful since it implies that, to verify that m_0^* is an equilibrium strategy for the persuader, we only need to check whether it maximizes the persuader's payoff conditional on (r,x) for each (r,x), where this conditional payoff is given by:

$$\begin{array}{l} E[q \mid r, x = m, s = 0] \quad \text{if } m = x \\ E[q \mid r, x = m, s = 0] - c \text{ if } m \neq x \end{array} \tag{3.4}$$

if the audience consists of Bayesians and by

$$\begin{array}{l} E^{C_i}[q \mid r, x = m, s = 0] \quad \text{if } m = x \\ E^{C_i}[q \mid r, x = m, s = 0] - c \text{ if } m \neq x \end{array} \tag{3.5}$$

where $C_i = \underset{C \in \{C_1, C_2\}}{\arg\max}\, p(s \in C \mid r, x = m)$ if the audience consists of coarse thinkers. In other words, this lemma establishes that it is unnecessary to compute consistent beliefs $\hat{p}^*$ and make sure that the persuader's strategy is a best response to those beliefs. Rather, we can treat the audience's beliefs as given by the *ex ante* probabilities in what follows.

Additionally, the lemma implies that our results extend to environments where messages are verifiable. The most literal interpretation of a message in situation *s = 0* is as unverifiable and potentially costly talk (e.g., *m* = "We try harder" for *s = Car Rental*), in which case $m \neq x$ implies that the persuader paid a weakly positive cost to send an altered message or to engage in advertising. However, the lemma establishes that messages may also be viewed as reflecting the inclusion of observable and verifiable product attributes (e.g., *m* = "contains silk" for *s = Shampoo*), in which case $m \neq x$ implies that the persuader paid some weakly positive cost to change an objective product

attribute and sophisticated individuals take messages at face value (as in the "persuasion games" of Milgrom and Roberts 1986b). [19]

Bayesians

To simplify the remaining analysis and get clear results, we make several additional assumptions. First, we assume that public information is uninformative in $s = 0$ in the sense that

$$E[q \mid r, s = 0] = E[q \mid s = 0] \tag{3.6}$$

for all r. Second, we assume that the expected quality conditional on each situation equals zero:

$$E[q \mid s] = 0 \tag{3.7}$$

These two additional assumptions are not important for our qualitative results and serve to simplify the algebra.

Together with the uninformative persuasion assumption, these assumptions imply that the Bayesian expectation of quality in situation $s = 0$ satisfies:

$$E[q \mid r, m, s = 0] = E[q \mid s = 0] = 0 \tag{3.8}$$

for all r and m.

---

[19] A somewhat subtle question is how coarse thinkers should compute $\hat{p}(s \mid r, m)$ when messages are verifiable. However, given categorization rule (2.3) together with update rule (2.6), it does not matter whether they calculate $\hat{p}(s \mid r, m) = \dfrac{\hat{p}(r, m \mid s)p(s)}{\sum_{s' \in \{0,1,2\}} \hat{p}(r, m \mid s')p(s')}$ or $\hat{p}(s \mid r, m) = \dfrac{p(r, m \mid s)p(s)}{\sum_{s' \in \{0,1,2\}} p(r, m \mid s')p(s')}$, where $\hat{p}(r, m \mid s)$ is defined as in footnote 9.

Under these assumptions, we first characterize the persuader's optimal (i.e., equilibrium) strategy when the audience consists of Bayesians:

Proposition 1 (Bayesian baseline): Suppose individuals are Bayesians. Then an optimal strategy of the persuader in situation $s = 0$ is to always report $m = x$ and to never pay the cost to send some other message $m \neq x$. If $c > 0$, this strategy is unique.

Proof: Recall that all messages are uninformative in situation $s = 0$. As a result, fixing r, the persuader receives payoff $E[q \mid r, m, s = 0] = 0$ (from (3.8)) if he sends $m = x$ and receives payoff $E[q \mid r, m, s = 0] - c = -c$ if he sends $m \neq x$. Clearly, $0 \geq -c$ so an optimal strategy of the persuader is to always report $m = x$. Since this inequality is strict whenever $c > 0$, this strategy is uniquely optimal whenever $c > 0$. ■

According to Proposition 1, persuaders never fabricate uninformative messages (or change uninformative product attributes) for Bayesians when the cost of doing so is positive. Because private signals in situation $s = 0$ are assumed to be uninformative about quality, Bayesians do not update their beliefs from messages sent by the persuader. Thus the persuader receives no benefit from fabricating such messages and is unwilling to pay a positive cost to do so.

Coarse thinkers

We now examine the persuader's optimal strategy when the audience consists of coarse thinkers. In the next two propositions, we characterize the persuader's optimal strategy conditional on public information. Fix public information r and, without loss of generality, assume that $C(r) = C_1$ (recall from before that $C(r) \equiv C(r, \varnothing)$). First,

consider the case where messages are not pivotal given r. Also, without loss of generality, suppose that private signal $x = a$ is more favorable than private signal $x = b$ in $s = 1$ in the following sense:

$$E[q \mid r, a, s = 1] \geq E[q \mid r, b, s = 1] \tag{3.9}$$

We have the following proposition:

Proposition 2 (Transference): Suppose that individuals are coarse thinkers, that messages are not pivotal given public information r, and that condition (3.9) holds. Then an optimal strategy of the persuader in situation $s = 0$ may involve the creation of a message. Specifically, so long as

$$c < \left(E[q \mid r, a, s = 1] - E[q \mid r, b, s = 1]\right) p(s = 1 \mid C_1) = c^* \tag{3.10}$$

any optimal strategy of the persuader dictates reporting $m=a$ whenever $x = b$.

Proof: Fix the public information r and recall that neither message is assumed to be pivotal. If the persuader reports $m = x = b$ to a coarse thinker, his payoff is $E^{C_1}[q \mid r, b, s = 0] = E[q \mid r, b, s = 1] p(s = 1 \mid C_1)$. If he replaces $x = b$ with $m = a$, his payoff is $E^{C_1}[q \mid r, a, s = 0] - c = E[q \mid r, a, s = 1] p(s = 1 \mid C_1) - c$. Subtracting the first payoff from the second, the persuader optimally replaces $x = b$ with $m = a$ if this difference is positive or, equivalently, if $c < c^*$. ■

Condition (3.10) yields a corollary:

Corollary 1: Suppose the conditions of Proposition 2 hold and let c^* be defined as above. Then, c^* is weakly increasing in (i) the probability of situation $s = 1$ in category

C_1, $p(s=1 \mid C_1) = \dfrac{p(s=1)}{p(s=0)+p(s=1)}$ and (ii) the magnitude of reaction in the co-categorized situation, $E[q \mid r,a,s=1] - E[q \mid r,b,s=1]$.

Proof: From (3.10), $c^* = \left(E[q \mid r,a,s=1] - E[q \mid r,b,s=1]\right) p(s=1 \mid C_1)$. This expression is increasing in (i) $p(s=1 \mid C_1)$ since $E[q \mid r,a,s=1] \geq E[q \mid r,b,s=1]$ by (3.9) and (ii) in $E[q \mid r,a,s=1] - E[q \mid r,b,s=1]$ since $p(s=1 \mid C_1) \geq 0$. ■

Proposition 2 illustrates that persuaders may pay a cost to alter messages for coarse thinkers even when the message they send cannot be pivotal. Their decision depends on whether the gap (in terms of the improved assessment of quality) between the best possible message (or attribute) and the private signal (or original attribute) is large enough to offset the cost of persuasion. This provides a way of thinking about why Culver replaced *x* = "no silk" with *m* = "silk" by putting "silk in the bottle."

Corollary 1 highlights the fact that, with two co-categorized situations, a persuader is more likely to manufacture a message in situation *s = 0* if it has a lower probability within its category. Since the benefit for manufacturing messages is the transference of the informational content from other situations in a category to the current one, a higher probability of these other situations increases transference and this benefit.

This point may shed light on what advertisers refer to as consumer involvement, a notion closely related to but not identical to that of "stakes." A high involvement product occupies a huge probability space in its category (*p(s = 0)* is high), so the transference from other situations is small, and hence so is the benefit of non-informative advertising. Our model predicts, as the marketing research recommends, that advertising in these instances should be informative (Sutherland and Sylvester 2000). In contrast, low

involvement products are mixed up in consumers' minds with many similar situations and hence there is greater scope for persuasion, exactly as the marketing literature suggests. Alberto Culver Shampoo and Schwarzenegger's defense of free trade are both consistent with this point. Voting may be another low involvement activity, which encourages non-informative advertising. Our point is *not* that people are incapable of rational high involvement thinking, but rather that in many instances they do not engage in such thinking, perhaps because it is not worth it. It is precisely in those instances of coarse thinking that persuasion pays.

We now consider the case where categorization depends on the exact message the persuader sends, i.e., one message is pivotal. To limit the number of cases, suppose that message $m = a$ is pivotal in changing categorization from C_1 to C_2 and that the persuader weakly "prefers" sending pivotal message $m = a$ over non-pivotal message $m = b$:

$$E[q \mid r, a, s = 2]p(s = 2 \mid C_2) \geq E[q \mid r, b, s = 1]p(s = 1 \mid C_1) \tag{3.11}$$

Since message $m = a$ is pivotal, it does not matter whether (3.9) still holds.

We then have:

Proposition 3 (Framing): Suppose individuals are coarse thinkers, that message $m = a$ is pivotal given public information r, and that condition (3.11) holds. Then an optimal strategy of the persuader in situation $s = 0$ may involve the creation of a message. Specifically, so long as

$$c < \left(E[q \mid r, a, s = 2]p(s = 2 \mid C_2) - E[q \mid r, b, s = 1]p(s = 1 \mid C_1)\right) \tag{3.12}$$

any optimal strategy of the persuader dictates reporting $m=a$ whenever $x = b$.

Proof:[20] Fix some piece of public information r and recall the assumption that $m = a$ is pivotal given this information. If the persuader reports $m = x = b$ to a coarse thinker, his payoff is $E^{C_1}[q \mid r, b, s = 0] = E[q \mid r, b, s = 1]p(s = 1 \mid C_1)$. If he replaces $x = b$ with $m = a$, his payoff is $E^{C_2}[q \mid r, a, s = 0] - c = E[q \mid r, a, s = 2]p(s = 2 \mid C_2) - c$. Subtracting the first payoff from the second, the persuader optimally replaces $x = b$ with $m = a$ when this difference is positive or, equivalently, when condition (3.12) holds. ■

Corollary 2: Suppose the conditions of Proposition 3 hold and that any message the persuader sends is uninformative about quality within each categorization given the public information: $E^{C_i}[q \mid r, m, s = 0] = E^{C_i}[q \mid r, s = 0]$ for all $i \in \{1, 2\}$. So long as (3.12) still holds, any optimal strategy of the persuader dictates replacing $x = b$ with $m = a$: the optimal strategy of the persuader dictates paying a positive cost to send a message that is uninformative about quality within each categorization in order to desirably categorize situation $s = 0$ in light of public information r.

Examples from the introduction illustrate Corollary 2. With its ad "Be Bullish," Merrill Lynch frames itself as provider of opportunities during the internet bubble, without conveying any useful information. With its ad, "We try harder," Avis frames itself as an underdog, and encourages car renters to interpret its lagging status in this favorable light, again without conveying any useful information.

[20] The conditions of Proposition 3 imply that the persuader would never wish to pay a cost to send the non-pivotal message $m = b$. This result hinges on the assumption (3.11). With the alternative (and equally reasonable) assumption that the inequality in (3.11) is reversed, the persuader would pay a sufficiently low cost to replace $x = a$ (and avoid sending the pivotal message $m = a$) with non-pivotal message $m = b$.

Corollary 3 (Withholding Good Messages): Suppose the conditions of Proposition 3 hold and that $m = b$ is a favorable message within each categorization given the public information: $E^{C_i}[q \mid r, b, s = 0] > E^{C_i}[q \mid r, a, s = 0]$ for all $i \in \{1, 2\}$. So long as (3.12) still holds, any optimal strategy of the persuader dictates replacing private signal $x = b$ with message $m = a$: the optimal strategy of the persuader may dictate paying a positive cost to send a message that is universally less favorable in order to avoid undesirably categorizing situation $s = 0$ in light of public information r.

Corollary 3 presents a more subtle prediction of the model, which we return to in Section 5 where we study mutual fund advertising. It says that firms avoid presenting good news when such news create unattractive frames. In Section 5, we show that mutual funds avoid presenting favorable information about their *relative* returns during periods of *falling* stock prices, perhaps because the mere mention of returns invites co-categorization of investments as grabbing opportunities, a frame that is not compelling during declining markets.

The previous two propositions characterize the persuader's optimal strategy conditional on public information. It is also interesting to examine the persuader's full strategy to see how it might depend on the public information. From Proposition 1, we know that the persuader's optimal strategy is independent of public information when the audience consists of Bayesians. The same is not necessarily true when the audience consists of coarse thinkers:

Proposition 4. Suppose individuals are coarse thinkers. Then an optimal strategy of the persuader in situation *s = 0* may depend on public information. That is, there may exist some *x* such that the persuader reports $m \neq x$ when *r = u* but reports $m = x$ when $r = d$.

Proof (by example): *Example A.* Suppose $C(u) = C(d) = C_1$ and message *m = a* is pivotal given each piece of public information; that is, $C(u,a) = C(d,a) = C_2$. Further, suppose

$$E[q \mid u,a,s=2]p(s=2 \mid C_2) > E[q \mid u,b,s=1]p(s=1 \mid C_1)$$

and

$$E[q \mid d,b,s=1]p(s=1 \mid C_1) > E[q \mid d,a,s=2]p(s=2 \mid C_2)$$

These two inequalities imply that $E^{C_2}[q \mid u,a,s=0] > E^{C_1}[q \mid u,b,s=0]$ and $E^{C_1}[q \mid d,b,s=0] > E^{C_2}[q \mid d,a,s=0]$. So long as

$$c < \min\left\{E^{C_2}[q \mid u,a,s=0] - E^{C_1}[q \mid u,b,s=0], E^{C_1}[q \mid d,b,s=0] - E^{C_2}[q \mid d,a,s=0]\right\}$$

the persuader's optimal strategy dictates replacing *x = b* with *m = a* when *r = u* and replacing *x = a* with *m = b* when *r = d*.

Example B. Suppose $C(u) = C_1$, $C(d) = C_2$, and messages are not pivotal given either possible piece of public information. Further suppose that $E[q \mid u,a,s=1] > E[q \mid u,b,s=1]$ and $E[q \mid d,b,s=2] > E[q \mid d,a,s=2]$. These two inequalities imply that $E^{C_1}[q \mid u,a,s=0] > E^{C_1}[q \mid u,b,s=0]$ and $E^{C_2}[q \mid d,b,s=0] > E^{C_2}[q \mid d,a,s=0]$. So long as

$$c < \min\left\{E^{C_1}[q \mid u,a,s=0] - E^{C_1}[q \mid u,b,s=0], E^{C_2}[q \mid d,b,s=0] - E^{C_2}[q \mid d,a,s=0]\right\}$$

the persuader's optimal strategy dictates replacing $x = b$ with $m = a$ when $r = u$ and replacing $x = a$ with $m = b$ when $r = d$. ■

We supply two examples in the proof of Proposition 4 (though only one is necessary). Merrill Lynch's advertising campaigns over the course of the internet bubble illustrate both examples, but also show how specific advertisements take advantage of both framing and transference. In the first example, there is always a pivotal message, which allows the persuader to frame $s = 0$ differently depending on the information he wishes to frame. Merrill Lynch advertises with "be bullish" to frame itself as a provider of opportunities when market returns have been high, but with "a tradition of trust" to frame itself as a provider of services when market returns have been low.

The second example shows that the persuader may wish to send different, non-pivotal, messages depending on how $s = 0$ is categorized given the public information. Having framed public information about market returns, Merrill Lynch advertises its financial *analysts* upon framing itself as a provider of opportunities, but its financial *consultants* or *advisors* upon framing itself as a provider of services. These features of Merrill Lynch's ads tap into, rather than alter, the prevailing mental model, and take advantage of transference.

4. Product Branding

A major challenge for the fields of Industrial Organization and Marketing is to understand product branding. Consumers buy many branded products, often repeatedly, at higher prices than identical or nearly identical "generic" products (Tirole 1988). According to Peter and Olson (2005, p. 97), 71% of cigarette buyers, 65% of mayonnaise

buyers, 61% of toothpaste buyers, and 53% of bath soap buyers are loyal to their brands (i.e., claim in a survey of 2000 respondents to mostly buy the same brand). Although some brands are physically different from generic products, others, such as Clorox bleach, are identical. In this section, we show how firms may be able to differentiate their products and create brands through uninformative advertising.

To fix ideas, consider the case of California Burgundy. Burgundy is a French region that produces high quality and expensive wines from the pinot noir grape. California also produces expensive wines from the pinot noir grape. In California, they are called pinot noir, not burgundy. About 40 years ago, California wine producers started making inexpensive red wines called California Burgundy. These wines contain no pinot noir grape, only cheaper varietals. Even so, it appears that merchants tend to charge more for a 5-liter package of Peter Vella California Burgundy than for the same sized package of Peter Vella Delicious Red. In May 2007, out of the 48 stores listed on winesearcher.com as selling both wines, 27 charge more, 20 charge the same amount, and 2 charge less for California Burgundy. The average "premium" on California Burgundy is 73 cents ($12.60 for California Burgundy as opposed to $11.87 for Delicious Red). As a brand, California Burgundy appears to command a price premium even though its physical characteristics have nothing to do with burgundy wine (it is still possible, however, that the brand is superior for some other reason we cannot verify).

Our model provides a way to understand why California Burgundy sells for more than Delicious Red. Consider a consumer facing one of two similar situations: buying a bottle of American red wine ($s = US$) and buying a bottle of French red wine ($s = FR$). American bottles are all initially labeled Table ($x = Table$) but French bottles are either

initially labeled Table or Burgundy (x = *Burgundy*). We assume that neither label is informative in evaluating the quality of American wine, but that the label Burgundy is a positive signal of quality for French wine. Specifically[21]

$$\begin{array}{ccccc} E[q \mid x = Burgundy, s = FR] & > & 0 & > & E[q \mid x = Table, s = FR] \\ E[q \mid x = Burgundy, s = US] & = & 0 & = & E[q \mid x = Table, s = US] \end{array} \tag{4.1}$$

The audience is populated by measure one of consumers. Fraction β of consumers are coarse thinkers who co-categorize buying a bottle of American wine with buying a bottle of French wine no matter how the bottle is labeled: for all m, $C(m) = \{US, FR\} \equiv C$.[22] Fraction $1-\beta$ are Bayesians who differentiate the two situations. From (4.1) it follows that for coarse thinkers:

$$E^C[q \mid m = Burgundy, s = US] > 0 > E^C[q \mid m = Table, s = US] \tag{4.2}$$

and for Bayesians:

$$E[q \mid m = Burgundy, s = US] = E[q \mid m = Table, s = US]. \tag{4.3}$$

A monopolist wine producer in the United States sells two homogeneous (or perfectly substitutable) wines, $w = 1$ and $w = 2$. The producer can at zero cost label the wines "Table" or pay a fixed advertising cost $c \geq 0$ to label one or both wines "Burgundy".[23] For simplicity, the marginal cost of production is set to equal 0.

A consumer buys at most one bottle of wine. Decision utility from purchasing a bottle of wine w is given by:

[21] We ignore the possibility of public information in this section.
[22] To simplify matters, we only allow for situation s = *US* to be co-categorized with one other situation, s = *FR*, in this section. In other words, we abstract away from the possibility that the coarse thinker may group situation "buying a bottle of American red wine" with another distinct situation (e.g., "buying a bottle of Italian red wine") depending on the message sent by the producer.
[23] If it seems troubling to assume that it is costly to label wine "Burgundy", note that we allow $c = 0$.

$$U(w) = \overline{u} + q_w - p_w, \tag{4.4}$$

where q_w denotes the quality and p_w the price of wine w and $\overline{u} > 0$ is some constant. To limit the number of cases considered, assume that $\beta\overline{u} < -E^C[q \mid Table, US]$.[24] Utility from not buying wine is assumed to equal 0. Consumers maximize their expected (decision) utility.

Consider a game with two periods. In the first period, the monopolist simultaneously decides how to label and price its wines.[25] In the second, consumers observe the label and price of each wine and decide which wine (if any) to buy.

An *equilibrium* in this context is defined to be a tuple of strategies satisfying: (i) the producer's choice of labels and prices (his strategy) maximizes profits given consumers' strategies and (ii) for each set of labels and prices each consumer's strategy dictates making a purchase decision that maximizes expected utility.

We restrict attention to pure strategy equilibria in which the firm sells both wines to a positive fraction of the population on the equilibrium path. Defining

$$\overline{c} \equiv \beta\left(\overline{u} + E^C[q \mid Burgundy, US]\right) \tag{4.5}$$

we have the following results:

Proposition 5: Suppose all individuals are Bayesians ($\beta = 0$) and a monopolist wine producer sells two homogeneous wines in the U.S. Then there exists an equilibrium such that the producer does not label either wine Burgundy and charges the same price $\overline{u}$ for each of its wines. When $c > 0$, any equilibrium has this property.

---

[24] This eliminates the possibility that the monopolist could find it optimal to label both wines Table and charge $\overline{u} + E^C[q \mid, Table, US] < \overline{u}$ (recall that $E^C[q \mid, Table, US] < 0$) for each wine (selling to both Bayesians and coarse thinkers).

[25] We assume that buyers do not take price to be a potential signal of quality in either *s* = *US* or *s* = *FR*.

Proof: In Appendix 1.

Proposition 6: Suppose fraction $0 < \beta < 1$ of consumers are coarse thinkers who co-categorize buying U.S and French wine and there is a monopolist wine producer selling two homogeneous wines in the U.S. If $c < \overline{c}$, product differentiation through uninformative advertising emerges: in any equilibrium, the producer replaces x = *Table* with m = *Burgundy* for exactly one of its wines and charges $\overline{u} + E^C[q \mid m = Burgundy, s = US] > \overline{u}$ for that wine while charging $\overline{u}$ for the other. The higher priced wine is sold to the coarse thinkers. If $c > \overline{c}$, then in any equilibrium the producer labels each wine m = x = *Table* and charges $\overline{u}$ for each wine. In any such equilibrium, coarse thinkers do not buy either wine.

Proof: In Appendix 1.

Proposition 6 shows how product branding emerges in equilibrium for certain parameter values. In such an equilibrium, the branded good is sold for a higher price than the generic good. The reason that this is possible is that the coarse thinking audience (incorrectly) believes the branded good is superior because it is associated with an attribute that contains decision-relevant information in a co-categorized situation. This rendition of branding seems broadly consistent with the standard discussions in the marketing literature (Sutherland and Sylvester 2000; Peter and Olson 2005).

5. Mutual Funds

The mutual fund industry presents a major challenge to financial economics. It is enormous, supervising around $7 trillion of investor assets. It includes thousands of competitors, who nonetheless charge high fees and remain highly profitable. Perhaps

most strikingly, it appears to provide no economic value to investors, with virtually all mutual funds underperforming by a significant margin passive investment strategies offered by low-fee index funds (Swensen 2005). How can an industry be so successful while adding so little value and charging so much? Perhaps part of the answer is successful persuasion. Below we present a simple model of advertising to coarse-thinking mutual fund investors, and offer some evidence that it can help explain.

A Simple Model

Suppose that coarse thinking individuals do not sufficiently differentiate between choosing a mutual fund and selecting other professional advisors, such as doctors or lawyers, or grabbing great opportunities, such as buying stocks or finding jobs on tips. Individuals face one of three similar situations: selecting a mutual fund ($s = MF$), selecting a professional service ($s = PS$), and grabbing an opportunity ($s = GO$).

The data an individual sees are as follows. First, he observes publicly available data about general past performance $r \in \{u, d\}$, where $r = u$ stands for good past performance and $r = d$ stands for bad past performance. For a mutual fund, this could be the past market return. For another professional service, this could be the history of success of a particular surgery or type of lawsuit. For grabbing an opportunity, this could be his own or his friends' experience with chasing tips.

Second, he may receive a more specific hard message from a persuader about past performance $m_p \in \{a_p, b_p\}$, where $m_p = a_p$ stands for good past performance and $m_p = b_p$ stands for bad past performance. For a mutual fund, this could be some measure of its past relative or absolute return. For another professional service, this could be the history

of success of a particular doctor or lawyer. For grabbing an opportunity, this could be a measure of past success, such as return, as well.

We depart from the formal model presented earlier in one way. We assume that the mutual fund cannot fabricate hard information about past performance. In other words, it cannot send $m_p \neq x_p$. On the other hand, the mutual fund can at zero cost choose not to report information about past returns and send message $m_p = \varnothing$ no matter the realization of x_p. This departure does not introduce any new conceptual issues.

To apply the analysis of previous sections we make five assumptions. First, we expand the set of possible private signals to include the empty message and assume that it is in general uninformative about quality:[26]

Assumption 1: $$E[q \mid r, x_p = \varnothing, s] = E[q \mid r, s] \text{ for all } r \text{ and } s \tag{5.1}$$

Second, we assume that the availability of past performance data is associated with grabbing opportunities. Specifically, we assume that the joint distribution *p* is such that selecting a mutual fund is co-categorized with grabbing an opportunity if and only if the fund reports a non-null message about its past performance.[27] Formally,

Assumption 2: $$C(r, m_p) = \begin{cases} C_{GrabOpp} \equiv \{MF, GO\} & \text{if } m_p \neq \varnothing \\ C_{ProServ} \equiv \{MF, PS\} & \text{if } m_p = \varnothing \end{cases} \tag{5.2}$$

for all *r*.

[26] A realistic assumption is that mutual funds always have access to verifiable past performance data (so $p(\varnothing \mid s = MF) = 0$). However, we do not need to make any explicit assumptions on $p(x_p \mid s = MF)$ for the following analysis to hold.

[27] An implicit assumption here is that, unlike mutual funds, other professional services and opportunities sometimes do not have access to past performance data (or such data is unverifiable) and access is uncorrelated with quality.

Third, we assume that all data on past performance are uninformative in evaluating the quality of a mutual fund (Carhart 1997; Chevalier and Ellison 1997; Sirri and Tufano 1998; Swensen 2005)[28]:

Assumption 3: $$E[q \mid r, x_p, s = MF] = 0 \text{ for all } (r, x_p) \tag{5.3}$$

Assumption 3 implies that it is also the case that

$$E[q \mid r, m_p, s = MF] = 0 \text{ for all } (r, m_p) \tag{5.4}$$

Fourth, we assume that in evaluating other professional services or opportunities good past performance data constitute good news about quality:

Assumption 4: $$\begin{array}{rcll} E[q \mid u, x_p, s'] & \geq & E[q \mid d, x_p, s'] & \text{for all } x_p \\ E[q \mid r, a_p, s'] & \geq & E[q \mid r, b_p, s'] & \text{for all } r \end{array} \tag{5.5}$$

for each $s' \in \{PS, GO\}$.

Finally, we assume that past aggregate performance data are "sufficiently" more informative relative to individual performance data in evaluating the quality of opportunities than in evaluating the quality of other professional services:[29]

Assumption 5:

$$\begin{array}{rcl} E[q \mid u, b_p, s = GO]p(s = GO \mid C_{GrabOpp}) & > & E[q \mid u, s = PS]p(s = PS \mid C_{ProServ}) \\ E[q \mid d, a_p, s = GO]p(s = GO \mid C_{GrabOpp}) & < & E[q \mid d, s = PS]p(s = PS \mid C_{ProServ}) \end{array} \tag{5.6}$$

The above assumptions yield the following proposition:

---

[28] We are carrying over assumption (3.9) from the earlier sections that, for all s, $E[q \mid s] = 0$

[29] The conditions of Assumption 5 guarantee, respectively, that the mutual fund always reports performance when aggregate returns have been high and never reports it when aggregate returns have been low. An example might clarify the link between the intuition and the mathematical conditions of Assumption 5. The mathematical conditions would be met, for instance, if past aggregate performance data are uninformative in selecting another professional service but are "more informative" than even individual performance data in grabbing an opportunity in the sense that $E[q \mid u, b_p, s = GO] > 0$ and $E[q \mid d, a_p, s = GO] < 0$.

Proposition 7: Suppose individuals are coarse thinkers and Assumptions 1- 5 hold. When aggregate returns have been high ($r = u$), the optimal strategy of the mutual fund dictates always reporting message $m_p = x_p$ about past performance; i.e., always reporting hard information about past returns. When aggregate returns have been low ($r = d$), the optimal strategy of the mutual fund dictates always reporting message $m_p = \varnothing$ about past performance; i.e., never reporting hard information about past returns.

Proof: Fixing r and x_p, the persuader receives payoff $E[q \mid r, x_p, s = GO]p(s = GO \mid C_{GrabOpp})$ if he reports $m_p = x_p$ (by Assumptions 2 and 3) and receives payoff $E[q \mid r, s = PS]p(s = PS \mid C_{ProServ})$ if he reports $m_p = \varnothing$ (by Assumptions 1, 2, and 3). By Assumptions 4 and 5, the former payoff is strictly higher than the latter whenever $r = u$, and strictly lower whenever $r = d$. Consequentially, any optimal strategy of the persuader dictates always reporting $m_p = x_p$ when $r = u$ and always reporting $m_p = \varnothing$ when $r = d$. Uniqueness is immediate. ■

Proposition 7 yields two implications. First, since past performance data are informative for evaluating the quality of opportunities, a mutual fund will advertise with data about its past performance when aggregate returns have been high to co-categorize selecting the fund with grabbing such opportunities and thereby to maximize the reaction to the good aggregate returns. Second, to minimize the reaction to bad aggregate returns, a mutual fund will not report even universally favorable information about past returns when aggregate returns have been low, so as to avoid co-categorizing selecting the fund with grabbing such opportunities.

Before turning to the evidence, three comments about these implications are in order. First, the prediction that mutual funds advertise their own performance in rising but not in falling markets is difficult to reconcile with any plausible theory of rational persuasion. One could imagine a rational signal-extraction theory in which own past performance is informative about managerial ability in rising, but not in falling, markets. However, this does not strike us as plausible (in fact, the reverse seems more plausible). Moreover, this rational signal-extraction theory is difficult to reconcile with other facts about mutual fund advertising, such as omission of data about fees.

Second, the model does not deal with one additional important fact about mutual fund advertising, namely that it rarely includes data on management fees (Cronqvist 2005), data which is arguably crucial to assessing future returns. An extension of our model, which takes advantage of the fact that consumers often do not know fees for other professional services, such as doctors, might account for this finding. Such an extension was explored in the previous draft of this paper.

Third, another extension of the model might deal with the closely related question of which products to advertise. Specifically, firms should advertise products that co-categorize selecting a mutual fund with grabbing opportunities after returns have been high, and products that co-categorize selecting the fund with choosing other professional services after returns have been low. Mullainathan and Shleifer (2006) present evidence that the advertising of growth funds is highly pro-cyclical. Below, however, we focus on the predictions about the inclusion of past returns in mutual fund advertisements.

Evidence

We put together a data set of all financial advertisements from two magazines: *Business Week* and *Money*. *Business Week* is a weekly business newsmagazine. We examine all issues from January 1, 1994 to December 31, 2003. *Money* is monthly and more specifically directed at individual investors. We examine all the issues from January 1, 1995 to December 31, 2003. (The one year difference in coverage is due to hard copy availability in Harvard libraries.) We copy and date all financial advertisements in every issue, both to count them and to examine their content. We aggregate the information on both the number and the content of ads into quarterly series. In particular, we keep track of who the advertiser is, whether it is a mutual fund and, if so, of what kind, and whether the ads include information on the fund's own past returns. Since we are interested in the persuasion of investors, we eliminate from the database business to business ads (principally investment banking ads, or other ads explicitly directed at companies). Our total sample includes 1469 ads from *Business Week* and 4971 ads from *Money*.

Figure 2 shows, for *Business Week* and *Money* separately, the share of stock mutual fund ads containing information on own past returns (absolute or relative) in all stock mutual fund ads over the sample period. It also shows, for the same period, the rolling one quarter lagged return on the S&P 500 index, the most common indicator of broad market returns. As Figure 2 shows, on average only about 60 percent of the stock mutual fund ads present any data on own past returns, and the correlation between the share of ads including these returns and the past market return is over .7 for both *Money*

and *Business Week*.[30] Indeed, Figure 2 makes clear that, after the market crash, past returns data disappear from advertisements. This finding is broadly consistent with the predictions of our model, namely that the inclusion of past returns data is used to frame mutual fund investing as grabbing an opportunity rather than as hiring advice.

One might object that this evidence is best explained by a simpler theory that funds only like to report good news, and the news is bad in down markets. If this were correct, then stock funds with good *relative* return news to report should report them in down markets as well, since they can always do so. Corollary 3 predicts, in contrast, that even good news about relative returns should not be reported to avoid the grabbing opportunities categorization. Is this prediction borne out by the data? Do companies really avoid advertising good relative returns in down markets, as our model predicts?

Figure 3 addresses this hypothesis. It shows the relevant data for T. Rowe Price, a mutual fund complex that is the most frequent advertiser in our sample. We supplement our advertising data with a sample of T. Rowe Price stock mutual funds with assets over $300 million at the beginning of the sample period, so we can compute the number of large T. Rowe Price stock mutual funds that outperform the market. Figure 3 shows that T. Rowe Price places a lot of ads during this period, and that it has many funds outperforming S&P 500 after 1999. If anything, the number of stock funds with good relative performance rises sharply during 2001-2002. Nonetheless, both the number of stock mutual fund ads, and the number of such ads reporting returns, falls to near zero after the market declines. Even though T. Rowe Price has many funds with positive relative performance, it chooses not to advertise them. This finding is, again, broadly

[30] These correlations fall by about .2, but remain highly statistically significant, if we detrend all series using linear time trends.

consistent with our model, in which advertising returns primes the "opportunities" frame, which is unattractive to investors in down markets.

6. Conclusion

This paper has supplied a formal model of associative thinking reflecting ideas about inference and persuasion from such diverse fields as linguistics, psychology, politics, marketing, and advertising. The main idea is that individuals "think coarsely": they group situations into categories, and apply the same model of inference to all situations within a category. Coarse thinking exhibits two features that persuaders take advantage of: (i) transference, whereby individuals transfer the informational content of a given message from situations in a category where it is useful to those where it is not, and (ii) framing, whereby objectively useless information influences individuals' choice of category. The model includes full Bayesian rationality as a limiting case, in which each situation is evaluated as if in its own category.

The model sheds light on several phenomena. It explains how "soft" messages with little informational content can be persuasive, especially in low involvement situations such as choosing inexpensive goods or evaluating political candidates. It helps dissect the content of successful advertisements. It illuminates product branding. And it helps account for some features of mutual fund advertising, such as the pro-cyclical inclusion of returns, which are hard to rationalize in any conventional model of informative communication.

Our paper is just a first step in the analysis of uninformative persuasion. Although we have allowed for category choice, we have not allowed for fluid categories,

which can accommodate much more creativity on the part of the persuader (Lakoff 1987). We have also focused on associative thinking rather than on associative feeling; there are no automatic quick judgments in our model. In persuasion, such feelings are likely to play an important role as well.

Harvard University

Appendix 1: Proofs

Proof of Lemma:

First I will show that part (i) holds so long as

$$\hat{\sigma}(m \mid r, x, s') = \begin{cases} 1 \text{ if } m = x \\ 0 \text{ if } m \neq x \end{cases} \tag{A1.1}$$

and

$$\hat{p}(x \mid r, m, s') = \begin{cases} 1 \text{ if } x = m \\ 0 \text{ if } x \neq m \end{cases} \tag{A1.2}$$

for $s' \in \{1,2\}$. Note that conditions (A1.1) and (A1.2) hold whether the audience is sophisticated or takes messages at face value.

To establish part (i), I need to show that

$$C(r,m) = \underset{C \in \{C_1, C_2\}}{\arg\max}\, \hat{p}(s \in C \mid r, m) \tag{A1.3}$$

does not depend on $\hat{\sigma}(m \mid r, x, s = 0)$ and $\hat{p}(x \mid r, m, s = 0)$. To this end,

$$\begin{array}{rcll} C_1 & = & \underset{C \in \{C_1, C_2\}}{\arg\max}\, \hat{p}(s \in C \mid r, m) & \Leftrightarrow \\ \hat{p}(s \in C_1 \mid r, m) & \geq & \hat{p}(s \in C_2 \mid r, m) & \Leftrightarrow \\ \hat{p}(s = 1 \mid r, m) & \geq & \hat{p}(s = 2 \mid r, m) & \Leftrightarrow \\ \dfrac{\hat{p}(r, m \mid s = 1) p(s = 1)}{\sum_{s'} \hat{p}(r, m \mid s') p(s')} & \geq & \dfrac{\hat{p}(r, m \mid s = 2) p(s = 2)}{\sum_{s'} \hat{p}(r, m \mid s') p(s')} & \Leftrightarrow \\ \hat{p}(r, m \mid s = 1) p(s = 1) & \geq & \hat{p}(r, m \mid s = 2) p(s = 2) & \Leftrightarrow \\ p(r, m \mid s = 1) p(s = 1) & \geq & p(r, m \mid s = 2) p(s = 2) & \end{array} \tag{A1.4}$$

Note that this last condition is independent of $\hat{\sigma}(m \mid r, x, s = 0)$ and $\hat{p}(x \mid r, m, s = 0)$.

Now, to prove parts (ii) and (iii) of the lemma it is sufficient to show that $E[q \mid r, m, s = 0]$ and $E^{C(r,m)}[q \mid r, m, s = 0]$ do not depend on $\hat{\sigma}(m \mid r, x, s = 0)$ and $\hat{p}(x \mid r, m, s = 0)$.

First consider $E[q \mid r, m, s = 0]$:

$$\begin{aligned} E[q \mid r, m, s = 0] &= \sum_{x' \in \{a,b\}} E[q \mid r, x', s = 0] \hat{p}(x' \mid r, m, s = 0) \\ &= \sum_{x' \in \{a,b\}} E[q \mid r, s = 0] \hat{p}(x' \mid r, m, s = 0) \\ &= E[q \mid r, s = 0] \end{aligned} \tag{A1.5}$$

for any $\hat{\sigma}(m \mid r, x, s = 0)$ and $\hat{p}(x \mid r, m, s = 0)$, where the second equality follows from the uninformative persuasion assumption. Equation (A1.5) establishes the lemma for the case where the audience consists of Bayesians.

Now consider $E^{C(r,m)}[q \mid r, m, s = 0]$. Fixing $C(r, m) = C_i$:

$$\begin{aligned} E^{C_i}[q \mid r, m, s = 0] &= E[q \mid r, m, s = 0] p(s = 0 \mid C_i) + E[q \mid r, m, s = i] p(s = i \mid C_i) \\ &= E[q \mid r, s = 0] p(s = 0 \mid C_i) + \left[\sum_{x' \in \{a,b\}} E[q \mid r, x', s = i] \hat{p}(x' \mid r, m, s = i) \right] p(s = i \mid C_i) \\ &= E[q \mid r, s = 0] p(s = 0 \mid C_i) + E[q \mid r, x = m, s = i] p(s = i \mid C_i) \end{aligned} \tag{A1.6}$$

for any $\hat{\sigma}(m \mid r, x, s = 0)$ and $\hat{p}(x \mid r, m, s = 0)$. ■

Proof of Propositions 5 and 6:

First, note that a necessary and sufficient condition for any given strategy of a consumer (Bayesian or coarse) to be a best response to the producer's strategy is that it specifies buying wine $w = i \neq j$ with positive probability for prices (p_1, p_2) and labels (m_1, m_2) only if

$$\begin{array}{lcll} \bar{u} + q_i^e - p_i & \geq & \bar{u} + q_j^e - p_j & \text{IC} \\ \bar{u} + q_i^e - p_i & \geq & 0 & \text{IR} \end{array} \tag{A1.7}$$

and places positive probability on not buying either wine only if

$$0 \geq \max_{k \in \{1,2\}} \left\{ \bar{u} + q_k^e - p_k \right\} \tag{A1.8}$$

where $q_i^e = E[q \mid m_i, US]$ if the consumer is a Bayesian and $q_i^e = E^C[q \mid m_i, US]$ if the consumer is a coarse thinker.

The monopolist maximizes its profit, $\prod = p_1 D_1(p_1, p_2, \delta) + p_2 D_2(p_1, p_2, \delta) - \delta_1 c - \delta_2 c$, given consumers' strategies, where $D_i \in (0,1)$ denotes the total demand for wine $w = i$ given consumers' strategies and δ_i is an indicator variable taking on the value of 1 if and only if the monopolist labels wine $w = i$ "Burgundy".

Consider a possible equilibrium where both wines are labeled "Table". In any such equilibrium, it must be the case that $p_1 = p_2 \equiv p$ since we confine attention to equilibria in which both wines are sold with positive probability on the equilibrium path and wine $w = i$ would face zero demand due to consumers' incentive compatibility (IC) constraints if $p_i > p_j$. Also, in any such equilibrium, $p = \bar{u}$. To see this, suppose first

that $p > \bar{u}$. Then both wines would face zero demand by consumers' IR constraints and the producer would want to deviate and charge $p = \bar{u}$, thereby selling to the Bayesians and making positive profit. The producer would also not want to set $\bar{u} + E^C[q \mid Table, US] < p < \bar{u}$ or $p < \bar{u} + E^C[q \mid Table, US]$ because it would wish to deviate by setting $p = \bar{u}$ and $p = \bar{u} + E^C[q \mid Table, US]$ respectively. It is thus left to check that the producer also would not wish to set $p = \bar{u} + E^C[q \mid Table, US]$ and sell to both Bayesians and coarse thinkers. The producer's payoff is $\bar{u} + E^C[q \mid Table, US]$ if it sets $p = \bar{u} + E^C[q \mid Table, US]$ whereas it is $(1-\beta)\bar{u}$ if it sets $p = \bar{u}$. The second payoff is greater than the first since $\beta\bar{u} < -E^C[q \mid Table, US]$ by assumption.

We have established that in any equilibrium where both wines are labeled "Table" the producer charges $p = \bar{u}$ for each wine (selling only to Bayesians) and earns $(1-\beta)\bar{u} \equiv \prod(NoDiff)$. (*)

Now consider a possible equilibrium where the producer labels only one wine (say $w = 2$) "Burgundy". In any such equilibrium, the producer clearly charges $p_1 = \bar{u}$ and $p_2 = \bar{u} + E^C[q \mid Burgundy, US]$ since these are the highest prices it can charge while still satisfying consumers' IR constraints. In such an equilibrium, Bayesians buy $w = 1$ since $p_1 = \bar{u} < p_2 = \bar{u} + E^C[q \mid Burgundy, US]$ and coarse thinkers buy $w = 2$ since $\bar{u} + E^C[q \mid Burgundy, US] - p_2 = 0 > \bar{u} + E^C[q \mid Table, US] - p_1 = E^C[q \mid Table, US]$. The payoff to the producer in such an equilibrium is

$$\begin{aligned} (1-\beta)\bar{u} + \beta(\bar{u} + E^C[q \mid Burgundy, US]) - c &= \beta E^C[q \mid Burgundy, US] + \bar{u} - c \\ &\equiv \prod(Diff) \end{aligned} \quad \text{(A1.9)}$$

Thus, we have established that in any equilibrium where wine *w = 2* is labeled "Burgundy" the producer charges $p_1 = \overline{u}$ for wine *w = 1* (selling only to Bayesians), charges $p_2 = \overline{u} + E^C[q \mid Burgundy, US]$ for wine *w = 2* (selling only to coarse thinkers), and earns $\prod(Diff)$. (**)

By similar logic, the producer earns at most $\max\left\{\overline{u}, \beta\left(\overline{u} + E^C[q \mid Burgundy, US]\right)\right\} - c < \prod(Diff)$ if it labels both wines "Burgundy" so it will not do so in equilibrium.

Finally, comparing $\prod(NoDiff)$ with $\prod(Diff)$, we see that $\prod(Diff) > \prod(NoDiff)$ iff $\beta E^C[q \mid Burgundy, US] + \overline{u} - c > (1-\beta)\overline{u} \Leftrightarrow c < \beta\left(\overline{u} + E^C[q \mid Burgundy, US]\right) = \overline{c}$. It follows from this that (i) the producer labels exactly one wine "Burgundy" when $c < \overline{c}$, (ii) the producer labels both wines "Table" when $c > \overline{c}$, and (iii) there exists an equilibrium such that the producer labels both wines "Table" when $c = \overline{c}$. (***)

The statements of Propositions 5 and 6 follow from (*), (**), and (***).■

Appendix 2: "More Bayesian" Updating Rule

In this appendix, we explore the alternative, "more Bayesian," updating rule for coarse thinkers presented in equations (2.4)-(2.5). Recall that this update rule is different in one key way from that in our primary model: the coarse thinker conditions on the information received in weighing the expectation of quality in a situation. Under this update rule, conditional on choosing a category (which still depends on the data received), the coarse thinker can simply be thought of as a Bayesian whose information set contains only the knowledge of the category but not that of the situation. We state and prove natural analogs to Propositions 2 and 3 under these alternative assumptions and demonstrate that the results presented in Section 2 are largely robust.

Equilibrium is defined as in the main text for the case of a coarse, sophisticated audience except the persuader is now assumed to maximize

$$\begin{array}{ll} E[q \mid r, m, C(r,m)] & \text{if } m = x \\ E[q \mid r, m, C(r,m)] - c & \text{if } m \neq x \end{array} \tag{A2.1}$$

(rather than (3.2)) taking the audience's beliefs $\hat{p}$ as given. The persuader is allowed to use mixed strategies.

We now characterize some optimal (equilibrium) strategies of the persuader when the audience consists of sophisticated coarse thinkers under this alternative updating rule. Fix public information r and, without loss of generality, assume that $C(r) = C_1$. First, consider the case where messages are not pivotal given r. Also, without loss of generality, suppose that private signal $x = a$ is more favorable than private signal $x = b$ in $s = 1$ in the following sense:

$$E[q \mid r, a, s = 1]p(s = 1 \mid r, a, C_1) \geq E[q \mid r, b, s = 1]p(s = 1 \mid r, b, C_1) \tag{A2.2}$$

We have the following proposition:

Proposition A2.1 (Transference): Suppose individuals are sophisticated coarse thinkers under the modified updating rule, that messages are not pivotal given public information *r*, and that condition (A2.2) holds. Then an optimal strategy of the persuader in situation *s = 0* may involve the creation of a message. Specifically, so long as

$$c \quad < \quad E[q \mid r, a, s = 1]p(s = 1 \mid r, a, C_1) - E[q \mid r, b, s = 1]p(s = 1 \mid r, b, C_1) \qquad \text{(A2.3)}$$

it cannot be optimal for the persuader in situation *s = 0* to always report the private signal. Further, if the inequality in (A2.3) holds, then there always exists an optimal strategy such that the persuader replaces *x = b* with *m = a* with positive probability given *r*.

Proof: See online appendix. ■

We now consider the case where categorization depends on the exact message the persuader sends, i.e., one message is pivotal. To limit the number of cases, suppose that message *m = a* is pivotal and that the persuader weakly "prefers" private signal *x = a* over private signal *x = b*:

$$E[q \mid r, a, s = 2]p(s = 2 \mid r, a, C_2) \geq E[q \mid r, b, s = 1]p(s = 1 \mid r, b, C_1) \qquad \text{(A2.4)}$$

Since message *m = a* is pivotal, it does not matter whether or not (A2.2) still holds.

We then have:

Proposition A2.2 (Framing): Suppose individuals are sophisticated coarse thinkers under the modified updating rule, that message *m = a* is pivotal given public information *r*, and that condition (A2.4) holds. Then an optimal strategy of the persuader in situation *s = 0* may involve the creation of a message. Specifically, so long as

$$c \quad < \quad E[q \mid r, a, s = 1]p(s = 2 \mid r, a, C_2) - E[q \mid r, b, s = 1]p(s = 1 \mid r, b, C_1) \qquad \text{(A2.5)}$$

it cannot be optimal for the persuader in situation $s = 0$ to always report the private signal. Further, if the inequality in (A2.5) holds, then there always exists an optimal strategy such that the persuader replaces $x = b$ with $m = a$ with positive probability given r.

Proof: See online appendix. ■

References

Akerlof, George, "The Market for Lemons," *Quarterly Journal of Economics* 84 (1970), 488-500.

Aragones, Enriqueta, Itzhak Gilboa, Andrew Postlewaite, and David Schmeidler, "Rhetoric and Analogies," Tel Aviv University Mimeo, 2001.

_____, "Fact-Free Learning," *American Economic Review* 95 (2005), 1355-1368.

Barberis, Nicholas, and Andrei Shleifer, "Style Investing," *Journal of Financial Economics* 68 (2003), 161-199.

Becker, Gary, "Rational Indoctrination and Persuasion," University of Chicago Mimeo, 2001.

Becker, Gary, and Kevin Murphy, "A Simple Theory of Advertising as a Good or Bad," *Quarterly Journal of Economics* 108 (1993), 941-964.

Bertrand, Marianne, Dean Karlin, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman, "What's Psychology Worth? A Field Experiment in Consumer Credit Markets," NBER Working Paper No. 11892, 2006.

Carhart, Mark, "On Persistence in Mutual Fund Performance," *Journal of Finance* 52 (1997), 57-82.

Carpenter, Gregory, Rashi Glazer, and Kent Nakamoto, "Meaningful Brands from Meaningless Differentiation: The Dependence on Irrelevant Attributes," *Journal of Marketing Research* 31 (1994), 339-350.

Chevalier, Judith, and Glenn Ellison, "Risk Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy* 105 (1997), 1167-1200.

Crawford, Vincent, "Lying for Strategic Advantage," *American Economic*

Review 93 (2003), 133-149.

Crawford, Vincent, and Joel Sobel, "Strategic Information Transmission," *Econometrica* 50 (1982), 1431-1451.

Cronqvist, Henrik, "Advertising and Portfolio Choice," University of Chicago Mimeo, 2005.

Daniel, Kent, and Sheridan Titman, "Market Reactions to Tangible and Intangible Information," *Journal of Finance* 61 (2006), 1605-1644.

DellaVigna, Stefano, and Ethan Kaplan, "The Fox News Effect," *Quarterly Journal of Economics* 122 (2007), 1187-1234.

Dewatripont, Mathias, and Jean Tirole, "Advocates," *Journal of Political Economy* 107 (1999), 1-39.

Downs, Anthony, *An Economic Theory of Democracy*, (New York, NY: Harper and Row, 1957)

Edelman, Gerald, *Bright Air, Brilliant Fire: On the Matter of the Mind*, (New York, NY: BasicBooks, 1992)

Ellison, Glenn, "A Model of Add-On Pricing," *Quarterly Journal of Economics* 120 (2005), 585-638.

Ettinger, David, and Philippe Jehiel, "Toward a Theory of Deception," Mimeo, 2007.

Eyster, Erik, and Matthew Rabin, "Cursed Equilibrium," *Econometrica* 73 (2005), 1623 -1672.

Fryer, Roland, and Matthew Jackson, "A Categorical Model of Cognition and Biased Decision-Making," Working Paper, 2004.

Gabaix, Xavier, and David Laibson, "Shrouded Attributes, Consumer Myopia,

and Information Suppression in Competitive Markets," *Quarterly Journal of Economics* 121 (2006), 505-540.

Gentzkow, Matthew, and Jesse Shapiro, "Media Bias and Reputation," *Journal of Political Economy* 114 (2006), 280-316.

Gilboa, Itzhak, and David Schmeidler "Case-Based Decision Theory," *Quarterly Journal of Economics* 110, (1995), 605-639.

Gilovich, Thomas, "Seeking the Past in the Future: The Effect of Associations to Familiar Events on Judgments and Decisions," *Journal of Personality and Social Psychology* 40 (1981), 797-808.

Glaeser, Edward, "The Political Economy of Hatred," *Quarterly Journal of Economics* 120 (2005), 45-86.

Glaeser, Edward, Giacomo Ponzetto, and Jesse Shapiro, "Strategic Extremism: Why Republicans and Democrats Divide on Religious Values," *Quarterly Journal of Economics* 120 (2005), 1283-1330.

Glazer, Jacob, and Ariel Rubinstein, "Debates and Decisions: On a Rationale of Argument Rules," *Games and Economic Behavior* 36 (2001), 158-173.

_____, "On Optimal Rules of Persuasion," *Econometrica* 72 (2004), 1715-1736.

Goffman, Erving, *Frame Analysis*, (Cambridge, MA: Harvard University Press, 1974)

Grossman, Sanford, "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics* 24 (1981), 461-483.

Grossman, Sanford, and Oliver Hart, "Disclosure Laws and Takeover Bids," *Journal of Finance* 35 (1980), 323-334.

Holden, Richard, and Emir Kamenica, "How to Persuade a Bayesian," Mimeo, 2006.

Hong, Harrison, Jeremy Stein, and Jialin Yu, "Simple Forecasts and Paradigm Shifts," *Journal of Finance* 62 (2007), 1207-1242.

Jehiel, Philippe, "Analogy-Based Expectation Equilibrium," *Journal of Economic Theory* 123 (2005), 81-104.

Jehiel, Philippe, and Frederic Koessler, "Revisiting Games of Incomplete Information with Analogy-Based Expectations," Mimeo, UCL, 2006.

Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47 (1979), 263-292.

_____, *Judgment Under Uncertainty: Heuristics and Biases*, (New York, NY: Cambridge University Press, 1982)

Kartik, Navin, Marco Ottaviani, and Francesco Squintani, "Credulity, Lies, and Costly Talk," *Journal of Economic Theory* 134 (2007), 93-116.

Kihlstrom, Richard, and Michael Riordan, "Advertising as a Signal," *Journal of Political Economy* 92 (1984), 427-450.

Kreps, David, and Robert Wilson, "Sequential Equilibria," *Econometrica* 50 (1982), 863 -894.

Krueger, Joachim, and Russell Clement, "Memory-Based Judgments About Multiple Categories," *Journal of Personality and Social Psychology* 67 (1994), 35-47.

Lakoff, George, *Women, Fire, and Dangerous Things*, (Chicago, IL: The University of Chicago Press, 1987)

Luntz, Frank, *Frank Luntz Strategy Report*, Frank Luntz Research Companies, 2006.

Malt, Barbara, Brian Ross, and Gregory Murphy, "Predicting Features for Members of Natural Categories When Categorization is Uncertain," *Journal of Experimental Psychology* 21 (1995), 646-661.

McCloskey, Donald, and Arjo Klamer, "One Quarter of GDP is Persuasion," *American Economic Review Papers and Proceedings* 85 (1995), 191-195.

Milgrom, Paul, "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics* 12 (1981), 380-391.

Milgrom, Paul, and John Roberts, "Price and Advertising as Signals of Product Quality," *Journal of Political Economy* 94 (1986a), 796-821.

_____, "Relying on the Information of Interested Parties," *Rand Journal of Economics* 17 (1986b), 18-32.

Mullainathan, Sendhil, "Thinking through Categories," Mimeo, MIT, 2000.

Mullainathan, Sendhil, and Andrei Shleifer, "Media Bias," NBER Working Paper No. 9295, 2002.

_____, "The Market for News," *American Economic Review* 95 (2005), 1031-1053.

_____, "Persuasion in Finance," NBER Working Paper No. 11838, 2006.

Murphy, Gregory, and Brian Ross, "Predictions from Uncertain Categorizations," *Cognitive Psychology* 27 (1994), 148-193.

Murphy, Kevin, and Andrei Shleifer, "Persuasion in Politics," *American Economic Review Papers and Proceedings* 94 (2004), 435-439.

Nelson, Philip, "Advertising as Information," *Journal of Political Economy* 82 (1974), 729-754.

Okuno-Fujiwara, Masahiro, Andrew Postlewaite, and Kotaro Suzumura, "Strategic Information Revelation," *Review of Economic Studies* 57 (1990), 25-47.

Peski, Marcin, "Categorization," University of Chicago Working Paper, 2006.

Peter, Paul, and Jerry Olson, *Consumer Behavior and Marketing Strategy*, (New York: McGraw Hill, 2005)

Rapaille, Clotaire, *The Culture Code*, (New York, NY: Broadway Books, 2006)

Reuter, Jonathan, and Eric Zitzewitz, "Do Ads Influence Editors? Advertising and Bias in the Financial Media," *Quarterly Journal of Economics* 121 (2006), 197-228.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets," *Journal of Political Economy* 82 (1974), 34-55.

Shapiro, Jesse, "A 'Memory-Jamming' Theory of Advertising," Mimeo, University of Chicago, 2006.

Shin, Hyun Song, "News Management and the Value of Firms," *Rand Journal of Economics* 25 (1994), 58-71.

Sirri, Erik, and Peter Tufano, "Costly Search and Mutual Fund Flows," *Journal of Finance* 53 (1998), 1589-1622.

Smith, Eliot, "Mental Representations and Memory," in *Handbook of Social Psychology*, 4th ed., Daniel Gilbert, Susan Fiske, and Gardner Lindzey, eds., pp. 391-445 (New York: McGraw-Hill, 1998)

Stigler, George, "The Economics of Information," *Journal of Political Economy* 69 (1961), 213-225.

_____, *The Theory of Price*, 4th ed. (New York, NY: MacMillan Publishing, 1987)

Stigler, George, and Gary Becker, "*De Gustibus Non Est Disputandum*," *American Economic Review* 67 (1977), 76-90.

Sutherland, Max, and Alice Sylvester, *Advertising and the Mind of the Consumer*, 2nd ed. (St. Leonards, Australia: Allen & Unwin, 2000)

Swensen, David, *Unconventional Success: A Fundamental Approach to Personal Investment*, (New York: Free Press, 2005)

Tirole, Jean, *The Theory of Industrial Organization*, (Cambridge, MA: MIT Press, 1988)

Zaltman, Gerald, "Rethinking Market Research: Putting People Back In," *Journal of Marketing Research* 34 (1997), 424-437.