8/5/2014

Behavioral Finance

David Hirshleifer
Merage School of Business
University of California, Irvine
david.h@uci.edu
(614) 214-4304

When citing this paper, please use the following: Hirshleifer D, 2014. Title. Annu. Rev. Econ. 7.: Submitted. Doi: 10.1146/annurev-financial-092214-043752

Keywords: Investor psychology, heuristics, overconfidence, attention, feelings, reference dependence, social finance

JEL code: G02 - Behavioral Finance: Underlying Principles

Abstract:

Behavioral finance studies the application of psychology to finance, with a focus on individual-level cognitive biases. I describe here the sources of judgment and decision biases, how they affect trading and market prices, the role of arbitrage and flows of wealth between more rational and less rational investors, how firms exploit inefficient prices and incite misvaluation, and the effects of managerial judgment biases. There is need for more theory and testing of the effects of feelings on financial decisions and aggregate outcomes. Especially, the time has come to move beyond behavioral finance to social finance, which studies the structure of social interactions, how financial ideas spread and evolve, and how social processes affect financial outcomes.

I thank Nicholas Barberis, James Choi, Bing Han, Gilles Hilary, Narasimhan Jegadeesh, Danling Jiang, Sonya Lim, Siew Hong Teoh, Anjan Thakor, Ivo Welch, Wei Xiong, and Liyan Yang for helpful comments, and Lin Sun and Qiguang Wang for excellent research assistance and comments.
Contents

1. Introduction

2. Market mispricing, arbitrage, and financial agents
   a. Arbitrage
   b. Financial agents

3. Psychological foundations

4. Overconfidence and self-esteem maintenance
   a. Psychology of overconfidence
   b. Investor overconfidence and self-esteem maintenance
   c. Managerial and advisor overconfidence and overoptimism

5. Limited attention and cognitive processing
   a. Failure to process signals and features of the decision environment
   b. Neglecting basic features of the decision environment
   c. Financial theories of category thinking
   d. Reference-dependence and framing
   e. Conceptual discretizing, loss aversion, and probability weighting
   f. Mental accounting and realization preference
   g. Heuristic learning

6. Feelings
   a. Familiarity and liking
   b. Financial theories based on feelings
c. Evidence on financial effects of familiarity and in-group bias

d. Sentiment, shifting optimism and risk tolerance

7. Firm behavior: Exploiting versus inciting misvaluation

   a. Theories of exploitive advisors and firms

   b. Evidence on exploitive advisors and firms

   c. Misvaluation, new issues and repurchase, and post-event returns

8. Conclusion: Behavioral finance versus social finance
1. Introduction

The stock price of EntreMed jumped about 600% in one weekend upon the republishation of information that was already publicly available five months earlier about a new cancer drug (Huberman & Regev (2001)). This violated the Efficient Market Hypothesis, which asserts that prices accurately reflect publicly available information. The Efficient Market Hypothesis is based on the idea that most, or at least the most important, investors are rational in processing information. Behavioral finance, in contrast, studies how people fall short of this ideal in their decisions, and how markets are, to some degree, inefficient.

The rise of behavioral finance over the last three decades has been felt throughout finance and economics. Many scholars are now ready to entertain the consequences of either rational or irrational aspects of human judgment, as relevant for the particular application at hand. This readiness is greatest for errors by individual market participants; vigorous debate continues about how psychological bias affects price determination in large and liquid markets. Nevertheless, a modern understanding of the finance field requires grounding in psychological as well as rational approaches. Today many of the leading theories about such fundamental topics as investor behavior, the cross-section of returns, corporate investment, and money management, derive from psychological factors.

Psychology has identified various judgment biases that can affect financial decision-making. Since psychological bias is the distinctive feature of behavioral finance, I organize this review by the type of bias (see also Shiller (1999)). Also, rather than viewing the psychology of judgment and decisions as a congeries of inexplicable facts, I organize the discussion of biases
around a relatively small number of underlying evolutionary and psychological roots. Then, I discuss financial theories founded upon each type of bias, and the evidence bearing upon them.

Some fundamentals of behavioral finance do not inherently depend on the specific psychological source of bias. So I discuss separately the topics of how arbitrage and flows of wealth promote market efficiency, how firms induce or react to mispricing, and how investor sentiment affects security markets.

The main focus of this review is on the effects, individual or aggregate, of individual-level bias. The topic of social processes, discussed in the conclusion, deserve greater attention in finance, and a separate review. Also, I do not go deeply here into distinguishing the effects of psychological bias from rational risk effects (see, e.g., the review of Daniel et al. (2002)).

Some surveys focus more heavily on issues that cut across different psychological biases, such as limits to arbitrage (Gromb & Vayanos (2010)), noise trading (Shleifer (2000)), and how valuations affect corporate behavior (Baker (2009)). For a greater focus on prospect theory, see the excellent survey of Barberis & Thaler (2003); neurofinance, Bernheim (2009); experimental economics and asset markets, Smith (2008); investments and asset pricing, Hirshleifer (2001); behavioral corporate finance, Baker & Wurgler (2012); behavioral accounting, Libby et al. (2002) and Hirshleifer & Teoh (2009a); and policy, regulation, or field experiments, Thaler & Sunstein (2008), Hirshleifer (2008), and Card et al. (2011).

2. Market mispricing, arbitrage, and financial agents

   a. Arbitrage
Arbitrage is the purchase or sale of goods to profit from differences in effective prices across trading venues. The term is used broadly to refer to the exploitation of profit opportunities whenever some assets are overpriced relative to others, based on the idea that buying cheap assets and selling similar but expensive ones can yield a relatively low-risk return. In perfect markets, arbitrage opportunities are limited by the risk aversion of investors and the riskiness of trading the mispriced asset (DeLong et al. (1991)).

An oft-neglected fact is that arbitrage is a double-edged blade that can make prices either more or less efficient. In asset market equilibrium under disagreement, price reflects a weighted average of beliefs. So both the irrational impellers of mispricing and the more rational correctors of it believe that they are performing profitable arbitrage against inefficient market prices. Whether greater arbitrage capital reduces mispricing therefore depends on whether this capital is wielded by 'smart' investors—those who are both rational and, if money managers, not pandering to the mistaken beliefs of irrational investors about what is a profit opportunity.

A powerful argument for why markets are often highly efficient is that in the long run wealth tends to flow to smart arbitrageurs, who end up dominating the market. However, irrational investors can earn higher expected profits than rational ones by bearing higher risk (DeLong et al. (1991)), or by inducing self-validating feedback into fundamentals (Hirshleifer et al. (2006)). Alternatively, rationality can falter if investing success increases subsequent bias (Daniel et al. (1998); Gervais & Odean (2001)).

If wealth does flow to smart investors, their influence on prices increases, owing either to credit constraints or decreasing risk aversion. However, this process is often slow, as strategy
performance is typically a very noisy indicator of ability (Yan (2008)). Meanwhile, new naïve money flows into markets each day; the succession of generations reshuffles wealth and talent.

If irrational investors misvalue the idiosyncratic components of the fundamental payoffs of many securities, if markets are frictionless, and if rational and irrational investors all bet on many securities, then owing to the large number of bets, the flow of wealth becomes swift and almost sure. This causes rational investors to acquire all the wealth very quickly. However, if most investors only place active bets on subsets of securities, the rate of wealth flow can be modest, accommodating relatively substantial and persistent mispricing (Daniel et al. (2001)).

b. Financial agents

It is usually supposed that institutional money managers and professional investment advisors are smart arbitrageurs, acting on behalf of less sophisticated individual investors. Sophisticated investors perform careful analysis to learn about biases of investors or consequent mispricing, and the insight derived thereby can be used to educate clients or to deploy client funds to achieve high returns. However, owing to conflict of interest, or to imperfect rationality of investment professionals, employing agents is an imperfect remedy for ignorance and folly. Money managers often pander to investor irrationality, in order to attract inflows.

This does not make financial advice and delegation pure evils. For example, in the model of Gennaioli et al. (2014), `money doctors' skim off some of the gains from investment, but still increase welfare by encouraging otherwise-distrustful individuals to participate in the market.
As for whether the ability of irrational investors to hire exploitive agents improves the efficiency of prices, there is no general unambiguous answer. So optimism about the inevitability of reaching almost perfect market efficiency must be tempered by recognition that agents may exacerbate investor bias. Furthermore, when, by chance, mispricing gets worse, smart arbs lose money on their existing positions and have more trouble raising funds. So corrective arbitrage pressure on price is weakest when it is needed the most (Shleifer & Vishny (1997)).

Owing to heavier total pressure from irrational investors speculating about systematic factors, we typically expect greater mispricing of factors than of idiosyncratic payoff components, except for idiosyncratic opportunities that arbs simply do not notice (Daniel et al. (2001)). For example, the book-to-market and accrual characteristics are associated with return comovement (Fama & French (1993); Hirshleifer et al. (2012)), so if the value and accrual anomalies (both discussed later) represent mispricing, they are probably relatively hard to arbitrage away.

3. Psychological foundations

Since people need to make judgments and decisions quickly using limited cognitive resources, they necessarily use shortcuts (Simon (1956); Kahneman et al. (1982)), often called “heuristics.” All thinking builds upon cognitive algorithms that operate automatically below the level of consciousness. The term “heuristics” encompasses both innate and automatic processes, and learned or consciously selected rules of thumb.
Heuristics often work well within some domains and for some types of problems, but badly in others. *Heuristic simplification* implies more errors for decision problems that range farther from the types of problems that the human mind evolved to deal with in the ancestral past.

In dual process theories of cognition, an automatic, non-deliberative system quickly generates perceptions and judgments; a slower, more effortful system monitors and revises such judgments as time and circumstances permit (Stanovich (1999); Kahneman (2011)). Following Haidt and Kesebir (2010), I refer to the fast process as the *intuitive system*, and the slow process as the *reasoning system*.

Kahneman (2011) describes human thinking as largely intuitive, and heavily influenced by the associations that are triggered by the presentation of a decision problem. People are overconfident that their intuitive way of thinking about a problem is correct; information that does not immediately come to mind tends to be completely neglected, a phenomenon that Kahneman calls WYSIATI (What You See Is All There Is).

*Feelings* provide the value weights assigned to possible outcomes to motivate decisions and actions. Affective reactions can also facilitate making fast use of urgent information about the environment (as in the *affect heuristic*; Slovic et al. (2002)). For example, a risky investment opportunity may trigger fear and, thereby, useful hesitation.

However, feelings often short-circuit useful analysis, as with exiting the stock market in sudden panic, or buying a hot stock based on enthusiasm rather than critical evaluation. Such
affective short-circuiting can also create self-discipline problems, such as not saving for retirement.

In modern financial markets, there are great benefits to making decisions analytically rather than relying solely upon feelings and intuition. Intuition-generating mechanisms suited to the human ancestral environment provide poor guidance for decisions in modern markets and large economies.

Beliefs have a social-signaling as well as a decision-making role. In the theory of Trivers (1991), people overestimate their personal merits so as to be more persuasive to others about them. Such self-deception comes at the cost of errors deriving from overconfident beliefs.

The three abovementioned elements—heuristic simplification, affective short-circuiting, and self-deception—explain most of the psychological biases studied in behavioral finance. These elements also underlie the dynamic psychological updating processes that maintain biases despite having opportunities to learn from past errors.

4. Overconfidence and self-esteem maintenance

a. Psychology of overconfidence

An immediate consequence of self-deception is that people will be overconfident about their merits of various sorts. In overprecision, people think that their judgments are more accurate than they really are. Overconfidence tends to be stronger when correct judgments are hard to form, such as when uncertainty is high. The difficulty effect is the finding that overprecision is stronger for challenging judgment tasks.
Recent studies both of overplacement (overestimation of one’s rank in the population) in the psychological laboratory (Benoit et al. (2014)) and the field (Merkle & Weber (2011)), and of overprecision in financial field settings, confirm that overconfidence is very strong (Ben-David et al. (2013)).

Since high ability contributes to good outcomes, overestimation of one’s merits promotes overoptimism about one’s prospects. People do tend to be overoptimistic about their life prospects (Weinstein (1980)), which affects their economic and financial decisions (Puri & Robinson (2007)).

If overconfidence is to persist as new information about ability arrives, there must be biases in updating processes that favor a positive self-assessment. Such self-enhancing attribution bias is well documented (Langer & Roth (1975)).

People tend to shift their attitudes in favor of actions they have chosen or have been induced to engage in without compensation, a phenomenon that motivates the theory of cognitive dissonance (Festinger & Carlsmith (1959)). Such shifts help people reconcile their past choices with the perception that they are good decision-makers. Self-enhancing updating promotes escalation of commitment (sticking too stubbornly to a choice despite opposing information, Staw (1976)), including the sunk cost effect (reluctance to terminate costly activities after expending resources on them; Thaler (1980)); and rationalization of one’s past behaviors (Nisbett & Wilson (1977)).

b. Investor overconfidence and self-esteem maintenance
i. Overconfidence and trading aggressiveness in static settings

Overconfidence causes investors to trade more aggressively, which tends to reduce their welfare (Odean (1998)). Overconfidence therefore helps solve the active investing puzzle: that individual investors trade individual stocks despite losing money doing so (Barber & Odean (2000)), and invest in active funds instead of indexing to obtain better net performance. Consistent with overconfidence, in experimental markets, some investors overestimate the precision of their signals, are more subject to the winner’s curse, and do worse in trading (Biais et al. (2005)).

By promoting bets on individual securities, overconfidence reduces diversification. However, as discussed later, underdiversification has other sources as well. So greater confidence, by encouraging participation in otherwise-neglected asset classes, can also promote diversification. Indeed, greater feeling of competence about investing is associated with weaker home bias in investing (discussed later; Graham et al. (2009)).

ii. Overconfidence and price overreaction in static settings

Overconfidence about some value-relevant information signal causes overreaction in prices, and therefore long-run correction (Odean (1998)). This implies negative return autocorrelations.

Any psychological force that causes overreaction to information will tend to make high price be a proxy for overvaluation and low price for undervaluation. This leads naturally to the
size (market value) effect. For example, overextrapolation of fundamentals or prices can cause such effects (Lakonishok et al. (1994)).

Scaling by a proxy for fundamentals, such as book value, cleanses market price of variation not derived from mispricing. So in the overconfidence-based capital asset pricing model of Daniel et al. (2001), fundamental-to-price ratios predict returns even more strongly, if the fundamental proxy is not too noisy. Both beta and scaled price variables such as book-to-market predict returns. Since scaled price variables capture both risk and mispricing effects, they can sometimes dominate beta in return prediction regressions even when risk is priced. Empirically, high beta-stocks do underperform (Frazzini & Pedersen (2014)).

Book-to-market is an example of how mispricing can be proxied by the deviation of market price from a benchmark that is less subject to misvaluation. Empirically, stocks with low price relative to fundamental proxies on average experience high subsequent returns. Such proxies include book value, earnings or cash flow (the value effect), past price (the winner/loser effect), or a constant (the size effect). The value effect has been confirmed in many markets and asset classes (Asness et al. (2013)).

Short-term interest rates can act as a fundamental scaling for long-term rates. So overconfidence further implies that the forward premium for bonds denominated in different currencies can negatively predict exchange rate shifts, the forward premium puzzle (Burnside et al. (2011)).

Further implications of overconfidence derive from comparative statics on its determinants. For example, the difficulty effect implies stronger overconfidence effects for
hard-to-value stocks. Consistent with this, the value effect is stronger among high R&D stocks (Chan et al. (2001)); momentum is also stronger for hard-to-evaluate stocks (as indicated by uncertainty proxies; Jiang et al. (2005)).

iii. Bias in self-attribution and trading aggressiveness in dynamic settings

In models of the dynamics of overconfidence, profits on an investor’s existing long or short position increase confidence, resulting in greater subsequent trading aggressiveness (Daniel et al. (1998)). It follows that for securities that are in positive net supply, high past returns should be associated with greater subsequent trading (Gervais & Odean (2001)).

Consistent with bias in self-attribution, trading activity by individual investors increases after they experience high returns (Barber & Odean (2002)). Similarly, investor trading and market trading volume increase after high returns (Statman et al. (2006); Griffin et al. (2007)).

iv. Overconfidence, biased self-attribution, and price under- vs. over-reactions

Bias in self-attribution implies short-run continuation of stock returns and long-run reversal. When a stock has risen, for example, relative to other stocks, in the short run this overreaction tends to continue; and, on average, it later falls, but this correction is hindered, so the decline also tends to continue. So short-run return continuation and long-run reversal together are consistent with a process of continuing overreaction and then correction (Daniel et al. (1998)). This model also implies post-event return continuation (post-event abnormal returns of the same sign on average as the event-date reaction) if firms tend to select good
news actions in response to underpricing (as with issuing overpriced shares and repurchasing underpriced shares); and continuation after earnings surprises.

Empirically, a contrasting pair of stylized facts is the tendency of stock returns to continue in the short run (positive autocorrelations with conditioning period of several months- Jegadeesh & Titman (1993)) versus a tendency to reverse in the long run (negative autocorrelations with a conditioning period of several years; DeBondt & Thaler (1985)). The short-run effect is called momentum, which is present in many asset classes in the time series (Moskowitz et al. (2012)) and the cross-section. The long-run reversal of returns is called the winner/loser effect.

Event studies typically report post-event return continuation, i.e., average post-event abnormal returns of the same sign as the event-date reaction, as summarized in Hirshleifer (2001). For example, seasoned equity issues (and IPOs, and debt issues) tend to be followed by negative abnormal returns (the new issues puzzle; Loughran & Ritter (1995); Spiess & Affleck-Graves (1995)), and repurchase by high returns (Ikenberry et al. (1995)).

Equity issuance is followed by low average market returns in many countries (Henderson et al. (2006)). At the aggregate level as well, the share of equity issues in total new equity and debt issues has been a negative predictor of U.S. market returns (Baker & Wurgler (2000)).

Also consistent with overconfidence and bias in self-attribution, earnings surprises are associated with subsequent abnormal returns of the same sign (post-earnings announcement drift, discussed in Section 5).
The ability of overconfidence and its dynamic counterpart, self-attribution bias, to explain a wide range of major patterns of return predictability is notable, but does not prove that overconfidence is the cause. Indeed, later sections discuss alternative possible psychological explanations for several of these effects. Distinguishing theories requires homing in on their distinctive implications.

v. Overconfidence, short-sales constraints, and overpricing

In the model of Miller (1977), owing to short-sale constraints, only relatively optimistic beliefs are impounded into price, resulting in overvaluation. Investors stubbornly disagree, although rationally optimists should update pessimistically based on the knowledge that there are sidelined pessimists. Such disagreement can be explained by overconfidence on the part of optimists that their own analysis is superior, or that disagreeing investors are rare (as in WYSIATI).

Empirically, dispersion of analyst forecasts is negatively associated with subsequent abnormal returns (Diether et al. (2002)). Clear examples of overpricing derived from disagreement and short-selling constraints occurred during the millennial high-tech boom, when the market value of a parent firm was sometimes substantially less than the value of its holdings in one of its publicly-traded divisions (Lamont & Thaler (2003)). Also consistent with the Miller theory, stocks with tighter short-sale constraints have stronger return predictability anomalies (Nagel (2005)), and greater long-short asymmetry in the accrual anomaly (Hirshleifer et al. (2011)).
Volatility increases the scope for disagreement, implying greater overvaluation. Empirically, stocks with high idiosyncratic risk (Ang et al. (2006)) do underperform.

In markets with short sale constraints, investors may buy overvalued stocks in the expectation of selling at an even higher price to overconfident investors. Lower available float should exacerbate such bubbles (Hong et al. (2006)), as confirmed for a bubble in Chinese warrants (Xiong & Yu (2011)).

c. Managerial and advisor overconfidence and overoptimism

A manager who is overconfident of his ability will tend to be optimistic about his firm’s prospects as well. In the model of Bernardo & Welch (2001), overconfidence has a bright side, as it encourages entrepreneurs to engage in socially desirable experimentation. Survey evidence confirms that entrepreneurs tend to be overoptimistic about their future success.

Overconfidence and overoptimism have obvious costs, but can also help shareholders by encouraging risk averse managers to take good risky or innovative projects (Campbell et al. (2011)). This leads to a benefit to matching managerial optimism or confidence appropriately to firms (Goel & Thakor (2008)). Different degrees of optimism between entrepreneurs and outside investors can result in inefficient screening of projects, creating a role for rational banks to act as a bridge between these two groups (Coval & Thakor (2005)).

i. Evidence on overconfidence, optimism, and investment and financing decisions
Several strands of evidence display both the bright and dark sides of managerial overconfidence and overoptimism suggested by theoretical models. On the dark side, bidders on average earn low returns from takeovers, more optimistic managers are more likely to make acquisitions, and the market reacts more negatively to their bids (Malmendier & Tate (2008)).

Optimistic CEOs also use less external finance, especially equity (Malmendier et al. (2011)), and finance relatively more with short-term debt (Graham et al. (2013)). The investment of firms with overoptimistic managers (as proxied by voluntarily retaining equity-like claims in the firm), is more sensitive to cash flow (Malmendier & Tate (2005)). This suggests that such managers view their firm as undervalued, making external capital seem expensive to them.

Both overconfidence and overoptimism are associated with greater corporate investment (Ben-David et al. (2013)). Potentially on the bright side, overoptimistic managers spend more on R&D, and obtain more patents relative to their R&D spending, perhaps because of greater willingness to bear risk (Hirshleifer et al. (2012)).

The optimism of analyst forecasts at long horizons suggests either that analysts are overoptimistic, or that they forecast optimistically for agency reasons (Richardson et al. (2004)). The association of analyst political attitudes with forecast optimism suggests that psychological factors play a role (Jiang et al. (2014)).

\textit{ii. Dynamics of managerial and analyst confidence}
Turning to the dynamics of managerial bias, there is evidence suggesting that managers tend to attribute good performance excessively to their own abilities rather than luck. Bias in managerial self-attribution has been found in the contexts of repeated acquisitions (Billett & Qian (2008)) and in the issuance of management earnings forecasts after past successes (Hilary & Hsu (2011)).

5. Limited attention and cognitive processing

Owing to limited attention and processing power, people tend to neglect relevant information signals and strategic features of the decision environment. This is manifested in a variety of more specific effects to be described, such as evaluation based on categories, the influence of framing and reference points on judgments, conceptual discretizing of continuous quantities, flawed tracking of costs and benefits in mental accounting, and the heuristic updating of beliefs.

a. Failure to process signals and features of the decision environment

People tend to neglect low salience signals and overreact to salient or recent news. Owing to WYSIATI, they also tend to be unaware of such errors, and hence do often not correct them. People also neglect important features of their decision environments, such as strategic motives for the actions of others. Such neglect is reflected in cognitive hierarchy models and evidence in the experimental game theory literature (Camerer et al. (2004)), and other models of neglect of strategic motives (Hirshleifer & Teoh (2003); Eyster & Rabin (2005)).

i. Financial theories of information neglect
Information sources can be biased because of inherent psychological bias, infection by public excitement, or conflict of interest. When investors do not adjust appropriately for biased signal provision, trading mistakes and mispricing follow (see Section 7.b).

In the models of Hirshleifer & Teoh (2003), Peng & Xiong (2006), and Hirshleifer et al. (2011), a subset of investors neglect a value-relevant information signal, resulting in return predictability. Examples of such signals include the deviation between GAAP and pro forma earnings, footnotes in financial statements about option compensation to managers, the breakdown of earnings between components with different value relevance (cash flows versus accruals), and earnings surprises.

Limited attention theories imply positive abnormal returns after neglected good news and negative abnormal returns after neglected bad news. Firms can temporarily increase their stock prices through earnings management, and presumably do so when the gains from having a high stock price are large.

For two reasons, limited attention causes overreactions as well as underreactions. First, investors overreact to salient news. Second, neglect of earnings components implies overreaction to the less predictive component, accruals (Sloan (1996); Hirshleifer et al. (2011)).

Hong & Stein (1999) study the interaction between “news-watchers” who condition only on signals about future cash flows and “momentum traders” who condition only on a partial history of prices. The information possessed by news-watchers is gradually incorporated into prices, and naïve momentum trading causes trends to overshoot and later correct. This generates return under- and overreactions. Momentum is strongest among low-attention
stocks owing to slower diffusion of information. Consistent with this prediction, Hong et al. (2000) find that momentum is stronger for small stocks and stocks with low analyst coverage.

\textit{ii. Financial evidence on information neglect, salience, and distraction}

\textbf{A. Investor naiveté}

Many investors are naïve in their financial beliefs, and do not understand basic concepts such as equity or diversification (Lusardi & Mitchell (2011)). Notably, there are (short-lived) episodes of extreme trading in response to egregious confusions between the abbreviated names of firms and the ticker symbols of other firms (Rashes (2001)). Such episodes suggest that more subtle confusions are rife.

\textbf{B. Evidence of pricing effects of signal neglect and neglect of strategic motives}

The introduction gave an example of high influence of salient news announcements. At the opposite extreme, there is severe neglect of non-salient information, such as that contained in demographic predictors of shifts in product demand (DellaVigna & Pollet (2007)).

A venerable anomaly is the sluggish reaction of stock prices to earnings surprises and revisions in analyst forecasts of earnings, \textit{post-earnings announcement drift} or \textit{PEAD} (Foster et al. (1984); Bernard & Thomas (1989)). The fact that subsequent returns associated with earnings surprises are concentrated at later earnings announcements, and that market reactions reflect naïve seasonal random walk expectations, support a limited attention explanation.
Accruals, the accounting adjustments made to cash flows to obtain earnings, are less positive than cash flow as a predictor of profitability. Neglect of the distinction between these earnings components, and of the incentives of managers to manage earnings, cause accruals and their abnormal ‘managed’ component to be negative predictors of returns, the *accrual anomaly* (Sloan (1996); Teoh et al. (1998a,b)). Accruals are also associated with bias in analyst forecasts (Teoh & Wong (2002)).

The accrual anomaly is based on a comparison of two non-parallel quantities, earnings and cash flow. The cash analog to earnings is *Free Cash Flow*, which is net of investment expenditures (just as earnings is net of depreciation). So the deviation between cash and accounting profitability should be a better indicator than accruals of misvaluation. Cumulating the deviations over time yields Net Operating Assets, which turns out to be a much stronger return predictor than accruals (Hirshleifer et al. (2004)).

Salience and distraction, by modulating investor attention, affect trading and mispricing. Several data confirm that information that is more salient or easier to process is incorporated more sharply into prices. The prices of country funds underreact to changes in the value of underlying assets, except when the news appears in the front page of *The New York Times* (Klibanoff et al. (1998)). Industry information is impounded into prices more rapidly in simple pure-play firms than in conglomerates that operate across industries (Cohen & Lou (2012)).

Consistent with high salience of news media coverage and the Miller (1977) disagreement model, individual investors are net buyers of stocks that have recently gained media attention, as well as stocks with high abnormal trading volume or extreme one-day
returns (Barber & Odean (2008)). Suggestive of gradual growth in net demand for stocks that have become the focus of investor attention, stocks with unusually high trading volume over a day or a week on average earn a return premium during the next month (Gervais et al. (2001)).

There should generally be greater resort to intuitive, heuristic thinking when an investor’s attentional resources are depleted, such as when there is greater decision pressure or distracting news. The sensitivity of the market reaction to earnings surprises is weaker on Fridays when attention should be low (DellaVigna & Pollet (2009)), and when the number of distracting same-day earnings announcements is large (Hirshleifer et al. (2009)), resulting in correspondingly larger post-earnings announcement drift.

b. Neglecting basic features of the decision environment

Even professionals have cognitive constraints and rely on heuristics. For example, a survey of CFOs found use of naïve capital budgeting approaches such as the payback criterion, and the use of a single discount rate to evaluate very different kinds of projects (Graham & Harvey (2001)).

In narrow framing (Kahneman & Lovallo (1993)), a decision problem is viewed in isolation from some of the factors that are relevant for it. For example, in Choi et al. (2009), individuals neglected the employer matching feature of contributions to their retirement plans, unless the decision problem was designed to force them to make integrated decisions. Under narrow framing, the addition of each asset to a portfolio is evaluated based upon whether it is viewed as inherently good or bad instead of in terms of its diversifying contribution to the overall portfolio.
In fact, people do tend to invest in excessively narrow sets of assets and asset classes. A notable stylized fact is that investors tend to eschew foreign securities, home bias (French & Poterba (1991); Tesar & Werner (1995)). This effect is stronger for investors with lower cognitive abilities and financial literacy (Grinblatt et al. (2011)). Sections 4 and 6 discuss other reasons for underdiversification.

c. Financial theories of category thinking

Behavioral explanations for comovement involve either irrational amplification of fundamental comovement, or other kinds of misperceptions. In the first approach, overconfident investors who overreact to information about fundamental factors induce return comovement (Daniel et al. (2001)).

In the model of Hirshleifer & Jiang (2010), a factor portfolio is built by going long and short on misvalued firms, and a stock’s factor loading measures the extent to which the firm inherits investor overreaction to fundamental factors. Such loadings are therefore proxies for firm-level misvaluation. Empirically, there is comovement in stock returns associated with a misvaluation factor based upon debt and equity issuance and repurchase; loadings on this factor are strong return predictors.

An alternative explanation for comovement in excess of fundamentals is that investors think heuristically about security categories. A basic mechanism of thought is classification, so that instances can be evaluated based on features of their categories (see, e.g., Ashby & Maddox (2005)). Such a heuristic is powerful, but flawed when categories are non-uniform.
In the style investing model of Barberis & Shleifer (2003), assets that share a style comove more than would be implied by fundamentals. Shifting the category of an asset raises its correlation with its new style. Owing to style-based trading, style-level momentum and value strategies are predicted to be more profitable than their asset-level counterparts. Related implications can be derived in a model that focuses explicitly on constraints on investor’s attention (Peng & Xiong (2006)).

Style investing can explain the temporary high returns of stocks upon S&P inclusion (Harris & Gurel (1986); Shleifer (1986)), comovement of stocks that share styles such as size and book-to-market, and increased comovement of stocks that are added to the S&P 500 with existing index members (Barberis et al. (2005)).

Both overreaction to fundamental factor signals, and style investing, imply comovement in excess of what would be expected rationally. Consistent with this implication, presumably-naïve retail investor trading is associated with return comovement (Kumar & Lee (2006)).

d. Reference-dependence and framing

Cognitive processes are to some extent specific to the domain of the decision problem (Cosmides & Tooby (2013)), and to the modality of presentation (graphical, numerical, or verbal; probabilities versus frequencies; see, e.g., Gigerenzer & Hoffrage (1995)). Even for given type of decision problem and modality, alternative descriptions of logically identical decision problems, such as the highlighting of a different reference for comparison of outcomes, have large effects on choices, a phenomenon known as framing (Tversky & Kahneman (1981)). Optimizing based on deviations of payoffs from reference points (a key feature of prospect
theory, discussed later in this section) implies framing effects, and therefore choices that become inconsistent as changing presentations or circumstances cause the reference point to shift.

There is extensive evidence that seemingly irrelevant reference points matter to investors and firms. Firms manage earnings to meet salient thresholds (forecasts or past earnings; DeGeorge et al. (1999)), and stock prices react sharply to even a small shortfall. Firms’ borrowing rates seem unduly influenced by previous rates (Dougal et al. (2014)). Past stock price highs affect firm and investor behavior and predict future stock and market returns (George & Hwang (2004); Baker et al. (2012)).

When individuals do not have an answer to a decision problem, they often substitute the solution to a related simpler problem, *attribute substitution* (Kahneman & Frederick (2002)). This can explain money illusion (Fisher (1928)), wherein nominal instead of real prices are used for investment decisions. In this spirit, Ritter & Warr (2002) argue that mistaken discounting at nominal interest rates induced long U.S. bear and bull markets as inflationary trends shifted.

e. **Conceptual discretizing, loss aversion, and probability weighting**

Expected utility theory cannot explain, with plausible levels of aversion to large risks, the degree to which people avoid small gambles (Rabin (2000)). This phenomenon, called *loss aversion* (Kahneman & Tversky (1979)), has been modeled as a distaste for gambles whose payoffs sometimes fall slightly short of a reference point. This suggests a kink in the value function at the reference point (as in prospect theory, discussed later; but see also Gal (2006)).
Empirically, loss aversion affects the trading decisions of professional investors (Coval & Shumway (2005)). Economists have long strived to understand the high estimated premium of equity expected returns over bonds (\citeN{mehra/prescott:85}). By increasing effective risk aversion, loss aversion offers a possible explanation for the equity premium and nonparticipation puzzles; shifts in loss aversion owing to the house money effect additionally can explain high equity return volatility and the value effect in the cross-section of returns (Benartzi & Thaler (1995) and Barberis & Huang (2001), but see also Beshears et al. (2012)). The equity premium over long-term bond yields has, however, been small for the last four decades (Welch & Levi (2013)), which is consistent with this explanation if investors over time have started to understand that their loss aversion was excessive.

Loss aversion may reflect the use of a heuristic of discretizing continuous variables so that even a small loss is perceived to be essentially different from a small gain. I call this phenomenon *conceptual discretizing*.

Conceptual discretizing can also explain why individuals overweight fairly unlikely events yet underweight extremely unlikely ones (treated as “virtually impossible”); such *probability weighting* is a key ingredient of prospect theory. In the model of Barberis & Huang (2008), probability weighting induces a demand for positively skewed “lottery stocks.” Alternatively, social interactions can induce such a demand even if investors have no direct preference for skewness (Han & Hirshleifer (2014)). These approaches can explain the high investor demand for, and low future returns experienced by positively skewed stocks (Boyer et al. (2010); Eraker & Ready (2014)).
f. Mental accounting and realization preference

*Mental accounting* is the system that people use to track their gains and losses relative to a reference point, and feel rewarded or punished for them. It involves narrow framing, wherein people separately optimize different kinds of gains and losses that are placed in different mental accounts. Investors reexamine each account intermittently for occasional action. Under mental accounting, people care about the labeling of payoffs by account, even when completely fungible across accounts, as this affects attribution as a gain or a loss.

Narrow framing, reference-dependence, loss aversion, and mental accounting are efficiently modeled as nontraditional preferences. However, all can be viewed as reflecting mistakes of analysis or belief, as with an investor who decides whether to sell a stock by focusing on its marginal return distribution without thinking about why he should care about covariance with his portfolio.

i. Realization preference

If selling a stock makes the incremental payoff in its mental account more salient, investors should become more willing to realize as the net gain increases *realization preference*. Under loss aversion, this applies even to small gains and losses, implying a jump at zero, *sign realization preference*. Such behavior can enhance self-esteem, if it is easier to pretend that mere “paper” losses will be regained.

In the model of Grinblatt & Han (2005), a greater willingness to sell above than below the purchase price causes price underreaction to news. Empirically this effect helps explain
return momentum. However, pure underreaction theories do not explain the evidence that momentum reverses in the long-run (Griffin et al. (2003); Jegadeesh & Titman (2011)).

In a test focusing directly on realizations, Lim (2006) finds that individual investors are more likely to sell losers on the same day than winners on the same day. This is consistent with the dual risk attitudes of prospect theory (risk loving in the loss domain, risk averse in the gain domain) together with realization preference.

A. The disposition effect

The *disposition effect* is the strong and widespread regularity that the probability of an investor selling an asset conditional upon a gain is greater than conditional upon a loss (Shefrin & Statman (1985)). The disposition effect is often appealed to as strong evidence that psychological bias affects trading, yet it is not known what bias causes it.

Experimental and field evidence reveals a *reverse* disposition effect (selling losers) for delegated holdings in mutual funds. The reversal of the disposition effect when investors can assign blame to others suggests that the urge to maintain self-esteem is a key driver of the effect (Chang et al. (2014)).

A direct realization preference explanation for the disposition effect was suggested by Shefrin & Statman (1985) and modeled by Barberis & Xiong (2012). Other possible explanations derive from the dual risk preference feature of prospect theory; Barberis & Xiong (2009) point out limitations of this approach, whereas Henderson (2012) and Li & Yang (2013) describe conditions under which the prospect theory explanation can work.
There is evidence of neurological processes associated with realization preference (Frydman et al. (2014)). However, discontinuity tests on U.S. investor trades do not support sign realization preference, and show that it is not the source of the disposition effect. Furthermore, the empirical V-shape in probability of both selling and buying as functions of gains or losses suggests that realization preference is not the dominant motive for selling decisions in general (Ben-David & Hirshleifer (2012)).

Contrary to common discussions, there is currently no strong empirical indication as to whether preference-based models or explicit belief bias models will offer a better explanation for the disposition effect. In empirical papers, explanations have typically been discussed in a static fashion; recent models derive predictions that reflect the dynamics of trading with realization preference (Barberis & Xiong (2012), Ingersoll & Jin (2013)).

ii. Prospect theory

Reference dependence and loss aversion are ingredients of prospect theory (Kahneman & Tversky (1979); Tversky & Kahneman (1992)), wherein individuals maximize a weighted sum across states of the world of value functions (utilities), value depends on gains or losses rather than levels, and where the weights are functions of probabilities (in a fashion discussed earlier). Value is an S-shaped function of gain/loss (dual risk attitudes), resulting in risk aversion in the gain domain and risk seeking in the loss domain. Loss aversion is reflected in a kink in the value function at zero gain or loss. Financial theories and evidence based upon the different ingredients of prospect theory were discussed in earlier sections.

g. Heuristic learning
i. Representativeness, hyperactive pattern-recognition, and overextrapolation

According to the representativeness heuristic (Kahneman & Tversky (1973)), people assess the probability of a state of the world based on how typical of that state the evidence seems to be. This is reasonable if typicality proxies for the conditional probability of the evidence given the state of the world. However, rationally one should adjust for the prior probabilities of the outcomes. In reality people tend to underweight verbal statements about unconditional population frequencies in updating beliefs—base-rate underweighting. This is another symptom of WYSIATI.

Furthermore, perceptions of how typical a piece of evidence is of a state of the world often reflect its conditional probability poorly. For example, error management theory holds that the human mind evolved to overweight the probabilities of opportunities or dangers when the potential cost of neglect is high (Haselton & Nettle (2006)). This suggests that people are subject to what may be called hyperactive pattern recognition. For example, people tend to overweight small samples in drawing inferences about distributions (the law of small numbers, Tversky & Kahneman (1971)). However, they also rely too little on large samples.

In financial markets, overextrapolation of security returns implies positive feedback trading. In the model of DeLong et al. (1990b), exogenous positive feedback trading causes overreaction and long-run return reversal, and potentially short-run momentum as well.

In the model of Barberis et al. (1998), conservatism bias (Edwards (1968)), in which individuals hold too tightly to estimates based upon early observations, causes short-term underreaction to earnings news (consistent with the PEAD anomaly). Owing to the
representativeness heuristic, if sequences of good earnings news occur, investors fixate on this pattern and overreact. This combination of effects generates return momentum and reversal, and an overreaction/reversal pattern in response to trends in public value signals (e.g., earnings news sequences).

Empirically, investors do naively extrapolate in experimental markets, survey, and field studies; and in various kinds of investments (e.g., Smith et al. (1988)). There is less support for overreaction to trends in public financial signals (Chan et al. (2004); Daniel & Titman (2006)).

**ii. Reinforcement learning**

Under *reinforcement learning*, an individual only extrapolates from his own direct experience, and without properly reflecting the informativeness of the data. There is financial evidence that investors learn to make financial decisions by naïve reinforcement. Investors overextrapolate their own past performance in making investment choices (Choi et al. (2009); Chiang et al. (2011)). Furthermore, past life experiences also affect both investor and managerial decisions (Greenwood & Nagel (2009), Malmendier, Tate & Yan (2011)).

**iii. Inertia and habits**

People easily lock into habits, and rely on them with little thought. This leads to big mistakes when circumstances change. When there is memory loss about the reasons for past decisions, and if the environment is reasonably stable, it is, nevertheless, constrained-optimal to rely on habits (Hirshleifer & Welch (2002)). Action-induced attitude changes, as with
cognitive dissonance and the sunk cost fallacy, can also induce inertia. Empirically, retirement investors seldom update their portfolios as conditions change (Choi et al. (2004)).

The status quo bias (Samuelson & Zeckhauser (1988)), a preference for the default choice among a set of options, also economizes on the reasoning system’s slow, effortful cognition. For example, defaults for pension plan contributions and allocations have large effects on investment decisions (e.g., Madrian & Shea (2001)).

6. Feelings

Feelings are a key source of the quick assessments provided by the intuitive system, and can overwhelm cooler analysis. For example, people who plan to consume sparingly are later tempted to consume heavily, resulting in time-inconsistent choices. This shows how immediacy can intensify the effects of feelings. People who foresee this can gain by imposing consumption rules upon themselves (Ainslie (1975)).

Present-biased decision-making (quasi-hyperbolic discounting; Laibson (1997)) has been applied in models of savings, liquidity premia and the equity premium puzzle. To resolve the time-inconsistency of such preferences in favor of saving more, people impose personal rules such as consuming only out of interest and dividends, not principal (Thaler & Shefrin (1981)). This can explain the preference of investors for cash dividends (Shefrin & Statman (1984)).

People often misattribute arousal and other transient feelings to other sources, biasing their judgments (Schwarz & Clore (1983)). Good mood increases optimism and risk-taking (Kuhnen & Knutson (2011)). The kind of feeling matters, not just its valence. For example,
when fearful, people tend to be more pessimistic and risk averse; when angry, more optimistic and risk tolerant (Lerner & Keltner (2001)).

a. Familiarity and liking

Exposure to an unreinforced stimulus tends to make people like it more, the *mere exposure effect* (Bornstein & D'Agostino (1992)). The evolutionary basis for this may be that what is familiar tends to be understood better, reducing risk; or that experience of a stimulus without adverse consequences indicates low risk. Indeed, familiarity reduces feelings of risk (Weber et al. (2005)). However, the familiarity heuristic can go astray, as when people prefer to bet on a matter about which they feel expert over another precisely equivalent gamble (Heath & Tversky (1991)).

The *endowment effect* (Kahneman et al. (1990)) is a preference for retaining what one has over exchanging for a better alternative (as with refusing to swap a lottery ticket for an equivalent one plus cash). A possible explanation is loss aversion. Alternatively, an already-owned good may be affectively attractive by virtue of sense of ownership.

*Ambiguity aversion* is a distaste for layered gambles relative to single-stage gambles with identical payoff distributions (Ellsberg (1961); Bossaerts et al. (2010)). For example, investors may dislike uncertainty about the structure of a financial market, as distinguished from the effect of the future state realization given that structure.

b. Financial theories based on feelings
Financial theorizing about feelings has been mostly informal (but see Mehra & Sah (2002)), which is surprising given their psychological importance. A basic theme is that mood swings affect optimism, risk tolerance, and market prices. Owing to misattribution of transient mood to long-term prospects, mood swings associated with weather or sports events can affect prices (as documented by Saunders (1993); Hirshleifer & Shumway (2003); Edmans et al. (2007)). Seasonal shifts in length of day can induce Seasonal Affective Disorder, and are correlated with market returns (Kamstra et al. (2003)).

Skepticism about the foreign and unfamiliar offers an explanation for the failure of investors to participate in important asset classes. Models of ambiguity aversion can help explain non-participation, familiarity bias, and their effects on asset pricing (Chen & Epstein (2002); Cao et al. (2011)). Such models potentially have an affective interpretation.

Feelings of envy may help explain the attractiveness of investments with lottery payoffs, as individuals hear about high payoffs obtained by others. In the model of Goel & Thakor (2010), the takeovers decisions of managers are influenced by feelings of envy toward other managers, resulting in merger waves.

c. Evidence on financial effects of familiarity and in-group bias

People prefer local investments and familiar ones, such as firms that they are customers of (Grinblatt & Keloharju (2001); Huberman (2001)). One reason is that investors may have superior information about local or familiar firms (Coval & Moskowitz (1999)). However, this does not seem to be the only reason for local bias. For example, at the cost of poor diversification, employees invest in their own firms without showing signs of superior
information (Benartzi (2001)). Furthermore, informational superiority seems an unlikely explanation for home bias exhibited by great masses of unsophisticated investors.

In-group bias (belief in the superior merits of one’s own group), which is relatively neglected in analytical modeling, implies bias in financial investing and economic exchange in favor of own-culture. Several studies provide supporting evidence (Grinblatt & Keloharju (2001)).

Consistent with in-group bias and with theories based on aversion to uncertainty or unfamiliarity, distrust is an important barrier to participation in the stock market (Guiso et al. (2008)) and exchange and investment between countries (Guiso et al. (2009)). More generally, familiarity and in-group biases are sources of underdiversification, a problem to which unsophisticated investors are especially subject (Goetzmann & Kumar (2008)).

d. Sentiment, shifting optimism and risk tolerance

Investor sentiment is the fluctuating general attitude toward investment categories, such as growth stocks or long-term bonds. It can be associated with shifts in assessments of expected returns or of risk. Waves of irrational enthusiasm for, or abhorrence of, certain investment characteristics derive from shifts in the salience of emotional or cognitive triggers in the economic environment. Such shifts can be magnified by self-reinforcing social processes induced by media bias or conformity effects.

In the model of DeLong et al. (1990a), irrational noise trading induces fluctuations in the price of an asset with riskfree dividends. Short horizons of rational risk averse investors prevent
full arbitrage between this asset and an asset with identical dividends that is not subject to noise trading. The theory implies that on average the speculative asset trades at a discount relative to fundamentals as compensation for its excess volatility.

Lee et al. (1991) more broadly suggest that closed-end funds, like other small stocks, are subject to noise trading, so that irrational trading induces premia or discounts relative to the price of their underlying assets. Consistent with a risk discount for stochastic fund premia, on average funds trade at discounts relative to their holdings. Furthermore, discounts and premia comove across funds and with the returns on small stocks in general, which suggests a common influence of sentiment among naïve individual investors.

If sentiment induces mispricing, then sentiment measures should predict future abnormal returns. Empirically, U.S. closed end funds discounts and premia predict future small stock returns (Swaminathan (1996)). However, in distinguishing the pricing effects of sentiment from other hypotheses, it is useful to employ measures of sentiment that are not based on market prices (Qiu & Welch (2006)). When several sentiment proxies are low, stocks that are hard to value and arbitrage earn high subsequent returns (Baker & Wurgler (2006)). High sentiment increases the profitability of the short legs but not the long legs of cross-sectional return anomalies (Stambaugh et al. (2012)).

Measures of global sentiment negatively predict country-level returns. Both global and local sentiment are stronger return predictors for stocks that are hard to value and to arbitrage (Baker et al. (2012)).
Shifts in market sentiment create incentives for interested parties to incite misvaluation. In the theory of Baker & Wurgler (2004), managers cater to investor preferences for or against dividends. When the stock price premium on payers is high, firms start paying dividends in order to incite higher valuation. Consistent with this prediction, when sentiment favors dividends more, nonpayers tend to initiate dividends.

7. Firm behavior: Exploiting versus inciting misvaluation

A distinction that is fundamental for firm behavior in inefficient markets is between exploiting mispricing, defined as an action taken in response to a preexisting level of mispricing, and inciting, an action designed to shift the level of mispricing (Hirshleifer (2001)). Inciting takes advantage of the function describing the relation between market price and the firm’s action.  

Inciting encompasses actions taken to shift mispricing either upward or downward. In contrast, “catering” (Baker & Wurgler (2012)) is defined as an action taken to increase price above fundamental value.

Also, it is common to distinguish inciting or catering from timing, wherein the firm is sure to undertake the action, but uses discretion as to when. However, this is not an exhaustive partition of cases; a firm can exploit in its choice of whether rather than when to take an action. Post-event return drift is often interpreted as timing without consideration of this very plausible possibility. More importantly, the possibility of incitement of misvaluation is often ignored in favor of timing in response to preexisting misvaluation.

\[ \text{Inciting encompasses actions taken to shift mispricing either upward or downward. In contrast, “catering” (Baker & Wurgler (2012)) is defined as an action taken to increase price above fundamental value.} \]

\[ \text{Also, it is common to distinguish inciting or catering from timing, wherein the firm is sure to undertake the action, but uses discretion as to when. However, this is not an exhaustive partition of cases; a firm can exploit in its choice of whether rather than when to take an action.} \]

\[ \text{Post-event return drift is often interpreted as timing without consideration of this very plausible possibility. More importantly, the possibility of incitement of misvaluation is often ignored in favor of timing in response to preexisting misvaluation.} \]
To illustrate this distinction, consider a firm that issues equity to exploit preexisting overvaluation. Owing to the negative average reaction to the announcement, there tends to be a reduction in overvaluation, but this will normally be an unavoidable adverse side-effect from the firm’s viewpoint, in which case this is not incitement. In contrast, a repurchase can be incitement if its purpose is to induce higher valuation (rather than merely distributing cash, or profiting from purchasing underpriced shares).

Upward earnings management designed to induce overvaluation (or eliminate undervaluation) is also incitement. Most financial executives in one survey reported that they would sacrifice economic value in order to avoid missing quarterly earnings forecasts (Graham et al. (2005)). Similarly, managing earnings downward with the purpose of reducing the stock price (e.g., to persuade potential competitors that the business is unprofitable, or to reduce the cost of share repurchase), is downward incitement. Verbal communication can also be used to incite misvaluation, as with misleading disclosures, and discussions with media and analysts (typically upward “hype”).

a. Theories of exploitive advisors and firms

Section 5 points out that neglect of public signals results in return predictability based upon the accounting information, and therefore that manipulation of disclosures can incite over- or undervaluation (Hirshleifer & Teoh (2003); Hirshleifer et al. (2011)).

Stein (1996) models the exploitation of exogenous stock market mispricing by firms in their financing and investment decisions. In Stein’s model, misvaluation affects real investment decisions more when managers have short time horizons, and firms should sometimes
paternalistically discount using beta even when beta is not a return predictor. In Daniel et al. (1998), new issues and repurchase amounts are selected by a firm as a function of mispricing to exploit investor overconfidence. This implies positive abnormal returns after repurchase and negative after new issues.

Ljungqvist et al. (2006) model the exploitation of individual investor optimism in initial public offerings. Cornelli et al. (2006) provide evidence that institutional investors and underwriters exploit misvaluation of IPOs by individual investors.

Investors with limited attention will sometimes overlook opportunism. One way to exploit customers is to add complexity; in the model of Carlin (2009), intentionally added complexity of financial products results in equilibrium price dispersion among competing providers.

Exploitation and incitement can have adverse macroeconomic effects as well. In the theory of Gennaioli et al. (2012a), intermediaries design securities that seem nearly riskfree to take advantage of investor neglect of nonsalient risks. This results in booms and crashes.

b. Evidence on exploitive advisors and firms

Evidence suggests that investors are overly credulous about the strategic incentives of information sources, leaving them vulnerable to manipulation by firms, advisors, and intermediaries (such as analysts, brokers, and money managers). Daniel et al. (2002) argue that credulity derives from limited attention and overconfidence, and that it explains a wide range
of financial behaviors and pricing anomalies. Jensen (2005) argues, for example, that firm overvaluation promotes exploitive behavior on the part of managers.

For example, evidence suggests that investors are naïve about strategic behavior by firms in their financial reporting. Issuers manage earnings upward at the time of IPO and seasoned issue; greater upward management is associated with worse post-event average abnormal returns (Teoh et al. (1998a,b)). This suggests that firms successfully incite overvaluation prior to issue, rather than just exploiting preexisting misvaluation.

As mentioned earlier, analyst forecasts do not discount adequately for earnings management. Furthermore, evidence suggests that investors are naïve about analyst incentives to bias forecasts (Richardson et al. (2004)) and recommendations (Malmendier & Shanthikumar (2007)). Investors seem to be credulous about the strategic motives of managers in various other contexts as well, such as trusting that name changes are indicative of firm and fund policies (Cooper et al. (2005)), that fund marketing expenses are unimportant (Barber et al. (2005)), and that broker recommended funds are superior (Guercio & Reuter (2013)).

The theoretical models of financing in inefficient markets discussed above predict abnormal returns after new issues and repurchase owing to firms selling their shares when overpriced and buying back when they are underpriced. Consistent with security issuance being associated with overvaluation, there is return continuation after new issues and repurchase (Section 4). In general, the occurrence of an event can predict subsequent abnormal returns either because of exploitation of existing mispricing, or because it incites mispricing. So post-event abnormal return evidence does not, in itself, establish whether overvaluation causes
issuance, whether issuance causes overvaluation, or whether other actions associated with
issuance cause overvaluation (e.g., earnings management inciting overvaluation at the time of
issue). These distinctions are often overlooked.

c. Misvaluation, new issues and repurchase, and post-event returns

Several studies point more specifically to exploitation of preexisting overpricing as part
of the explanation. Surveys of U.S. CFOs find that misvaluation of their firms’ stocks is an
important factor in deciding whether to issue equity, and that CFOs try to time interest rates in
issuing debt (Graham & Harvey (2001)). Furthermore, measures of prior misvaluation based
upon the deviation of price from contemporaneous fundamentals are associated with
subsequent new issuance of debt and especially equity, especially among overvalued firms
(Dong et al. (2012)).

Investment and growth-related measures are negative predictors of abnormal stock
returns (Titman et al. (2004); Cooper et al. (2008); Polk & Sapienza (2009)). Such evidence does
not resolve whether investment induces overvaluation (either as incitement, or as an
unintended side-effect), or whether investment choices exploit preexisting misvaluation.
Evidence that higher discretionary accruals is associated with greater investment is consistent
with incitement. However, consistent with exploitation also playing a role, proxies for prior
misvaluation predict investment (Gilchrist et al. (2005)).

Misvaluation can also affect takeover behavior. In the model of Shleifer & Vishny (2003),
overvalued bidders use equity and undervalued bidders pay cash. Potentially consistent with
(but not proof of) misvaluation affecting takeover behavior, Loughran & Vijn (1997) find
negative post-event abnormal returns to stock acquirers. Proxies for misvaluation are also associated with the use of equity as payment, transaction characteristics, and market reactions to announcement in ways largely consistent with the Shleifer & Vishny (2003) model (Ang & Cheng (2006); Dong et al. (2006)); Rhodes-Kropf et al. (2005) also provide evidence of valuation (though not necessarily mispricing) effects.

8. Conclusion: Behavioral finance and social finance

I close with suggestions for future research. First, given the large grab bag of possible behavioral biases to choose from, building a financial model by just assuming some behavior that seems plausible, or even by invoking a documented psychological bias, is not always compelling. A healthy nascent trend in behavioral economics and finance has been to run laboratory and field experiments that closely match the decision environment assumed in the financial model.

Second, the affective revolution in psychology of the 1990s, which elucidated the central role of feelings in decision-making, has only partially been incorporated into behavioral finance. More theoretical and empirical study is needed of how feelings affect financial decisions, and the implications of this for prices and real outcomes. This topic includes moral attitudes that infuse decisions about borrowing/saving, bearing risk, and exploiting other market participants.

Third, behavioral finance should continue its evolution from broad descriptions of imperfect rationality and its consequences, such as noise trading or sentiment, toward analysis of particular psychological biases or categories of effects (e.g., overestimation of mean payoff,
underestimation of risk, or shifting risk preferences). Doing so will naturally draw more focused attention to specific pathways of causality, thereby helping to address endogeneity issues in some tests of the effects of sentiment or media.

Most importantly, there is a need to move from behavioral finance to social finance (and social economics). Social finance includes the study of how social norms, moral attitudes, religions and ideologies affect financial behaviors (Hilary & Hui (2009), Hong et al. (2009), Kumar (2009), Kumar et al. (2011), Mcguire et al. (2012), Hong & Kostovetsky (2012), Hutton et al. (2013)), and how ideologies that affect financial decisions form and spread. This enterprise will draw on social psychology and sociology as well as cognitive psychology and decision theory, and will require focused attention to the microstructure of social transactions.

Previous research has documented the spread of investment and managerial behaviors through observation of public behaviors or through social networks (see, e.g., the review of Hirshleifer & Teoh (2009b)). However, mere contagion is consistent with the spread of almost any behavior. To derive richer implications, it will be crucial to understand the transmission biases and amplification processes that make some investment ideas spread more easily than others. An initial set of leads is provided in the survey evidence and discussions of Robert Shiller (e.g., Shiller (2000)). Recent research has begun to model social transmission biases (Han & Hirshleifer (2014)) and test for their financial effects (Simon & Heimer (2012); Kaustia & Knüpf (2012)).

Analysis of social interactions promises to provide greater insight into where heuristics come from (since they are far from entirely innate), and to offer a foundation for understanding
shifts in investor sentiment. As such, it can potentially offer a deeper basis for understanding the causes and consequences of financial bubbles and crises. Even more fundamentally, understanding how financial ideas spread from person to person may eventually suggest theories of how investment and corporate ideologies, such as value versus growth philosophies, or the belief that indebtedness is bad, evolve.

Behavioral finance has primarily focused on individual level biases. Social finance promises to offer equally fundamental insight, and to be a worthy descendant of behavioral finance.
References


