

CHAPTER 2

Psychology-Based Models of Asset Prices and Trading Volume[★]

Nicholas Barberis^{*,†}

^{*}Yale School of Management, New Haven, CT, United States of America

[†]National Bureau of Economic Research, Cambridge, MA, United States of America

E-mail address: nick.barberis@yale.edu

Contents

1.	Introduction	80
2.	Empirical Facts	82
2.1	Aggregate asset classes	83
2.2	The cross-section of average returns	85
2.3	Bubbles	88
3.	Limits to Arbitrage	90
4.	Beliefs: Extrapolation	93
4.1	Return extrapolation	94
4.2	Extrapolation of fundamentals	103
4.3	Sources of extrapolative beliefs	104
4.4	Experience effects	110
4.5	Extrapolative beliefs: Summary	112
5.	Beliefs: Overconfidence	113
5.1	Disagreement with a short-sale constraint	117
6.	Other Belief-Based Approaches	119
6.1	Sticky beliefs	119
6.2	Models of under- and over-reaction	121
6.3	Beliefs about rare events	124
6.4	Feelings and beliefs	125
6.5	Herding and social interaction	126
6.6	Psychology-free approaches	128
7.	Preferences: Gain-Loss Utility and Prospect Theory	130
7.1	The elements of prospect theory	131
7.2	Prospect theory and the cross-section: Static models	137

^{*} This article has been informed by many discussions over the years with Robin Greenwood, Ming Huang, Lawrence Jin, Matthew Rabin, Andrei Shleifer, Richard Thaler, and Wei Xiong, as well as with my students and my colleagues in the fields of behavioral finance and behavioral economics. I am grateful to Douglas Bernheim, Stefano DellaVigna, and David Laibson for their comments on an early draft, and to Daniel Benjamin, Pedro Bordalo, Erik Eyster, Shane Frederick, Sam Hanson, Philipp Krueger, Alan Moreira, Tobias Moskowitz, Charles Nathanson, Cameron Peng, David Thesmar, and Baolian Wang for their help with questions that came up during the writing process.

7.3	Prospect theory and the cross-section: Dynamic models	140
7.4	Prospect theory and the aggregate stock market	144
7.5	Prospect theory applications: Summary	148
7.6	Other alternatives to Expected Utility	149
8.	Preferences: Ambiguity Aversion	151
9.	Bounded Rationality	154
9.1	Inattention	156
9.2	Categorization	157
10.	Discussion and Conclusion	159
	Appendix A. Derivation of Eqs. (3), (4), and (5)	162
	References	164

1. INTRODUCTION

The modern era of finance research began in the 1950s with the work of Markowitz (1952) on portfolio choice and Modigliani and Miller (1958) on capital structure. For the next four decades, research in finance – especially research on asset prices – was dominated by a single framework, the “traditional” framework, which is based on two assumptions about individual psychology. The first is that people have rational beliefs: when new information arrives, they immediately update their beliefs about future outcomes, and do so correctly, as prescribed by Bayes’ rule. The second is that they make decisions according to Expected Utility: given their beliefs, they choose the action with the highest Expected Utility for a utility function that is defined over consumption outcomes and that is increasing and concave.

In the 1980s, a new paradigm, behavioral finance, began to emerge. Its goal is to make sense of facts in finance using models that are *psychologically more realistic* than those that came before. The field has grown rapidly over the past 30 years. In this article, I review research in behavioral finance, focusing on applications to asset prices and trading volume.¹

Research in behavioral finance has tried to improve the psychological realism of the traditional model along three dimensions. First, through *more realistic assumptions about individual beliefs* – in particular, that people do not update their beliefs in a fully rational manner, thereby deviating from Bayes’ rule. Second, through *more realistic assumptions about individual preferences* – for example, by rethinking what it is that

¹ The models of asset prices and volume that I discuss also make predictions about individual portfolio choice, but I do not cover that topic in much detail; for a fuller treatment of it, see Guiso and Sodini (2013) and Beshears et al. (2018). For behavioral finance approaches to corporate finance, see Baker and Wurgler (2012) and Malmendier (2018). For other surveys of psychology-based models of asset prices, see Shleifer (2000), Hirshleifer (2001), and Barberis and Thaler (2003).

people derive utility from and what form their utility function takes, or by replacing Expected Utility with an alternative framework such as prospect theory. And third, by taking account of *cognitive limits* – by recognizing that people are unlikely to be able to immediately process all of the information that is relevant to their financial situation, given how much information of this kind arrives every week. This article is structured around these three dimensions: I review, in turn, models of beliefs, models of preferences, and models of cognitive limits.

Why did behavioral finance emerge when it did, in the 1980s, and then gather steam in the 1990s? At least three factors played a role. First, by the late 1980s, there was a growing sense that some basic facts about financial markets were hard to reconcile with the traditional finance framework. In a 1981 paper – a paper seen by many as the starting point of modern behavioral finance – Robert Shiller argued that fluctuations in stock market prices are unlikely to be the result of rationally-varying forecasts of firms' future cash flows. Other papers, among them De Bondt and Thaler (1985), showed that some investment strategies earn average returns that are higher than can be justified by simple measures of risk. In the view of many researchers, these findings called for a new generation of models – for example, models that allow for less than fully rational beliefs – and this spurred the growth of behavioral finance. The technology stock mania of the late 1990s and the more recent real estate boom and subsequent financial crisis brought additional attention to the field.

A second factor is the work on limits to arbitrage in the 1990s. An old critique of behavioral finance, the “arbitrage critique,” posits that irrational investors cannot have a substantial and long-lived impact on the price of an asset, because, if they did, this would create an attractive opportunity for *rational* investors who would trade aggressively against the mispricing and remove it. This argument proved compelling to economists for many years and slowed the development of behavioral finance. In the 1990s, however, some researchers pushed back against the arbitrage critique, noting that, in reality, rational traders face risks and costs that limit their ability to correct a mispricing (De Long et al., 1990a; Shleifer and Vishny, 1997). This work on limits to arbitrage, which I discuss in more detail later, has been influential and was an important factor in the rise of behavioral finance.

The third reason for the growth of behavioral finance in the 1990s was the dramatic progress in the 1970s and 1980s in an area of psychology known as “judgment and decision-making.” This field, which was dominated for years by Daniel Kahneman and Amos Tversky, seeks to paint a more realistic picture of how people form beliefs and make decisions. For financial economists who were looking for guidance on how to make their models more psychologically accurate, this research was invaluable. Much of the conceptual progress in behavioral finance over the past 30 years has come from incorporating ideas from the field of judgment and decision-making into more traditional finance models.

Behavioral finance research on asset prices has an ambitious agenda. It argues that the traditional model does not offer a complete account of even the most basic aspects of the data – facts about asset market fluctuations, average returns on asset classes, trading volume, and bubbles – and that a behavioral finance perspective is essential for a full understanding of the evidence. On some dimensions, behavioral finance can already claim success: it has shown that models based on a few simple assumptions about individual psychology can explain a wide range of empirical facts, and can make concrete, testable predictions, some of which have already been confirmed in the data.

As indicated above, this article is organized around the three approaches researchers have taken to develop psychologically realistic models of investor behavior. In Sections 4 through 6, I review models based on more psychologically accurate assumptions about individual beliefs; the main ideas here are extrapolation and overconfidence. In Sections 7 and 8, I discuss models that focus on individual preferences; the key concepts here are gain-loss utility and prospect theory, and ambiguity aversion. And in Section 9, I cover models that take into account people's cognitive limits. In Section 10, I evaluate progress in the field and conclude. Before embarking on the discussion of behavioral finance models in Section 4, I first cover two important background topics. In Section 2, I review the main empirical facts that are the focus of study in the field of asset prices. In Section 3, I summarize the research on limits to arbitrage and relate the themes of this article to the concept of "efficient markets."

Over the years, researchers in behavioral finance have pursued a number of different psychology-based approaches. One might therefore worry that the field consists of many scattered ideas. Fortunately, this is not the case. The center of gravity in behavioral finance lies in just three frameworks: the extrapolation framework (Section 4), the overconfidence framework (Section 5), and a gain-loss utility framework inspired by prospect theory (Section 7). These frameworks are not in competition with one another, in part because they have somewhat different applications: extrapolation is most helpful for explaining fluctuations in financial markets, overconfidence for understanding trading volume, and gain-loss utility for thinking about assets' average returns. It is true that there is as yet no "unified" model in behavioral finance – no single, parsimonious, widely-used model that makes psychologically realistic assumptions about both beliefs and preferences. However, the research to date is beginning to point to the form that such a model might take. In the concluding section, I attempt to sketch its outlines.

2. EMPIRICAL FACTS

Much of the research on asset prices is aimed at understanding a specific set of widely agreed-upon empirical findings. These were first documented in the context of the stock market: academic researchers have long had access to high quality data on stock market prices. However, an important finding of recent years is that many of the patterns we observe in the stock market are also present in other asset classes. This raises

the appeal of behavioral finance approaches: the most prominent psychology-based assumptions about investor behavior apply in a natural way in all asset classes, not just the stock market.

Below, I review three groups of empirical facts: facts about aggregate asset classes; facts about the cross-section of average returns; and facts about bubbles. In the case of aggregate asset classes and the cross-section, I first describe the facts in the context of the stock market and then summarize what is known about other asset classes.²

2.1 Aggregate asset classes

There are three central facts about the returns on the overall U.S. stock market: these returns are predictable in the time series; they display “excess volatility”; and their average level is high.

Time-series predictability. Stock market ratios of price to fundamentals – the market’s price-to-earnings (P/E) ratio or its price-to-dividend (P/D) ratio – predict the market’s subsequent excess return – its return in excess of the risk-free rate – with a negative sign (Campbell and Shiller, 1988; Fama and French, 1988). This time-series predictability is widely viewed as *the* essential fact about stock market returns that needs to be understood. The best-known rational models of it are based on rationally-changing forecasts of future risk – specifically, the “rare disasters” framework and the “long-run risk” framework; changing investor risk aversion, specifically, the “habit formation” framework; and rational learning.³ For example, under the rare disasters framework, if investors rationally decide that an economic disaster is less likely going forward, they lower the risk premium that they use to discount firms’ future cash flows. As a consequence, the stock market’s P/E ratio rises today and its subsequent return in excess of the risk-free rate is lower, on average. The P/E ratio is then negatively related to the subsequent excess return, as in the data.

Excess volatility. Shiller (1981) and LeRoy and Porter (1981) show that aggregate stock market prices are “excessively” volatile, in the sense that it is hard to justify their fluctuations on the basis of rationally-varying forecasts of the future cash flows paid to investors. To see why, suppose that variation in the P/D ratio of the stock market is driven by rationally-changing forecasts of future dividend growth: in some years, investors rationally expect higher dividend growth in the future and push the stock market price up relative to current dividends; in other years, they rationally expect lower dividend growth and push the price down relative to current dividends. Since

² See Ilmanen (2011) for a detailed review of empirical evidence on asset prices. I do not discuss facts about volume in this section, but instead introduce them at the appropriate time later in the article.

³ On rare disasters, see Gabaix (2012) and Wachter (2013); on long-run risk, see Bansal and Yaron (2004) and Bansal et al. (2012); on changing risk aversion, see Campbell and Cochrane (1999); and on rational learning, see Timmermann (1993) and Pastor and Veronesi (2009). Campbell (2018) provides a comprehensive review of rational models of asset prices.

these forecasts of dividend growth are taken to be rational, this framework implies that, in a long time series of data, the P/D ratio will predict subsequent dividend growth with a positive sign. In historical U.S. data, however, it does not (Campbell and Shiller, 1998). Rationally-changing forecasts of future cash flows are therefore unlikely to be the main source of stock market fluctuations.⁴

Time-series predictability and excess volatility are now seen as the same phenomenon. To see why, note that, in historical U.S. data, the stock market's P/D ratio is stationary; loosely put, years of high P/D ratios are followed by years with moderate P/D ratios. Mathematically, this can happen in one of two ways: either dividends D must rise or prices P must fall. From the discussion of excess volatility above, we can rule out the first channel: in the data, high P/D ratios are not followed by higher dividend growth. High P/D ratios must therefore be followed by lower average returns. But this is exactly the finding known as time-series predictability.

Excess volatility and time-series predictability are not confined to the stock market. They have also been documented in other major asset classes, including real estate and long-term bonds. For example, it is hard to explain real estate price fluctuations on the basis of rational forecasts of income growth or rent growth, and the price-to-rent ratio predicts subsequent housing returns with a negative sign (Glaeser and Nathanson, 2015).

Given that excess volatility and time-series predictability are the same phenomenon, it is not surprising that the rational approaches to thinking about excess volatility are the same as those that have been used to address time-series predictability: rationally-changing forecasts of future risk, changing risk aversion, and rational learning.

Equity premium. Over the past century, the average return of the U.S. stock market has exceeded the average return of Treasury Bills by over 5% per year. Such a large equity premium has proven hard to explain in a rational model, a puzzle known as the “equity premium puzzle.” For example, a simple rational model – one with a representative investor who has power utility preferences over lifetime consumption – predicts an equity premium of less than 0.5% per year when the investor's risk aversion is set at levels suggested by simple thought experiments (Mehra and Prescott, 1985). In this model, the risk of the stock market is measured by the covariance of stock market returns with consumption growth. Empirically, this quantity is very small: the standard deviation of aggregate consumption growth, in particular, is low. The representative investor therefore does not view the stock market as very risky, and requires a very low equity premium.

In recent years, the equity premium puzzle has received less attention. One possible reason is that, after many years of effort, researchers have run out of ideas for solving the

⁴ Giglio and Kelly (2017) and Augenblick and Lazarus (2018) show that some price movements in financial markets are hard to explain based on rationally-varying forecasts of either cash flows or discount rates.

Table 1 Firm-level characteristics with predictive power for stock returns in the cross-section. The “+” and “−” signs indicate whether the characteristic has positive or negative predictive power

Characteristic	
Past three-year return	−
Past six-month return	+
Past one-month return	−
Earning surprise	+
Market capitalization	−
Price-to-fundamentals ratio	−
Issuance	−
Systematic volatility	−
Idiosyncratic volatility	−
Profitability	+

puzzle and have turned their attention to other topics. Another is a belief that the historical equity premium is not as high as it once was, and is therefore less anomalous. And another possibility is that the extant rational and behavioral explanations of the puzzle are seen as resolving it in a satisfactory way. I discuss a prominent psychology-based explanation of the puzzle in Section 7.4.⁵

2.2 The cross-section of average returns

Why do some financial assets have higher average returns than others? The benchmark rational model for thinking about this, the Capital Asset Pricing Model, or CAPM, predicts that the average return of an asset is determined by the asset’s “beta” – the covariance of the asset’s return with the return on the overall market, scaled by the variance of the market return – and by beta alone. In this framework, assets with higher betas are riskier and compensate investors by offering a higher average return.

The CAPM’s predictions have been roundly rejected. Stocks with higher betas do not have higher average returns (Fama and French, 1992). At the same time, several other stock-level characteristics do have significant predictive power for the cross-section of average returns. Understanding why these variables predict returns is the focus of a major research effort. Table 1 lists some of the best-known predictors; the “+” and “−” signs indicate whether the variable has positive or negative predictive power. To be clear, if a variable F has, say, negative predictive power, this means that, on average, stocks with high values of F have a lower return than stocks with low values of F in a way that is not captured by beta and that remains statistically significant even after

⁵ See Mehra (2008) for a review of different approaches to tackling the puzzle.

controlling for the other main predictor variables. I now briefly describe the findings summarized in the table.

Past long-term return. A stock's return over the past three to five years predicts the stock's subsequent return with a negative sign in the cross-section (De Bondt and Thaler, 1985). This is known as "long-term reversal."

Past medium-term return. A stock's return over the past six months or one year predicts the stock's subsequent return with a positive sign in the cross-section (Jegadeesh and Titman, 1993). This is known as "momentum." Notice the contrast with long-term reversal. A long-standing challenge is to build a parsimonious model that captures both of these patterns.

Past short-term return. A stock's return over the past week or month predicts the stock's subsequent return with a negative sign in the cross-section (Lehmann, 1990). This is often referred to as "short-term reversal."

Earnings surprise. The size of the surprise in a firm's most recent earnings announcement predicts the subsequent return of the firm's stock with a positive sign (Bernard and Thomas, 1989). Informally, if a firm announces earnings that are better than expected, its stock price naturally jumps up on the day of the announcement, but, more interestingly, keeps rising in the weeks after the announcement. This is known as "post-earnings announcement drift."

Market capitalization. A firm's market capitalization predicts the firm's subsequent stock return with a negative sign in the cross-section (Banz, 1981).

Price-to-fundamentals ratio. A stock's price-to-earnings, price-to-cash flow, and price-to-book ratios predict the stock's subsequent return with a negative sign (Basu, 1983; Rosenberg et al., 1985; Fama and French, 1992). The difference in the average returns earned by "value" stocks with low price-to-fundamentals ratios and "growth" stocks with high price-to-fundamentals ratios is known as the value premium.

Issuance. Stocks of firms that issue equity, whether in an initial public offering or a seasoned equity offering, earn a lower average return than a control group of firms (Loughran and Ritter, 1995). Stocks of firms that repurchase equity have a higher average return than a control group (Ikenberry et al., 1995).

Systematic volatility. The average raw return of high beta stocks is similar to the average raw return of low beta stocks. This stands in contrast to the prediction of the CAPM, namely that high beta stocks should have a higher average return, and is known as the "beta anomaly" (Black, 1972; Frazzini and Pedersen, 2014).

Idiosyncratic volatility. The volatility of a stock's daily idiosyncratic returns over the previous month predicts the stock's subsequent return with a negative sign (Ang et al., 2006).

Profitability. Measures of a firm's profitability – for example, its gross margin scaled by asset value – predict the subsequent return of the firm's stock with a positive sign (Novy-Marx, 2013; Ball et al., 2015).⁶

Data mining is a concern in the context of the cross-section of average returns (Harvey et al., 2016). Academic researchers and investment managers have strong incentives to find variables that predict stock returns in a statistically significant way – the academics so that they can publish a paper that might further their careers, and the managers so that they can pitch a new stock-selection strategy to potential clients. Given these incentives, researchers have likely gone through hundreds of firm characteristics in their search for predictors. Even if none of these variables has true predictive power for stock returns, some of them will predict returns in sample in a statistically significant way, thereby giving the appearance of a genuine relationship. While over one hundred variables have been shown, in published studies in academic journals, to predict returns in a significant way, it is likely that, in some cases, there is no true predictive relationship.

The variables in Table 1 are thought to have genuine predictive power for future returns because they forecast returns out of sample; while many of the relationships in Table 1 were first documented using U.S. stock market data from the 1960s to the 1990s, they also hold in many international stock markets or in the U.S. market before the 1960s.

There is no consensus explanation for any of the empirical findings in Table 1. Some researchers have tried to understand these patterns using rational frameworks where differences in average returns across stocks are due to differences in risk; since the CAPM does not explain the findings in the table, the risks considered by these frameworks are necessarily something other than beta. Other researchers have pursued behavioral explanations for these patterns, and I discuss several of these later in the article.

The facts summarized in this section have drawn attention from hedge funds. These funds implement strategies that, in the case of a characteristic F with negative predictive power for returns, buy stocks with low values of F and short stocks with high values of F . For several of the characteristics, such strategies have historically earned a higher average return than would be expected for their risk, as judged by the CAPM, even after taking transaction costs into account, at least for the low transaction costs incurred by sophisticated investors (Frazzini et al., 2015; Novy-Marx and Velikov, 2015). Hedge funds' pursuit of these strategies may explain why the predictive power of some of the characteristics in Table 1 has weakened over time. However, perhaps because of the

⁶ Other well-known predictors of stock returns that I do not discuss in detail in this article are asset growth (Cooper et al., 2008); investment (Titman et al., 2004); and accruals (Sloan, 1996). All three predict subsequent returns with a negative sign in the cross-section.

“limits to arbitrage” that I discuss in Section 3, most of the empirical findings in the table remain robust even in recent data.⁷

I have presented the predictive relationships in Table 1 in the context of the stock market because that is the context in which they were first discovered. However, several of these patterns also hold in other asset classes. For example, momentum and long-term reversal are present not only in the U.S. stock market, but also in the stock markets of several other developed countries, and across country equity indices, government bonds, currencies, and commodities (Asness et al., 2013). Similarly, the beta anomaly is present not only in the U.S. stock market but also in many international stock markets, Treasury bonds, corporate bonds, and futures (Frazzini and Pedersen, 2014).

2.3 Bubbles

One definition of a bubble is that it is an episode in which an asset becomes substantially overvalued for a significant period of time, where “overvalued” means that the price of the asset exceeds a **sensible** present value of the asset’s future cash flows. In some settings, this definition can be used productively. For example, in experimental markets, it is often possible to state exactly what the rational present value of an asset’s future cash flows is (Smith et al., 1988). Even in actual financial markets, we can sometimes compute this present value with a high degree of confidence, or at least put sharp bounds on it (Xiong and Yu, 2011). However, in many real-world contexts, it is hard to determine the sensible price of an asset, making it difficult to decide whether a particular episode constitutes a bubble and leading some observers to question whether bubbles, as defined above, even exist (Fama, 2014).

In light of this, it may be more productive to use an *empirical* definition of a bubble – to define a bubble as an episode with a set of concrete empirical characteristics.⁸ One candidate set of characteristics is: (i) the price of an asset rises sharply over some period of time and then declines, reversing much of the rise; (ii) during the price rise, there are many reports in the media and from sophisticated investors that the asset is overvalued; (iii) as the price of the asset rises and reaches its peak, there is an abnormally high volume of trading in the asset; (iv) during the episode, many investors have highly extrapolative expectations about returns, in that their expectation about the asset’s future return is strongly positively related to its recent past return; (v) during the

⁷ McLean and Pontiff (2016) show that the predictive power of characteristics such as those in Table 1 goes down after a study documenting the predictability is published in a journal – and that the decline is larger, the weaker are the limits to arbitrage. Their interpretation is that the publication of a finding brings it to the attention of hedge funds, which then exploit it.

⁸ Almost every other asset price phenomenon studied by economists – and *all* the phenomena described in Sections 2.1 and 2.2 – is defined in an empirically concrete way. For example, “momentum” is the finding that a stock’s past six-month return predicts the stock’s subsequent return with a positive sign. It therefore seems reasonable to ask that the term “bubble” be defined in a similarly concrete way.

price rise, some sophisticated investors increase their exposure to the asset; and (vi) in the early stages of the price rise, there is positive news about the asset's future cash flows.

A bubble can be defined as an episode that exhibits characteristics (i) and (ii) – these are the most essential features of a bubble – and at least two, say, of characteristics (iii) through (vi). Characteristics (iii) through (vi) are motivated, respectively, by evidence that, for assets that appear to be overvalued, we typically observe heavy trading (Hong and Stein, 2007); extrapolative investor expectations (Case et al., 2012); some sophisticated investors increasing their exposure to the asset (Brunnermeier and Nagel, 2004); and strong fundamental news in the early stages of the asset's rise in price (Kindleberger, 1978).

Given that all of characteristics (i) through (vi) are either already concrete or can be made so, it should be straightforward, given the right data, for researchers to check and agree whether a particular episode constitutes a bubble by the above empirical definition. I suggest that the historical episodes that would be categorized as bubbles by this definition would match up well with the episodes that have been informally viewed as bubbles. At the very least, it is clear that, by this definition, bubbles exist. For example, stocks in the technology sector in the U.S. in the late 1990s exhibit all six characteristics.

The challenge for economists is to write down a model of asset prices and trading volume that, on occasion and under some circumstances, generates episodes that feature most or all of the characteristics laid out above. The best-known rational framework for thinking about bubbles is the aptly-named “rational bubble” model (Blanchard and Watson, 1982; Tirole, 1985). This framework can generate characteristics (i) and (ii), but, at least in its basic form, does not generate any of characteristics (iii) through (vi). Rational bubble models also face other challenges, both theoretical and empirical. For example, Giglio et al. (2016) use data from the U.K. and Singapore to check for a rational-bubble component in the price of real estate, but find no such component.⁹

According to the empirical definition suggested here, an episode can only be categorized as a bubble with a high degree of confidence “after the fact,” once both the price rise *and* price decline have been observed. However, Greenwood et al. (2018) show that, by conditioning on other observables, it is possible to say, even while the price of an asset is still rising, that there is an increased likelihood that the episode constitutes a bubble – in other words, that the price rise will eventually be followed by a sharp decline.

⁹ See Brunnermeier and Oehmke (2013) for a detailed discussion of the rational bubble model and of other frameworks for thinking about bubble episodes.

3. LIMITS TO ARBITRAGE

A common theme in behavioral finance is that investors who are less than fully rational can affect the prices of financial assets. There is a challenge to this view, sometimes called the arbitrage critique. According to this critique, irrational investors cannot have a substantial, long-lived impact on prices. The argument is that, if irrational investors affect the price of an asset – if they push it up too high or down too low – this creates an attractive opportunity for rational investors, or “arbitrageurs,” who trade aggressively against the mispricing, causing it to disappear. As recently as the late 1980s, most financial economists found the arbitrage critique persuasive, and this slowed the development of behavioral finance.

In the 1990s, a group of researchers pushed back against the arbitrage critique in a convincing way, and this is one reason why behavioral finance began to develop rapidly in that decade. In a nutshell, the response to the arbitrage critique is this. According to the critique, if irrational investors push the price of an asset up too high or down too low, this creates an attractive opportunity for rational investors. The flaw in this argument is that, in reality, trading against a mispricing is not as attractive an opportunity as it seems. Real-world arbitrageurs – hedge funds, for example – face risks and costs when they attack a mispricing. These risks and costs lead them to trade less aggressively, which, in turn, allows the mispricing to survive. In short, there are limits to arbitrage, and this means that irrational investors *can* have a substantial impact on prices.

What are the risks and costs that arbitrageurs face? I now briefly discuss the most important of these; see Shleifer (2000) and Gromb and Vayanos (2010) for more comprehensive reviews.

Fundamental risk. A hedge fund that takes a position in a misvalued asset runs the risk that there will be adverse news about the asset’s fundamentals. Suppose that the “fair” price of a stock – the sum of its expected future cash flows, carefully forecasted using all available information and discounted at a rate that properly accounts for risk – is equal to \$20, but that excessively pessimistic traders push the market price down to \$15. A hedge fund that tries to take advantage of this by buying at \$15 runs the risk that there will be bad news about the fundamentals of the underlying company, news that pulls the fair price down to \$10, say. If the market price of the stock then converges to the fair price of \$10, as it is eventually likely to do once irrational investors correct their mistaken perceptions, the fund will lose money. If hedge funds and other arbitrageurs are risk averse, fundamental risk of this kind can be enough to make them trade less aggressively against the mispricing.

Fundamental risk can be partially hedged by taking an offsetting position in a “substitute” asset. For example, a hedge fund that buys shares of General Motors because it perceives them to be undervalued can simultaneously short shares of Ford; this protects the fund against adverse fundamental news about the economy as a whole or about the automobile sector in particular. However, it is difficult to avoid fundamental risk

entirely: in this example, the hedge fund is still exposed to bad idiosyncratic news about General Motors which does not affect Ford. Moreover, some assets – for example, aggregate asset classes – do not have good substitutes.

Noise-trader risk. Another risk that hedge funds face when they trade against a mispricing is that the mispricing may get worse in the short run: the irrational investors causing the mispricing may become even more irrational in the short term (De Long et al., 1990a). This constitutes a risk for hedge funds and other arbitrageurs because these funds typically manage other people's money (Shleifer and Vishny, 1997). If a mispricing worsens, a fund trading against the mispricing will post poor returns. Seeing this, the fund's outside investors may decide that the fund manager is not very skilled, leading them to withdraw their money and forcing the manager to close out his trade at a loss. If the fund manager recognizes this possibility in advance, he will trade less aggressively against the mispricing.

Hedge funds' use of leverage, or borrowed money, compounds the problem. If a fund borrows money to buy an undervalued asset and the mispricing then worsens, so that the asset falls further in value, the lender, seeing the value of his collateral decline, may call the loan, again forcing the fund to close out its trade at a loss. Similarly, if the fund shorts an overvalued asset and the asset then goes up in price, the asset lender will demand additional margin. If the fund cannot meet this demand, the short position will be closed out at a loss. Once again, a fund manager who foresees these potential outcomes will trade less aggressively against the mispricing.

Synchronization risk (Abreu and Brunnermeier, 2002). Suppose that an asset is overvalued and that its price continues to rise. A hedge fund manager who detects the overvaluation may be reluctant to trade against it because he does not know how many other hedge funds are aware of the mispricing. If other funds are not aware of the mispricing, then, if he trades against it, he will not be able to affect the price of the asset and will instead lose out on profits as the price continues to rise. Due to this "synchronization risk," he delays trading against the mispricing until he is confident that enough other funds are aware of the mispricing. During this waiting period, irrational investors continue to have a substantial impact on the price of the asset.

Aside from these risks, arbitrageurs also face costs that hinder their ability to correct a mispricing: the obvious costs of trading, such as commissions, bid-ask spreads, and short-selling fees, but also, importantly, the cost of discovering a mispricing and understanding the risks of exploiting it.

Thus far, we have responded to the "static" form of the arbitrage critique – the argument that, if irrational investors cause a mispricing, this creates an attractive opportunity that arbitrageurs will aggressively exploit. However, there is a second, "dynamic" version of the critique, which posits that, because they trade in suboptimal ways, the irrational investors will eventually lose most of their money and will therefore play a much smaller role in financial markets. As a consequence, there will be much less mispricing of assets.

These are several counter-arguments to this dynamic version of the critique. First, it is likely that, every year, many new inexperienced investors enter financial markets, thus replenishing the stock of irrational investors and preventing such investors from becoming “extinct.” Second, many of the unsophisticated traders in financial markets also earn labor income from their day jobs, and this allows them to pursue even unprofitable investment strategies for a long time. Third, it can take irrational investors years if not decades to lose a substantial fraction of their wealth (Yan, 2008).

In this section, I have focused on the theoretical work on limits to arbitrage. However, there is also very useful empirical research on this topic which studies specific market situations that are widely viewed as mispricings. One example is “twin shares”: shares that are claims to the same cash-flow stream but that trade at different prices (Froot and Dabora, 1999). Another is “negative equity stubs”: cases where the market value of a company’s shares is lower than the value of the company’s stake in a subsidiary firm (Mitchell et al., 2002; Lamont and Thaler, 2003). Yet another is “index inclusions”: stocks that are added to the S&P 500 index jump up a lot in price upon inclusion, even though a stock’s inclusion in the index is not intended to convey any information about the present value of the stock’s future cash flows (Harris and Gurel, 1986; Shleifer, 1986).

The existence of these situations is evidence that there are limits to arbitrage: if there were no such limits, these mispricings would not arise. These examples also serve as laboratories for understanding which of the various limits to arbitrage are the most important. In the case of twin shares, noise trader risk is of primary importance; for negative equity stubs and index inclusions, both fundamental risk and noise-trader risk play a role.

How does the concept of “efficient markets” relate to the ideas in this section, and to the themes of the article more generally? A market is efficient if the prices of the financial assets in the market “properly reflect all available information,” where “all available information” is most commonly taken to mean “all public information” (Fama, 1991). In a framework with rational investors and no frictions, financial markets are efficient: since investors are rational, they know what the “proper” price of an asset is, and since there are no frictions, they are able to freely trade the asset until its price reaches this proper level. The research described in this section indicates that, if some investors are not fully rational, or if they are subject to a friction of some kind – a short-sale constraint, say – then financial markets may be inefficient: the irrationality or friction can generate a mispricing that, due to limits to arbitrage, rational investors are unable to correct.

Market efficiency is a fundamental and historically significant concept. However, the term is now seldom used in academic conferences, and it will not appear again in the remaining sections of this article. The reason is that the terms “efficient market” and “inefficient market” are too broad to capture the debate at the frontier of finance

research today. If an economist says that he believes that markets are efficient, we understand this to mean that his preferred model of the world is one with rational investors and no frictions. However, at this point in the evolution of finance as a field, we care about the details: Which specific rational frictionless model does the researcher have in mind? For example, what form do investor preferences take? Can he write down the model and show that it explains a range of facts and makes testable predictions? Similarly, if an economist says that he believes that markets are inefficient, we understand this to mean that his preferred model of the world is one where some investors are not fully rational or where there is a friction. However, at this point, we want to know exactly which model the researcher has in mind: What is the specific irrationality or friction? Again, can he write the model down and show that it explains a range of facts and makes testable predictions? The battle that is being fought today is not between broad concepts like “efficient markets” and “inefficient markets” but between specific, precisely-defined models: long-run risk vs. extrapolation, say, or habit formation vs. gain-loss utility.

Over the past few decades, there has been a remarkable shift in the views of financial economists. Until the late 1980s, most of them embraced the arbitrage critique and thought that it was unlikely that irrational investors could have a substantial impact on asset prices. Now, at the time of writing, many if not most finance researchers accept that there are limits to arbitrage and that irrational investors can therefore affect asset prices. If there is disagreement, it is about how strong these limits are and how much irrational investors matter. The success of the work on limits to arbitrage helped to usher in a new era of intense research in behavioral finance, one where economists began to take seriously models of asset prices in which some investors are not fully rational. In Sections 4 through 9, I discuss these models in detail.

4. BELIEFS: EXTRAPOLATION

Psychology-based models of financial markets aim to improve our understanding of the data by making more accurate assumptions about people’s beliefs, preferences, and cognitive limits. In Sections 4 through 6, I review models where the focus is on the first of these, namely beliefs.

One of the most useful ideas in behavioral finance is that people have *extrapolative* beliefs: their estimate of the future value of a quantity is a positive function of the recent past values of that quantity. This idea is typically applied to beliefs about returns or growth in fundamentals: we work with models where investors’ expectation of the future return of an asset is a positive function of the asset’s recent past returns, or where their expectation of a firm’s future earnings growth is a positive function of the firm’s recent earnings growth rates. However, it can also be usefully applied to estimates of other quantities – future volatility, say, or future crash risk.

In this section, I focus primarily on *return* extrapolation because it has proved to be a particularly helpful assumption. In Section 4.1, I use a simple model to show how return extrapolation can explain a wide range of facts about asset prices. More briefly, in Section 4.2, I discuss models where investors extrapolate past fundamentals. In Section 4.3, I explore the possible sources of extrapolative beliefs. Finally, in Section 4.4, I review models of “experience effects” which posit a type of heterogeneity in extrapolation across individuals that has additional implications for portfolio holdings, volume, and prices.

4.1 Return extrapolation

Return extrapolation is the idea that people’s expectation of the future return of an asset, asset class, or fund is a weighted average of the past returns of the asset, asset class, or fund, where the weights on the past returns are positive and larger for more recent past returns. Models based on this assumption can explain, in straightforward and intuitive ways: (i) medium-term momentum, long-term reversal, and the value premium in the cross-section of average returns; (ii) excess volatility and time-series predictability in aggregate asset classes; and (iii) the formation and collapse of bubbles. Understanding these phenomena is a major goal of research on asset prices, so it is striking that a single, simple assumption can address all of them. Another appealing feature of the extrapolation approach is that it can be applied in a natural way in any asset class, and can therefore explain why empirical patterns like excess volatility, momentum, and reversal are present in many asset classes, not just the stock market.

The idea that investors have extrapolative beliefs can be found in decades-old qualitative accounts of investor behavior. The first wave of formal research on the topic appeared in the 1990s and includes papers such as Cutler et al. (1990), De Long et al. (1990b), Frankel and Froot (1990), Hong and Stein (1999), and Barberis and Shleifer (2003). Recently, there has been a second wave of research on the topic, including papers such as Barberis et al. (2015, 2018), Adam et al. (2017), Glaeser and Nathanson (2017), Cassella and Gulen (2018), DeFusco et al. (2018), Jin and Sui (2018), and Liao and Peng (2018). This second wave has been spurred in part by renewed attention to a neglected but potentially very useful type of data, namely survey data on the beliefs of real-world investors about future asset returns (Bacchetta et al., 2009; Amromin and Sharpe, 2014; Greenwood and Shleifer, 2014). Several surveys ask investors, both individual and institutional, to forecast the return on the stock market over the next six months or one year. These data provide direct evidence of extrapolative beliefs: the average belief of the surveyed investors about the future stock market return is a positive function of recent past stock market returns. But the data also point to *over-extrapolation*: the average belief of the surveyed investors is *negatively* related to the subsequent realized return, suggesting that the extrapolative beliefs are incorrect.

I now explain, with the help of a simple model, how return extrapolation can generate facts (i), (ii), and (iii) about asset prices listed at the start of this section. Consider an economy with $T + 1$ dates, $t = 0, 1, \dots, T$, and two assets: a risk-free asset whose net return is zero, and a risky asset which has a fixed supply of Q shares and is a claim to a single cash flow \tilde{D}_T to be paid at time T . The value of \tilde{D}_T is given by

$$\begin{aligned}\tilde{D}_T &= D_0 + \tilde{\varepsilon}_1 + \tilde{\varepsilon}_2 + \dots + \tilde{\varepsilon}_T, \\ \tilde{\varepsilon}_t &\sim N(0, \sigma_\varepsilon^2), \text{ i.i.d. over time.}\end{aligned}\tag{1}$$

Here, D_0 is realized and publicly announced at time 0, while $\tilde{\varepsilon}_t$ is realized and publicly announced at time t . The expected value of \tilde{D}_T at time 0 is therefore D_0 , while its expected value at time t is $D_t \equiv D_0 + \sum_{j=1}^t \varepsilon_j$. The price of the risky asset at time t , P_t , is determined in equilibrium.

Models of return extrapolation usually feature two types of investors, and I follow this approach here. The first type is the extrapolators themselves. At time t , their belief about the future price change of the risky asset is

$$E_t^e(P_{t+1} - P_t) = X_t \equiv (1 - \theta) \sum_{k=1}^{t-1} \theta^{k-1} (P_{t-k} - P_{t-k-1}) + \theta^{t-1} X_1, \tag{2}$$

in words, a weighted average of past price changes that puts more weight on the more recent past: the parameter θ is in the $(0, 1)$ interval. The “e” superscript indicates that this is the expectation of extrapolators, and X_1 denotes the extrapolators’ belief at time 1; in the analysis below, I set it to a neutral, steady-state value. For brevity, I refer to the beliefs in (2) as “return extrapolation,” even though “price-change extrapolation” would be a more accurate label.

In the Appendix, I show that if, at each date t , extrapolators maximize a utility function with constant absolute risk aversion γ that is defined over next period’s wealth and also, for simplicity, take the conditional distribution of the next period’s price change to be Normal with variance σ_ε^2 , then their per-capita share demand for the risky asset at time t is

$$N_t^e = \frac{X_t}{\gamma \sigma_\varepsilon^2}, \tag{3}$$

in words, their expectation of the future price change scaled by their risk aversion and by their estimate of the risk they are facing.

The second type of investor is arbitrageurs, whom I refer to as “fundamental traders.” For simplicity, I assume that these traders are boundedly rational: they do not have a full understanding of extrapolator demand; rather, they believe that, in future periods, the extrapolators will hold the risky asset in proportion to their weight in the population.

Table 2 Parameter values for a model of extrapolative beliefs

Parameter	Value
D_0	100
σ_ε	3
Q	1
T	20
μ^f	0.3
μ^e	0.7
γ	0.1
θ	0.7

In the Appendix, I show that, under this assumption, the per-capita share demand of the fundamental traders at time t is

$$N_t^f = \frac{D_t - (T - t - 1)\gamma\sigma_\varepsilon^2 Q - P_t}{\gamma\sigma_\varepsilon^2}, \quad (4)$$

where the “f” superscript denotes “fundamental traders.” Notice that this demand is higher, the lower is the price P_t relative to the expected cash flow D_t . By buying when the price is low relative to fundamentals and selling when it is high, these traders ensure that the price does not deviate too far from a sensible present value of the final cash flow. For our purposes, the assumption of bounded rationality is innocuous. As I explain below, if the fundamental traders were more fully rational in that they understood how extrapolators form beliefs, the level of mispricing would actually be larger than in the simpler economy I study here.

If the fraction of extrapolators and of fundamental traders in the population are $\mu^e \in [0, 1)$ and $\mu^f = 1 - \mu^e$, respectively, then, by substituting (3) and (4) into the market-clearing condition $\mu^e N_t^e + \mu^f N_t^f = Q$, we obtain the equilibrium price of the risky asset:

$$P_t = D_t + \frac{\mu^e}{\mu^f} X_t - \gamma\sigma_\varepsilon^2 Q \left(T - t - 1 + \frac{1}{\mu^f} \right), \quad t = 1, \dots, T - 1. \quad (5)$$

The first term on the right-hand side of (5) shows that the asset’s price is anchored to the expected value of the final cash flow. The second term shows that the price is a positive function of X_t : if past price changes have been strongly positive, extrapolators become more bullish about the future price change and therefore increase their demand for the risky asset, pushing its price higher. The third term is a price discount which compensates investors for the risk they are bearing.

To show how this simple model captures facts (i), (ii), and (iii) above, I use (5) to compute the price of the risky asset following a large, positive cash-flow shock. Specifically, there are $T = 20$ periods, and the remaining parameters take the values in Table 2. I set ε_1 to zero, ε_2 to 6 – a large, two-standard deviation cash-flow shock – and

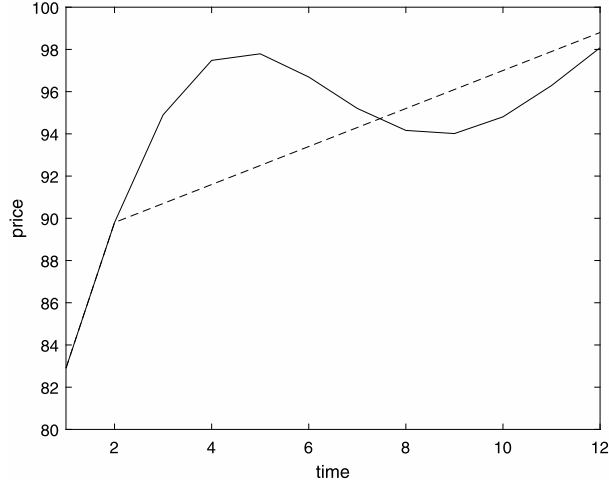


Figure 1 Asset price dynamics under extrapolative beliefs. The solid line plots the price of a risky asset following good news at time 2 about its future cash flow. Some investors in the economy, “extrapolators,” form beliefs about the asset’s future price change as a weighted average of its past price changes, while other investors, “fundamental traders,” base their demand on the difference between price and the present value of the asset’s future cash flow. The dashed line plots the price of the risky asset in an economy where all investors are fundamental traders.

ε_3 through ε_{20} to zero. The solid line in Fig. 1 plots the price of the risky asset from time 1 to time 12. The dashed line plots the “fundamental value” of the asset – the price the asset would trade at if all investors were fundamental traders. At time t , this equals $D_t - \gamma\sigma_\varepsilon^2 Q(T - t)$.¹⁰

Why does the price behave in the way shown? At date 2, there is a large positive cash-flow shock that pushes the price up. At date 3, the extrapolators, seeing the large price increase at date 2, become more optimistic about the future price change of the risky asset and therefore push the price up higher. At dates 4 and 5, they become even more optimistic, pushing the price higher still. At date 6, however, the price falls. By this point, the largest past price increases – those at dates 2 and 3 – have receded some way into the past. From (2), this means that their contribution to extrapolators’ beliefs is smaller than before. This reduces extrapolators’ enthusiasm and causes the price to drop.

From Fig. 1, we can see, informally, how a model where some investors extrapolate past price changes can explain: (i) medium-term momentum, long-term reversal, and the value premium in the cross-section; (ii) excess volatility and time-series predictability in aggregate asset classes; and (iii) bubbles.

¹⁰ In this example, the economy starts in a steady state in which $P_0 = D_0 - \gamma\sigma_\varepsilon^2 QT$ and $X_1 = \gamma\sigma_\varepsilon^2 Q$.

We can see momentum in the way the positive price change at date 2 is followed by another positive price change at date 3: the price change at date 2 leads extrapolators to become more bullish at date 3 and hence to push the price even further up on that date. We can see long-term reversal in the way the high return between date 1 and date 5 is followed by a poor return over the next few periods. If the asset has had a good long-term past return, this is a sign that extrapolators have been buying it aggressively, causing it to become overpriced; the overvaluation is then followed by low returns. And we can see excess volatility in the way that, following the good cash-flow news at date 2, the solid line rises above the dashed line: due to aggressive buying by extrapolators at date 3, the price of the risky asset is more volatile than in the absence of extrapolators.

Fig. 1 also shows that, in an economy with extrapolators, the difference between price and fundamental value predicts the subsequent return with a negative sign; for example, the high price relative to fundamental value at date 5 is followed by a low return. When the price is high relative to fundamentals, this is a sign that extrapolators have been buying the asset aggressively, causing it to be overpriced; this overvaluation is then followed by a low return. This mechanism generates time-series predictability in aggregate asset classes and a value premium in the cross-section.

Finally, the swift rise in the asset's price between date 1 and date 5 followed by a decline in the periods thereafter captures the essential characteristic of an asset bubble. If, instead of the single cash-flow shock at date 2, there were two or three substantial cash-flow shocks in close succession, the rise and fall in price would be even more dramatic and even more reminiscent of a bubble. The theory of bubbles suggested here is in essence the one articulated by Kindleberger (1978). Bubbles are initiated by strongly positive cash-flow news that cause the price of an asset to rise. This price increase leads extrapolators to become bullish, generating further price increases and even greater extrapolator enthusiasm. Eventually, as described above, extrapolator enthusiasm abates and the bubble begins to deflate.

At certain points in the cycle in Fig. 1, the extrapolators' trades are profitable: at date 3, for example, extrapolators increase their holdings of the risky asset, a move that earns them a positive return over the next period. Overall, however, they underperform the fundamental traders: at date 5, for example, extrapolators have a large exposure to the risky asset, but the asset's subsequent return is negative.

At this point, the reader may have the following question in mind: "Return extrapolation is presented as a mistake. But it sounds similar to momentum trading, which is viewed as a smart strategy. How can the same thing be both a mistake *and* smart?" The answer is that extrapolation, as defined here, and momentum trading are *not* the same thing; there is a difference in timing between the two. If there is a positive price change between time $t - 1$ and time t , a momentum trader buys immediately, at time t . For

example, his demand for shares of the risky asset at time t can be written as

$$N_t^m = \frac{P_t - P_{t-1}}{\gamma \sigma_\varepsilon^2}, \quad (6)$$

where the “m” superscript stands for “momentum traders.” The extrapolators, however, do not buy at time t ; they buy at time $t + 1$: notice, from Eq. (2), that extrapolators’ expectations at time t do not depend on the most recent price change from time $t - 1$ to t , but only on the price change from $t - 2$ to $t - 1$ and on price changes before that. For example, in Fig. 1, following the positive price change at date 2, the momentum trader buys immediately at date 2, while the extrapolator buys only at date 3. In this framework, momentum trading is profitable precisely because it “front runs” the extrapolators (Haghani and McBride, 2016).

Why do we not include the most recent price change, $P_t - P_{t-1}$, on the right-hand side of Eq. (2)? In other words, why do we not specify extrapolator beliefs as

$$E_t^e(P_{t+1} - P_t) = X_t \equiv (1 - \theta) \sum_{k=0}^{t-2} \theta^k (P_{t-k} - P_{t-k-1}) + \theta^{t-1} X_1? \quad (7)$$

At a practical level, the reason is that we want a model that generates momentum; after all, momentum is a robust feature of actual returns. A model where extrapolators form beliefs according to Eq. (7) can generate excess volatility, time-series predictability, long-term reversal, and a value premium, but struggles to generate momentum. If, as in the example above, there is good cash-flow news at date 2, extrapolators with the beliefs in (7) become more bullish *immediately* and hence push the price up a lot on that date. By date 3, however, their enthusiasm may have already started to wane: the large price increase at date 2 is now one period further back in the past and therefore has less impact on their beliefs.

What is an *economic* justification for excluding the most recent price change from Eq. (2)? One possibility is that the right-hand side of (2) actually represents the extrapolators’ belief at time $t - 1$ rather than at time t , but that, due to frictions and constraints, the extrapolators do not get around to translating this belief into a portfolio adjustment until time t . Another possibility is that some investors observe past price changes with a delay. These channels are particularly relevant in the housing market: it takes time for a transaction price to be incorporated into the Case–Shiller house price index, and it takes time to buy a house. If specification (2) – a specification that generates momentum – is especially applicable in the housing market, this may explain why there is more short-term positive autocorrelation in housing returns than in stock market returns (Glaeser and Nathanson, 2015, 2017).

The parameter θ , which determines the relative weight extrapolators put on recent as opposed to distant past returns, plays a key role in the model presented above. We

can estimate the value of θ using the survey data studied by Greenwood and Shleifer (2014) and others. In quarterly data, $\theta \approx 0.9$, suggesting that investor expectations are affected primarily by the past two years of returns. However, Cassella and Gulen (2018) use rolling regressions to document a striking fact: the value of θ estimated from the survey data varies significantly over time. For example, in the late 1990s, the value of θ fell, suggesting that the typical investor was forming beliefs based only on very recent returns.

Why would the average investor's θ vary over time? One possibility is that there is a within-individual change in θ . For example, if investors decide that there has been a shift in the forces driving the economy, they may start to put more weight on recent returns. Another possibility is that the composition of the investor population changes over time. After a sharp rise in the stock market, investors with low θ are likely to enter the market, precisely because they put heavy weight on recent observations. As a result, the average θ across investors will fall. In Section 4.4, I discuss a framework that predicts that young people are more likely to have a low value of θ .

As noted above, the extrapolation framework can generate time-series predictability, the finding that a high P/D ratio in the stock market is followed by a low return, on average. Cassella and Gulen (2018) draw a new prediction out of the extrapolation framework, namely that a high P/D ratio in the stock market will be followed by an especially low average return over the next year if, at the time of the high P/D ratio, the typical investor's θ is low: when θ is low, investors more quickly "forget" the positive price changes that caused them to become excited in the first place, and the overvaluation therefore corrects faster. Cassella and Gulen (2018) confirm this prediction in the data: years of high P/D ratios and low θ are followed by lower returns on average than are years of high P/D ratios and high θ .

The simple model presented above shows how extrapolation of past price changes can explain momentum and reversal in the cross-section, time-series predictability and excess volatility in aggregate asset classes, and the formation and collapse of bubbles. I now briefly discuss some of the more sophisticated models of extrapolative beliefs that have been used to study these phenomena.

Aggregate stock market. Barberis et al. (2015) use a model of return extrapolation to think about aggregate stock market prices. In this model, there are two assets – a risk-free asset and a risky asset – and two types of infinitely-lived investors who maximize lifetime consumption utility. One type is extrapolators whose expectation about the future price change of the stock market is a weighted average of the market's past price changes. The second type is rational traders who know the stochastic process that prices follow in equilibrium.

The model generates the important features of aggregate stock market prices: the price of the risky asset exhibits excess volatility, and the difference between the price and a measure of fundamentals predicts subsequent price changes with a negative sign.

Extrapolators have a large effect on prices despite the presence of rational traders. Why is this? Suppose that there is good cash-flow news that pushes up the stock market price, causing extrapolators to become more bullish and hence to push the price up even further. At this point, the rational traders recognize that the extrapolators are likely to stay bullish for some time – the beliefs in Eqs. (2) and (7) are persistent, particularly for higher values of θ – and that it therefore does not make sense to trade aggressively against them. This allows the extrapolators to have a significant effect on prices.

In Barberis et al.'s (2015) model, investors have constant absolute risk aversion. While this makes the model more tractable, it also makes it hard to evaluate its quantitative predictions: ratio-based quantities such as the P/D ratio are not well-defined. Jin and Sui (2018) present a model of asset prices in which the representative investor has both extrapolative beliefs about returns and preferences of the Epstein–Zin form. This model's quantitative predictions are easier to evaluate. When the model is calibrated to survey data, these predictions match important moments of actual stock market returns.¹¹

An appealing feature of models of the stock market based on return extrapolation is that they are consistent with the survey data on the beliefs of actual investors about future returns: in line with the survey evidence, some of the investors in Barberis et al.'s (2015) model have extrapolative beliefs about returns, as does the representative investor in Jin and Sui's (2018) framework. By contrast, the prominent rational models of stock market fluctuations – models based on habit formation, long-run risk, or rare disasters – are not consistent with the survey data. In these models, high stock market returns are associated with a decline in risk aversion or perceived risk, at which point investors expect lower, not higher, returns going forward.

The cross-section. Hong and Stein (1999) and Barberis and Shleifer (2003) use models in which some investors extrapolate past price changes to think about the cross-section of average returns.¹² In Hong and Stein (1999), the extrapolative beliefs are derived from other assumptions, rather than posited exogenously; in Section 4.3, I discuss the micro-foundation they propose. In Barberis and Shleifer's (2003) model, investors put assets into categories and form expectations about the future price change of each category as a weighted average of its past price changes. One advantage of this framework is that it generates not only momentum and reversal in the cross-section, but also comovement in returns that is unrelated to fundamentals, a phenomenon that is observed in the data. I discuss this framework in more detail in Section 9.

Bubbles. The model of return extrapolation summarized in Eqs. (1)–(5) can generate the most essential feature of a bubble, namely a sharp increase in an asset's price followed

¹¹ See Adam et al. (2017) for a related analysis.

¹² See Da et al. (2018) for direct evidence of extrapolative expectations about the returns of individual stocks.

by a steep decline. However, there are aspects of bubbles that this model does not capture and that call for a richer framework.

As a bubble grows, sophisticated investors often increase, rather than decrease, their exposure to the risky asset in question: Brunnermeier and Nagel (2004) show that, during the U.S. technology-stock bubble of the late 1990s, the weight of technology stocks in hedge fund portfolios was higher than the weight of technology stocks in the overall market portfolio. We cannot explain this using the two-agent model outlined above: given that extrapolators increase their holdings of the risky asset as the bubble forms, the more sophisticated fundamental traders necessarily reduce their holdings. De Long et al. (1990b) show that it is possible to generate riding of a bubble by sophisticated investors in a three-agent model that features extrapolators, boundedly-rational fundamental traders like those described in Eq. (4), and fully rational traders who understand how extrapolators form beliefs. In this case, as the bubble forms, the extrapolators increase their holdings of the risky asset; the fully rational traders also increase their holdings in order to profit from future buying by extrapolators; and the boundedly-rational fundamental traders reduce their holdings.

Most bubbles feature very high trading volume (Hong and Stein, 2007). The model in Eqs. (1)–(5) can generate high volume during periods of overvaluation, but not in a fully satisfactory way: it says that, during a bubble, the high volume results from extrapolators steadily increasing their holdings of the risky asset and fundamental traders steadily decreasing their holdings. This theory of volume faces empirical challenges: for example, during bubble episodes, few investors monotonically increase or decrease their holdings of the bubble asset; more commonly, they alternately increase and decrease their exposure over time (Liao and Peng, 2018).

Barberis et al. (2018) put forward a new theory of volume during bubble periods. They propose that investors who extrapolate past price changes do not do so blindly, but rather also pay attention to what fundamental traders focus on, namely how prices compare to fundamentals. Specifically, their model features I types of extrapolators, where the demand of extrapolators of type i for shares of the risky asset is¹³

$$N_t^{e,i} = w_{i,t} \left(\frac{D_t - (T - t - 1)\gamma\sigma_\varepsilon^2 Q - P_t}{\gamma\sigma_\varepsilon^2} \right) + (1 - w_{i,t}) \frac{X_t}{\gamma\sigma_\varepsilon^2}, \quad (8)$$

where X_t , defined in (2), is a weighted average of past price changes. Notice that this demand function depends not only on past price changes, but also, to some extent, on the fundamental trader demand in (4). Crucially, the relative weight each extrapolator puts on past price changes as opposed to fundamental trader demand varies slightly over time, and independently so across extrapolators, an idea that the authors call “wavering”

¹³ The full model incorporates a short-sale constraint, but I leave it out of Eq. (8) for simplicity.

and motivate based on time-varying attention. Specifically,

$$\begin{aligned} w_{i,t} &= \bar{w} + u_{i,t} \\ u_{i,t} &\sim N(0, \sigma_u^2), \text{ i.i.d., } \forall i, t, \end{aligned} \quad (9)$$

where \bar{w} is a constant in the $(0, 0.5)$ interval, and where $w_{i,t}$ is truncated so that its value is always between 0 and 1.

Barberis et al. (2018) show that this model generates high volume during bubble periods in a way that is more consistent with observed trading behavior. During a bubble, each extrapolator is torn between two powerful but conflicting investment signals. On the one hand, prices have recently been rising, which makes the extrapolator think that they will keep rising and encourages him to buy (the $X_t/\gamma\sigma_\varepsilon^2$ term in (8)). On the other hand, prices are very high relative to fundamentals, which makes him think that there might be a crash and encourages him to sell (the first term in parentheses on the right-hand side of (8)). During a bubble, these signals are so strong that even a small change in the relative weight the extrapolator assigns to them leads to a large portfolio adjustment, and hence to trading volume. The model predicts that, during a bubble, volume will be strongly positively related to the asset's past return. The authors confirm this prediction using data on four historical bubble episodes.¹⁴

4.2 Extrapolation of fundamentals

Section 4.1 focused on models of return extrapolation. I now turn to models where investors extrapolate past fundamentals – for example, where their expectation about the future cash-flow growth rate of a firm, or of the aggregate corporate sector, is a positive function of recent cash-flow growth rates.¹⁵

Models based on cash-flow extrapolation can explain a range of important facts about asset prices, including excess volatility and return predictability in aggregate asset classes, and long-term reversal and the value premium in the cross-section of stock returns. Suppose that the true mean growth rate of an asset's cash flows is constant, but that, at each moment of time, investors' expectation of the future cash-flow growth rate is a positive function of recent past growth rates. If cash flows have recently been rising quickly, investors become optimistic about future cash-flow growth and push the asset's price up relative to current fundamentals. Since, in reality, the mean cash-flow growth rate is no higher than before, extrapolators are disappointed by subsequent cash flows, on average, and the asset earns a low return. This price rise followed by a low average return captures excess volatility and time-series return predictability in the aggregate

¹⁴ See DeFusco et al. (2018) and Liao and Peng (2018) for alternative extrapolation-based models of volume, as well as evidence that supports the mechanisms they describe.

¹⁵ I use “cash flow” as an umbrella term for the various types of fundamentals that investors may extrapolate, whether dividends, earnings, or revenues.

stock market, and, in an economy with many assets, long-term reversal and a value premium. Moreover, if, as suggested above in the discussion of Eq. (2), there is a gap between the time at which extrapolators form beliefs and the time at which they trade, then the cash-flow extrapolation framework also generates momentum in asset returns.

Several papers formalize the intuition in the previous paragraph. An early model of cash-flow extrapolation is that of Barberis et al. (1998), who show how this assumption can explain long-term reversal and a value premium in the cross-section. I describe this model in Section 6.2. More recent models include Fuster et al. (2012), Choi and Mertens (2013), Alti and Tetlock (2014), Hirshleifer et al. (2015), and Bordalo et al. (2017, 2018a).¹⁶

Models of cash-flow extrapolation predict that the past growth in an asset's fundamentals will be negatively related to the asset's subsequent return. The evidence on this prediction is mixed. On the one hand, Daniel and Titman (2006) find that, in the cross-section, the past growth in a firm's fundamentals is not related to the subsequent return of the firm's stock. On the other hand, in an in-depth study of the shipping industry, Greenwood and Hanson (2015) show that, here, the level of fundamentals does forecast a low subsequent return. Moreover, in the context of corporate bonds, Greenwood and Hanson (2013) show that past default rates predict subsequent bond returns with a positive sign, a finding that is consistent with investors over-extrapolating default rates. Taken together, these results suggest that investors extrapolate past fundamentals only under certain conditions. However, we do not yet have a good understanding of what these conditions are.

4.3 Sources of extrapolative beliefs

In Sections 4.1 and 4.2, we saw that a simple assumption – that some investors form beliefs about future returns or cash-flow growth based on recent past returns or cash-flow growth – can explain a wide range of facts about asset prices. While this is encouraging, it immediately raises another question: Why would investors form beliefs in this way? I now discuss some possible sources of extrapolative beliefs. These fall into two categories: mechanisms based on concepts from psychology, and mechanisms based on bounded rationality, the idea that investors do not have the cognitive capacity to process all relevant information in a timely and accurate way. I describe the psychological approaches first and then turn to the bounded-rationality approaches. I focus primarily on return extrapolation but also discuss cash-flow extrapolation.

A long-standing idea, articulated by Barberis et al. (1998), is that extrapolative beliefs stem from the “representativeness heuristic,” a concept introduced by Kahneman and Tversky (1972, 1973a). According to this heuristic, when people face questions such as

¹⁶ See also Barsky and De Long (1993) and Lakonishok et al. (1994) for early discussions of cash-flow extrapolation.

“What is the probability that object A belongs to class B?,” or “What is the probability that data A are generated by model B?,” their answers are based on the extent to which A is representative of B, in other words, the extent to which A reflects the essential characteristics of B. In many situations, this heuristic generates good probability judgments. In some situations, however, it leads the individual astray. In particular, it can cause an error known as base-rate neglect.

The following example, based on Kahneman and Tversky’s experiments, illustrates base-rate neglect. Consider this description of a person named Steve:

“Steve is very shy and withdrawn, invariably helpful, but with little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.”

Is Steve more likely to be a lawyer or a librarian? Many people are tempted to say “librarian.” Kahneman and Tversky argue that they come to this answer by applying the representativeness heuristic: their judgment is based on the fact that the description is more representative of a librarian than a lawyer; shyness, for example, is more common among librarians than lawyers. However, in this case, the representativeness heuristic is leading people to the wrong answer: it is causing them to neglect the “base rate,” the fact that there are far more lawyers in the population than librarians, so that, even if Steve *sounds* like a librarian, he is actually more likely to be a lawyer. In mathematical terms, Bayes’ rule states that

$$\frac{p(\text{librarian}|\text{data})}{p(\text{lawyer}|\text{data})} = \frac{p(\text{data}|\text{librarian})}{p(\text{data}|\text{lawyer})} \frac{p(\text{librarian})}{p(\text{lawyer})}. \quad (10)$$

The original question – Is Steve more likely to be a lawyer or a librarian? – asks people to evaluate the left-hand side of Eq. (10). One interpretation of the representativeness heuristic is that it leads people to base their judgment on the first ratio on the right-hand side. In doing so, they neglect the second term on the right-hand side, the relative likelihood of librarians and lawyers in the population. In many situations, the first ratio on the right-hand side is a good approximation to the left-hand side. In this case, it is not.

The representativeness heuristic and the base-rate neglect it gives rise to offer a foundation for extrapolative beliefs. To see the argument in the case of return extrapolation, consider an asset that has posted several periods of high returns. An investor who uses the representativeness heuristic will over-estimate the asset’s true mean return: his judgment is affected by the fact that the high realized returns are more representative of an asset with a high true mean return than of an asset with a low or moderate true mean return. However, by reasoning in this way, he is neglecting the base rate – the fact that it is relatively rare for an asset to have a high true mean return. An asset that posted a few periods of high returns is actually more likely to be an asset with a moderate true

mean that just happened to have a few good draws than to be an asset with a high true mean.

A similar argument can be used to motivate extrapolation of fundamentals. Following a few periods of high cash-flow growth for an asset, investors who use the representativeness heuristic believe that the true mean cash-flow growth rate is high: the data are more representative of an asset whose true mean growth rate is high than of an asset whose true mean growth rate is moderate or low. However, they are over-estimating the growth rate because they are neglecting the base rate – the fact that it is rare for an asset’s true mean growth rate to be high for an extended period of time. An early formalization of this argument appears in Barberis et al. (1998).

Bordalo et al. (2017) and Bordalo et al. (2018a) link the representativeness heuristic to extrapolative beliefs using a related, but distinct, argument. They build on a model of representativeness laid out in Gennaioli and Shleifer (2010) and Bordalo et al. (2014) in which a type t is representative of a group G if the ratio

$$\frac{p(t|G)}{p(t|-G)} \quad (11)$$

is high, in other words, **if type t is more common within group G than outside it.** For example, “old people” (t) are representative of a Florida resident (G), not because most people in Florida are old – they are not – but because it is old people, rather than other age groups, that are over-represented in Florida relative to other states: the ratio

$$\frac{p(t|\text{Florida resident})}{p(t|\text{non-Florida resident})}$$

is higher for $t = \text{“old person”}$ than for $t = \text{“middle-aged person”}$ or $t = \text{“young person.”}$ Bordalo et al. (2014) propose that people overweight representative types when forming beliefs, so that their distorted belief $\hat{p}(t|G)$ is given by

$$\hat{p}(t|G) = \frac{1}{Z} p(t|G) \left(\frac{p(t|G)}{p(t|-G)} \right)^\theta, \quad (12)$$

where $\theta > 0$ and Z is a normalizing factor. **As such, people over-estimate the likelihood that a randomly-chosen Florida resident is old. While the distorted beliefs in (12) are incorrect, they contain a “kernel of truth”: they take a true feature of the data – for example, that old people are over-represented in Florida – and exaggerate it.**

Bordalo et al. (2017) and Bordalo et al. (2018a) incorporate this belief distortion into a dynamic framework they call “diagnostic expectations.” Let ω_t be the state of the economy at time t . The individual’s assessment, at time t , of the probability that the state of the economy at time $t + 1$ will be ω_{t+1} is

$$\hat{p}(\omega_{t+1}|\omega_t) = \frac{1}{Z} p(\omega_{t+1}|\omega_t) \left(\frac{p(\omega_{t+1}|\omega_t)}{p(\omega_{t+1}|E_{t-1}(\omega_t))} \right)^\theta. \quad (13)$$

Here, $p(\omega_{t+1}|\omega_t)$ is again the rationally-assessed probability. As in (12), the individual assigns greater likelihood to states ω_{t+1} that are more representative of the current state ω_t . An important modeling choice is to specify what “ $-G$ ” corresponds to in a dynamic setting. In (13), it is the time t state that the individual expected at time $t-1$, in other words, the time t state in which he receives no news, relative to his prior expectation. Once again, the beliefs in (13) contain a kernel of truth: they take the time $t+1$ states whose true likelihood has gone up in light of the time t news, and inflate their probability.

The diagnostic expectations framework can generate extrapolation of fundamentals. Suppose that ω_t is the state of a firm’s fundamentals at time t , and that the time t state is good, so that ω_t is higher than expected. Good fundamentals at time $t+1$ are more representative of the good time t state: under the true process for fundamentals, the ratio

$$\frac{p(\omega_{t+1}|\omega_t)}{p(\omega_{t+1}|E_{t-1}(\omega_t))} \quad (14)$$

is larger for high values of ω_{t+1} than for low values of ω_{t+1} . Since investors overweight representative states when forming beliefs about the future, they are overly optimistic about future fundamentals. In this sense, they over-extrapolate the good news at time t . Bordalo et al. (2017, 2018a) incorporate diagnostic expectations into models of individual stock returns and credit cycles, respectively. The extrapolative mechanism described above plays an important role in both models.

Another idea from psychology that has been used to motivate extrapolative beliefs is a belief in a “law of small numbers” (Tversky and Kahneman, 1971). The law of large numbers is a mathematical fact which, informally, states that a large sample of data will reflect the properties of the model that generated it. For example, the ratio of Heads to Tails in a large sample of tosses from a fair coin is very likely to be close to one. Tversky and Kahneman (1971) propose that people have an incorrect belief in a law of small numbers: they think, incorrectly, that even a small sample of data will reflect the properties of the model that generated it. One motivation for this idea is the “gambler’s fallacy”: the tendency, after observing five tosses of Heads from a fair coin, to predict that the next toss will be Tails (Benjamin et al., 2017). An interpretation of this is that people expect even a small sample of tosses to reflect the essential characteristic of a fair coin, namely an even mix of Heads and Tails. If the next toss is to bring the sample closer to this expected 50:50 mix, it must be Tails.¹⁷

An individual who believes in a law of small numbers draws overly strong inferences from small samples: since he believes that even a short sample reflects the properties of

¹⁷ While people incorrectly believe in a law of small numbers, they do not believe in the law of large numbers, even though the latter is a true mathematical fact (Tversky and Kahneman, 1971). Benjamin et al. (2015) explore some economic implications of this non-belief in the law of large numbers.

the model that generated it, he thinks that he can infer a lot about the model from the sample. This provides a basis for extrapolative beliefs. To illustrate the idea in the case of return extrapolation, consider again an asset that has posted a few periods of high returns. If an investor believes that even this short sample reflects the properties of the model that generated it, he will over-estimate the asset's true mean return and will therefore forecast a high return on the asset going forward. A similar argument can be used to motivate extrapolation of fundamentals. Rabin (2002) and Rabin and Vayanos (2010) present models of a belief in a law of small numbers and show how it can lead to extrapolative beliefs.¹⁸

The behavior in the gambler's fallacy, where people predict Tails after five Heads, thereby expecting a trend to reverse, appears inconsistent with the evidence of return extrapolation in the stock market, where investors expect a trend in returns to continue. However, there is a way to reconcile these results using the belief in a law of small numbers (Rabin and Vayanos, 2010). In the case of the gambler's fallacy, people *know* the model generating the data: the coin is known to be fair. In this case, a belief in a law of small numbers leads them to predict a reversal: only through a reversal will the data come closer to reflecting the essential property of a fair coin. In the case of the stock market, investors know much less about the model generating the data; for example, they are unsure about the true mean stock return. Here, a belief in a law of small numbers leads them to extrapolate: they make an overly strong inference about the mean from the sample and therefore expect any trend they observe to continue.

The sources of extrapolative beliefs described above are based on psychological concepts – on base-rate neglect, representativeness, or a belief in a law of small numbers. A number of papers argue instead that extrapolative beliefs stem from certain kinds of bounded rationality. I now discuss some of the ideas in this line of work.

Hong and Stein (1999) propose that investors extrapolate assets' past returns because they are unable to observe the assets' underlying fundamentals – here, news about the assets' future cash flows – and view returns as signaling something about these fundamentals. Specifically, the authors build a model with a risk-free asset and a risky asset, and two sets of investors, “newswatchers” and “momentum traders.” Because of cognitive processing limits, each type of investor can attend to only one kind of information. Newswatchers process only private information about the risky asset's fundamentals; they do not condition their demand on past price changes. Momentum traders, by contrast, do not observe the private information; their demand at time t is a function only of the price change from time $t - 2$ to time $t - 1$. Whether this demand is a

¹⁸ A belief in a law of small numbers has been used in the same way to interpret the so-called “hot-hand fallacy,” whereby people think that, given two basketball players with the same average ability, the one who has made more of his last few shots is significantly more likely to make his next one, even though there is not enough predictability in the data to justify such a belief (Gilovich et al., 1985).

positive or negative function of the past price change is determined endogenously in equilibrium.

In the Hong and Stein (1999) model, private information is assumed to diffuse slowly through the newswatcher population. It is therefore optimal for momentum traders' demand for the risky asset to be a positive, rather than negative, function of the asset's past price change: if the past price change is positive, this is a sign that good private information is diffusing through the newswatcher population; since this information is likely to keep diffusing through the newswatcher population in the near future, the price of the asset is likely to keep rising, making it optimal for momentum traders to increase their demand today. As such, momentum traders extrapolate past returns. In particular, since they do not observe the private information, they keep buying the risky asset even after the good news has diffused across all the newswatchers. This, in turn, causes the asset's price to overshoot. Hong and Stein (1999) show that the interaction of newswatchers and momentum traders generates medium-term momentum and long-term reversal in the risky asset price. Hong et al. (2000) provide evidence in support of the model's predictions.

Fuster et al. (2010) present an alternative bounded-rationality foundation for extrapolative beliefs. They propose that, when fitting a model to data for the purpose of making a forecast, people have a preference for *simple* models, and that this, in turn, can lead to over-extrapolation. Suppose that an economic variable follows a complex process – an autoregressive $AR(p)$ process with $p = 20$, say, under which the variable exhibits positive autocorrelation in the short run and partial mean-reversion in the long run. If an individual analyzes the past realizations of this variable using a simpler model – an $AR(q)$ process with $q < 20$, say – he will detect the short-term positive autocorrelation but will miss the long-term mean-reversion. As a consequence, he will extrapolate a past trend too far into the future. Fuster et al. (2010) propose a model of belief formation that they label “natural expectations” under which an individual's beliefs are a weighted average of the rational forecast and the forecast obtained using his preferred simple model. Fuster et al. (2012) show formally how such expectations can generate over-extrapolation and, in an asset pricing context, excess volatility and return predictability.

Glaeser and Nathanson (2017) propose another bounded-rationality foundation for extrapolative beliefs. Under some assumptions, the price P_t of a risky asset is approximately equal to $D_t/(r - g_t)$, where D_t , r , and g_t are the current dividend level, discount rate, and expected dividend growth rate, respectively. In Glaeser and Nathanson's model, investors incorrectly believe that the market price P_t of a risky asset equals $D_t/(r - g)$, in other words, that the price reflects a constant, rather than time-varying, expected dividend growth rate. As a result, if the asset's price rises sharply, investors incorrectly attribute this to a large rise in D . This leads them to over-estimate the dividend growth rate, which, when coupled with their belief that $P = D/(r - g)$, means that they also

over-estimate the future growth of prices. As a consequence, they effectively engage in return extrapolation: they buy the asset after good past returns in anticipation of good future returns.

Finally, Greenwood and Hanson (2015) argue that a type of bounded rationality called competition neglect contributes to over-extrapolation of fundamentals. A firm with high past cash-flow growth attracts competition, reducing the firm's ability to grow its cash flows at the same rate in the future. If investors fail to take this into account, they will over-estimate the firm's future cash-flow growth.

4.4 Experience effects

In the extrapolation-based models of Section 4.1, an investor's demand for an asset depends on a weighted average of the asset's past returns, where recent returns receive more weight. Malmendier and Nagel (2011) propose a framework that is related, but also distinct in an important way. They suggest that an investor's demand for a risky asset is based on a weighted average of the returns that the investor has experienced *during his lifetime*. Specifically, the demand of investor i at time t for a risky asset depends on the quantity

$$\frac{1}{A} \sum_{k=1}^{\text{age}_{i,t}-1} (\text{age}_{i,t} - k)^{\lambda} R_{t-k}, \quad (15)$$

where $\text{age}_{i,t}$ is investor i 's age at time t and R_{t-k} is the asset's return in period $t - k$. Notice that this quantity is a function only of returns the investor has personally lived through, and that, when $\lambda > 0$, he puts more weight on more recent past returns. The constant A scales the weights on the past returns so that they sum to one.

The specification in (15) generates a type of heterogeneity in extrapolation across investors, one that leads to new predictions, especially about the relative behavior of the young and the old. For example, it predicts that, following two or three years of strongly positive stock market returns, the young will increase their exposure to the stock market more than the old: in (15), the past two or three years are a bigger determinant of investor demand for the young than for the old. Similarly, after two or three years of poor returns, the young will decrease their exposure more than the old.

Malmendier and Nagel (2011) show that the specification in (15) has explanatory power for the portfolio holdings of real-world investors. Using data from the Survey of Consumer Finances, they document that, whether an individual participates in the stock market, and the fraction of wealth that he invests in the stock market if he does participate, depend in a positive and significant way on the market returns he has experienced during his lifetime, in other words, on the quantity in (15). Their estimate of λ is approximately 1.5, which implies substantial weight on even distant past returns. Additional evidence suggests that the demand in (15) reflects beliefs rather than preferences: a survey-based measure of an investor's expectation about the future stock

market return is positively correlated with the weighted average of returns the investor has experienced over his lifetime.¹⁹

Malmendier et al. (2017) present a model of asset prices, volume, and portfolio choice in which investors extrapolate based on data they have personally experienced. There are two assets: a risk-free asset and a risky asset, which, at time t , pays a dividend $d_t \sim N(\mu, \sigma^2)$, i.i.d. over time. The economy is populated by investors of different ages who do not know the value of μ . Investor i 's estimate of μ at time t is a weighted average of the dividends paid during *his* lifetime, with more weight on more recent dividends:

$$\frac{1}{A} \sum_{k=0}^{\text{age}_{i,t}} (\text{age}_{i,t} + 1 - k)^\lambda d_{t-k}. \quad (16)$$

As such, this specification couples extrapolation of fundamentals with a focus on personal experience. The authors show that, under these assumptions, young investors increase their portfolio holdings more than old investors following a good cash-flow realization. The model also generates higher levels of volume during periods of sustained good or bad cash-flow news as these news lead to increased disagreement and hence trading across cohorts of different ages. The authors present evidence consistent with the model's predictions.²⁰

In Section 4.3, I noted that extrapolative beliefs are typically motivated by the representativeness heuristic, a belief in a law of small numbers, or cognitive limits that make it hard for investors to monitor fundamentals or use complex models. While the specifications in (15) and (16) are mathematically similar to those in Sections 4.1 and 4.2, they are often motivated using a different psychological concept, namely, the “availability heuristic” (Kahneman and Tversky, 1973b). According to this heuristic, an individual assesses the probability of an event by the extent to which he can recall instances of the event from memory; the easier it is to recall such instances, the more likely he judges the event to be. This is often a reasonable heuristic, but, in some situations, it can lead the individual astray because of biases in recall: more recent events and events with more emotional impact come to mind more easily. Consider an individual trying to judge the likelihood of being mugged in, say, the Wicker Park neighborhood of Chicago. If there have been some recent muggings in Wicker Park, or if a friend of the individual was mugged at some point in Wicker Park, it will be easier to recall instances of Wicker Park muggings, and this may lead the individual to overestimate the likelihood of a mugging in that neighborhood.

¹⁹ There are now several papers that document the effect of personal experience on beliefs about economic outcomes; see, for example, Das et al. (2017) and Kuchler and Zafar (2017).

²⁰ See Collin-Dufresne et al. (2016), Ehling et al. (2018), and Nagel and Xu (2018) for other ways of incorporating experience effects into models of asset prices.

The availability heuristic can be used to motivate the expression in (15). Suppose that an investor is trying to judge how good the future stock market return is likely to be. To make this judgment, he may bring to mind the past stock market returns that he can remember. Since he is more likely to recall returns that he personally observed in his lifetime, these may affect his judgment more than those he did not observe; in the extreme, the returns he did not live through may receive a weight of zero, as in (15). Similarly, since he is more likely to recall recent returns than more distant ones, recent returns may receive more weight in his judgment, as in (15) when $\lambda > 0$.

The availability heuristic is a plausible explanation of why investors put more weight on returns they have personally observed. However, it may not be the best explanation for why investors put more weight on more recent returns. Empirically, investor demand for assets depends more heavily on recent past returns even when the investment decision is not based on information recalled from memory, so that the availability heuristic is unlikely to play a role. For example, retail investors' propensity to buy individual stocks and mutual funds is a positive function of these assets' recent returns. However, when people invest in a mutual fund or individual stock, they rarely do so based on their recollection of the assets' past returns, simply because they have no memory of these returns. More likely, they read a news report or have a conversation with a friend that draws their attention to the specific stock or fund. They then look up the asset's past returns and what they see leads them to invest in it. Representativeness, a belief in a law of small numbers, and cognitive limits may therefore offer more parsimonious accounts than the availability heuristic for why recent returns are weighted more heavily: they apply both in situations where past returns are recalled from memory *and* in situations where they are obtained from an external source.

4.5 Extrapolative beliefs: Summary

The concept of extrapolative beliefs offers a simple and intuitive way of understanding a host of important facts about asset prices. It has other economic applications as well. For example, Bordalo et al. (2018a) use extrapolative beliefs as the basis for a model of credit cycles – of lending to firms that is by turns excessive and insufficient – while Gennaioli et al. (2016) use them to think about corporate investment behavior.

While we have learned a lot from the research on extrapolation, many questions remain. For example: What determines θ , in other words, how far back people look when forming beliefs about the future? Why does θ vary over time? Which quantity are investors forming beliefs about? Is it returns, fundamentals, or both? Does extrapolation affect only forecasts of the mean of the distribution investors care about, or forecasts of higher moments too? And most important, what is the root cause of extrapolative demand for assets? Section 4.3 lists a number of possibilities, but we do not know whether they are the true sources of extrapolation. Other psychological, neural, or biological mechanisms may play just as large a role. Throughout their evolutionary

history, humans have had to make forecasts about the sources of reward they encounter – sources of food, for example. It is likely that deep mechanisms have evolved for making these judgments. A better understanding of these mechanisms may help us make sense of people's forecasts about the stock market, a very modern source of reward.

5. BELIEFS: OVERCONFIDENCE

An important line of research in behavioral finance builds on the idea that people are overconfident. The term “overconfidence” refers to a number of related but distinct psychological phenomena (Moore and Healy, 2008). I focus on two of these: “overprecision” and “overplacement.”

Overprecision is the finding that people overestimate the precision of their judgments. In a typical study, an individual is asked to estimate various quantities – the number of gas stations in Connecticut, say – but also to give a 90% confidence interval for each estimate. If the individual's judgments are well calibrated, the point estimates will fall into the intervals approximately 90% of the time. In practice, they fall into the intervals only about 50% of the time, suggesting that the individual is too confident in his estimates (Klayman et al., 1999; Soll and Klayman, 2004).

Overplacement is the finding that people have overly rosy views of their abilities relative to others. In surveys, for example, considerably more than 50% of respondents rank themselves above the 50th percentile on a variety of positive traits (Dunning et al., 1989, 2004).

A growing body of research argues that the concept of overconfidence can help explain a number of puzzling facts in finance, most notably the very high trading volume in financial markets. To give just one data point: since 1998, turnover in the U.S. stock market has exceeded 100% on an annual basis (French, 2008).

There are several non-speculative reasons why someone might trade a financial asset – reasons that have nothing to do with beliefs about future price changes. These include trading to meet a liquidity need, trading to rebalance a portfolio, and trading to minimize taxes. However, it is unlikely that these motives explain much of the volume we observe. Most trading is likely driven by investors' beliefs about the future price of the asset – specifically, by disagreement among investors about what this future price will be.

Why, in general, would people disagree? Suppose that two individuals have the same prior belief about the future value of some quantity, that they then both observe the same information signal about this future value, and that they are both fully rational. It follows that they will have the same posterior belief about the future value of the quantity. This implies that, if two people are to disagree, one of three things must be true: (i) they have different prior beliefs; (ii) they observe different information; or (iii) one or both of them is not fully rational.

To build a model of trading volume based on disagreement, there are therefore three approaches we can try. We can assume that investors have different prior beliefs; that they observe different information, for example, that some information is “private” to a specific investor rather than publicly available; or that some investors are not fully rational. These three approaches make different predictions about trading volume. An informal summary is this. Models where rational investors observe different information tend to predict relatively low volume – and in extreme cases, no volume at all. Models where investors are not fully rational – specifically, models where investors are overconfident – can generate substantial volume, and this is one reason why overconfidence has become an important concept in behavioral finance. Models where investors have different priors are, loosely speaking, somewhere in between: they typically generate more trading than models with private information but less trading than models with overconfident investors. I now discuss the three cases in more detail.

When investors are rational and their rationality is common knowledge, it is hard to generate substantial trading volume simply on the basis that investors observe different information, an insight that was fully appreciated only in the 1980s (Milgrom and Stokey, 1982). Suppose that an investor does some research to estimate an asset’s future cash flows. Based on this research, he decides that a sensible price for the asset is \$20. When he checks the market price, however, he sees that it is \$15. A rational investor does not react to this by buying many shares of the seemingly cheap stock. Instead, he infers that other investors have probably unearthed less positive information than he. This leads him to lower his own valuation of the stock and hence to be less inclined to trade it.

A model in which investors are overconfident generates substantially more trading volume (Odean, 1998a; Eyster et al., 2018). More precisely, following Eyster et al. (2018), I will say that an investor is “overconfident” if he overestimates the precision of his *own* information signal, and “dismissive” if he underestimates the precision of *other people’s* information signals. Consider an economy with many investors, each of whom receives a private signal. Both in the case where each investor is overconfident and in the case where each investor is dismissive, there is substantial volume. The intuition is straightforward. Each investor can tell from the market price that other investors have received signals that differ from his own. However, as a result of either overconfidence or dismissiveness, he underestimates the precision of these other signals relative to the precision of his own signal, and therefore does not update his beliefs very much. Since none of the investors alters his beliefs very much, there is substantial disagreement among them about the value of the asset, and this leads to trading.²¹

An intuitive prediction of the overconfidence framework is that more overconfident people will trade more. Grinblatt and Keloharju (2009) test this using data from Fin-

²¹ See Banerjee and Kremer (2010) and Banerjee (2011) for other dismissiveness-based models of trading volume.

land. At the age of 19 or 20, every Finnish man goes into the military. As he begins his service, each conscript takes a series of tests – some psychological tests, and some tests of intellectual aptitude. One of the psychological tests asks him how confident he is, on a scale from 1 to 9. Grinblatt and Keloharju (2009) measure the individual's overconfidence as his self-reported confidence minus how confident he should be based on his performance on the aptitude tests. Strikingly, this measure of overconfidence, computed when the individual is 19 or 20 years old, predicts how frequently he trades stocks several years later when he opens and uses a brokerage account.

Another test of the overconfidence hypothesis builds on the finding that, on average, men are more overconfident than women (Lundeberg et al., 1994). In the trading context, this predicts that, on average, men trade more and earn lower returns due to the transaction costs they incur. Barber and Odean (2001) confirm this prediction using data from a large U.S. discount brokerage firm on the trading of 78,000 individuals between 1991 and 1997.²²

Eyster et al. (2018) argue that a different departure from rationality, “cursedness,” may also be helpful for thinking about trading volume. In models based on overconfidence or dismissiveness, each investor infers from market prices that other investors have signals that differ from his own, but, because he underestimates the precision of these signals relative to the precision of his own signal, he does not update his beliefs very much. By contrast, the cursedness hypothesis posits that, when an investor sees that the market price of an asset differs from his own valuation of the asset, he fails to understand what this implies, namely that other investors have signals that differ from his. As a result, he does not update his beliefs. If many investors exhibit such cursed thinking, there will be substantial disagreement among them and hence heavy trading. The implications of cursedness for volume are similar to those of extreme dismissiveness – the case where each investor recognizes that other investors have signals that differ from his own but believes that those signals have zero precision.

Earlier, I listed three sources of investor disagreement: different priors, different information, and a departure from rationality. So far, I have discussed the latter two approaches. Morris (1996) analyzes the case of different priors. In his model, investors start with different priors, but then observe the *same* sequence of public news and update their beliefs in a rational manner. The model predicts substantial trading volume. This is a surprising result: one might have thought that different priors would not generate much trading volume. Suppose that investor A has a more optimistic prior than investor B. One might have thought that, as news is publicly released over time, investor A would remain more optimistic than investor B, and that, as a result, trading volume would be muted. Morris (1996) shows that, in fact, even though the two investors observe the same information, the identity of the more optimistic investor can vary a lot

²² See Glaser and Weber (2007) for additional evidence linking overconfidence to trading activity.

over time. As a consequence, there is substantial trading. Nonetheless, the different-priors framework does not lead to as much volume as the overconfidence framework. In the different-priors framework, after many periods of public news, investors' beliefs converge and volume declines.

So far, we have used the concept of overconfidence to think about trading volume. Daniel et al. (1998, 2001) show that overconfidence also offers an explanation of some basic facts about asset prices. In the simplest version of their model, there are three dates: $t = 0, 1$, and 2. There is a risk-free asset and a risky asset that is a claim to a cash flow at time 2. The representative investor is risk-neutral. At time 1, by doing some research, he obtains a private signal about the value of the cash flow. If he is overconfident – if he overestimates the precision of the signal – the asset will be misvalued: overvalued if the signal is good and undervalued if it is bad. At time 2, when the cash flow is announced, there is a price reversal, on average, as the misvaluation is corrected. The excessive movement in the asset price at time 1 followed by the correction at time 2 offers a way of thinking about several empirical patterns: excess volatility and return predictability in aggregate asset classes, and long-term reversal and the value premium in the cross-section of average returns.

Daniel et al. (1998) show that an extension of the above model can capture other aspects of the data as well – in particular, medium-term momentum. Suppose that there are now four dates, $t = 0, 1, 2$, and 3. At time 1, the investor again observes a private signal. At time 2, public information is released. Finally, at time 3, the value of the cash flow is announced. Suppose that the investor is not only overconfident but that his degree of confidence varies over time: if the public information at time 2 is consistent with the private signal at time 1, then, at time 2, the investor becomes even more confident in his private signal; but if the public signal is not consistent with the private one, his confidence in the private signal remains unchanged. This asymmetric updating is motivated by “self-attribution bias” – the tendency, driven by a desire to maintain a positive self-image, to give oneself credit for a good outcome but to blame a bad outcome on extraneous bad luck. The idea, in the case where the public signal is inconsistent with the private one, is that, to maintain a positive view of himself, the investor is reluctant to accept that there might be a flaw in the research he conducted to generate the private signal.

Biased updating of this kind generates both medium-term momentum and long-term reversal. Suppose that the private information at time 1 is good, so that the price of the risky asset goes up at that time. If the public signal at time 2 is also good, the price rises further because of the investor's increased confidence in his favorable private signal. If the public signal at time 2 is bad, the price remains largely unchanged as the investor maintains his confidence in his private signal. On average, then, the price rises at time 2, generating momentum. The announcement of the cash flow at time 3 generates a reversal, on average.

5.1 Disagreement with a short-sale constraint

The models we have encountered so far in this section posit that investors are overconfident, but do not impose any trading frictions. An important framework in behavioral finance couples overconfidence-driven disagreement with a specific friction, namely, a constraint on short sales.

Models that combine investor disagreement with a short-sale constraint are appealing because they can explain why an asset might become significantly overvalued. There are two mechanisms through which this overpricing can occur: a static mechanism and a dynamic one, described in the late 1970s by Miller (1977) and Harrison and Kreps (1978), respectively. These papers, now well known, were largely ignored for two decades after they appeared, probably because they departed from the rational frictionless framework which dominated finance thinking at the time. They were “discovered” only in the late 1990s when researchers realized that they offered a useful way of thinking about the U.S. technology-stock bubble.

To see the static channel through which disagreement and a short-sale constraint can generate overpricing, suppose that one group of investors has optimistic expectations about an asset’s future cash flows, while another group has pessimistic expectations. If it were possible to sell short, the pessimistic investors would do so; the price of the asset would then reflect the expectations of both groups of investors, and there would not be any mispricing. However, if short sales are not possible, the pessimists do not take a position in the asset. The price of the asset then reflects only the expectations of the optimists. As such, the asset is overpriced.

Harrison and Kreps (1978) describe a dynamic mechanism through which disagreement, coupled with a short-sale constraint, can lead to overvaluation. If there is a lot of disagreement among investors about an asset’s future cash flows, then, today, each investor is willing to pay more for the asset than his own (discounted) expectation of the asset’s future cash flows because he reasons that, at some point in the future, after new information is released and investors update their beliefs, there is a chance that some of the other investors will be more optimistic than he and that he will therefore be able to sell the asset to them at a premium. In other words, the existence of a “resale option” means that, when there is significant disagreement, the asset is overvalued relative to the present value of its future cash flows as perceived by the asset’s holders. The short-sale constraint plays an important role in this logic. If short sales were possible, investor A could exploit investor B’s greater optimism by waiting for investor B to turn optimistic, shorting the stock, and selling it to him. If short sales are not possible, however, the only way investor A can exploit investor B’s greater optimism is by buying the asset *today*, so that, when investor B becomes optimistic, he can sell it to him. This extra demand for the asset today causes the overvaluation.

Scheinkman and Xiong (2003) present a model that builds on Harrison and Kreps’ (1978) framework while incorporating a number of new elements. They explicitly in-

introduce overconfidence as the source of investor disagreement and study the model's implications for volume as well as prices. In the model, there is a risky asset that is a claim to a dividend stream whose mean f is unobserved. There are two groups of risk-neutral investors, group A and group B. Both groups observe two public signals that provide information about f ; call them signal A and signal B. Group A overestimates the informativeness of signal A while group B overestimates the informativeness of signal B. Think of the two signals as the editorial pages of two newspapers, N_A and N_B , and suppose that group A overestimates the informativeness of newspaper N_A while group B overestimates the informativeness of newspaper N_B . Every day, the investors update their beliefs based on what they read in the two newspapers, with each group putting more weight on what its preferred newspaper says. Finally, short sales are not allowed.

Scheinkman and Xiong (2003) confirm Harrison and Kreps' result: because of the resale option, the price of the asset is higher than the present value of the asset's future cash flows, as perceived by the asset's holders. For example, group A is willing to pay more than its present value of the future cash flows because there is a chance that, at some point in the future, group B will read something in its preferred newspaper that causes it to become more bullish, allowing group A to sell the asset on at a premium. Group B investors reason in the same way.²³

The framework I have outlined – one that couples overconfidence-driven disagreement with a short-sale constraint – has become influential not just because it offers a theory of overvaluation, but also because it can explain why overvaluation is often accompanied by high trading volume, an empirical regularity emphasized by Hong and Stein (2007). Most historical bubble episodes feature very high trading volume. During the technology-stock bubble of the 1990s, for example, it was not just that technology stocks attained high prices; they were also heavily traded. Similarly, growth stocks with high prices relative to fundamentals have higher turnover than value stocks with low prices relative to fundamentals.

Scheinkman and Xiong (2003) show how overconfidence-based disagreement and a short-sale constraint can explain this coincidence of overvaluation and high volume. If the investors in their model become more overconfident, or if signals A and B become more informative, this leads to larger fluctuations in the relative optimism of the two investor groups. This, in turn, leads to higher trading volume, but also greater overvaluation: the resale option is now more valuable as each investor sees that, tomorrow, he has a chance of selling to someone significantly more optimistic than himself.²⁴

²³ See Section IV of Xiong (2013) for a numerical example that illustrates this logic.

²⁴ Hong and Sraer (2013) use the disagreement framework to think about “quiet” bubbles – the minority of bubbles, often associated with debt securities, that do not feature high volume – while Hong and Sraer (2016) use it to explain the relationship between beta and average return in the cross-section of stocks.

A prediction of the framework of this section is that assets that people disagree about more will be more overvalued and therefore have lower average returns. Diether et al. (2003) document a simple fact consistent with this prediction: stocks that analysts disagree about more – stocks for which analysts’ forecasts of future earnings are more dispersed – have a lower average return.

The research reviewed in this section shows that overconfidence-based disagreement can shed light on a number of facts about trading volume and asset prices. However, one aspect of this framework limits its scope: while overconfidence can lead people to hold quite different beliefs, it does not say exactly what these beliefs will be; for example, the identity of signals A and B in Scheinkman and Xiong’s (2003) model is left unspecified. On this dimension, the extrapolation-based models of Section 4 go one step beyond the models of Section 5. Models of extrapolative beliefs implicitly incorporate overconfidence-based disagreement: in a model with extrapolators and fundamental traders, each type of trader fails to learn anything from the fact that the other type is willing to trade with him. However, these models also explicitly specify each trader’s beliefs. This, in turn, allows them to make more precise predictions which can be brought to the data.

6. OTHER BELIEF-BASED APPROACHES

The effort to build psychologically-realistic models of investor beliefs has focused primarily on extrapolation of the past (Section 4) and overconfidence (Section 5). However, economists have also explored other belief specifications. In this section, I review some of these.

6.1 Sticky beliefs

Investors with extrapolative or overconfident beliefs adjust their views too much in response to information and thereby cause excessive movements in prices. For example, Fig. 1 shows that, following good cash-flow news that pushes the price of an asset up, extrapolators become bullish about future price changes and push the price even higher, a price increase that is excessive and that is later reversed. However, some phenomena – for example, momentum and post-earnings announcement drift – may instead be the result of investors adjusting their beliefs *too little* in response to new information (Bernard and Thomas, 1989; Jegadeesh and Titman, 1993).

How can sluggish belief adjustment generate post-earnings announcement drift and momentum? Consider a firm with a stock price of \$40 which subsequently announces earnings that are much better than expected. If investors are rational, they fully update their beliefs about the firm’s future prospects and immediately push the price of its stock up to the appropriate level – up to \$50, say. However, if investors instead have “sticky” beliefs, they do not update their beliefs sufficiently when the news is announced; as a

result, the stock price moves up only to a limited extent on the day of the announcement – up to \$45, say. Only over the next few weeks or months, as investors realize that their initial reaction was insufficient, does the price finally move up to the appropriate level of \$50. In short, when investors have sticky beliefs, the stock price jumps up on the announcement date and keeps rising in the weeks thereafter. But this is precisely the pattern known as post-earnings announcement drift. A similar intuition links sticky beliefs to momentum.

An early attempt to model sticky beliefs is Barberis et al. (1998), a paper I discuss in the next section. A more recent effort is Bouchaud et al. (2018), who embed a framework used by Coibon and Gorodnichenko (2012) to model inflation expectations into a model of asset prices. Consider a firm whose cash flow or profit at time t is π_t . The representative investor's expectation at time t about the next period's cash flow, denoted $F_t\pi_{t+1}$, is given by

$$F_t\pi_{t+1} = (1 - \lambda)E_t\pi_{t+1} + \lambda F_{t-1}\pi_{t+1}, \quad (17)$$

where $E_t\pi_{t+1}$ is the *rational* time t expectation about the future cash flow. When $\lambda = 0$, the investor has rational expectations, but when $\lambda > 0$, her beliefs are sticky, in that they depend in part on her expectations at time $t - 1$. Bouchaud et al. (2018) also assume that, at time t , the investor receives a signal s_t about the next period's cash flow, and that this signal is persistent over time:

$$\begin{aligned} \pi_{t+1} &= s_t + \varepsilon_{t+1} \\ s_{t+1} &= \rho s_t + u_{t+1}, \end{aligned} \quad (18)$$

where ε and u are noise terms. The signal s_t can be simply the time t cash flow π_t .

Bouchaud et al. (2018) show analytically that their model can explain the profitability anomaly described in Section 2.2 – the fact that, in the cross-section, a firm's profitability predicts the firm's subsequent stock return with a positive sign. To see the intuition, suppose that the time t signal s_t is the time t cash flow π_t . If the cash flow at time t is high, this indicates that cash flow at time $t + 1$ will also be high. However, since the investor's beliefs are sticky, she does not fully incorporate this information into her expectations at time t . As a result, she is positively surprised, on average, at time $t + 1$. Cash flow therefore predicts the subsequent stock return with a positive sign. The model also generates momentum and post-earnings announcement drift, and the intuition is similar.

The authors test their explanation of the profitability anomaly using analysts' forecasts of future earnings as a proxy for investors' expectations about those earnings. For each firm, they measure the stickiness of the consensus analyst forecast for the firm by regressing the forecast error, $\pi_{t+1} - F_t\pi_{t+1}$, on the prior forecast revision, $F_t\pi_{t+1} - F_{t-1}\pi_{t+1}$.

They find that, consistent with their model, the profitability anomaly is stronger for firms covered by analysts with stickier beliefs.

Bouchaud et al. (2018) also compute, for each analyst in their sample, a measure of the stickiness of the analyst's beliefs. They find that an analyst's beliefs are stickier, the more industries she covers. This suggests that belief stickiness stems in part from cognitive limits – from the fact that it takes time and effort to process news, especially when these news are about firms in different industries. I discuss cognitive limits in more detail in Section 9.

Psychological aspects of belief updating also offer a basis for sticky beliefs. One of these is known as “conservatism” (Edwards, 1968), and is illustrated by the following experiment. There are two bags: Bag 1 contains 700 blue chips and 300 red chips; Bag 2 contains 300 blue chips and 700 red chips. The experimenter draws 12 chips, with replacement, from one of the two bags. Of the 12 chips, 8 are blue and 4 are red. Participants in the experiment are then asked to estimate the probability that the chips were drawn from Bag 1. The correct answer is 0.97. However, most participants estimate the probability to be between 0.7 and 0.8. As such, their beliefs are reacting too little to the information they are given.²⁵

Slow updating of beliefs may also stem from the “anchoring heuristic” (Kahneman and Tversky, 1974). According to this heuristic, when people estimate the value of some quantity, they often start with an initial, possibly arbitrary, estimate, and then adjust away. In principle, this is a reasonable approach. The problem is that, in practice, the adjustment away is typically insufficient. In one experiment, participants are asked to estimate the percentage of United Nations' countries that are African. Before giving a percentage, they are asked whether their guess is higher or lower than a randomly-generated number between 0 and 100. Their subsequent estimates are affected by this random number. Those who are asked to compare their estimate to 10 subsequently estimate 25%, while those who compare to 60, estimate 45%.

Anchoring of this kind can lead to sticky beliefs in financial settings. When a firm announces surprisingly good earnings, an analyst or investor will often start with her prior view about the firm and then seek to adjust it in a positive direction. If, as the evidence on anchoring suggests, this adjustment is insufficient, the investor's beliefs will be sticky.

6.2 Models of under- and over-reaction

In Sections 4 and 5, we saw that some facts about asset prices are plausibly the result of investor *over*-reaction – an overreaction driven by extrapolative beliefs or by overconfidence. In Section 6.1, we noted that other facts about asset prices are suggestive

²⁵ Two other psychological phenomena that are related to conservatism and that may also cause stickiness in beliefs are “belief perseverance” and “confirmation bias” (Rabin and Schrag, 1999; Baron, 2000; Pouget et al., 2017).

of investor *under*-reaction. A long-standing challenge is to build a model that features both under-reaction and over-reaction and makes testable predictions about the circumstances in which each will occur.²⁶

An early model of under- and over-reaction is that of Barberis et al. (1998). I first describe the model and then discuss its psychological foundations. There is a risk-free asset and a risky asset which is a claim to a firm's future earnings, all of which are paid out as dividends. In reality, the firm's earnings follow a random walk, so that changes in earnings are uncorrelated over time. However, the representative investor thinks that, at any time, earnings are driven by one of two regimes: a "mean-reverting" regime in which earnings changes are negatively correlated over time, or a "trending regime" in which earnings changes are positively correlated over time. The investor also believes that, in each period, there is a small chance of a switch in the identity of the regime that is generating earnings.

The investor uses past earnings changes to infer which of the two regimes is currently driving earnings and then forecasts future earnings based on her incorrect understanding of how they evolve. For example, if the past five earnings changes have all been positive, she believes that, with high probability, the trending regime is currently generating earnings. But if the past five earnings changes have alternated in sign, she believes that, with high probability, the mean-reverting regime is currently driving earnings.

The model captures several facts about the cross-section of stock returns, including post-earnings announcement drift, medium-term momentum, long-term reversal, and the value premium. To see this, suppose that, at time t , there is a positive shock to earnings that is *not* part of a sequence of positive shocks, so that the investor thinks that the mean-reverting regime is more likely to be driving earnings. Since she believes that the positive shock at time t is likely to be followed by a negative shock at time $t + 1$, she does not push the stock price up very much at time t . But since the earnings change at time $t + 1$ is actually equally likely to be positive or negative, the investor is on average positively surprised at time $t + 1$ and the stock price rises again on that date. The good earnings news at time t is therefore followed by a positive return at time $t + 1$, as in post-earnings announcement drift. The same mechanism generates momentum in stock returns.

Now suppose that, at time t , there is a positive shock to earnings that is part of a sequence of positive shocks going back to time $t - 4$, say, so that the investor thinks that the trending regime is more likely to be driving earnings. Since she believes that the positive shock at time t is likely to be followed by another positive shock at time $t + 1$, she pushes the stock price up a lot at time t . However, since the earnings change at

²⁶ In a laboratory study of expectation formation, Landier et al. (2017) document both underreaction and overreaction in individual forecasts, but also find that overreaction, in the form of extrapolative expectations, predominates.

time $t + 1$ is actually equally likely to be positive or negative, the investor is on average disappointed at that time and the stock price falls. The good long-term return from time $t - 4$ to time t is therefore followed by a poor return at time $t + 1$, capturing the long-term reversal phenomenon. The same mechanism generates a value premium.

Barberis et al. (1998) motivate the model's assumptions using concepts from psychology that we have already encountered: conservatism (see Section 6.1) and representativeness (see Section 4.3). The mean-reverting regime captures the effect of conservatism – in this regime, the investor reacts “conservatively” to a given change in earnings because she believes that it is likely to be followed by an opposite change in earnings in the next period – while the trending regime captures the effect of representativeness, the idea that the investor over-estimates future earnings growth after high past growth in earnings because the data are representative of a firm whose true growth rate is high.

The model can also be motivated based on representativeness alone. After a sequence of earnings changes that alternate in sign, the investor believes that earnings changes are driven by a mean-reverting regime because the data are representative of such a regime. And after a sequence of earnings changes that all have the same sign, the investor believes that earnings changes are driven by a trending regime because the data are representative of *that* regime. The investor's mistake is to neglect the base rates – the fact that the true likelihoods of the mean-reverting and trending regimes are zero, and that actual earnings changes are driven by a third regime in which these changes are uncorrelated.

Rabin (2002) presents a model of under- and over-reaction based on an incorrect belief in a law of small numbers. Consider an investor who is trying to judge, based on a firm's realized earnings growth over time, whether the firm's true earnings growth rate is high, moderate, or low. Suppose that her initial belief is that the growth rate is moderate, and that she then observes a single realization of high earnings growth. She is now less confident that the true growth rate is moderate, but still views this as the most probable case. Since she believes in a law of small numbers, she predicts that next period's earnings growth will be low: since she expects the firm's earnings growth to look moderate even in a small sample, low earnings growth is required to “balance out” the high earnings growth in the current period. But precisely because she predicts low earnings growth in the next period, she underreacts to the high earnings growth in the current period.

Now suppose that the firm posts several periods of high earnings growth. Since the investor believes in a law of small numbers, she is now too confident that the firm's true growth rate is high. As a result, she over-estimates future earnings growth and thereby overreacts to the recent news. Similar to the Barberis et al. (1998) model, then, Rabin's (2002) framework predicts overreaction to a *sequence* of good or bad news, but underreaction to a single piece of good or bad news that is *not* part of a sequence of similar news.

Bordalo et al. (2018b) bring new evidence and theory to the question of whether people under- or over-react to information. They obtain professional analyst forecasts of 20 macroeconomic variables. By running a regression of forecast error on forecast revisions, they find that *individual* analysts tend to overreact to information, but that the *consensus* forecast underreacts.

To explain these results, the authors couple a rational-inattention framework of the kind commonly used in macroeconomics with Bordalo et al.'s (2018a) diagnostic expectations framework, described in Section 4.3. In this model, each analyst receives a noisy signal of the current value of the quantity she is forecasting. Since she processes this signal according to diagnostic expectations, she overreacts to it. However, since each analyst does not observe others' signals, the consensus forecast underreacts. Bordalo et al.'s (2018b) focus is on macroeconomic forecasts; it is an open question how well their framework explains empirical patterns in asset prices.

6.3 Beliefs about rare events

An individual's investment decisions are likely to be particularly affected by her beliefs about rare extreme outcomes – outcomes where she suffers a large financial loss or enjoys a big financial gain. Two of the most established judgment heuristics predict that, following good economic news, people under-estimate the likelihood of a very bad future outcome and that, following bad news, they over-estimate this likelihood. Such beliefs can explain a number of empirical facts, including excess volatility and time-series predictability in aggregate asset classes.

The availability heuristic is particularly relevant to beliefs about rare events. As discussed in Section 4.4, under this heuristic, people judge the likelihood of an event by how easy it is to recall instances of the event. This implies that, if there has been a financial crash in recent memory, people over-estimate the likelihood of another such crash. Conversely, after many years with no crash, people under-estimate the likelihood that one will occur (Goetzmann et al., 2017). Jin (2015) constructs a model of asset prices in which some investors' beliefs have this feature. Aside from excess volatility and time-series predictability, the model also captures an idea emphasized by Marks (2011) and Gennaioli et al. (2012), namely that the risk of a crash is greatest following a long quiet period with no significant downturns: at such a moment, the availability heuristic leads people to under-estimate the chance of a crash and hence to take large, levered positions in risky assets; this leaves financial markets vulnerable to bad fundamental news and makes a crash more, rather than less, likely.

Through a different mechanism, the availability heuristic predicts that an individual who invests in a positively-skewed asset will over-estimate her chance of a right-tail outcome; for example, when investing in an IPO stock, she will over-estimate her chance of an outsize return of the kind delivered by Microsoft and Google after their IPOs. The reason is that the media write about firms whose stocks post spectacularly good

returns after their IPOs, but not about the many companies with mediocre post-IPO returns. The individual therefore finds it easier to recall examples of stocks that performed well after their IPOs, which leads her to over-estimate her own chance of a good outcome from investing in an IPO. In the same way, when starting a company, an entrepreneur will over-estimate her chance of success: the media focus on the few entrepreneurial ventures that succeed wildly and ignore the many that fail. Entrepreneurs' beliefs indeed exhibit an optimistic bias of this kind (Astebro et al., 2014). Such beliefs help explain the low average return to entrepreneurship, and the low average returns on positively-skewed assets, an empirical fact we return to in Section 7.2.

The representativeness heuristic also predicts that people will under-estimate the likelihood of a major downturn following good economic news and over-estimate it following bad news. There are two reasons for this. First, good economic news are not representative of a model of the economy that features big crashes. Second, as discussed in Section 4.3, representativeness leads an individual who sees good economic data to over-estimate the mean of the distribution generating these data. If she does not change her estimate of the volatility of the distribution, she will under-estimate the likelihood of a left-tail outcome (Bordalo et al., 2018a).

6.4 Feelings and beliefs

A person's mood or emotional state can distort her beliefs. Specifically, there is clear evidence from both the laboratory and the field that an exogenous stimulus that improves (worsens) mood leads to more positive (negative) beliefs about unrelated events.

Johnson and Tversky (1983) conduct an experiment in which each participant reads one of a number of short newspaper articles. Some of these articles describe a happy story; others, a sad one. After reading their assigned articles, the participants are asked to judge the likelihood of various undesirable events – for example, the frequency of some specific causes of death. The participants who read sad articles make more negative judgments than a control group, while those who read happy articles make more positive judgments.

Several studies document analogous results in the field. Edmans et al. (2007) find that, if a national soccer team loses a World Cup match, the country's stock market performs poorly the following day. Their interpretation is that the disappointing sports news worsens the national mood, leading people to form more negative judgments about something unrelated, namely the economy's future prospects. Similarly, Hirshleifer and Shumway (2003) show that, if, on a particular day, it is sunnier than expected at the location of a national stock exchange, the country's stock market earns an above-average return on that day.

Kamstra et al. (2003) predict that the stock markets of countries in the northern hemisphere will trend downward as the winter solstice approaches: the fewer hours of daylight induce a type of depression known as Seasonal Affective Disorder in many

people, and depression leads to risk aversion. Using the same logic, they predict that stock markets in the southern hemisphere will trend downward as the *summer* solstice approaches. They find evidence consistent with these predictions.

While these results are specific in nature, they point to a more general mechanism that may amplify market fluctuations. If the stock market goes up, this improves the mood of many investors, leading them to form more positive judgments about the future prospects of the economy. This, in turn, causes them to push the stock market up still higher.

6.5 Herding and social interaction

In most of the models discussed so far, each investor observes past fundamentals or returns and uses them, independently of other investors, to form beliefs about the future. Since many investors are assumed to form beliefs in the same way, these beliefs have a significant impact on asset prices. In reality, however, an individual's beliefs are probably based not only on past economic outcomes but also on what she observes other people doing or saying, and, in particular, on her inferences about the beliefs that underlie other people's actions and statements.

To fix ideas, consider N individuals, $i = 1, \dots, N$, each of whom has to choose one of two options, A or B. They make these choices in sequence: individual 1 chooses first, then individual 2, and so on. Each person sees a private, equally strong, signal about which of the two options is better, and also observes the choices of all the people who went before her. Banerjee (1992) and Bikhchandani et al. (1992) draw a remarkable result out of this simple setting. Even if all the individuals are fully rational, and this rationality is common knowledge, the individuals may herd on the inferior choice: they may all choose option A even if their private signals in aggregate strongly indicate that B is the better choice.²⁷ If we take A and B to be two investment strategies, this suggests that, even in an economy where everyone is fully rational, people may pile into assets whose true prospects are weak. This, in turn, can cause these assets to become overpriced.

While this result is striking and important, subsequent research has shown that it holds only under fairly narrow conditions. Summarizing this work, Eyster and Rabin (2010) write: "...it is probably fair to say that the full-rationality model predicts a relatively limited form of herding and does so in a relatively limited set of domains... While tempting to use this literature to help understand dramatic instances of social pathology or mania... this is not something that the rational-herding literature can readily deliver."

²⁷ For example, suppose that individuals 1 and 2 receive signals indicating that A is the better choice, while individuals 3 through N receive signals indicating that B is better. In addition, assume that, when indifferent, an individual follows her private signal. In this case, everyone chooses A: for $i \geq 3$, individual i understands that the previous choices reveal two signals favoring A; even though her own signal points to B, she chooses A.

A more robust explanation for why people sometimes herd into suboptimal choices comes from a framework where some individuals are not fully rational. Eyster and Rabin (2010) return to the setting where N individuals choose in sequence between options A and B. As before, each individual sees a noisy private signal as well as the choices of the people who came before her. However, the individuals are not fully rational: when individual k makes her choice, she fails to recognize that individual j 's choice, where $j < k$, was based in part on individual j 's observation of the people who came before *her*. Instead, individual k thinks that individual j 's choice was based purely on her private signal.

Eyster and Rabin (2010) show that this error – a failure to understand what is driving other people's choices – can generate inefficient herding in a robust way. The intuition is clear: if I see ten people choose option A, and I believe that each of those people chose A because they had private information that A was better, then I will believe that there is strong evidence for A's superiority and will choose A myself, even if my own private signal suggests that B is better. My error is that I fail to recognize that some of the ten people may have chosen A not because they had private information in its favor but because they observed others choosing A. If we now think of A and B as investment opportunities, this framework can help us make sense of episodes where many investors, en masse, pursue a similar investment strategy, even in the absence of much objective information in favor of the strategy.

In financial markets, people make decisions not only by observing what others do, but also by talking to other people about their investments. Han et al. (2018) build a model of such investor-to-investor communication. There are N investors, each of whom follows one of two investment strategies. In each period, two randomly-chosen investors meet; one of them is randomly designated the "sender," and the other, the "receiver." With probability $s(R_i)$, the sender tells the receiver what the return R_i on her chosen investment strategy was in the previous period; $s(R_i)$ is a positive linear function of R_i ,

$$s(R_i) = \beta R_i + \gamma, \quad \text{where } \beta, \gamma > 0,$$

capturing the idea that, because she wants to make a good impression, an individual is more likely to talk about her investments, the better their performance. If the sender communicates this information, then, with probability $r(R_i)$, the receiver switches to the sender's investment strategy; $r(R_i)$ is an increasing and convex function of R_i ,

$$r(R_i) = aR_i^2 + bR_i + c, \quad \text{where } a, b, c > 0,$$

capturing extrapolative beliefs: if the receiver hears that a strategy performed well recently, she believes that it will continue to perform well, a belief that is all the stronger, the higher the strategy's recent return. Importantly, receivers do not understand the se-

lected nature of what they are hearing: they do not realize that senders are more likely to report a strategy's return when it is high.

Han et al. (2018) draw a number of predictions out of this framework. For example, they show that strategies that invest in volatile and positively-skewed assets spread more rapidly across the population. In equilibrium, such assets are overpriced. As such, the model offers a way of understanding the low average return of volatile and skewed stocks.²⁸

6.6 Psychology-free approaches

The approaches discussed in Sections 4 and 5, and thus far in Section 6, have mostly been rooted in a psychological concept – for example, in the representativeness heuristic, the availability heuristic, or overconfidence. Some behavioral finance models are able to shed light on the data without appealing to any specific psychology. This “psychology-free” approach was more common in the early years of behavioral finance research, when economists were wary of committing to any particular psychological concept. However, even today, it is deployed in useful ways.

A common structure for a behavioral finance model of asset prices, one laid out by De Long et al. (1990a) and Shleifer and Summers (1990), is one where less than fully rational investors, often called “noise traders,” are met in the market by more rational traders known as “arbitrageurs.” Even with no additional psychological assumptions, this bare-bones framework already yields testable predictions.

One prediction is that the level of mispricing will be higher for assets whose characteristics make arbitrage more difficult. A large amount of evidence is consistent with this. For example, most of the cross-sectional anomalies described in Section 2.2 are stronger in the subset of stocks where arbitrage is more limited – for example, among stocks with lower market capitalizations, lower liquidity, or higher idiosyncratic volatility.

To exploit an undervaluation, a rational investor typically buys the mispriced asset. To exploit an overvaluation, she typically shorts the asset in question. Going short is riskier than going long: an investor who shorts is exposed to very large potential losses, and the asset loan may be recalled before the mispricing corrects. This implies that more of the misvaluation in financial markets will involve *over-* rather than *under-*valuation. Consistent with this prediction, investment strategies designed to exploit the anomalies of Section 2.2 are typically more profitable on the short side than on the long side (Stambaugh et al., 2012). And while bubbles have for decades been a much-discussed phenomenon, there is far less talk of “negative bubbles,” in other words, of episodes of substantial *under-*valuation, suggesting that such episodes are less common.

²⁸ See Shiller (1984) for an early discussion of investor communication, and Hong et al. (2004, 2005) for evidence that social interaction affects investment decisions.

Stambaugh et al. (2015) combine two of the points noted above – that idiosyncratic volatility is a limit to arbitrage and that shorting is riskier than going long – to come up with an explanation for a puzzling fact from Section 2.2, namely that stocks with high idiosyncratic volatility have a lower average return. A stylized version of their argument goes as follows. Divide all stocks into two groups: those with low idiosyncratic volatility and those with high idiosyncratic volatility. Since it is easier to correct mispricing in less volatile stocks, assume, for simplicity, that there is no mispricing at all among these stocks. By contrast, since it is hard to correct mispricing in more volatile stocks, there is significant mispricing among these stocks. However, because it is riskier for an arbitrageur to short than to go long, this mispricing will primarily take the form of *over*-pricing rather than underpricing. As a result, high-volatility stocks will earn an average return that is lower than that of low-volatility stocks.

In an early psychology-free model of asset prices, De Long et al. (1990a) consider an economy with two assets, asset A and asset B, which are claims to the same cash-flow stream. Asset B differs from asset A in that it is traded in part by noise traders whose demand is subject to i.i.d. shocks. De Long et al. (1990a) show that, in equilibrium, asset B's price is lower, and its average return higher, than those of asset A. The reason is that noise traders' fluctuating demand for asset B causes the price of that asset to be more volatile. This higher volatility is an additional risk for investors in asset B, one they are compensated for through a lower price and a higher average return.

Lee et al. (1991) use the De Long et al. (1990a) model to address a classic puzzle in finance: the fact that, on average, the shares of a closed-end fund trade at a discount to "net asset value," the market value of the assets that the fund holds.²⁹ The idea is that the assets held by closed-end funds correspond to asset A in the De Long et al. (1990a) model, while the fund shares themselves correspond to asset B, the asset that is subject to noise trader demand. Closed-end funds are more likely to be traded by unsophisticated retail investors than are the underlying fund assets; as such, the prices of closed-end fund shares are affected by the fluctuating optimism and pessimism of these investors. This time-varying sentiment constitutes an additional risk for holders of fund shares. To compensate for this risk, closed-end funds trade at a discount to net asset value, just as, in the De Long et al. (1990a) framework, asset B trades at a discount to asset A.

Baker and Wurgler (2006, 2007) pursue a different psychology-free approach. Rather than using psychology to identify markers of misvaluation, they proceed empirically: they propose six measures of the extent to which stock market investors are displaying excessive exuberance, and combine these measures to create a "sentiment index." The measures, computed on an annual basis, are: the number and average first-day

²⁹ A closed-end fund is a fund that, at inception, raises money from investors and allocates this money to stocks or other assets. From this point on, the fund's shares are traded on an exchange – investors wanting to buy or sell fund shares do so there at the prevailing market price.

return of IPOs, stock market turnover, the closed-end fund discount, the equity share (the fraction of equity issues among all new issues), and the dividend premium (the relative valuation of dividend-paying firms as compared to non-dividend payers).

Baker and Wurgler (2007) show that, consistent with their index capturing excessive exuberance or pessimism, the value of the index predicts the subsequent stock market return with a negative sign. Baker and Wurgler (2006) further predict that, when their index takes a high value, stocks that are harder to arbitrage or harder to value will be more overpriced and earn a lower average return than stocks that are easier to arbitrage or to value: the former group of stocks is more affected by investor exuberance. They test this by identifying hard-to-arbitrage and hard-to-value stocks as stocks with low market capitalizations, stocks of young and unprofitable firms, more volatile stocks, and stocks that do not pay dividends. The data are consistent with the prediction: when the value of the index is high (low), these stocks earn a lower (higher) subsequent return, on average.³⁰³¹

7. PREFERENCES: GAIN-LOSS UTILITY AND PROSPECT THEORY

Sections 4 to 6 were about belief-based models of asset prices and trading volume. In Sections 7 and 8, I turn to preference-based models – models that try to make sense of the data using psychologically richer specifications of individual preferences, in other words, of how people evaluate the distribution of future outcomes that they perceive. Most models of asset prices assume that investors evaluate risk according to the Expected Utility framework, with a utility function that is increasing, concave, and defined over consumption outcomes. Such models offer compelling explanations for some aspects of the data – for example, the diversified portfolios held by many households – but struggle to explain others.

In Sections 7 and 8, we will see that, by drawing on ideas from psychology to refine our assumptions about individual preferences, we can make sense of a wider range of facts and also generate new predictions, some of which have already found empirical support. Much of the progress has come from incorporating ideas in Kahneman and Tversky's (1979, 1992) "prospect theory" into models of investor behavior. I discuss this research in Sections 7.1 through 7.5. In Section 7.6, I review other alternatives to Expected Utility. In Section 8, I turn to another important idea about preferences, namely, ambiguity aversion.

³⁰ Building on evidence that, on average, a person's mood improves as the week progresses from Monday to Friday, Birru (2018) predicts that speculative stocks – stocks that are hard to value or to arbitrage – will perform well (poorly) relative to non-speculative stocks on Fridays (Mondays). The data line up with this prediction.

³¹ Other psychology-free tests of behavioral finance models can be found in La Porta et al. (1997) and Engelberg et al. (2017).

7.1 The elements of prospect theory

Over the past few decades, researchers have accumulated a large amount of data on attitudes to risk in experimental settings. These data show that the Expected Utility framework is not an accurate description of individual decision-making under risk. This has spurred the development of “non-EU” models which try to capture the experimental evidence in parsimonious ways. The most influential of these non-EU theories, and the one that has been most widely applied in finance, is prospect theory.

There are two versions of prospect theory. The “original” prospect theory is laid out in Kahneman and Tversky (1979). While this version contains all of the theory’s essential insights, it also has some limitations: it can be applied only to gambles with at most two non-zero outcomes, and, in some situations, predicts that an individual will choose a dominated gamble. Tversky and Kahneman (1992) present a modified version of the theory, “cumulative prospect theory,” that addresses these weaknesses, and this is the version I present below.

Consider the gamble

$$(x_{-m}, p_{-m}; x_{-m+1}, p_{-m+1}; \dots; x_{-1}, p_{-1}; x_0, p_0; x_1, p_1; \dots; x_n, p_n), \quad (19)$$

to be read as “gain x_{-m} with probability p_{-m} , x_{-m+1} with probability p_{-m+1} , and so on,” where $x_i < x_j$ for $i < j$; where $x_0 = 0$, so that x_{-m} through x_{-1} are losses and x_1 through x_n are gains; and where $\sum_{i=-m}^n p_i = 1$. For example, a 50:50 bet to win \$110 or lose \$100 is written as

$$\left(-\$100, \frac{1}{2}; \$110, \frac{1}{2}\right). \quad (20)$$

In the Expected Utility framework, an individual evaluates the gamble in (19) as

$$\sum_{i=-m}^n p_i U(W + x_i), \quad (21)$$

where W is her initial wealth and $U(\cdot)$ is typically an increasing and concave function. By contrast, under cumulative prospect theory, the individual assigns the gamble the value

$$\sum_{i=-m}^n \pi_i v(x_i), \quad (22)$$

where

$$\pi_i = \begin{cases} w(p_i + \dots + p_n) - w(p_{i+1} + \dots + p_n) & \text{for } 0 \leq i \leq n \\ w(p_{-m} + \dots + p_i) - w(p_{-m} + \dots + p_{i-1}) & \text{for } -m \leq i < 0 \end{cases}, \quad (23)$$

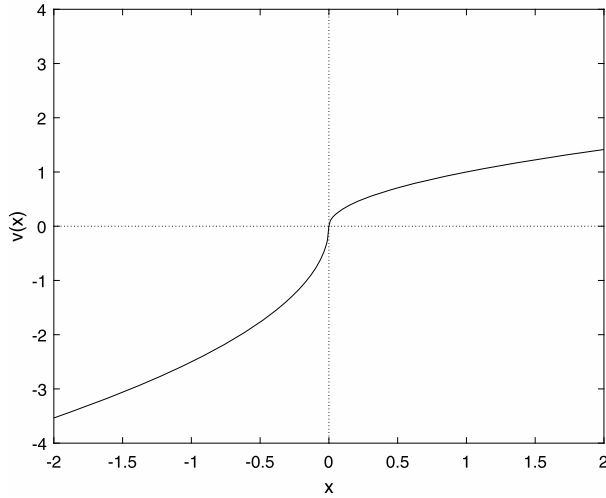


Figure 2 The prospect theory value function. The figure plots the value function proposed by Tversky and Kahneman (1992) as part of their cumulative prospect theory, namely, $v(x) = x^\alpha$ for $x \geq 0$ and $v(x) = -\lambda(-x)^\alpha$ for $x < 0$, for $\alpha = 0.5$ and $\lambda = 2.5$.

and where $v(\cdot)$ is known as the value function and $w(\cdot)$ as the probability weighting function. Tversky and Kahneman (1992) propose the functional forms

$$v(x) = \begin{cases} x^\alpha & \text{for } x \geq 0 \\ -\lambda(-x)^\alpha & \text{for } x < 0 \end{cases} \quad (24)$$

and

$$w(P) = \frac{P^\delta}{(P^\delta + (1 - P)^\delta)^{1/\delta}}, \quad (25)$$

where $\alpha, \delta \in (0, 1)$ and $\lambda > 1$. Fig. 2 plots the value function in (24) for $\alpha = 0.5$ and $\lambda = 2.5$. Fig. 3 plots the weighting function in (25) for $\delta = 0.4$ (the dashed line), for $\delta = 0.65$ (the solid line), and for $\delta = 1$ (the dotted line). Note that $v(0) = 0$, $w(0) = 0$, and $w(1) = 1$.

Prospect theory has four components: reference dependence, loss aversion, diminishing sensitivity, and probability weighting. I take each one in turn.

Reference dependence. In prospect theory, the individual derives utility not from final wealth levels but rather from *gains and losses* measured relative to some reference point: the argument of $v(\cdot)$ in (22) is x_i , not $W + x_i$. To motivate reference dependence, Kahneman and Tversky point to explicit experimental evidence – see Problems 11 and 12 in their 1979 paper – but also to the design of the human perceptual system: we are more sensitive to *changes* in brightness, loudness, or temperature than to the absolute levels of these attributes.

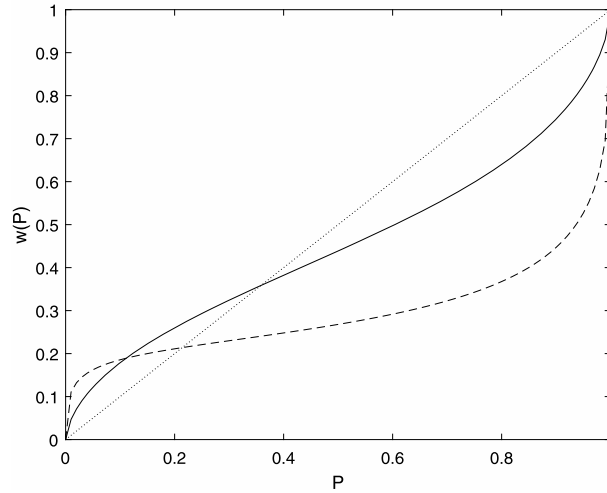


Figure 3 The prospect theory probability weighting function. The figure plots the probability weighting function proposed by Tversky and Kahneman (1992) as part of their cumulative prospect theory, namely, $w(P) = P^\delta / (P^\delta + (1 - P)^\delta)^{1/\delta}$, for three different values of δ . The dashed line corresponds to $\delta = 0.4$, the solid line to $\delta = 0.65$, and the dotted line to $\delta = 1$.

Loss aversion. Loss aversion is the idea that people are significantly more sensitive to losses – even small losses – than to gains of the same magnitude. It is captured by the steeper slope of the value function $\nu(\cdot)$ in the region of losses than in the region of gains, a feature that, loosely speaking, creates a “kink” at the point of zero gain. Kahneman and Tversky (1979) infer loss aversion from the widespread aversion to gambles like the one in (20), a 50:50 bet to win \$110 or lose \$100. Rabin (2000) shows that, in the Expected Utility framework, it is difficult, if not impossible, to explain why someone would turn this gamble down: an Expected Utility individual is almost risk-neutral over small-stakes gambles, and would therefore want to take a small-stakes gamble with positive expected value like the one in (20). Loss aversion explains the aversion to this gamble as follows: instead of focusing on final wealth outcomes, the individual pays attention to the potential gains and losses, \$110 and $-\$100$; since she is much more sensitive to the potential loss of \$100 than to the potential gain of \$110, she turns the gamble down. In (24), the degree of loss aversion is determined by the parameter λ . Tversky and Kahneman (1992) estimate $\lambda = 2.25$ for the median participant in their experiments.

Diminishing sensitivity. While the Expected Utility function $U(\cdot)$ is typically concave everywhere, the prospect theory value function $\nu(\cdot)$ is concave only over gains; in the region of losses, it is convex. This combination of concavity and convexity is known as diminishing sensitivity. Kahneman and Tversky (1979) infer the concavity over gains from the fact that people are typically risk averse over moderate-probability gains: they

usually prefer a \$500 gain to a 50% chance of \$1000. Similarly, they infer the convexity over losses from the fact that people are typically risk-seeking over moderate-probability losses: they prefer a 50% chance of losing \$1000 to a certain loss of \$500. In (24), the degree of diminishing sensitivity is controlled by the parameter α . Tversky and Kahneman (1992) estimate $\alpha = 0.88$ for their median participant, which implies a mild degree of curvature. More recent studies have estimated a stronger degree of diminishing sensitivity; the average estimate of α across these studies is approximately 0.7 (Booij et al., 2010).

Probability weighting. In prospect theory, the individual does not use objective probabilities, but rather, transformed probabilities obtained from objective probabilities using a weighting function $w(\cdot)$ such as the one in (25), plotted for specific values of δ in Fig. 3. In the original prospect theory, the weighting function is applied to the probability density function – to the probability of gaining exactly \$100, or of losing exactly \$200. Eq. (23) shows that, in cumulative prospect theory, the weighting function is instead applied to the cumulative distribution function: to the probability of gaining *at least* \$100, say, or of losing \$200 *or more*.

Fig. 3 shows that, for low values of P , $w(P) > P$. In the original prospect theory, this implies that the individual overweights low-probability outcomes. In cumulative prospect theory, it instead implies that she overweights the *tails* of any distribution she is considering: from (23), the tail outcomes x_{-m} and x_n receive weights of $w(p_{-m}) > p_{-m}$ and $w(p_n) > p_n$, respectively. One motivation for the overweighting of tails – or indeed, the overweighting of low-probability outcomes – is the fact that people typically like both lottery tickets *and* insurance: they prefer a 1 in 1000 chance of \$5000 to a certain \$5, but also prefer to pay \$5 than to face a 1 in 1000 chance of losing \$5000. This combination of behaviors is hard to explain under Expected Utility because it entails both risk aversion *and* risk seeking over a wide range of wealth levels. Prospect theory captures it through the overweighting of tail outcomes: if the individual sufficiently overweights the 1 in 1000 chance of winning the \$5000 jackpot, she finds the lottery attractive, and if she sufficiently overweights the 1 in 1000 chance of the \$5000 loss, she finds the insurance policy attractive too.

The simultaneous preference for lottery tickets and insurance motivates the shape of the weighting function $w(\cdot)$ in Fig. 3 in the range of low probabilities. How can we understand the shape of the function over the full range of probabilities? Gonzalez and Wu (1999) ask participants, for several values of p , to state the dollar amount such that they would be indifferent between receiving that dollar amount in cash and taking a gamble that offers \$100 with probability p . Since, under prospect theory, an individual evaluates the gamble as $w(p)v(100)$, people's responses allow us to sketch out the weighting function $w(\cdot)$. The resulting function looks like the one in Fig. 3. For example, when $p = 0.05$, the median cash equivalent across participants in Gonzalez and Wu's study is \$10, and when $p = 0.9$, it is \$63.

The transformed probabilities in (23) do not represent erroneous beliefs; rather, they are decision weights. In the framework of prospect theory, someone who is offered a 0.001 chance of winning \$5000 knows exactly what it means for something to have a 0.001 probability of occurring; however, when evaluating the gamble, she weights the \$5000 outcome by more than 0.001.

Eq. (25) shows that the degree of probability weighting is determined by the parameter δ ; a lower value of δ implies more overweighting of tails. Tversky and Kahneman (1992) estimate $\delta = 0.65$ for their median subject.³²

The implementation of probability weighting in cumulative prospect theory, laid out in (23), is known as a “rank-dependent” formulation. As noted earlier, this formulation gives cumulative prospect theory two advantages relative to the original prospect theory: it allows the theory to be applied to gambles with more than two nonzero outcomes, and it ensures that the decision-maker does not choose a dominated gamble. However, it also has at least one drawback: direct tests of rank dependence find little support for it (Bernheim and Sprenger, 2016). For financial economists, the advantages of cumulative prospect theory outweigh the disadvantages: finance research almost always uses probability distributions that have more than two nonzero outcomes, and, in a financial setting, the choice of dominated options can lead to riskless arbitrage opportunities. Not surprisingly, then, cumulative prospect theory is the version of prospect theory more commonly encountered in finance. However, both intuition and explicit modeling indicate that most of the applications of probability weighting in finance hold under both the cumulative *and* original versions of prospect theory – in other words, they do not rely on rank dependence.³³

Prospect theory is sometimes implemented in conjunction with “narrow framing.” In the traditional framework where the utility function is defined over total wealth or lifetime consumption, an individual evaluates any new risk by merging it with her pre-existing risks – her other financial risks, house price risk, or labor income risk – and checking if the combination is an improvement. However, experimental evidence suggests that, when presented with a new risk, people sometimes evaluate it in isolation, separately from other concurrent risks. This is known as narrow framing.

³² Prospect theory was developed in a setting where people make decisions “from description,” in other words, where the gambles they are facing are described to them – as (\$5000, 0.001), say. Hertwig et al. (2004) study decisions “from experience,” where people learn a gamble’s distribution by sampling from it. The overweighting of low-probability tail outcomes is present for decisions from description, but not for decisions from experience. The root cause of this is not fully understood, and its implications for finance are unclear. However, at the very least, it indicates that a person’s mental representation of a gamble affects her attitude toward it.

³³ The intuition is this. As we will see later in Section 7, most of the applications of probability weighting in finance depend on investors having a strong preference for a positively-skewed return distribution. Such a preference arises both under the original prospect theory, where people overweight low-probability outcomes, *and* under cumulative prospect theory, where they overweight the tails of distributions.

Tversky and Kahneman (1981) present examples of narrow framing in the laboratory. More recently, Barberis et al. (2006) argue that even as simple a phenomenon as the widespread rejection of the gamble

$$\tilde{G} = \left(-\$100, \frac{1}{2}; \$110, \frac{1}{2} \right)$$

is evidence not only of loss aversion, but of narrow framing as well. An individual who is offered the gamble \tilde{G} almost certainly has some pre-existing risks. If she is loss averse but does *not* engage in narrow framing – in other words, if she merges \tilde{G} with her pre-existing risks and is loss averse only over the resulting gains and losses in her *overall* wealth – then, as Barberis et al. (2006) show, she is very likely to accept \tilde{G} : her pre-existing risk moves her away from the kink in the value function $v(\cdot)$, and when she is away from the kink, she is no longer averse to \tilde{G} .³⁴ A plausible explanation for why people do typically turn down \tilde{G} is that they engage in narrow framing: an individual who is loss averse over the gains and losses in \tilde{G} itself will find the gamble aversive.

In the context of financial markets, a prospect theory individual who exhibits narrow framing will decide whether and how much to invest in an asset by applying prospect theory to her potential gains and losses in the asset itself rather than to the gains and losses in overall wealth that would result from taking a position in the asset. We do not have a full understanding of why people engage in narrow framing. One possibility is that it is a heuristic used to simplify complex decisions. It can be difficult for an individual to compute the distribution of gains and losses in overall wealth that would result from combining a new risk with her pre-existing risks. She therefore adopts a heuristic whereby she takes on the new risk if its *own* distribution of gains and losses is appealing, something that is easier to determine.

Prospect theory contains insights that seem very relevant to financial decision-making. However, any attempt to incorporate these insights into a more traditional finance framework faces a fundamental difficulty. In prospect theory, the individual derives utility from “gains” and “losses.” This is certainly consistent with the way many people talk about their investment decisions: their willingness to buy an asset appears to depend heavily on how much money they think they can make or lose by doing so. The problem is that it is not clear exactly which gains and losses people have in mind. Is it gains and losses in consumption, in financial wealth, or in the value of specific components of wealth? Is it annual gains and losses or gains and losses measured at some

³⁴ An example may be helpful. Consider an individual who is loss averse with a piecewise-linear value function $v(\cdot)$ that is twice as steep in the region of losses as in the region of gains, who is facing the pre-existing risk $(\$30000, \frac{1}{2}; -\$10000, \frac{1}{2})$, and who is offered gamble \tilde{G} . This individual will *accept* \tilde{G} because the utility of the combined gamble $(30100, \frac{1}{4}; 29900, \frac{1}{4}; -9890, \frac{1}{4}; -10100, \frac{1}{4})$, namely 5007, is higher than the utility of the pre-existing risk, namely 5000.

other frequency? Does a gain mean a return that exceeds zero, or one that exceeds the risk-free rate or the return the investor expected to earn?

There are no consensus answers to these questions. Researchers are exploring several plausible definitions of “gain” and “loss” and trying to determine which of them explains the widest range of facts and makes predictions that find support. One promising specification takes the gains and losses to be annual dollar changes in financial wealth. In this specification, then, the “reference point” that the individual compares her current wealth to in order to determine her most recent gain or loss is her wealth one year in the past.

In Sections 7.2 and 7.3, I discuss applications of prospect theory to the cross-section of average returns; Section 7.2 focuses on static models and Section 7.3 on dynamic models. In Section 7.4, I discuss applications of prospect theory to aggregate asset classes.

7.2 Prospect theory and the cross-section: Static models

Why do some financial assets have higher average returns than others? In a model with rational investors and no frictions, differences in average returns across assets are due to differences in risk. In the benchmark rational model of the cross-section of returns, the CAPM, the risk of an asset is measured by its beta; in this model, then, an asset’s average return is determined by its beta and by its beta alone. This prediction has not fared well: as noted in Section 2.2, several firm-level characteristics predict stock returns in the cross-section, even after controlling for beta. In light of the poor performance of the CAPM, it is natural to ask: Can we improve our understanding of the cross-section of returns using a model that makes more realistic assumptions about the way people evaluate risk – specifically, one in which investors evaluate risk according to prospect theory, rather than Expected Utility?

Barberis and Huang (2008) incorporate prospect theory into an otherwise traditional model of the cross-section. There are two dates, date 0 and date 1, and $J + 1$ assets: a risk-free asset with a fixed return and J risky assets whose payoffs have a multivariate Normal distribution. If investors had Expected Utility preferences, assets would be priced in the way described by the CAPM. Barberis and Huang (2008) assume instead that investors evaluate risk according to cumulative prospect theory: for each possible portfolio strategy, they use Eqs. (22)–(25) to compute the prospect theory value of the distribution of gains and losses that would result from that portfolio strategy, and then choose the strategy with the highest prospect theory value. Here, an investor’s gain or loss is her wealth at time 1 minus her wealth at time 0 scaled up by the gross risk-free rate. Investors therefore do not engage in narrow framing: they evaluate gains and losses in overall wealth, not gains and losses in the value of individual stocks. The scaling by the risk-free rate makes the model more tractable, but may also be psychologically accurate: investors may only think of their change in wealth at time 1 as a true “gain” if it is greater than what they would have earned by allocating all of their time 0 wealth to

the risk-free asset. Finally, investors have identical preferences – the values of α , δ , and λ in (24) and (25) are the same for all of them.

It turns out that this initial prospect theory model does not deliver any novel predictions: as in the case where investors have Expected Utility preferences, the CAPM holds and average returns are determined by beta and beta alone.³⁵

To generate a more useful prediction, Barberis and Huang (2008) move away from the multivariate Normal payoff structure. They introduce one additional asset into the economy: a positively-skewed asset in small supply whose return is independent of all other asset returns. If investors had Expected Utility preferences, the expected return on this new asset would slightly exceed the risk-free rate. Barberis and Huang (2008) find that, when investors have prospect theory preferences, a different result obtains: the positively-skewed asset earns an average return substantially below the risk-free rate. As such, the prospect theory framework generates a novel prediction: that an asset's *own* skewness – even its idiosyncratic skewness – is priced. This contrasts with the Expected Utility framework, where, in the presence of skewed assets, a different quantity, an asset's coskewness with the market portfolio, is priced.

The equilibrium in the economy with prospect theory investors involves heterogeneous holdings. It is easier to think about this in the case where there is a short-sale constraint, although the main results do not depend on such a constraint. In equilibrium, some investors hold the J multivariate Normal assets and no position at all in the positively-skewed asset, while other investors hold both the J multivariate Normal assets and a large, undiversified position in the skewed asset. These heterogeneous holdings are not the result of heterogeneous preferences – all investors have the same preferences – but rather reflect non-unique global optima. On the one hand, holding a large position in the skewed asset decreases an investor's utility: the undiversified position increases the volatility of her portfolio, which, due to loss aversion, is unappealing. On the other hand, holding a large position in the skewed asset increases the investor's utility by making the distribution of her future gains and losses in wealth more positively skewed: since, under probability weighting, she overweights tail outcomes, such a distribution is very attractive. In equilibrium, the expected return of the skewed asset adjusts so that these opposing forces cancel out, leaving the holders of the skewed asset with the same utility as the non-holders and allowing the market in the asset to clear.

In summary, the equilibrium expected return of the skewed asset in Barberis and Huang's (2008) model is low because, by holding a sufficient amount of the skewed asset, some investors are able to add skewness to the distribution of their gains and losses in wealth. Due to probability weighting, this is very appealing. The investors are therefore willing to pay a high price for the skewed asset and to accept a low average return on it.

³⁵ See Levy et al. (2011) for a detailed discussion of this result.

Researchers have used this framework to make sense of a number of empirical facts. In Section 2.2, we noted that the long-term average return of IPO stocks is lower than that of a control group of stocks. Why is this? One interesting feature of IPO stock returns is that they are highly positively skewed: most IPO stocks deliver unimpressive returns in the years after the offering, but a small handful perform extraordinarily well. According to the prospect theory framework, then, IPO stocks *should* earn a low average return. Green and Hwang (2012) present evidence consistent with this theory of average IPO returns.

Another fact noted in Section 2.2 is that stocks with high past idiosyncratic return volatility have a lower average return than stocks with low past idiosyncratic return volatility. One explanation of this builds on the fact that stocks with high past idiosyncratic volatility have higher idiosyncratic skewness in their subsequent returns; under the prospect theory framework, this explains why these stocks have a lower average return. Boyer et al. (2010) present evidence in support of this view.

In the same way, prospect theory can help us make sense of the low average return on distressed stocks, stocks traded off the main exchanges, and out-of-the-money options on individual stocks: all of these assets have positively-skewed returns. It can also explain differences in valuations between conglomerate firms and matched portfolios of single-segment firms: empirically, these differences line up with differences in the skewness of conglomerates and single-segment firms.³⁶

Some papers directly test the prediction that stocks with more positively-skewed returns have a lower average return. Boyer et al. (2010) use a regression model to predict a stock's future return skewness, while Conrad et al. (2013) infer a stock's expected skewness from the prices of the options that are traded on the stock. Both papers find support for the prediction.³⁷

The more robust prediction of the Barberis and Huang (2008) model is that positively-skewed assets have a low average return. If investors can sell short, the model further predicts that negatively-skewed assets earn a high average return. This prediction can be used to make sense of the high average return on merger arbitrage strategies (Wang, 2016), high-quality corporate bonds, and so-called catastrophe bonds, all of which have negatively-skewed returns.

The prospect theory model can also be used to think about the concentrated positions that some households have in specific assets: in the model, some investors take

³⁶ See Mitton and Vorkink (2010), Boyer and Vorkink (2014), Conrad et al. (2014), and Eraker and Ready (2015) on conglomerate firms, stock options, distressed stocks, and stocks traded off-exchange, respectively.

³⁷ In related work, Kumar (2009) shows that stocks with low prices, high past idiosyncratic return volatility, and high past idiosyncratic return skewness have a low average return, while Bali et al. (2011) show that stocks with a high maximum daily return over the past month have a low average return. One interpretation of these findings is that investors perceive stocks with these characteristics as having positively-skewed returns.

an undiversified position in a positively-skewed asset so as to give themselves a chance of a big gain in wealth, something that, under probability weighting, is very appealing. Mitton and Vorkink (2007) provide evidence for this view using data on the stock holdings of individual investors: they find that the stocks held by investors with less diversified portfolios are more positively skewed than the typical stock.

The Barberis and Huang (2008) model assumes broad framing, as do traditional models of asset prices. However, *narrow* framing may also play a role in the pricing of stocks: investors' demand for a stock may depend, to some extent, on their evaluation of the stock's *own* return distribution. A model that incorporates both prospect theory and narrow framing delivers a prediction similar to that of the Barberis and Huang (2008) model, namely that assets with more positively-skewed returns have a lower average return: a stock with positively-skewed returns is attractive to a prospect theory investor who frames narrowly because, as a consequence of probability weighting, the investor overweights the right tail of the stock's return distribution. As such, a model with narrow framing can likely also address the empirical patterns discussed above. Barberis and Huang (2001), Grinblatt and Han (2005), Li and Yang (2013), and Barberis et al. (2016) investigate the implications of narrow framing for asset prices; I come back to some of these papers in the next section.

The prospect theory framework is not the only one in which an asset's idiosyncratic skewness is priced. A similar prediction can be derived from a model where investors manipulate their beliefs (Brunnermeier and Parker, 2005); from the Expected Utility framework, if the utility function has a convex segment (Mitton and Vorkink, 2007); and from another non-EU model, salience theory, which I come back to in Section 7.6.

The predictions discussed in this section are driven by preferences, not beliefs: the investors in Barberis and Huang's (2008) model have correct beliefs about returns – for example, they correctly perceive the return distribution of the typical IPO stock – but, because they overweight tail outcomes in their decision-making, they pay a premium for positively-skewed assets. An alternative view is that these assets are overpriced because investors over-estimate the *likelihood* of a right-tail outcome. In Section 6.3, I noted that there is a psychological basis for such beliefs. An open question asks how much of the average return on skewed assets is driven by investor preferences, as opposed to beliefs.

7.3 Prospect theory and the cross-section: Dynamic models

In the applications of prospect theory discussed in Section 7.2, it was probability weighting that played the most important role. Loss aversion also entered the picture, but there was no mention at all of diminishing sensitivity. Diminishing sensitivity does have inter-

esting implications for asset prices, but to see them, we have to move beyond one-period models to dynamic models.³⁸

The applications of diminishing sensitivity to asset prices build on a link between prospect theory and an aspect of trading known as the disposition effect. This is the finding that individual investors, as well as some institutional investors, have a greater propensity to sell stocks trading at a gain relative to purchase price than stocks trading at a loss (Shefrin and Statman, 1985; Odean, 1998b). It is difficult to explain this as the result of fully rational thinking. For example, one rational hypothesis is that individuals sell prior winner stocks because they have information that these stocks will subsequently perform poorly, and hold on to prior loser stocks because they have information that these stocks will subsequently perform well. The data reject this view, however: Odean (1998b) finds that the prior winners that individuals sell subsequently perform better than the prior losers they hold on to.

A long-standing idea, first articulated by Shefrin and Statman (1985), is that prospect theory, and specifically its diminishing sensitivity component, can explain the disposition effect. An informal summary of the argument goes as follows. If an investor buys a stock which then goes up, this brings her into the “gain” region of the value function in Fig. 2; if the stock instead goes down, this brings her into the “loss” region. Since the value function is concave over gains but convex over losses, it appears that the investor is more risk averse after a gain than after a loss, and hence more likely to sell after a gain. Notice that this argument assumes stock-level narrow framing, as do all the applications of diminishing sensitivity I discuss below.³⁹

Barberis and Xiong (2009) try to formalize the argument in the previous paragraph and find that this is not as easy as one might think. In their model, an investor trades a risk-free asset and a single risky asset over $T + 1$ dates, $t = 0, 1, 2, \dots, T$, and derives prospect theory utility at time T from the difference between her time T wealth, W_T , and her initial wealth scaled up by the gross risk-free rate, $W_0 R_f^T$. They find that, while the diminishing sensitivity component of prospect theory does make it more likely that the investor will exhibit a disposition effect, the loss aversion component pushes in the opposite direction: it makes it more likely that she will exhibit the reverse of the disposition effect, in other words, display a greater propensity to realize *losses* rather than

³⁸ The reason for this is that diminishing sensitivity becomes important when an investor evaluates a distribution that consists primarily of gains, or primarily of losses. This is not the case in a one-period model that involves only one decision, the initial purchase decision: when a risky asset is first purchased, it exposes the investor to both potential gains *and* potential losses.

³⁹ A very common error, both in popular accounts of behavioral finance but also in academic studies, is to say that loss aversion is at the root of Shefrin and Statman’s (1985) explanation of the disposition effect. Their argument is based on diminishing sensitivity, not loss aversion: it relies on the prospect theory value function’s concavity over gains and convexity over losses, not on the kink in the value function at the point of zero gain. The statement “people don’t realize losses because of loss aversion” may *sound* plausible, but it is completely incorrect as a summary of Shefrin and Statman’s hypothesis.

gains. The model predicts a disposition effect only if the diminishing sensitivity effect overcomes the loss aversion effect. In a two-period model ($T = 2$), and for Tversky and Kahneman's (1992) estimates of the parameters of the value function in (24), namely $\alpha = 0.88$ and $\lambda = 2.25$, this is not the case: $\alpha = 0.88$ represents a mild degree of diminishing sensitivity while $\lambda = 2.25$ represents a strong degree of loss aversion, and so the loss aversion effect dominates. The investor therefore exhibits the reverse of the disposition effect.

Why, in Barberis and Xiong's model, does loss aversion lead to the reverse of the disposition effect? Consider an investor whose value function exhibits *only* loss aversion and no diminishing sensitivity, so that it is piecewise-linear with $\alpha = 1$ and $\lambda > 1$. For simplicity, suppose also that, in each period, the return on the risky asset, a stock, can take one of two equiprobable values. Because the investor is loss averse, she is willing to buy the stock at time 0 only if it has a high expected return – if it offers a 50% chance of a gain of \$10, say, and a 50% chance of a loss of \$5. For an investor with a piecewise-linear value function, how risk averse she is after a gain or loss depends on how far she is from the kink at the origin. Since, after the gain, she is twice as far from the kink as she is after a loss, she is actually more risk averse after the loss. This, in turn, means that it is after a loss that she is more likely to sell shares of the stock.

Subsequent to Barberis and Xiong's (2009) findings, researchers have identified a number of ways in which prospect theory *can* generate a disposition effect. First, it may be that the degree of diminishing sensitivity for the average person is stronger than that estimated by Tversky and Kahneman (1992). Indeed, more recent experimental work has typically generated estimates of α lower than 0.88 (Booij et al., 2010). Real-world investors may also be less loss-averse than Tversky and Kahneman's (1992) laboratory participants: very loss-averse people are unlikely to open a brokerage account in the first place. Second, in Barberis and Xiong (2009), the reference point that the investor uses to define her “gain” or “loss” is her (scaled) initial wealth. If the reference point instead adjusts gradually over time, or if it is based on *expected* wealth rather than initial wealth, prospect theory leads to a disposition effect more readily (Meng and Weng, 2017; Andrikogiannopoulou and Papakonstantinou, 2018). Finally, it may be that, rather than deriving utility from her trading profit over some fixed interval of time, the investor instead derives utility directly from *realized* gains and losses: if she buys shares of a stock at \$40 and then, a few months later, sells some of the shares, she derives utility from the realized gain or loss at the moment of sale (Barberis and Xiong, 2012; Frydman et al., 2014). Ingersoll and Jin (2013) show that, when coupled with prospect theory, such “realization utility” can generate a robust disposition effect.

Overall, it is reasonable to say that there *is* a link between prospect theory and the disposition effect, although this link needs to be constructed carefully. The applications of diminishing sensitivity to asset prices build on this link.

Grinblatt and Han (2005) observe that, if many investors display a disposition effect, there will be momentum in stock returns. Suppose that a stock goes up a lot in value.

Since many investors exhibit a disposition effect, the stock is subject to significant selling pressure as it rises; this slows the price rise and causes the stock to become undervalued. Subsequently, the stock earns a high average return as the undervaluation corrects. A high past return therefore predicts a high future return, consistent with the empirical finding of momentum.

If many investors exhibit a disposition effect, then, this generates momentum in stock returns. But we also noted that, under some conditions, prospect theory can generate a disposition effect. Putting these two insights together, it follows that, if many investors make decisions according to prospect theory, this can explain why we see momentum in stock returns. The full logic is: if a stock goes up in price, this brings its holders into the gain region of the value function, while if the stock goes down, this brings its holders into the loss region of the value function. Since the value function is concave over gains and convex over losses, the prospect theory investors sell the stock more heavily after a gain than after a loss. The stock is therefore more undervalued after a prior gain. As a consequence, its subsequent return is higher, on average, after a prior gain.

Grinblatt and Han (2005) do not explicitly derive momentum from prospect theory. Instead, they derive it from the disposition effect and then point out that prospect theory is a possible source of the disposition effect. Li and Yang (2013) formalize the link between prospect theory and momentum: they show that, in an economy where some investors make decisions according to prospect theory, asset returns can indeed exhibit momentum. Specifically, there is momentum in asset returns so long as investors' degree of diminishing sensitivity is sufficiently strong. This is broadly consistent with the partial equilibrium results of Barberis and Xiong (2009): only when the degree of diminishing sensitivity is sufficiently strong will the prospect theory investors exhibit the disposition effect that is needed to generate momentum.

Wang et al. (2017) propose a different application of diminishing sensitivity to asset prices. Using a numerical example, they argue that, if we take the subsample of stocks that are trading at a gain relative to purchase price for their average investor, we should see a positive relationship between volatility and average return in this subsample: if a stock is trading at a gain, this brings its investors into the concave region of the value function; since the investors are risk averse in this region, they require a higher average return on more volatile stocks. Conversely, among stocks trading at a loss relative to purchase price for their average investor, there should be a negative relationship between volatility and average return: if a stock is trading at a loss, this brings its holders into the convex region of the value function; since the investors are risk-seeking in this region, they require a lower average return on more volatile stocks. Wang et al. (2017) test these predictions using two measures of whether a stock is trading at a gain or at a loss – one developed by Grinblatt and Han (2005) and another developed by Frazzini (2006). The data line up in the predicted way: under both measures of gain or loss, the relationship between volatility and average return is positive for stocks trading at a gain and negative for stocks trading at a loss.

7.4 Prospect theory and the aggregate stock market

The models discussed in Sections 7.2 and 7.3 suggest that prospect theory is useful for thinking about the cross-section of average returns. But can it help us understand the behavior of the *aggregate* stock market?

In a famous application of prospect theory in finance, Benartzi and Thaler (1995) argue that this theory can provide an explanation of the equity premium puzzle described in Section 2.1. They propose that, when considering an investment in the stock market, people bring to mind the historical distribution of annual stock market returns; and similarly that, when considering an investment in the bond market, people bring to mind the historical distribution of annual bond market returns. Benartzi and Thaler (1995) focus on the annual frequency on the grounds that people track returns on an annual basis, perhaps because they do their taxes once a year or because they receive their most comprehensive brokerage statements at the end of the year.

Benartzi and Thaler's main result is that the prospect theory value of the historical distribution of annual U.S. stock market returns, computed by applying Eqs. (22)–(25) for the preference parameters estimated by Tversky and Kahneman (1992), is approximately equal to the prospect theory value of the historical distribution of annual U.S. bond market returns. On the one hand, the greater volatility of stock market returns serves to lower their prospect theory value relative to that of bond market returns: a loss-averse individual strongly dislikes volatility. On the other hand, the much higher average return of the stock market – the high historical equity premium – serves to increase the prospect theory value of the stock market relative to that of the bond market. Benartzi and Thaler's result shows that these two forces cancel out, suggesting that the stock market earns a high average return so that it can be competitive with the bond market in the eyes of prospect theory investors: if the equity premium were any lower, these investors would strictly prefer the bond market and would be unwilling to hold the available supply of stocks.

Benartzi and Thaler (1995) do not incorporate the insights of prospect theory into an equilibrium model, one where asset prices are determined endogenously. Barberis et al. (2001) take up this task. They consider an economy with a riskless asset and a risky asset in which an infinitely-lived representative investor has the preferences⁴⁰

$$E \sum_{t=0}^{\infty} \left[\rho^t \frac{C_t^{1-\gamma}}{1-\gamma} + b_0 \bar{C}_t^{-\gamma} \rho^{t+1} \nu(X_{t+1}) \right], \quad (26)$$

where

⁴⁰ Eq. (26) is a special case of the preferences considered by Barberis et al. (2001). I describe the more general specification they use later in this section.

$$X_{t+1} = S_t(R_{t+1} - R_{f,t}) \quad (27)$$

$$v(X) = \begin{cases} X & \text{for } X \geq 0 \\ \lambda X & \text{for } X < 0 \end{cases} \quad (28)$$

The first term in (26) is the traditional specification of investor preferences, one based on lifetime consumption utility. The second term is new, and is inspired by the reference dependence element of prospect theory, whereby people derive utility from gains and losses; specifically, this term captures the idea that the investor also derives utility from annual gains and losses in her financial wealth. The gain or loss between time t and $t + 1$, denoted X_{t+1} , is defined as the investor's financial wealth at time $t + 1$ minus her financial wealth at time t scaled up by the gross risk-free rate. To see how this leads to (27), suppose that the value of the investor's holdings of the risk-free asset and the risky asset at time t are B_t and S_t , respectively, and that the gross returns of the two assets between t and $t + 1$ are $R_{f,t}$ and R_{t+1} . The investor's gain or loss in financial wealth between t and $t + 1$ is then

$$\begin{aligned} X_{t+1} &= (B_t R_{f,t} + S_t R_{t+1}) - (B_t + S_t) R_{f,t} \\ &= S_t (R_{t+1} - R_{f,t}). \end{aligned} \quad (29)$$

The scaling by the risk-free rate makes the model more tractable but may also be psychologically accurate.

The specification in (26) captures only the reference dependence and loss aversion components of prospect theory: the value function $v(\cdot)$ in (28) is piecewise-linear; for tractability, again, diminishing sensitivity and probability weighting are ignored. The parameter ρ is the time discount rate and $\overline{C}_t^{-\gamma}$ is a scaling term based on aggregate consumption \overline{C}_t which ensures that the two components of preferences remain similarly important even as wealth in the economy grows. Finally, b_0 is the weight the investor puts on the gain-loss utility term.

Barberis et al. (2001) embed the specification in (26) in an endowment economy. They find that, for values of loss aversion λ drawn from experimental studies, and for values of b_0 that put substantial weight on the gain-loss utility term, the equilibrium equity premium is large – as high as 6% per year for high values of b_0 . The investor recognizes that her stock market holdings will lead to large annual fluctuations in financial wealth, which, due to loss aversion, will be unpleasant. As compensation, she requires a high average return on the stock market.⁴¹

⁴¹ Gneezy and Potters (1997), Thaler et al. (1997), and Beshears et al. (2017) use laboratory and field experiments to explore the link between loss aversion and investment decisions. Easley and Yang (2015) and Guo and He (2017a) build models where investors with traditional preferences interact with prospect theory investors.

De Giorgi and Legg (2012) incorporate both loss aversion *and* probability weighting into an analysis of aggregate stock market prices.⁴² The introduction of probability weighting leads to an equity premium that is even higher than that generated by loss aversion alone. The reason is that the aggregate stock market is negatively skewed: it is subject to occasional large crashes. Stock market investors are therefore exposed to a small chance of a large drop in financial wealth. Under probability weighting, they overweight this left-tail outcome, which makes the stock market all the less appealing. As a result, they require a high equity premium in order to hold the market supply of stocks.

Barberis et al. (2001) show how a modification of the preferences in (26) can address not only the equity premium puzzle, but also the excess volatility and predictability of stock market returns. Motivated by experimental evidence of Thaler and Johnson (1990), they propose that an investor's degree of loss aversion λ is not constant, but rather varies over time based on the past gains and losses that the investor has experienced: if she has experienced gains in recent years, she is less loss averse going forward, while if she has experienced losses, she is more loss averse. Barberis et al. (2001) capture this by redefining the value function $v(\cdot)$ in (28) as

$$v(X, z_t) = \begin{cases} X & \text{for } X \geq 0 \\ \lambda(z_t)X & \text{for } X < 0 \end{cases}, \quad (30)$$

where

$$\lambda(z_t) = \lambda + k(z_t - 1), \quad k > 0, \quad (31)$$

and where z_t is the time t measure of the investor's past gains and losses, with higher values of z_t indicating larger prior losses.⁴³

One way to think about the specification in (30)–(31) is that it captures the investor's capacity for dealing with bad news. Following substantial gains in financial wealth, the investor may feel psychologically stronger – more able to deal with any bad news that comes along in the near future. In effect, she is less loss averse. Conversely, following substantial losses in financial wealth, she may feel psychologically shaken up, and hence less able to handle any additional bad news that comes along. In effect, she is more loss averse.

Such shifts in loss aversion offer a way of thinking about the volatility and predictability of stock market returns. After good cash-flow news that pushes the value of the stock market up, the investor becomes less loss averse; she is therefore willing to pay

⁴² De Giorgi and Legg (2012) do not incorporate probability weighting into the specification in (26), but rather into a different gain-loss utility specification proposed by Barberis and Huang (2009) and recently improved by Guo and He (2017b).

⁴³ Eqs. (30) and (31) are a slightly simplified version of Barberis et al.'s (2001) specification.

an even higher price for the stock market, thereby generating excess volatility, and to accept a lower average return on the stock market going forward, thereby generating time-series predictability. After bad cash-flow news that depresses the value of the stock market, the investor becomes more loss averse; she pushes the stock market even further down and requires a higher average return on it going forward.⁴⁴

To motivate the convex segment of the prospect theory value function, Kahneman and Tversky (1979) present evidence of *risk-seeking* in the domain of losses: a 50% chance of a \$1000 loss is typically preferred to a certain loss of \$500. Barberis et al. (2001), guided in part by evidence of Thaler and Johnson (1990), assume *risk-aversion* in the domain of losses: after experiencing a loss, the individual becomes more averse to potential future losses. These ideas are not contradictory. Kahneman and Tversky (1979) show that someone who is facing the *possibility* of a loss will take risk in an effort to avoid the loss. Barberis et al. (2001) propose that someone who has *already* taken a loss – someone who has admitted to herself and to others that she has taken a loss and has experienced the pain of this admission – will subsequently be more risk averse.

Imas (2016) articulates this view and tests it experimentally. Participants in the experiment are offered a sequence of gambles. In one condition, the participants settle their accounts in cash after each gamble is played out; in another condition, they settle their accounts at the end, after all the gambles have been played out. The idea is that a participant who has to settle her account along the way is forced to “accept” any intermediate loss she experiences. Consistent with the view outlined in the previous paragraph, Imas (2016) finds that a participant who experiences an intermediate loss in the round-by-round settlement condition subsequently takes less risk, as compared to a participant in the settle-at-the-end condition who also experiences an intermediate loss.⁴⁵

When the term “prospect theory” is mentioned, it is often concepts like loss aversion that come to mind. However, the most radical feature of the prospect theory-inspired specification in (26) may not be loss aversion, but rather that, in contrast to the vast majority of dynamic models of stock market prices, the investor derives utility not only from consumption, but also from gains and losses in financial wealth.

Models in which investors derive utility from gains and losses in financial wealth, or even from gains and losses in specific components of their wealth, appear to be helpful for understanding the data. But how can we motivate such specifications? One motivation is based on the concept of “mental accounting,” which refers to the way people think about, keep track of, and evaluate their financial performance (Thaler,

⁴⁴ See Cohn et al. (2015) for experimental evidence of countercyclical risk aversion similar to that assumed by Barberis et al. (2001).

⁴⁵ Lian et al. (2018) present evidence of another risk-taking phenomenon, “reaching for yield,” whereby investors take more risk when interest rates are low. Using a model of He and Zhou (2011), they show that prospect theory can capture this behavior.

1999). When an individual invests in financial markets, her primary goal is typically to achieve good consumption outcomes for herself in the future. To check that she is on track for good consumption outcomes, she monitors her financial wealth on a regular basis – every year, say. If, at the end of the year, she sees that her financial wealth has gone up in value, she treats this as good news – an indication that she *is* on track for good consumption outcomes in the future. And precisely because it is good news, it is natural that this increase in financial wealth generates a burst of positive utility for her at the moment she observes it. Conversely, if, at the end of the year, the individual sees that her financial wealth has gone down in value, this is bad news – a sign that she is not on track for good consumption outcomes in the future – and precisely because it is bad news, it plausibly generates a burst of negative utility at the moment she observes it. If changes in financial wealth do create these bursts of utility, the individual may take them into account when allocating her portfolio. This, in turn, motivates the specification in (26).

7.5 Prospect theory applications: Summary

Prospect theory offers a more accurate description of risk attitudes than Expected Utility, at least in experimental settings. As such, it has the potential to advance our understanding of asset prices and investor behavior. To some extent, it is delivering on this promise. The research reviewed in Sections 7.2 through 7.4 shows how prospect theory sheds light on a wide range of facts and makes novel testable predictions. It is particularly helpful for thinking about *average* returns on financial assets. For example, it offers a simple explanation for both the high average return on the aggregate stock market *and* the low average return on positively-skewed assets.

In the 1990s, prospect theory research in finance focused on the loss-aversion element of the theory. The applications of loss aversion – to understanding the high equity premium and why many households do not participate in the stock market – are influential and have stood the test of time. However, in recent years, there have not been many new finance applications of loss aversion. Over the past decade, it is instead probability weighting and diminishing sensitivity that have received more attention.

Most implementations of prospect theory in finance use a “backward-looking” reference point: they posit that, when computing her gain or loss, the investor compares her current financial wealth to her past financial wealth. This historical reference point is often assumed to adapt over time: at time t , the investor compares her current wealth to wealth at time $t - 1$, but at time $t + 1$, she compares her current wealth to wealth at time t .

Koszegi and Rabin (2006, 2007) advocate “forward-looking” reference points. In the finance context, this could mean that an investor computes her gain or loss at time t by comparing her current wealth to the wealth she *expected* to have at that point. Such reference points have not been extensively explored in finance. It is not clear how plau-

sible they are: the high degree of uncertainty about assets' expected returns may mean that expectation-based reference points are not commonly used by real-world investors. Nonetheless, recent work by Pagel (2016) and Meng and Weng (2017) suggests that they may be relevant in some circumstances.⁴⁶

Specifying the reference point is just one of the difficulties facing prospect theory implementations in finance. A broader challenge is defining the “gains” and “losses” that are the carriers of value. Notice that the models discussed in Section 7 adopt somewhat different definitions. For example, in some models, investors derive utility from gains and losses in financial wealth, while in other models, they derive utility from gains and losses in individual assets. There may be good reasons for these different assumptions. However, there is a clear need for a theory that can give guidance on which gains and losses investors focus on in any particular setting. Developing such a theory is an important open challenge.

7.6 Other alternatives to Expected Utility

The effort to model investor preferences in a more accurate way has been influenced primarily by prospect theory. However, other non-EU theories have also been applied in finance in useful ways. These include disappointment aversion, rank-dependent expected utility, and salience theory.

In disappointment aversion, a framework due to Gul (1991), the individual puts more weight on outcomes that lie below the certainty equivalent of her future wealth distribution. Informally, she is loss averse relative to a specific reference point, namely this certainty equivalent. Disappointment aversion has been used to think about the high historical equity premium and household non-participation in the stock market (Epstein and Zin, 2001; Ang et al., 2005; Routledge and Zin, 2010; Dahlquist et al., 2016).

Rank-dependent expected utility, due to Quiggin (1982) and Yaari (1987), modifies the Expected Utility framework to allow for probability weighting: the individual does not use objective probabilities, but rather transformed probabilities obtained from a weighting function $g(\cdot)$. Specifically, the general gamble in (19) is evaluated as

$$\sum_{i=-m}^n \pi_i U(W + x_i), \quad (32)$$

where

$$\pi_i = g(p_{-m} + p_{-m+1} + \dots + p_i) - g(p_{-m} + p_{-m+1} + \dots + p_{i-1}), \quad (33)$$

⁴⁶ See O'Donoghue and Sprenger (2018) for a comprehensive discussion of reference points in economic decision-making.

and where the most common forms for $g(\cdot)$ are $g(P) = P^\phi$, for $\phi \in (0, 1]$, and the form used in cumulative prospect theory, namely

$$g(P) = \frac{P^\delta}{(P^\delta + (1 - P)^\delta)^{1/\delta}}. \quad (34)$$

As with disappointment aversion, this specification has been used to think about the equity premium and non-participation in the stock market, but also about the concentrated positions some individuals have in skewed assets (Epstein and Zin, 1990; Polkovnichenko, 2005).

Salience theory is described in Bordalo et al. (2012). In this theory, context matters: the value an individual assigns to a gamble depends on which other gambles she is comparing it to. Suppose that there are N gambles, indexed by i , and S states, indexed by s , and that the payoff of gamble i in state s is $x_{i,s}$. To evaluate gamble i , the individual ranks the payoffs $\{x_{i,s}\}_{s=1,\dots,S}$ by their salience. The salience of $x_{i,s}$ can be measured as

$$\frac{|x_{i,s} - \bar{x}_s|}{|x_{i,s}| + |\bar{x}_s| + \theta}, \quad (35)$$

where $\theta > 0$ and $\bar{x}_s = \sum_{i=1}^N x_{i,s}/N$. This captures the idea that state s is more salient for gamble i if the gamble's payoff in that state is very different from the mean payoff of the other gambles in that state. Suppose that, for gamble i , state s has salience rank $k_{i,s}$, where $k_{i,s}$ is an integer that ranges from 1 for the most salient state to S for the least salient state. The individual then evaluates the gamble as

$$\frac{1}{\Delta} \sum_{s=1}^S p_s \delta^{k_{i,s}} U(x_{i,s}), \quad (36)$$

where p_s is the probability of state s , $\delta \in (0, 1]$, and Δ is a normalizing factor that ensures that the sum of the weights on the S possible outcomes equals one. The weight on the most salient state is therefore adjusted by δ/Δ , while the weight on the least salient state is adjusted by δ^S/Δ ; as such, more salient states receive more weight. One immediate consequence is that a lottery-like gamble is appealing: the state in which the gamble pays a jackpot is overweighted because, in this state, the gamble's payoff is very different from that of the comparison gambles.

Bordalo et al. (2013) present a model of asset prices in which investors evaluate risk according to salience theory. A type of narrow framing is assumed: salience theory is applied at the level of individual assets, rather than at the portfolio level. One of the model's predictions is that an asset's idiosyncratic skewness will be priced. In Section 7.2, I noted that prospect theory makes a similar prediction and used this prediction to make sense of a range of facts about asset prices. Salience theory can likely also address these empirical patterns. Other potential applications are the equity premium and value

premium puzzles. Cosemans and Frehen (2017) test some of the predictions of salience theory for asset prices, with supportive results.⁴⁷

8. PREFERENCES: AMBIGUITY AVERSION

Economists have long distinguished between situations of “risk,” where the decision-maker does not know the future payoff of an investment but is able to assign probabilities to the various possible outcomes, and situations of “ambiguity,” where she does not feel able to assign probabilities to these outcomes. An important idea is that people are averse to situations of ambiguity in a way that they are not to situations of mere risk.

Research on ambiguity aversion has its origin in the Ellsberg paradox (Ellsberg, 1961). In a simple version of the paradox, there are two urns: Urn C, where “C” stands for Certain, and Urn U, where “U” stands for Uncertain. Urn C has 100 balls in it; 50 of them are red and 50 are black. Urn U also has 100 balls in it; each ball in this urn is either red or black, but the proportion of balls of each color is not known. Participants in an experiment are asked to choose between the following two bets:

Bet R1: A ball will be drawn from Urn C. You will receive \$100 if it is red.

Bet R2: A ball will be drawn from Urn U. You will receive \$100 if it is red.

Participants are separately asked to choose between the following two bets:

Bet B1: A ball will be drawn from Urn C. You will receive \$100 if it is black.

Bet B2: A ball will be drawn from Urn U. You will receive \$100 if it is black.

The modal participant choices are R1 and B1 (Becker and Brownson, 1964). This is labeled a paradox because these choices are not consistent with any beliefs a participant might have about the proportion of red balls in Urn U: the choice of R1 over R2 is consistent with a belief that fewer than 50% of the balls in Urn U are red, while the choice of B1 over B2 is consistent with a belief that more than 50% of the balls in Urn U are red. Subsequent experiments have checked that participants are not simply indifferent between R1 and R2, or between B1 and B2.

The leading hypothesis for understanding participant behavior in the above experiment is the ambiguity aversion hypothesis, which states that people are averse to situations of ambiguity – again, situations where they do not feel able to assign probabilities to future outcomes. Urn U is an example of such a situation: it is not known what fraction of balls in the urn are red. Due to ambiguity aversion, then, participants are reluctant to bet on it.

⁴⁷ Other preference specifications that researchers have explored are horizon-dependent risk aversion, whereby the individual is more averse to an imminent risk than to a distant one (Eisenbach and Schmalz, 2016; Andries et al., 2018); and regret aversion (Michenaud and Solnik, 2008; Frydman and Camerer, 2016). Hartzmark and Shue (2017) present evidence of a perceptual distortion known as a “contrast effect”: investors perceive a given earnings surprise to be less impressive if the previous day’s earnings surprises were particularly positive.

Over the past 30 years, a lot of effort has gone into developing mathematical models of ambiguity aversion, and there are now several models to choose from, as well as vigorous debates about their relative merits. Three of the best-known frameworks are models with multiple priors (Gilboa and Schmeidler, 1989); smooth ambiguity models (Klibanoff et al., 2005); and models of robust control (Hansen and Sargent, 2001).

Suppose that there is an uncertain outcome X , and that the individual does not feel able to describe X by a single probability distribution. In the multiple-priors framework, she brings to mind many candidate probability distributions – many “models” of X . She then chooses the action that maximizes the minimum Expected Utility she could obtain under any of these candidate probability distributions:

$$\max_{\text{action}} \min_{\text{models}} EU(X). \quad (37)$$

In the case of the Ellsberg paradox, an individual may bring to mind 101 different models of Urn U , where, under model $i \in \{0, 1, \dots, 100\}$, there are i black balls and $100 - i$ red balls in the urn. The framework in (37) then leads her to choose $R1$ over $R2$ but also $B1$ over $B2$: for action $R1$, $\min EU(X)$ is $0.5U(100)$, while for action $R2$, $\min EU(X)$ is zero, with the minimum attained for $i = 100$; similarly, for action $B1$, $\min EU(X)$ is $0.5U(100)$, while for action $B2$, $\min EU(X)$ is zero, with the minimum attained for $i = 0$. The multiple-priors framework therefore captures the observed behavior in the Ellsberg paradox.

In the multiple-priors framework, the model that is “worst” for the individual – in the case of action $R2$, this is model $i = 100$, for which none of the balls in Urn U is red – has an outsize impact on the individual’s choice, perhaps more than is plausible. The smooth ambiguity and robust control frameworks both try to address this by having the individual put less weight on models that are, in some sense, more extreme.

Ambiguity aversion has been applied in finance in a number of ways (Epstein and Schneider, 2010). It may explain why many households do not participate in the stock market: the return of the stock market is more ambiguous than the return on a bank deposit or a Treasury Bill; if an individual is sufficiently ambiguity averse, she may not invest in the stock market at all (Dow and Werlang, 1992). A similar logic provides a way of thinking about the high historical equity premium: if investors are ambiguity averse, they will require a much higher average return on the stock market than on Treasury Bills (Maenhout, 2004).

Ambiguity aversion has also been used to think about the under-diversification in many household portfolios (Goetzmann and Kumar, 2008). Individuals tilt their portfolios toward domestic stocks rather than foreign stocks (French and Poterba, 1991); toward stocks of locally-headquartered firms rather than stocks of firms located further away (Ivkovic and Weisbenner, 2005; Seasholes and Zhu, 2010); and toward the stock of the company they work for (Benartzi, 2001). One explanation for these findings is

that individuals are ambiguity averse and view the returns of the domestic stock market, locally-headquartered stocks, and their own company stock as less ambiguous than the returns of foreign stock markets, distant stocks, and non-company stocks, respectively (Uppal and Wang, 2003).

Ambiguity aversion may also amplify financial crises (Krishnamurthy, 2010). After bad economic news that pushes asset prices down, investors may feel that the level of ambiguity about future economic outcomes has gone up – in informal terms, that the future is now more uncertain. This increase in perceived ambiguity pushes asset prices even further down.

Ambiguity aversion is a topic where there has been enormous effort on the theoretical front – on developing formal, mathematical models – but less progress on the empirical side. The paucity of empirical research is unfortunate; it means that there is still relatively little direct evidence that ambiguity aversion matters for investor decisions or asset prices. The applications described above are promising, but there are other plausible explanations of non-participation, under-diversification, and crisis amplification.

One way to make empirical progress is to leverage developments in psychology. While ambiguity aversion remains the leading explanation of the Ellsberg paradox, another hypothesis, the “competence hypothesis,” can also address it (Heath and Tversky, 1991). According to this framework, an individual is ambiguity averse when she does not feel “competent” to evaluate the probability distribution of future outcomes, but ambiguity-seeking when she does feel competent. Here, “competence” refers to how much the individual feels she knows, relative to what could be known. The competence hypothesis explains the Ellsberg paradox by positing that, when thinking about the uncertain Urn U, the individual does not feel competent: there is something important that could be known about the urn that she does not know. As a consequence, she is ambiguity averse.

Heath and Tversky (1991) and Fox and Tversky (1995) present extensive evidence in support of the competence hypothesis – for example, that the same individual can be ambiguity-averse in one situation but ambiguity-seeking in another, depending on her feelings of competence. One of their most striking findings is how easy it is to alter a person’s attitude to ambiguity by manipulating her feelings of competence, either by contrasting a situation the person is thinking about with other, more ambiguous situations, or by reminding her of other people who are more or less competent than she at evaluating a particular source of uncertainty. The richness of the competence hypothesis may inspire new predictions in the realm of finance that can be brought to the data.

Some of the applications of ambiguity aversion can also be addressed using a related but distinct notion from psychology, namely that people have a preference for things that are familiar (Huberman, 2001). A concrete version of this is the “mere exposure effect” – the finding that mere exposure to someone or something makes us like that person

or thing more than justified based on informational considerations alone (Zajonc, 1968; Bornstein, 1989). In one experimental demonstration of this effect, participants with no knowledge of the Chinese language are shown a sequence of Chinese characters. Some of the characters appear just once in the sequence, others two times, others five times, and others ten times. For each unique character, the experimenter asks participants whether the character connotes something positive. The striking finding is that, on average, participants associate characters that appear more often with a more positive meaning (Zajonc, 1968).

The potential applications of the mere exposure effect overlap with those of ambiguity aversion. In particular, it offers an explanation for the instances of underdiversification where people tilt their portfolios toward domestic stocks, locally-headquartered stocks, and their own company stock. We are exposed, on a daily basis, to our home country, our local region, and the company we work for. The mere exposure effect may then lead us to view investments that are linked to these more favorably than we should.

How does ambiguity aversion differ from the mere exposure effect in its predictions? Consider an investor in, say, Brazil. The ambiguity aversion hypothesis predicts that this investor will tilt her portfolio toward the U.S. stock market: from a statistical perspective, there is less ambiguity about the distribution of U.S. stock market returns because of the many years of high quality data about it. By contrast, the mere exposure effect predicts that, since the Brazilian investor is exposed more often to her home country, she will tilt toward the Brazilian stock market. The available evidence is more consistent with the prediction of the mere exposure effect.

9. BOUNDED RATIONALITY

Sections 4 through 6 reviewed belief-based models of asset prices, while Sections 7 and 8 covered preference-based models. I now discuss a third approach to improving the psychological realism of our models of financial markets, one that takes account of human cognitive limits.

The traditional framework for thinking about asset prices assumes that, as soon as a piece of news is released, investors attend to it immediately, determine straight away what it implies for the prices of financial assets, and adjust their portfolios without delay. Given the large volume of news that is released on a weekly or even daily basis, and given how difficult it can be to discern the implications of this news, it is unlikely that, for all its impressive abilities, the human brain can do what the traditional framework assumes that it does. There are surely limits to how quickly a person can gather, process, and react to complex information.

The bounded-rationality approach to economic decision-making takes these cognitive limits seriously, and proposes a framework where people optimize subject to

constraints on their mental processing capacity. According to this framework, when people act the way they do, they are doing the best they can, given their cognitive limits; these limits prevent them from doing any better.

In this section, I describe some bounded-rationality approaches to understanding asset prices. The most prominent idea in this area is **inattention** – the notion that, because of the brain’s finite processing capacity, investors are unable to attend promptly to all relevant information – but I also discuss another concept, categorical thinking.⁴⁸

While the bounded-rationality approach is helpful for understanding investor behavior, it also faces a challenge. **According to this framework, investors are optimizing subject to cognitive constraints; given these constraints, it is not possible for them to act in a more rational way. The difficulty with this view is that, in many cases, real-world investors *could* do better without imposing any additional demands on their mental processing capacity: the more rational thing to do is often just as easy to implement, and sometimes easier to implement, than what they are currently doing.**

An example may be helpful. The data suggest that some investors have poor market timing: they increase their exposure to the stock market following good returns, just when, perhaps because it is overvalued, the market has a low expected return; similarly, they decrease their exposure to the stock market following low returns, just when, possibly because it is undervalued, it has a high expected return (Friesen and Sapp, 2007). These investors could do better by implementing a strategy that is no more cognitively demanding than what they are currently doing, namely, the *opposite* of what they are currently doing: increasing their exposure to the stock market following poor returns and decreasing their exposure following good returns. For another example, recall that many investors allocate substantial sums to actively-managed mutual funds, even though these funds underperform index funds on average after taking fees into account. These investors could do better without imposing any additional burden on their mental processing capacity: the more rational strategy, investing in index funds, is easier to implement than what they are currently doing.

The bounded-rationality framework may therefore not be the best way to think about certain aspects of investor behavior: investors often fail to take actions that would not only improve their financial outcomes but that are actually easier to implement than what they are currently doing. The psychology-based models of Sections 4 through 8 are a more promising way of understanding these aspects of behavior: according to these models, people have deeply-held psychological heuristics which make it hard for them to embrace rational courses of action and instead lead them astray. For example, the representativeness heuristic may lead an investor to increase her exposure to the stock

⁴⁸ Other kinds of bounded rationality in finance that I do not discuss in detail here are nominal illusion (Modigliani and Cohn, 1979); confusion about dividends (Hartzmark and Solomon, 2018); and additive as opposed to proportional thinking (Shue and Townsend, 2018).

market after good returns and to reduce it after poor returns, while overconfidence may lead her to believe that, even if index funds beat actively-managed funds on average, *she* has the ability to pick out in advance the few active funds that will beat the index.

Despite this challenge, the bounded-rationality approach *is* helpful for thinking about some empirical facts in finance. I now describe some of the research in this area.

9.1 Inattention

The brain's finite processing capacity suggests that real-world investors will not be able to immediately attend to every piece of news that is relevant to the prices of financial assets. Rather, it will take them some time to gather, process, and react to this information. A basic prediction of this "inattention" framework is that asset prices will underreact to news.

Inattention-driven underreaction offers a simple explanation for post-earnings announcement drift – the fact, described in Section 2.2, that, if a firm announces unexpectedly good (poor) earnings, its stock price not only moves up (down) on the day of the announcement, but keeps rising (falling) in the weeks thereafter. According to the inattention view, investors do not all have the cognitive capacity to immediately figure out the implications of an earnings announcement for the firm's future prospects; it can take them days if not weeks to complete the task. As a result, on a day with good earnings news, the stock price reacts in a muted way; only over the next few weeks, once more investors have had a chance to think through the implications of the announcement, does the stock price finally rise to the level that properly reflects the news.

The inattention hypothesis has found support in empirical tests. DellaVigna and Pollet (2009) show that, on average, there is more post-earnings announcement drift when a firm announces earnings on a Friday; the interpretation is that, with the weekend coming up, investors are less attentive. Firms appear to exploit this: a larger fraction of the earnings news announced on Fridays is bad news. Hirshleifer et al. (2009) show that, on average, there is more post-earnings announcement drift when a firm announces earnings on the same day as many other firms; a natural interpretation is that, due to the simultaneous arrival of news from many different firms, it takes investors more time to process each individual announcement.

Earnings announcements are not the only kind of news that investors appear to underreact to because of inattention. Others are: information about a firm's customers (Cohen and Frazzini, 2006); news about demographic shifts (DellaVigna and Pollet, 2007); information about the quality of a firm's R&D investments (Cohen et al., 2013); the absence of news about the status of a merger bid (Giglio and Shue, 2014); foreign-market news (Huang, 2015); changes to a firm's 10-k statement (Cohen et al., 2016);

and, more broadly, news that is harder to process (Cohen and Lou, 2012) or that is released on a gradual basis (Da et al., 2014a).

The work of DellaVigna and Pollet (2007) on demographic shifts is instructive. For a stylized example that illustrates the idea in their study, suppose that, in 2018, a large number of babies are born in the U.S. This implies that, six years later, say, there will be higher than usual demand for toys. If investors are attentive, they will push up the stock prices of toy companies in 2018, when the information pointing to higher future demand first appears. If they are not attentive, however, they will push up the stock prices of toy companies only later – perhaps only in 2024 when the increased demand for toys becomes too obvious to ignore. DellaVigna and Pollet (2007) show that stock market investors are inattentive: a strategy that, when demographic news is released, buys (shorts) the stocks of firms that stand to benefit (lose) from the demographic shift earns positive excess returns. Their study also gives us a sense of how inattentive investors are: stock market prices do properly reflect the implications of demographic shifts for firms' cash flows over the next five years; it is the implications of these shifts for cash flows more than five years out that are not fully incorporated into prices.⁴⁹

9.2 Categorization

A basic feature of the way humans think – one that likely evolved to help us navigate a complex world – is that we put things into categories and form beliefs at the level of these categories (Rosch and Lloyd, 1978). It is plausible that such category-based thinking occurs in financial markets, too. There are thousands of individual securities and investment funds, and it is impossible for investors to have an informed opinion about each and every one. Investors may therefore simplify their decision-making by putting financial assets into categories and forming beliefs about the future performance of these assets at the category level. The frequent use of labels such as “value stocks,” “growth stocks,” “small-cap stocks,” and “quality stocks” is consistent with such category-based thinking.

Barberis and Shleifer (2003) study asset prices in an economy where some investors put assets into categories and form beliefs about the assets' future returns at the category level. These category-based investors are also extrapolators: their belief about the future return of an asset category is a positive function of its recent past return. The model makes several predictions – for example, it predicts medium-term momentum and long-term reversal in category-level returns. However, its more novel contribution is to offer a behavioral theory of comovement.

⁴⁹ Da et al. (2011, 2014b) construct measures of investor attention and relate these to trading activity and asset returns. Attention is likely strongly affected by media coverage, and a sizable literature examines the impact of the media on financial markets (Tetlock, 2007; Engelberg and Parsons, 2011; Tetlock, 2015). While I have focused here on empirical work, there is growing interest in *models* of inattention; see Gabaix (2018) for a review.

In an economy with rational investors and no frictions, the price of an asset is the rational expectation of its future cash flows, discounted at a rate that properly accounts for the risk of these cash flows. Comovement in the returns of a subset of all risky assets is then due to one of two things: correlated news about the level of the assets' future cash flows, or correlated news about the riskiness of these cash flows.⁵⁰

The sources of comovement identified by the rational, frictionless framework likely explain a lot of the comovement in actual asset returns. Nonetheless, there are numerous instances where asset returns appear to comove more than can be explained by these fundamental sources. For example, the returns of value stocks with low price-to-earnings ratios comove strongly, as do the returns of small-cap stocks, but it has proven difficult to link these return correlations to comovement in fundamentals (Fama and French, 1995). Commodity returns comove in ways that are not explained by changes in current or expected future values of macroeconomic variables (Pindyck and Rotemberg, 1990). The returns of closed-end mutual funds comove with the returns of small-cap stocks even when the funds' holdings consist of large-cap stocks (Lee et al., 1991). And the returns of closed-end country funds – funds whose shares are traded in the U.S. but which invest in foreign firms – comove as much with the returns of the market where they are traded (the U.S.) as with the returns of the market where their investments are located (Hardouvelis et al., 1994; Bodurtha et al., 1995).

The model of comovement proposed in Barberis and Shleifer (2003) can explain some of these patterns. In this model, returns of assets in the same category comove more than can be explained by fundamentals. Suppose that “small-cap stocks” is a category in the minds of many investors and that this category has recently performed well. Since these investors are also extrapolators, they become bullish about the future return of the small-cap category and therefore increase their holdings of the assets in this category. By hitting all small-cap stocks in tandem, this demand shock generates comovement in the returns of small-cap stocks over and above any comovement caused by fundamentals. In the same way, this framework can explain the common factor in the returns of value stocks, small-cap stocks, and commodities.

Barberis et al. (2005) point out that, under the category-based view of comovement, after a stock is added to the S&P 500 index, its returns should start comoving more with the returns of the other stocks in the S&P 500. Under the traditional view, however, there will be an increase in return comovement only if, after inclusion, there is an increase in the comovement of the included firm's *fundamentals* with the fundamentals of firms already in the S&P 500, something that there is little evidence of. Barberis et

⁵⁰ News about future interest rates or changes in risk aversion, the two other forces that affect asset prices in rational models with no frictions, generate comovement in the returns of *all* risky assets, rather than just a subset of them.

al. (2005) test their prediction using a bivariate regression – specifically, a regression of the return of the included stock on the return of the S&P 500 and the return of stocks not in the S&P 500. They find that, after inclusion, consistent with category-based comovement, there is an increase in the coefficient on the first independent variable (the S&P 500 return) and a decrease in the coefficient on the second (the non-S&P 500 return).

Other papers also provide evidence of category-level comovement. For example, Green and Hwang (2009) show that there is comovement in the returns of stocks with similar prices. There is continued debate about whether the findings in these papers truly reflect “excessive” comovement or whether they should instead be attributed to comovement in fundamentals (Chen et al., 2016).

10. DISCUSSION AND CONCLUSION

Research on psychology-based approaches to thinking about asset prices began in earnest in the 1990s. In relative terms, then, the field is young: the traditional approach to studying asset prices, based on models with fully rational investors, has been pursued intensively since the 1960s. As such, it is too early to draw firm conclusions about the value of behavioral models of asset prices. Still, on some dimensions, these models have clearly been useful: they have shown that a few simple assumptions about investor psychology explain a wide range of important facts – facts about asset market fluctuations, trading volume, bubbles, and the performance of investment strategies – and lead to concrete new predictions.

In this survey, I have emphasized three frameworks that, based on the research to date, appear particularly helpful for thinking about asset prices and trading volume. These are the extrapolation framework (Section 4), the overconfidence framework (Section 5), and the gain-loss utility framework inspired by prospect theory (Section 7). The three frameworks are not necessarily in competition with one another, in part because they have somewhat different applications. The extrapolation framework explains excess volatility and time-series predictability in aggregate asset classes; momentum, long-run reversal, and the value premium in the cross-section; and the formation and collapse of asset bubbles. Overconfidence, by generating disagreement across investors, can address the high volume of trading in financial markets; when coupled with a short-sale constraint, it further explains the coincidence of high valuations and heavy trading that we see in the data. Gain-loss utility and prospect theory help us understand assets’ *average* returns: the high average return on the aggregate stock market, but also the low average return on positively-skewed assets such as IPOs, out-of-the-money options, and volatile stocks.

The psychological assumptions that are prominent in finance overlap partially but not fully with those invoked in other areas of economics. Reference dependence,

loss aversion, overconfidence, and inattention have all been applied both in finance and elsewhere in economics. But while the representativeness heuristic and probability weighting have been heavily exploited in finance, their use in other areas of economics is more limited. And while hyperbolic discounting has been applied extensively outside finance, it has had little traction in the field of asset prices. These differences are due to the pre-eminent role of risk in finance. The focus on risk means that probability weighting – an aspect of how people process risky outcomes – is relevant in finance, while hyperbolic discounting – a feature of time, not risk, preferences – is less so.

The research reviewed in this article takes a “positive” approach: it uses psychology-based models to make sense of observed facts. But what are the “normative” implications of these models? If the models are on the right track, what should a rational investor do?

The models described here appear to point to an active trading strategy for rational investors, one where these investors tilt their portfolios toward low price-to-earnings stocks and gently time the stock market to take advantage of return predictability. However, most financial economists, including behavioral finance specialists, instead advise households to follow a passive strategy – for example, to invest in index funds. Why is this? One reason is that, until recently, most of the investment vehicles that would allow households to take advantage of mispricing have been unappealing: after fees, the average actively-managed fund underperforms index funds. However, in just the past few years, new financial products have appeared that exploit mispricing in more cost-effective ways – many of these products mechanically buy and sell assets with the characteristics listed in Section 2.2 – and economists are beginning to recommend them as good investment options. These new products are attracting large flows from institutional investors and are drawing interest from households too.

Researchers entering behavioral finance today are often unaware that the field was controversial when it emerged in the 1980s and 1990s. The controversy was due in part to academic politics – to resistance from people who had spent their lives working on the traditional framework only to see their legacy threatened by an upstart paradigm. However, some of the controversy was rooted in reasonable scientific critiques of the new field. One of these is the arbitrage critique, which has now been addressed by the work on limits to arbitrage discussed in Section 3.

A second critique of behavioral finance which was regularly voiced in the 1990s is the “lack of discipline critique.” It posits that, because there are many ways in which people depart from full rationality, it is too easy to find “explanations” for observed facts by flipping through the pages of a psychology textbook. In a 1998 paper, Eugene Fama predicted that we would soon see a profusion of psychological assumptions in finance – 30 different assumptions to explain 30 different facts, say (Fama, 1998).

Twenty years later, it is striking to see that Fama’s prediction has proven false, and his concerns unfounded. While researchers are investigating a variety of psychological

assumptions, the center of gravity in behavioral finance lies in a small number of ideas, primarily the three I picked out above – extrapolation, overconfidence, and gain-loss utility. The lack of discipline critique is heard less often these days. One reason for this is that there has been a concerted effort to not simply posit explanations for known facts, but to test the new predictions of these explanations. But another reason is precisely that researchers have kept their focus on a small set of psychological concepts.

A long-term goal of behavioral finance research is to converge on a “unified” psychology-based model of investor behavior – a model that, in a parsimonious way, makes psychologically realistic assumptions about both beliefs and preferences, and has broad explanatory power. It is too early to know what this model will look like, but the research to date makes it possible to speculate about its form. Relative to the traditional framework based on rational beliefs and Expected Utility preferences, one behavioral alternative posits extrapolative rather than rational beliefs, and modifies preferences to incorporate gain-loss utility and elements of prospect theory. In this framework, an individual thinking about investing in an asset first forecasts the gains and losses in financial wealth that could result from doing so, and bases this forecast on the asset’s past returns, especially those in the more recent past; in this way, her behavior combines extrapolation with a focus on potential gains and losses. The individual then evaluates this perceived distribution of future gains and losses as suggested by prospect theory, putting more weight on the tails of the distribution and on potential losses as opposed to gains. There is, as yet, very little work on models that couple extrapolative beliefs with gain-loss utility.⁵¹ However, the research discussed in this article suggests that such models will explain a broad set of facts about average returns, volatility, and predictability.⁵²

While extrapolative beliefs and gain-loss utility are both helpful for understanding the data, they both also raise fundamental questions that have not been fully answered: Why do people extrapolate, and how do they extrapolate? For example, how far back do people look when forming judgments about the future, and why? And if, when making investment decisions, people think about the gains and losses that could result, how do they define these potential gains and losses?

Over the past three decades, the effort to build psychologically accurate models of investor behavior and asset prices has followed one particular approach: incorporating ideas from the area of psychology known as judgment and decision-making – ideas such as representativeness, availability, overconfidence, and prospect theory – into otherwise traditional finance models. I have argued in this article that this approach has

⁵¹ One exception is Liao and Peng (2018), who propose, and test, a theory of volume based on the interaction of extrapolative beliefs and realization utility.

⁵² It might seem that this unified framework neglects the insights of research on overconfidence, but this is not the case. As noted in Section 5.1, the extrapolation framework implicitly incorporates overconfidence, in that extrapolators fail to learn anything from the fact that other investors are willing to trade with them.

been fruitful. However, it is also unnecessarily narrow in scope: it focuses largely on the work of Daniel Kahneman and Amos Tversky – work that, while very influential, represents only a sliver of the field of psychology. Beyond the work on judgment and decision-making, there are large tracts of research in psychology on attention, memory, and emotion, not to mention a rapidly growing field known as decision neuroscience, that economists have not connected to very much. It seems important to do so, and I expect that researchers will take up this task more energetically in the years ahead. A related goal is to unify the models of human behavior that are emerging across the social and natural sciences. Researchers in economics, finance, psychology, and neuroscience are proposing models of how people form representations of the world and of how they make choices based on these representations. These models differ from one another. It is important to understand why they differ – and if there is no good reason, to unify them.

In the 1990s, when behavioral finance became an active area of inquiry, conferences for both academics and practitioners staged debates between proponents of traditional models and advocates of behavioral models. These debates were memorable and often entertaining, but they did little to advance the cause of science. Recognizing this, researchers on both sides gave up the verbal jousting and focused instead on writing down models of their ideas and taking these models to the data. To the extent that behavioral finance has flourished, it is because of this approach: it has produced models that explain important facts in simple, intuitive ways and make predictions that have found support in the data. I expect that the field will continue to follow this approach in the years ahead, and that this effort will further advance our understanding of investor behavior in financial markets.

APPENDIX A. DERIVATION OF EQS. (3), (4), AND (5)

At time t , each extrapolator's objective is

$$\max_{N_t^e} E_t^e(-e^{-\gamma(W_t + N_t^e(P_{t+1} - P_t))}).$$

If, for simplicity, the extrapolator assumes that the conditional distribution of $P_{t+1} - P_t$ is Normal with variance σ_ϵ^2 , his first-order condition implies

$$N_t^e = \frac{E_t^e(P_{t+1} - P_t)}{\gamma \sigma_\epsilon^2}. \quad (38)$$

Coupling (38) with the beliefs in Eq. (2) gives Eq. (3).

At time t , each fundamental trader's objective is

$$\max_{N_t^f} E_t^f(-e^{-\gamma(W_t + N_t^f(P_{t+1} - P_t))}).$$

As for the extrapolator, the first-order condition leads to

$$N_t^f = \frac{E_t^f(P_{t+1} - P_t)}{\gamma\sigma_\varepsilon^2}.$$

At time $T - 1$, the fundamental trader sets $E_{T-1}^f(P_T) = D_{T-1}$, so that

$$N_{T-1}^f = \frac{D_{T-1} - P_{T-1}}{\gamma\sigma_\varepsilon^2}.$$

From the market-clearing condition

$$\mu^e N_{T-1}^e + \mu^f N_{T-1}^f = Q$$

we obtain

$$\begin{aligned} P_{T-1} &= D_{T-1} + \frac{\mu^e}{\mu^f} \gamma \sigma_\varepsilon^2 N_{T-1}^e - \frac{\gamma \sigma_\varepsilon^2 Q}{\mu^f} \\ &= D_{T-1} + \frac{\mu^e}{\mu^f} X_{T-1} - \frac{\gamma \sigma_\varepsilon^2 Q}{\mu^f}. \end{aligned} \tag{39}$$

Recall that fundamental traders are boundedly rational: they do not understand how extrapolators form beliefs but instead assume that, in future periods, extrapolators will hold the risky asset in proportion to their weight in the population. From (39), this implies

$$E_{T-2}^f(P_{T-1}) = D_{T-2} + \frac{\mu^e}{\mu^f} \gamma \sigma_\varepsilon^2 Q - \frac{\gamma \sigma_\varepsilon^2 Q}{\mu^f},$$

so that

$$N_{T-2}^f = \frac{D_{T-2} - \gamma \sigma_\varepsilon^2 Q - P_{T-2}}{\gamma \sigma_\varepsilon^2}.$$

From the market-clearing condition

$$\mu^e N_{T-2}^e + \mu^f N_{T-2}^f = Q$$

we obtain

$$\begin{aligned} P_{T-2} &= D_{T-2} + \frac{\mu^e}{\mu^f} \gamma \sigma_\varepsilon^2 N_{T-2}^e - \gamma \sigma_\varepsilon^2 Q \left(1 + \frac{1}{\mu^f}\right) \\ &= D_{T-2} + \frac{\mu^e}{\mu^f} X_{T-2} - \gamma \sigma_\varepsilon^2 Q \left(1 + \frac{1}{\mu^f}\right). \end{aligned}$$

Continuing the backward induction in this way leads to Eqs. (4) and (5).

REFERENCES

- Abreu, D., Brunnermeier, M., 2002. Synchronization risk and delayed arbitrage. *Journal of Financial Economics* 66, 341–360.
- Adam, K., Marcet, A., Beutel, J., 2017. Stock price booms and expected capital gains. *American Economic Review* 107, 2352–2408.
- Alti, A., Tetlock, P., 2014. Biased beliefs, asset prices, and investment: a structural approach. *Journal of Finance* 69, 325–361.
- Amromin, G., Sharpe, S., 2014. From the horse's mouth: economic conditions and investor expectations of risk and return. *Management Science* 60, 845–866.
- Andries, M., Eisenbach, T., Schmalz, M., 2018. Horizon-Dependent Risk Aversion and the Timing and Pricing of Uncertainty. Working paper.
- Andrikogiannopoulou, A., Papakonstantinou, F., 2018. History-Dependent Risk Preferences: Evidence from Individual Choices and Implications for the Disposition Effect. Working paper.
- Ang, A., Bekaert, G., Liu, J., 2005. Why stocks may disappoint. *Journal of Financial Economics* 76, 471–508.
- Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61, 259–299.
- Asness, C., Moskowitz, T., Pedersen, L., 2013. Value and momentum everywhere. *Journal of Finance* 68, 929–985.
- Astebro, T., Herz, H., Nanda, R., Weber, R., 2014. Seeking the roots of entrepreneurship: insights from behavioral economics. *Journal of Economic Perspectives* 28, 46–69.
- Augenblick, N., Lazarus, E., 2018. Restrictions on Asset-Price Movements Under Rational Expectations: Theory and Evidence. Working paper.
- Bacchetta, P., Mertens, E., van Wincoop, E., 2009. Predictability in financial markets: what do survey expectations tell us? *Journal of International Money and Finance* 28, 406–426.
- Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61, 1645–1680.
- Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives* 21, 129–151.
- Baker, M., Wurgler, J., 2012. Behavioral corporate finance: an updated survey. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), *Handbook of the Economics of Finance*, vol. 2. North-Holland, Amsterdam.
- Bali, T., Cakici, N., Whitelaw, R., 2011. Maxing out: stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99, 427–446.
- Ball, R., Gerakos, J., Linnainmaa, J., Nikolaev, V., 2015. Deflating profitability. *Journal of Financial Economics* 117, 225–248.
- Banerjee, A., 1992. A simple model of herd behavior. *Quarterly Journal of Economics* 107, 797–817.
- Banerjee, S., 2011. Learning from prices and the dispersion in beliefs. *Review of Financial Studies* 24, 3025–3068.
- Banerjee, S., Kremer, I., 2010. Disagreement and learning: dynamic patterns of trade. *Journal of Finance* 65, 1269–1302.
- Bansal, R., Yaron, A., 2004. Risks for the long run: a potential resolution of asset pricing puzzles. *Journal of Finance* 59, 1481–1509.
- Bansal, R., Kiku, D., Yaron, A., 2012. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review* 1, 183–221.
- Banz, R., 1981. The relation between return and market value of common stocks. *Journal of Financial Economics* 9, 3–18.
- Barber, B., Odean, T., 2001. Boys will be boys: gender, overconfidence, and common stock investment. *Quarterly Journal of Economics* 116, 261–292.

- Barberis, N., Huang, M., 2001. Mental accounting, loss aversion and individual stock returns. *Journal of Finance* 56, 1247–1292.
- Barberis, N., Huang, M., 2008. Stocks as lotteries: the implications of probability weighting for security prices. *American Economic Review* 98, 2066–2100.
- Barberis, N., Huang, M., 2009. Preferences with frames: a new utility specification that allows for the framing of risks. *Journal of Economic Dynamics and Control* 33, 1555–1576.
- Barberis, N., Shleifer, A., 2003. Style investing. *Journal of Financial Economics* 68, 161–199.
- Barberis, N., Thaler, R., 2003. A survey of behavioral finance. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), *Handbook of the Economics of Finance*, vol. 1. North-Holland, Amsterdam.
- Barberis, N., Xiong, W., 2009. What drives the disposition effect? An analysis of a long-standing preference-based explanation. *Journal of Finance* 64, 751–784.
- Barberis, N., Xiong, W., 2012. Realization utility. *Journal of Financial Economics* 104, 251–271.
- Barberis, N., Huang, M., Santos, T., 2001. Prospect theory and asset prices. *Quarterly Journal of Economics* 116, 1–53.
- Barberis, N., Huang, M., Thaler, R., 2006. Individual preferences, monetary gambles, and stock market participation: a case for narrow framing. *American Economic Review* 96, 1069–1090.
- Barberis, N., Mukherjee, A., Wang, B., 2016. Prospect theory and stock returns: an empirical test. *Review of Financial Studies* 29, 3068–3107.
- Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. *Journal of Financial Economics* 49, 307–345.
- Barberis, N., Shleifer, A., Wurgler, J., 2005. Comovement. *Journal of Financial Economics* 75, 283–317.
- Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015. X-CAPM: an extrapolative capital asset pricing model. *Journal of Financial Economics* 115, 1–24.
- Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2018. Extrapolation and bubbles. *Journal of Financial Economics* 129, 203–227.
- Baron, J., 2000. *Thinking and Deciding*. Cambridge University Press.
- Barsky, R., De Long, J.B., 1993. Why does the stock market fluctuate? *Quarterly Journal of Economics* 108, 291–311.
- Basu, S., 1983. The relationship between earnings yield, market value and return for NYSE common stocks: further evidence. *Journal of Financial Economics* 12, 129–156.
- Becker, S., Brownson, F., 1964. What price ambiguity? Or the role of ambiguity in decision-making. *Journal of Political Economy* 72, 62–73.
- Benartzi, S., 2001. Excessive extrapolation and the allocation of 401(k) accounts to company stock. *Journal of Finance* 56, 1747–1764.
- Benartzi, S., Thaler, R., 1995. Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics* 110, 75–92.
- Benjamin, D., Moore, D., Rabin, M., 2017. Biased Beliefs About Random Samples: Evidence from Two Integrated Experiments. Working paper.
- Benjamin, D., Rabin, M., Raymond, C., 2015. A model of non-belief in the law of large numbers. *Journal of the European Economic Association* 14, 515–544.
- Bernard, V., Thomas, J., 1989. Post-earnings announcement drift: delayed price response or risk premium? *Journal of Accounting Research (Supplement)*, 1–36.
- Bernheim, D., Sprenger, C., 2016. Direct Tests of Cumulative Prospect Theory. Working paper.
- Beshears, J., Choi, J., Laibson, D., Madrian, B., 2017. Does aggregated returns disclosure increase portfolio risk-taking? *Review of Financial Studies* 30, 1971–2005.
- Beshears, J., Choi, J., Laibson, D., Madrian, B., 2018. Behavioral household finance. In: Bernheim, D., DellaVigna, S., Laibson, D. (Eds.), *Handbook of Behavioral Economics*, vol. 1. North-Holland, Amsterdam.
- Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy* 100, 992–1026.

- Birru, J., 2018. Day of the week and the cross-section of returns. *Journal of Financial Economics*. <https://doi.org/10.1016/j.jfineco.2018.06.008>. Forthcoming.
- Black, F., 1972. Capital market equilibrium with restricted borrowing. *Journal of Business* 45, 444–455.
- Blanchard, O., Watson, M., 1982. Bubbles, rational expectations, and financial markets. In: Wachtel, P. (Ed.), *Crisis in the Economic and Financial Structure: Bubbles, Bursts, and Shocks*. Lexington Books.
- Bodurtha, J., Kim, D., Lee, C.M., 1995. Closed-end country funds and U.S. market sentiment. *Review of Financial Studies* 8, 879–918.
- Booij, A., van Praag, B., van de Kuilen, G., 2010. A parametric analysis of prospect theory's functionals for the general population. *Theory and Decision* 68, 115–148.
- Bordalo, P., Gennaioli, N., Shleifer, A., 2012. Salience theory of choice under risk. *Quarterly Journal of Economics* 127, 1243–1285.
- Bordalo, P., Gennaioli, N., Shleifer, A., 2013. Salience and asset prices. *American Economic Review: Papers and Proceedings* 103, 623–628.
- Bordalo, P., Gennaioli, N., Shleifer, A., 2018a. Diagnostic expectations and credit cycles. *Journal of Finance* 73, 199–227.
- Bordalo, P., Coffman, K., Gennaioli, N., Shleifer, A., 2014. Stereotypes. *Quarterly Journal of Economics* 131, 1753–1794.
- Bordalo, P., Gennaioli, N., La Porta, R., Shleifer, A., 2017. Diagnostic Expectations and Stock Returns. Working paper.
- Bordalo, P., Gennaioli, N., Ma, Y., Shleifer, A., 2018b. Overreaction in Macroeconomic Expectations. Working paper.
- Bornstein, R., 1989. Exposure and affect: overview and meta-analysis of research, 1968–1987. *Psychological Bulletin* 106, 265–289.
- Bouchaud, J-P., Kruger, P., Landier, A., Thesmar, D., 2018. Sticky expectations and the profitability anomaly. *Journal of Finance*. Forthcoming.
- Boyer, B., Vorkink, K., 2014. Stock options as lotteries. *Journal of Finance* 69, 1485–1527.
- Boyer, B., Mitton, T., Vorkink, K., 2010. Expected idiosyncratic skewness. *Review of Financial Studies* 23, 169–202.
- Brunnermeier, M., Nagel, S., 2004. Hedge funds and the technology bubble. *Journal of Finance* 59, 2013–2040.
- Brunnermeier, M., Oehmke, M., 2013. Bubbles, financial crises, and systemic risk. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), *Handbook of the Economics of Finance*, vol. 2. North-Holland, Amsterdam.
- Brunnermeier, M., Parker, J., 2005. Optimal expectations. *American Economic Review* 95, 1092–1118.
- Campbell, J.Y., 2018. *Financial Decisions and Markets*. Princeton University Press, Princeton, New Jersey.
- Campbell, J.Y., Cochrane, J., 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107, 205–251.
- Campbell, J.Y., Shiller, R., 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1, 195–228.
- Campbell, J.Y., Shiller, R., 1998. Valuation ratios and the long-run stock market outlook. *Journal of Portfolio Management* 24, 11–26.
- Case, K., Shiller, R., Thompson, A., 2012. What have they been thinking? Homebuyer behavior in hot and cold markets. *Brookings Papers on Economic Activity* 2012, 265–315.
- Cassella, S., Gulen, H., 2018. Extrapolation bias and the predictability of stock returns by price-scaled variables. *Review of Financial Studies*. <https://doi.org/10.1093/rfs/hhx139>. Forthcoming.
- Chen, H., Singal, V., Whitelaw, R., 2016. Comovement revisited. *Journal of Financial Economics* 121, 624–644.
- Choi, J., Mertens, T., 2013. Extrapolative Expectations and the Equity Premium. Working paper.
- Cohen, L., Frazzini, A., 2006. Economic links and predictable returns. *Journal of Finance* 63, 1977–2011.

- Cohen, L., Lou, D., 2012. Complicated firms. *Journal of Financial Economics* 104, 383–400.
- Cohen, L., Diether, K., Malloy, C., 2013. Misvaluing innovation. *Review of Financial Studies* 26, 635–666.
- Cohen, L., Malloy, C., Nguyen, Q., 2016. Lazy Prices. Working paper.
- Cohn, A., Engelmann, J., Fehr, E., Marechal, M., 2015. Evidence for countercyclical risk aversion: an experiment with financial professionals. *American Economic Review* 105, 860–885.
- Coibon, O., Gorodnichenko, Y., 2012. What can survey forecasts tell us about inflation rigidities? *Journal of Political Economy* 120, 116–159.
- Collin-Dufresne, P., Johannes, M., Lochstoer, L., 2016. Asset pricing when ‘this time is different’. *Review of Financial Studies* 30, 505–535.
- Conrad, J., Dittmar, R., Ghysels, E., 2013. Ex ante skewness and expected stock returns. *Journal of Finance* 68, 85–124.
- Conrad, J., Kapadia, N., Xing, Y., 2014. Death and jackpot: why do individual investors hold overpriced stocks? *Journal of Financial Economics* 113, 455–475.
- Cooper, M., Gulen, H., Schill, M., 2008. Asset growth and the cross-section of stock returns. *Journal of Finance* 63, 1609–1651.
- Cosemans, M., Frehen, R., 2017. Saliency Theory and Stock Prices: Empirical Evidence. Working paper.
- Cutler, D., Poterba, J., Summers, L., 1990. Speculative dynamics and the role of feedback traders. *American Economic Review: Papers and Proceedings* 80, 63–68.
- Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. *Journal of Finance* 66, 1461–1490.
- Da, Z., Engelberg, J., Gao, P., 2014b. The sum of all FEARS: investor sentiment and asset prices. *Review of Financial Studies* 28, 1–32.
- Da, Z., Gurun, U., Warachka, M., 2014a. Frog in the pan: continuous information and momentum. *Review of Financial Studies* 27, 2171–2218.
- Da, Z., Huang, X., Jin, L., 2018. Extrapolative Beliefs in the Cross-Section: What Can We Learn from the Crowds? Working paper.
- Dahlquist, M., Farago, A., Tedongap, R., 2016. Asymmetries and portfolio choice. *Review of Financial Studies* 30, 667–702.
- Daniel, K., Titman, S., 2006. Market reactions to tangible and intangible information. *Journal of Finance* 61, 1605–1643.
- Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. *Journal of Finance* 53, 1839–1885.
- Daniel, K., Hirshleifer, D., Subrahmanyam, A., 2001. Overconfidence, arbitrage, and equilibrium asset pricing. *Journal of Finance* 56, 921–965.
- Das, S., Kuhnen, C., Nagel, S., 2017. Socioeconomic Status and Macroeconomic Expectations. Working paper.
- De Bondt, W., Thaler, R., 1985. Does the stock market overreact? *Journal of Finance* 40, 793–808.
- De Giorgi, E., Legg, S., 2012. Dynamic portfolio choice and asset pricing with narrow framing and probability weighting. *Journal of Economic Dynamics and Control* 36, 951–972.
- De Long, J.B., Shleifer, A., Summers, L., Waldmann, R., 1990a. Noise trader risk in financial markets. *Journal of Political Economy* 98, 703–738.
- De Long, J.B., Shleifer, A., Summers, L., Waldmann, R., 1990b. Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance* 45, 375–395.
- DeFusco, A., Nathanson, C., Zwick, E., 2018. Speculative Dynamics of Prices and Volume. Working paper.
- DellaVigna, S., Pollet, J., 2007. Demographics and industry returns. *American Economic Review* 97, 1667–1702.
- DellaVigna, S., Pollet, J., 2009. Investor inattention and Friday earnings announcements. *Journal of Finance* 64, 709–749.
- Diether, K., Malloy, C., Scherbina, A., 2003. Stock prices and differences of opinion: empirical evidence that stock prices reflect optimism. *Journal of Finance* 57, 2113–2141.

- Dow, J., Werlang, S., 1992. Uncertainty aversion, risk aversion, and the optimal choice of portfolio. *Econometrica* 60, 197–204.
- Dunning, D., Heath, C., Suls, J., 2004. Flawed self-assessment: implications for health, education, and the workplace. *Psychological Science in the Public Interest* 5, 69–106.
- Dunning, D., Meyerowitz, J., Holzberg, A., 1989. Ambiguity and self-evaluation: the role of idiosyncratic trait definitions in self-serving assessments of others. *Journal of Personality and Social Psychology* 57, 1082–1090.
- Easley, D., Yang, L., 2015. Loss aversion, survival, and asset prices. *Journal of Economic Theory* 160, 494–516.
- Edmans, A., Garcia, D., Norli, O., 2007. Sports sentiment and stock returns. *Journal of Finance* 62, 1967–1998.
- Edwards, W., 1968. Conservatism in human information processing. In: Kleinmütz, B. (Ed.), *Formal Representation of Human Judgment*. John Wiley and Sons, New York.
- Ehling, P., Graniero, A., Heyerdahl-Larsen, C., 2018. Asset prices and portfolio choice with learning from experience. *Review of Economic Studies* 85, 1752–1780.
- Eisenbach, T., Schmalz, M., 2016. Anxiety in the face of risk. *Journal of Financial Economics* 121, 414–426.
- Ellsberg, D., 1961. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics* 75, 643–669.
- Engelberg, J., Parsons, C., 2011. The causal impact of media in financial markets. *Journal of Finance* 66, 67–97.
- Engelberg, J., McLean, R.D., Pontiff, J., 2017. Anomalies and news. *Journal of Finance*. Forthcoming.
- Epstein, L., Schneider, M., 2010. Ambiguity and asset markets. *Annual Review of Financial Economics* 2, 315–346.
- Epstein, L., Zin, S., 1990. First-order risk aversion and the equity premium. *Journal of Monetary Economics* 26, 387–407.
- Epstein, L., Zin, S., 2001. The independence axiom and asset returns. *Journal of Empirical Finance* 8, 537–572.
- Eraker, B., Ready, M., 2015. Do investors overpay for stocks with lottery-like payoffs? An examination of the returns of OTC stocks. *Journal of Financial Economics* 115, 486–504.
- Eyster, E., Rabin, M., 2010. Naive herding in rich-information settings. *American Economic Journal: Microeconomics* 2, 221–243.
- Eyster, E., Rabin, M., Vayanos, D., 2018. Financial markets where traders neglect the informational content of prices. *Journal of Finance*. Forthcoming.
- Fama, E., 1991. Efficient capital markets: II. *Journal of Finance* 46, 1575–1617.
- Fama, E., 1998. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics* 49, 283–306.
- Fama, E., 2014. Two pillars of asset pricing. *American Economic Review* 104, 1467–1485.
- Fama, E., French, K., 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22, 3–25.
- Fama, E., French, K., 1992. The cross-section of expected stock returns. *Journal of Finance* 47, 427–465.
- Fama, E., French, K., 1995. Size and book-to-market factors in earnings and returns. *Journal of Finance* 50, 131–155.
- Fox, C., Tversky, A., 1995. Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics* 110, 585–603.
- Frankel, J., Froot, K., 1990. Chartists, fundamentalists, and trading in the foreign exchange market. *American Economic Review* 80, 181–185.
- Frazzini, A., 2006. The disposition effect and underreaction to news. *Journal of Finance* 61, 2017–2046.
- Frazzini, A., Pedersen, L., 2014. Betting against beta. *Journal of Financial Economics* 111, 1–25.
- Frazzini, A., Israel, R., Moskowitz, T., 2015. Trading Costs of Asset Pricing Anomalies. Working paper.

- French, K., 2008. Presidential address: the cost of active investing. *Journal of Finance* 63, 1537–1573.
- French, K., Poterba, J., 1991. Investor diversification and international equity markets. *American Economic Review* 81, 222–226.
- Friesen, G., Sapp, T., 2007. Mutual fund flows and investor returns: an empirical examination of fund investor timing ability. *Journal of Banking and Finance* 31, 2796–2816.
- Froot, K., Dabora, E., 1999. How are stock prices affected by the location of trade? *Journal of Financial Economics* 53, 189–216.
- Frydman, C., Camerer, C., 2016. Neural evidence of regret and its implications for investor behavior. *Review of Financial Studies* 29, 3108–3139.
- Frydman, C., Barberis, N., Camerer, C., Bossaerts, P., Rangel, A., 2014. Using neural data to test a theory of investor behavior: an application to realization utility. *Journal of Finance* 69, 907–946.
- Fuster, A., Hebert, B., Laibson, D., 2012. Natural expectations, macroeconomic dynamics, and asset pricing. *NBER Macroeconomics Annual* 26, 1–48.
- Fuster, A., Laibson, D., Mendel, B., 2010. Natural expectations and macroeconomic fluctuations. *Journal of Economic Perspectives* 24, 67–84.
- Gabaix, X., 2012. Variable rare disasters: an exactly solved framework for ten puzzles in macro-finance. *Quarterly Journal of Economics* 127, 645–700.
- Gabaix, X., 2018. Behavioral inattention. In: Bernheim, D., DellaVigna, S., Laibson, D. (Eds.), *Handbook of Behavioral Economics*, vol. 2. North-Holland, Amsterdam. Forthcoming.
- Gennaioli, N., Shleifer, A., 2010. What comes to mind? *Quarterly Journal of Economics* 125, 1399–1433.
- Gennaioli, N., Ma, Y., Shleifer, A., 2016. Expectations and investment. *NBER Macroeconomics Annual* 30, 379–431.
- Gennaioli, N., Shleifer, A., Vishny, R., 2012. Neglected risks, financial innovation, and financial fragility. *Journal of Financial Economics* 104, 452–468.
- Giglio, S., Kelly, B., 2017. Excess volatility: beyond discount rates. *Quarterly Journal of Economics* 133, 71–127.
- Giglio, S., Shue, K., 2014. No news is news: do markets underreact to nothing? *Review of Financial Studies* 27, 3389–3440.
- Giglio, S., Maggiori, M., Stroebel, J., 2016. No-bubble condition: model-free tests in housing markets. *Econometrica* 84, 1047–1091.
- Gilboa, I., Schmeidler, D., 1989. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics* 18, 141–153.
- Gilovich, T., Vallone, R., Tversky, A., 1985. The hot hand in basketball: on the misperception of random sequences. *Cognitive Psychology* 97, 295–314.
- Glaeser, E., Nathanson, C., 2015. Housing bubbles. In: Duranton, G., Henderson, V., Strange, W. (Eds.), *Handbook of Regional and Urban Economics*. North-Holland, Amsterdam.
- Glaeser, E., Nathanson, C., 2017. An extrapolative model of house price dynamics. *Journal of Financial Economics* 126, 147–170.
- Glaser, M., Weber, M., 2007. Overconfidence and trading volume. *The Geneva Risk and Insurance Review* 32, 1–36.
- Gneezy, U., Potters, J., 1997. An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics* 112, 631–645.
- Goetzmann, W., Kumar, A., 2008. Equity portfolio diversification. *Review of Finance* 12, 433–463.
- Goetzmann, W., Kim, D., Shiller, R., 2017. Crash Beliefs from Investor Surveys. Working paper.
- Gonzalez, R., Wu, G., 1999. On the shape of the probability weighting function. *Cognitive Psychology* 38, 129–166.
- Green, T.C., Hwang, B.-H., 2009. Price-based return comovement. *Journal of Financial Economics* 93, 37–50.
- Green, T.C., Hwang, B.-H., 2012. Initial public offerings as lotteries: skewness preference and first-day returns. *Management Science* 58, 432–444.

- Greenwood, R., Hanson, S., 2013. Issuer quality and corporate bond returns. *Review of Financial Studies* 26, 1483–1515.
- Greenwood, R., Hanson, S., 2015. Waves in ship prices and investment. *Quarterly Journal of Economics* 130, 55–109.
- Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. *Review of Financial Studies* 27, 714–746.
- Greenwood, R., Shleifer, A., You, Y., 2018. Bubbles for Fama. *Journal of Financial Economics*. Forthcoming.
- Grinblatt, M., Han, B., 2005. Prospect theory, mental accounting, and momentum. *Journal of Financial Economics* 78, 311–339.
- Grinblatt, M., Keloharju, M., 2009. Sensation seeking, overconfidence, and trading activity. *Journal of Finance* 64, 549–578.
- Gromb, D., Vayanos, D., 2010. Limits to arbitrage. *Annual Review of Financial Economics* 2, 251–275.
- Guiso, L., Sodini, P., 2013. Household finance: an emerging field. In: Constantinides, G., Harris, M., Stulz, R. (Eds.), *Handbook of the Economics of Finance*, vol. 2. North-Holland, Amsterdam.
- Gul, F., 1991. A theory of disappointment in decision making under uncertainty. *Econometrica* 59, 667–686.
- Guo, J., He, X., 2017a. Equilibrium asset pricing with Epstein–Zin and loss-averse investors. *Journal of Economic Dynamics and Control* 76, 86–108.
- Guo, J., He, X., 2017b. A New Preference Model that Allows for Narrow Framing. Working paper.
- Haghani, V., McBride, S., 2016. Return Chasing and Trend Following: Superficial Similarities Mask Fundamental Differences. Working paper.
- Han, B., Hirshleifer, D., Walden, J., 2018. Social Transmission Bias and Investor Behavior. Working paper.
- Hansen, L., Sargent, T., 2001. Robust control and model uncertainty. *American Economic Review* 91, 60–66.
- Hardouvelis, G., La Porta, R., Wizman, T., 1994. What moves the discount on country equity funds? In: Frankel, J. (Ed.), *The Internationalization of Equity Markets*. University of Chicago Press, Chicago, IL.
- Harris, L., Gurel, E., 1986. Price and volume effects associated with changes in the S&P 500 list: new evidence for the existence of price pressures. *Journal of Finance* 41, 815–829.
- Harrison, J.M., Kreps, D., 1978. Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics* 92, 323–336.
- Hartzmark, S., Shue, K., 2017. A tough act to follow: contrast effects in financial markets. *Journal of Finance*. <https://doi.org/10.1111/jofi.12685>. Forthcoming.
- Hartzmark, S., Solomon, D., 2018. The Dividend Disconnect. Working paper.
- Harvey, C., Liu, Y., Zhu, H., 2016. ... and the cross-section of expected returns. *Review of Financial Studies* 29, 5–68.
- He, X., Zhou, X.Y., 2011. Portfolio choice under cumulative prospect theory: an analytical treatment. *Management Science* 57, 315–331.
- Heath, C., Tversky, A., 1991. Preference and belief: ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty* 4, 5–28.
- Hertwig, R., Barron, G., Weber, E., Erev, I., 2004. Decisions from experience and the effect of rare events in risky choice. *Psychological Science* 15, 534–539.
- Hirshleifer, D., 2001. Investor psychology and asset pricing. *Journal of Finance* 56, 1533–1597.
- Hirshleifer, D., Shumway, T., 2003. Good day sunshine: stock returns and the weather. *Journal of Finance* 58, 1009–1032.
- Hirshleifer, D., Li, J., Yu, J., 2015. Asset pricing in production economies with extrapolative expectations. *Journal of Monetary Economics* 76, 87–106.
- Hirshleifer, D., Lim, S., Teoh, S.H., 2009. Driven to distraction: extraneous events and underreaction to earnings news. *Journal of Finance* 64, 2289–2325.

- Hong, H., Sraer, D., 2013. Quiet bubbles. *Journal of Financial Economics* 110, 596–606.
- Hong, H., Sraer, D., 2016. Speculative betas. *Journal of Finance* 71, 2095–2144.
- Hong, H., Stein, J., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance* 54, 2143–2184.
- Hong, H., Stein, J., 2007. Disagreement and the stock market. *Journal of Economic Perspectives* 21, 109–128.
- Hong, H., Kubik, J., Stein, J., 2004. Social interaction and stock market participation. *Journal of Finance* 59, 137–163.
- Hong, H., Kubik, J., Stein, J., 2005. Thy neighbor's portfolio: word-of-mouth effects in the holdings and trades of money managers. *Journal of Finance* 60, 2801–2824.
- Hong, H., Lim, T., Stein, J., 2000. Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55, 265–295.
- Huang, X., 2015. Thinking outside the borders: investors' underreaction to foreign operations information. *Review of Financial Studies* 28, 3109–3152.
- Huberman, G., 2001. Familiarity breeds investment. *Review of Financial Studies* 14, 659–680.
- Ikenberry, D., Lakonishok, J., Vermaelen, T., 1995. Market underreaction to open market share repurchases. *Journal of Financial Economics* 39, 181–208.
- Ilmanen, A., 2011. *Expected Returns: An Investor's Guide to Harvesting Market Rewards*. John Wiley & Sons, New York.
- Imas, A., 2016. The realization effect: risk-taking after realized versus paper losses. *American Economic Review* 106, 2086–2109.
- Ingersoll, J., Jin, L., 2013. Realization utility with reference-dependent preferences. *Review of Financial Studies* 26, 723–767.
- Ivkovic, Z., Weisbenner, S., 2005. Local does as local is: information content of the geography of individual investors' common stock investments. *Journal of Finance* 60, 267–306.
- Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance* 48, 65–91.
- Jin, L., 2015. *A Speculative Asset Pricing Model of Financial Instability*. Working paper.
- Jin, L., Sui, P., 2018. *Asset Pricing with Return Extrapolation*. Working paper.
- Johnson, E., Tversky, A., 1983. Affect, generalization, and the perception of risk. *Journal of Personality and Social Psychology* 45, 20–31.
- Kahneman, D., Tversky, A., 1972. Subjective probability: a judgment of representativeness. *Cognitive Psychology* 3, 430–454.
- Kahneman, D., Tversky, A., 1973a. On the psychology of prediction. *Psychological Review* 80, 237–251.
- Kahneman, D., Tversky, A., 1973b. Availability: a heuristic for judging frequency and probability. *Cognitive Psychology* 5, 207–232.
- Kahneman, D., Tversky, A., 1974. Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–291.
- Kamstra, M., Kramer, L., Levi, M., 2003. Winter blues: a SAD stock market cycle. *American Economic Review* 93, 324–343.
- Kindleberger, C., 1978. *Manias, Panics, and Crashes: A History of Financial Crises*. Basic Books.
- Klayman, J., Soll, J., Gonzalez-Vallejo, C., Barlas, S., 1999. Overconfidence: it depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes* 79, 216–247.
- Klibanoff, P., Marinacci, M., Mukherji, S., 2005. A smooth model of decision making under ambiguity. *Econometrica* 73, 1849–1892.
- Koszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. *Quarterly Journal of Economics* 121, 1133–1165.

- Koszegi, B., Rabin, M., 2007. Reference-dependent risk attitudes. *American Economic Review* 97, 1047–1073.
- Krishnamurthy, A., 2010. Amplification mechanisms in liquidity crises. *American Economic Journal: Macroeconomics* 2, 1–30.
- Kuchler, T., Zafar, B., 2017. Personal Experiences and Expectations About Aggregate Outcomes. Working paper.
- Kumar, A., 2009. Who gambles in the stock market? *Journal of Finance* 64, 1889–1933.
- La Porta, R., Lakonishok, J., Shleifer, A., Vishny, R., 1997. Good news for value stocks: further evidence on market inefficiency. *Journal of Finance* 52, 859–874.
- Lakonishok, J., Shleifer, A., Vishny, R., 1994. Contrarian investment, extrapolation, and risk. *Journal of Finance* 49, 1541–1578.
- Lamont, O., Thaler, R., 2003. Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy* 111, 227–268.
- Landier, A., Ma, Y., Thesmar, D., 2017. New Experimental Evidence on Expectations Formation. Working paper.
- Lee, C., Shleifer, A., Thaler, R., 1991. Investor sentiment and the closed-end fund puzzle. *Journal of Finance* 46, 75–109.
- Lehmann, B., 1990. Fads, martingales, and market efficiency. *Quarterly Journal of Economics* 105, 1–28.
- LeRoy, S., Porter, R., 1981. The present-value relation: tests based on implied variance bounds. *Econometrica* 49, 97–113.
- Levy, H., De Giorgi, E., Hens, T., 2011. Two paradigms and Nobel Prizes in economics: a contradiction or coexistence? *European Financial Management* 18, 163–182.
- Li, Y., Yang, L., 2013. Prospect theory, the disposition effect, and asset prices. *Journal of Financial Economics* 107, 715–739.
- Lian, C., Ma, Y., Wang, C., 2018. Low Interest Rates and Risk Taking: Evidence from Individual Investment Decisions. Working paper.
- Liao, J., Peng, C., 2018. Price and Volume Dynamics in Bubbles. Working paper.
- Loughran, T., Ritter, J., 1995. The new issues puzzle. *Journal of Finance* 50, 23–50.
- Lundeberg, M., Fox, P., Punccohar, J., 1994. Highly confident but wrong: gender differences and similarities in confidence judgments. *Journal of Educational Psychology* 86, 114–121.
- Maenhout, P., 2004. Robust portfolio rules and asset pricing. *Review of Financial Studies* 17, 951–983.
- Malmendier, U., 2018. Behavioral corporate finance. In: Bernheim, D., DellaVigna, S., Laibson, D. (Eds.), *Handbook of Behavioral Economics*, vol. 1. North-Holland, Amsterdam.
- Malmendier, U., Nagel, S., 2011. Depression babies: do macroeconomic experiences affect risk-taking? *Quarterly Journal of Economics* 126, 373–416.
- Malmendier, U., Pouzo, D., Vanasco, V., 2017. Investor Experiences and Financial Market Dynamics. Working paper.
- Markowitz, H., 1952. Portfolio selection. *Journal of Finance* 7, 77–91.
- Marks, H., 2011. *The Most Important Thing: Uncommon Sense for the Thoughtful Investor*. Columbia University Press.
- McLean, R.D., Pontiff, J., 2016. Does academic research destroy return predictability? *Journal of Finance* 71, 5–32.
- Mehra, R., 2008. The equity premium puzzle: a review. *Foundations and Trends in Finance* 2, 1–81.
- Mehra, R., Prescott, E., 1985. The equity premium: a puzzle. *Journal of Monetary Economics* 15, 145–161.
- Meng, J., Weng, X., 2017. Can prospect theory explain the disposition effect? A new perspective on reference points. *Management Science*. <https://doi.org/10.1287/mnsc.2016.2711>. Forthcoming.
- Michenaud, S., Solnik, B., 2008. Applying regret theory to investment choices: currency hedging decisions. *Journal of International Money and Finance* 27, 677–694.

- Milgrom, P., Stokey, N., 1982. Information, trade, and common knowledge. *Journal of Economic Theory* 26, 17–27.
- Miller, E., 1977. Risk, uncertainty, and divergence of opinion. *Journal of Finance* 32, 1151–1168.
- Mitchell, M., Pulvino, T., Stafford, E., 2002. Limited arbitrage in equity markets. *Journal of Finance* 57, 551–584.
- Mitton, T., Vorkink, K., 2007. Equilibrium under-diversification and the preference for skewness. *Review of Financial Studies* 20, 1255–1288.
- Mitton, T., Vorkink, K., 2010. Why do firms with diversification discounts have high expected returns? *Journal of Financial and Quantitative Analysis* 45, 1367–1390.
- Modigliani, F., Cohn, R., 1979. Inflation, rational valuation, and the market. *Financial Analysts Journal* 35, 24–44.
- Modigliani, F., Miller, M., 1958. The cost of capital, corporation finance, and the theory of investment. *American Economic Review* 48, 261–297.
- Moore, D., Healy, P., 2008. The trouble with overconfidence. *Psychological Review* 115, 502–517.
- Morris, S., 1996. Speculative investor behavior and learning. *Quarterly Journal of Economics* 111, 1111–1133.
- Nagel, S., Xu, Z., 2018. Asset Pricing with Fading Memory. Working paper.
- Novy-Marx, R., 2013. The other side of value: the gross profitability premium. *Journal of Financial Economics* 108, 1–28.
- Novy-Marx, R., Velikov, M., 2015. A taxonomy of anomalies and their trading costs. *Review of Financial Studies* 29, 104–107.
- O'Donoghue, T., Sprenger, C., 2018. Reference-dependent preferences. In: Bernheim, D., DellaVigna, S., Laibson, D. (Eds.), *Handbook of Behavioral Economics*, vol. 1. North-Holland, Amsterdam.
- Odean, T., 1998a. Volume, volatility, price, and profit when all traders are above average. *Journal of Finance* 53, 1887–1934.
- Odean, T., 1998b. Are investors reluctant to realize their losses? *Journal of Finance* 53, 1775–1798.
- Pagel, M., 2016. Expectations-based reference-dependent preferences and asset pricing. *Journal of the European Economic Association* 14, 468–514.
- Pastor, L., Veronesi, P., 2009. Learning in financial markets. *Annual Review of Financial Economics* 1, 361–381.
- Pindyck, R., Rotemberg, J., 1990. The excess comovement of commodity prices. *Economic Journal* 100, 1173–1189.
- Polkovnichenko, V., 2005. Household portfolio diversification: a case for rank-dependent preferences. *Review of Financial Studies* 18, 1467–1502.
- Pouget, S., Sauvagnat, J., Villeneuve, S., 2017. A mind is a terrible thing to change: confirmatory bias in financial markets. *Review of Financial Studies* 30, 2066–2109.
- Quiggin, J., 1982. A theory of anticipated utility. *Journal of Economic Behavior and Organization* 3, 323–343.
- Rabin, M., 2000. Risk aversion and expected utility theory: a calibration theorem. *Econometrica* 68, 1281–1292.
- Rabin, M., 2002. Inference by believers in the law of small numbers. *Quarterly Journal of Economics* 117, 775–816.
- Rabin, M., Schrag, J., 1999. First impressions matter: a model of confirmatory bias. *Quarterly Journal of Economics* 114, 37–82.
- Rabin, M., Vayanos, D., 2010. The gambler's and hot hand fallacies. *Review of Economic Studies* 77, 730–778.
- Rosch, E., Lloyd, B., 1978. *Cognition and Categorization*. Lawrence Erlbaum Associates, New Jersey.
- Rosenberg, B., Reid, K., Lanstein, R., 1985. Persuasive evidence of market inefficiency. *Journal of Portfolio Management* 11, 9–17.

- Routledge, B., Zin, S., 2010. Generalized disappointment aversion and asset prices. *Journal of Finance* 65, 4–32.
- Scheinkman, J., Xiong, W., 2003. Overconfidence and speculative bubbles. *Journal of Political Economy* 111, 1183–1219.
- Seasholes, M., Zhu, N., 2010. Individual investors and local bias. *Journal of Finance* 65, 1987–2010.
- Shefrin, H., Statman, M., 1985. The disposition to sell winners too early and ride losers too long. *Journal of Finance* 40, 777–790.
- Shiller, R., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 71, 421–436.
- Shiller, R., 1984. Stock prices and social dynamics. *Brookings Papers on Economic Activity* 1984, 457–510.
- Shleifer, A., 1986. Do demand curves for stocks slope down? *Journal of Finance* 41, 579–590.
- Shleifer, A., 2000. *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford University Press, Oxford.
- Shleifer, A., Summers, L., 1990. The noise trader approach to finance. *Journal of Economic Perspectives* 4, 19–33.
- Shleifer, A., Vishny, R., 1997. The limits of arbitrage. *Journal of Finance* 52, 35–55.
- Shue, K., Townsend, R., 2018. Can the Market Multiply and Divide? Non-proportional Thinking in Financial Markets. Working paper.
- Sloan, R., 1996. Do stock prices reflect information in accruals and cash flows about future earnings? *Accounting Review* 71, 289–315.
- Smith, V., Suchanek, G., Williams, A., 1988. Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica* 56, 1119–1151.
- Soll, J., Klayman, J., 2004. Overconfidence in interval estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30, 299–314.
- Stambaugh, R., Yu, J., Yuan, Y., 2012. The short of it: investor sentiment and anomalies. *Journal of Financial Economics* 104, 288–302.
- Stambaugh, R., Yu, J., Yuan, Y., 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. *Journal of Finance* 70, 1903–1948.
- Tetlock, P., 2007. Giving content to investor sentiment: the role of media in the stock market. *Journal of Finance* 62, 1139–1168.
- Tetlock, P., 2015. The role of media in finance. In: Anderson, S.P., Stromberg, D., Waldfogel, J. (Eds.), *Handbook of Media Economics*, vol. 1b. North-Holland, Amsterdam.
- Thaler, R., 1999. Mental accounting matters. In: Kahneman, D., Tversky, A. (Eds.), *Choice, Values and Frames*. Russell Sage Foundation, Cambridge.
- Thaler, R., Johnson, E., 1990. Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. *Management Science* 36, 643–660.
- Thaler, R., Tversky, A., Kahneman, D., Schwartz, A., 1997. The effect of myopia and loss aversion on risk-taking: an experimental test. *Quarterly Journal of Economics* 112, 647–661.
- Timmermann, A., 1993. How learning in financial markets generates excess volatility and predictability in stock prices. *Quarterly Journal of Economics* 108, 1135–1145.
- Tirole, J., 1985. Asset bubbles and overlapping generations. *Econometrica* 53, 1071–1100.
- Titman, S., Wei, K.C., Xie, F., 2004. Capital investments and stock returns. *Journal of Financial and Quantitative Analysis* 39, 677–700.
- Tversky, A., Kahneman, D., 1971. Belief in the law of small numbers. *Psychological Bulletin* 76, 105–110.
- Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science* 211, 453–458.
- Tversky, A., Kahneman, D., 1992. Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297–323.
- Uppal, R., Wang, T., 2003. Model misspecification and underdiversification. *Journal of Finance* 58, 2465–2486.

- Wachter, J., 2013. Can time-varying risk of rare disasters explain aggregate stock market volatility? *Journal of Finance* 68, 987–1035.
- Wang, B., 2016. Probability Weighting and Asset Prices: Evidence from Mergers and Acquisitions. Working paper.
- Wang, H., Yan, J., Yu, J., 2017. Reference-dependent preferences and the risk–return tradeoff. *Journal of Financial Economics* 123, 395–414.
- Xiong, W., 2013. Bubbles, crises, and heterogeneous beliefs. In: Fouque, J-P., Langsam, J. (Eds.), *Handbook for Systemic Risk*. Cambridge University Press.
- Xiong, W., Yu, J., 2011. The Chinese warrants bubble. *American Economic Review* 101, 2723–2753.
- Yaari, M., 1987. The dual theory of choice under risk. *Econometrica* 55, 95–115.
- Yan, H., 2008. Natural selection in financial markets: does it work? *Management Science* 54, 1935–1950.
- Zajonc, R., 1968. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology* 9, 1–27.