Junior Management Science 5(3) (2020) 262-294

Junior Management Science

journal homepage: www.jums.academy

Prospect Theory and Stock Returns During Bubbles

Maximilian Piehler

Ludwig-Maximilians-Universität München

Abstract

I test the hypothesis that investors evaluate stocks based on the prospect theory value of the distribution of past returns. Because some investors tilt towards stocks with high prospect theory value, these stocks become overvalued and earn low subsequent returns. During bubbles this effect should be stronger, due to rising limits to arbitrage and increased participation of individual investors. I do not find strong support for this prediction in the cross section of returns in U.S. stock markets. In contrast to other variables know to explain returns however, prospect theory value does not lose its predictive power during bubbles. Investors with prospect theory preferences seem to choose stocks whose returns optimally combine low standard deviation with high skewness.

Keywords: Prospect Theory; bubbles; limits to arbitrage; individual investors.

1. Introduction

1.1. Motivation

The decision whether to buy or sell a stock is risky. In order to understand the buying and selling behaviour in the stock market it is therefore important to investigate how investors make risky choices. In other words, how do investors evaluate a given stock and how can this be integrated into pricing models? In the past century, many researchers modelled investor's choices via expected utility. They assumed that investors maximize their expected utility, which is the sum of the utility of each outcome weighted by its probability. It relies on certain assumptions formalized by Neumann (1944). Expected utility has been a useful tool for research in investor behaviour and asset pricing. Nonetheless, the pricing models based on its assumptions do not explain stock prices with complete precision. Several anomalies seem to persist in the stock market, despite multi-factor pricing models. For example, Jegadeesh and Titman (1993) find that stocks that performed well during the past year subsequently outperform stock that did poorly during the past year. This effect is generally called momentum. De Bondt and Thaler (1985) report on long-term reversal: a stock's returns over the last two to five years on average negatively predict the subsequent return. Similarly, there is a short-term reversal in stock returns i.e. the return of the last month negatively predicts the following return, discovered by Lehmann (1990). More recently, Ang et al. (2006) find that the volatility of idiosyncratic returns has a negative effect on returns, while Novy-Marx (2013) shows how measures of firm profitability have a positive effect on returns, with a predictive power comparable to that of the book-to-market ratio. In light of these return patterns which are hard to make sense of using classical measures of risk and return, some researchers have focused on behavioural explanations. Maybe it is wrong to assume that investors make rational investment choices in the stock market. Kahneman and Tversky (1979) and Tversky and Kahnemann (1992) offer an alternative model of preferences and choice under risk: prospect theory. They specifically develop prospect theory to address violations of the underlying axioms of expected utility. Unlike expected utility, prospect theory is a descriptive model built on experimental evidence. In order to test the accuracy of this theory, it is necessary to apply it in real world settings. Do investors in the stock market behave according to prospect theory? Barberis et al. (2016) derive testable predictions of prospect theory for average returns in the stock market. I closely follow their work and first test their prediction in my sample. However, I argue that the magnitude of the predicted effect is time-varying; it should be different for bubble and non-bubble periods. But first I want to turn to prospect theory and the original prediction made by Barberis et al. (2016).

1.2. Approach

Prospect theory has not been tested in the stock market very often, maybe because there are some specifications that

DOI: https://doi.org/10.5282/jums/v5i3pp262-294

are hard to translate into empirical work. These specification issues come down to seemingly simple questions that arise when thinking about the choice making process in prospect theory. Generally, individuals evaluate a risky choice based on a certain set of outcomes and associated probabilities. Before getting to the evaluation though, individuals first need to form a representation of the risky choice they are facing. This representation should be formulated in terms of gains and losses, since prospect theory is defined over gains and losses and not absolute wealth. In contrast to the valuation process of a given set of gains and losses, there is no formal description of the representation process in prospect theory. In a short review I therefore give an overview of the research concerning two related questions: (1) What is the reference point investors use to measure gains and losses?, and (2) Where do investors get the information on returns and how could the presentation form influence their representation of stock returns?. A debate in the research on reference points is whether individuals look to the status quo, where the reference point is initial wealth, or the recently held expectations about an outcome to determine gains and losses. Although a reference point based on expectations might make more intuitive sense, I find research supporting both views. Regarding the sources of information and the impact of presentation form on the perceived return, it becomes clear that individuals could form representations that do not reflect the actually observed information. Nolte and Schneider (2018) for example find that the shape of price paths influences perceived attractiveness of an investment, even keeping the underlying return distribution equal. Accounting for these phenomena and integrating them into an empirical framework indeed is an issue when applying prospect theory. Once the process of representation is finished, Tversky and Kahnemann (1992) provide detailed formulae and parameter estimates to calculate the associated utility i.e. the prospect theory value.

My first assumption is that investors evaluate the attractiveness of a stock based on its returns in the past. From theoretical point of view looking at past returns is not ideal, since a stock's value should depend on future cash flows of the underlying firm, not to mention the anomalies where past returns negatively predict future returns. But in practice many investors might confront issues concerning the availability of resources, be it in the form of time, knowledge or data. Taking aside professional investors, which are not the investor group central to the analysis, the assumption of past return distributions as basis for the subsequent evaluation seems reasonable. Therefore I compute the prospect theory value for each stock using its past return distribution.

For the development of predictions I rely on the model of asset prices by Barberis et al. (2016). In their model, two types of investors populate the market. Rational investors construct the tangency portfolio that is based on the mean-variance framework of Markowitz (1952). Individual investors have a preference for stocks with high prospect theory value and do not like stocks with low prospect theory value. They therefore adjust the tangency portfolio positively (negatively) for high (low) prospect theory value stocks. This setup results in the model's prediction that stocks with high prospect theory value will earn low subsequent returns. By introducing a small extension to this model, I make a second prediction that the subsequent returns on stocks with high prospect theory value will be lower during bubble periods compared to non-bubble periods. The argument for the stronger effect during bubbles is twofold.

The first part of the argument is that the limits to arbitrage for rational investors should rise, conceding individual investors more influence on prices. There are several papers supporting this view. The most important concept is delayed arbitrage, introduced by Abreu and Brunnermeier (2003). They show that rational investors i.e. arbitrageurs do not necessarily have interest in correcting mispricing during bubbles, because they try to "ride" the bubble and exit right before it collapses. In addition, Scheinkman and Xiong (2003) find that overconfident investors with heterogeneous beliefs about fundamental value can buy stocks exceeding their own valuation. Finally, a combination of these heterogeneous beliefs about fundamental value and the workings of the market for borrowing stock reported by D'Avolio (2002) lead to higher short selling costs during bubbles. The second part of the argument is that the fraction of individual investors increases during bubbles, which leads to a stronger overvaluation of high prospect theory value stocks and lower subsequent returns. Barber and Odean (2008) find that individual investor often buy stocks that grab their attention. Therefore the media coverage of a bubble or stock market rally could incentivise new irrational investor to buy the set of stocks that drive the bubble.

In order to test the original prediction by Barberis et al. (2016) and my own prediction it is necessary to define the exact format of past return distribution and what constitutes a bubble in stock markets. Coming back to the debate over reference points for gains and losses, it is best to test three different benchmarks to measure past returns against: raw returns and returns in excess of both the risk-free rate and the market. While raw returns relate to initial wealth, market returns could be a proxy for investor's return expectations. The information source should determine the frequency and the look back window for the returns in the distribution. The sample I use could be loosely split in half pre-internet and half internet era. In the first half it is reasonable to adopt the convention that investors use investment booklets as the main information source[1]. These booklets often show monthly price charts with a lookback period of three to five years. Past return distribution therefore refers to monthly returns over the past five years in my framework. The important change in the internet era is the sheer amount of information available, although price charts still seem to be prominent[2]. Hence I decide not to switch the definition of return distributions, although it is an interesting thought. Regarding the definition of bubbles I rely on a "rational bubble" model and an econometric method developed in a series of papers by Phillips et al.

[1]See Alexander et al., 1998, Table 3, p.307.
[2]See Glaser et al., 2019, Appendix A Table 1, p.7-8.

(2015a,b) and Phillips et al. (2011). They use right-tailed unit root tests to infer explosive price behaviour by working with the stochastic processes of the underlying dividends and fundamentals.

The empirical analysis begins with the detection and time stamping of bubbles in the S&P 500, which is assumed to be a good proxy for U.S. stock markets in general. I detect the existence of a bubble and the time stamping process results in two phases of explosive prices, from February to August 1987 and from December 1995 to September 2000. These time periods coincide with the bubble leading up to Black Monday and the tech bubble at the end of the last century. The main empirical test consists of Fama and MacBeth (1973, just Fama-MacBeth hereafter) regressions, in order to allow good comparability of my results with the analysis of Barberis et al. (2016). First, the prospect theory value is computed for every month using a stock's returns in excess of the market over the past five years. Then the prospect theory value, together with control variables known to predict returns, serves as the independent variable in a monthly regression on returns. The result of the Fama-MacBeth test is the time-series average of these cross-sectional regressions. I find support for the original prediction that stocks with high prospect theory value earn low returns in U.S. stock markets. In order to test whether this effect is different during bubbles, the sample is split into two sub periods. As a result I cannot confirm that high prospect theory value stocks earn lower returns during bubbles compared to non-bubble periods. Depending on the number of control variables used, the prospect theory value has higher predictive power for returns during bubbles. But the most sophisticated model shows that the effect of the prospect theory value is very similar to non-bubble periods. Portfolio sorts on the prospect theory value point into the same direction. On the surface, the low minus high prospect theory value portfolio earns higher abnormal returns during bubbles. Conducting robustness tests shows that the returns to this strategy remain high even when using different returns, return distribution construction windows, subsamples or excluding low price stocks. However, the returns to the low minus high portfolio are less robust during bubbles, because the returns in the high prospect theory value deciles are not significantly different from zero.

To test the plausibility of the argument that limits to arbitrage are related to the relationship between returns and prospect theory value, I analyse the interaction of prospect theory value with several proxies for the limits to arbitrage. These tests show that the negative effect of prospect theory value on returns is generally stronger for stocks experiencing higher limits to arbitrage. On the other hand they show that rising limits to arbitrage have no impact on the predictive power of the prospect theory value during bubbles. This is at least true for the different proxies for the limits to arbitrage I employ.

In the later part of the paper I take a closer look into the mechanism behind the predictive power of the prospect theory value for returns. Each Fama-MacBeth regression turns off two components of prospect theory by setting their parameters to one. In addition, the sample is split into bubble and non-bubble periods once again. Loss aversion is found to be the most significant predictor of returns by itself. This indicates that investors with prospect theory preferences should prefer stocks with low standard deviation. In a relevance hierarchy probability weighting comes in second. The prospect theory variable constructed with probability weighting only is not a significant predictor of returns, but in conjunction with loss aversion it contributes positively to the significance. Since probability weighting seems to play at least some role, investors under prospect theory should also prefer skewness. Sorting stocks into deciles based on prospect theory value provides support for these findings. The results indicate that investors with prospect theory preferences like stocks that optimally trade of low standard deviation and high skewness.

This study mainly contributes to the work of Barberis et al. (2016) and shows that the relationship between the prospect theory value and returns is strong and predicts returns almost equally well during bubbles and non-bubble periods. The analysis of the power of known return predictors also shows that behavioural explanations, like prospect theory, can be valuable in research on asset pricing. In addition, my paper provides further evidence on the reliability of the tests for explosive price behaviour developed by Phillips et al. (2015a). The paper is organized as follows. In chapter 2 prospect theory and the construction of return distributions is presented. This entails the presentation and discussion of each component relevant to prospect theory. Subsequently chapter 3 introduces the model of asset price by Barberis et al. (2016). There I develop a small extension to the model and form the hypothesis. Chapter 4 presents the evidence backing the hypothesis, together with an overview of the research on the limits to arbitrage. Chapter 5 consists of a discussion of the history of bubbles and different measurement models of explosive behaviour in prices. In addition, it presents the model for detection and time stamping of bubbles by Phillips et al. (2015a) leading up to the application of the method to the time series of the S&P 500. In chapter 6, I analyse the results from several empirical tests designed to provide evidence for my hypothesis. This part of the paper starts off with the data description and summary statistics and then continues over to portfolio sorts and the Fama-MacBeth regressions designed to be the central empirical test. In later parts of chapter 6 I look more closely into the limits to arbitrage argument and the mechanism behind prospect theory. In the final chapter of the paper I summarize the results and main conclusion and give a direction for future research in related areas.

2. Prospect theory

2.1. Representation and valuation

In the framework of this study, investors evaluate stocks in a two-step process. The first step is the preliminary analysis of the presented stock, in which some representation of outcomes is formed. How do investors mentally represent a stock? Which sources do they use as a basis for the representation? In order to answer these questions in the context of prospect theory and the stock market, it is important to define what gains and losses are in the stock market, while also keeping in mind the information sources used by investors. I therefore present a short overview of the related research in section 2.3. But first it is useful to take a closer look at the second step, valuation, during which investors evaluate the representation of the stock they formed earlier. For this part of the process, Tversky and Kahnemann (1992) provide a formal framework and parameter estimates.

2.2. Valuation

2.2.1. Valuation methodology

Based on a critique of the expected utility framework, Kahneman and Tversky (1979) develop a descriptive model of decision making under risk. They analyse the preferred choices of individuals confronted with a set of outcomes and associated probabilities. Because these "prospects" are restricted to two non-zero outcomes, the applicability of the original version of prospect theory is limited. Further, individuals sometimes choose dominated prospects. Tversky and Kahnemann (1992) address these problems in their cumulative formulation of prospect theory. I first present the functional form of this theory and then discuss the components and their implications for decision making under risk.

Let X be a set of uncertain outcomes x, where all outcomes except one neutral outcome are gains or losses. The function that assigns each uncertain outcome a probability p is the prospect to be considered by the individual:

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

The subscripts $-m$ and n stand for the most negative and most positive outcome, respectively. Before individuals determine the value of a prospect they transform probabilities into decision weights:

$$\pi_i = \begin{cases} w(p_i + \cdots + p_n) - w(p_{i+1} + \cdots + p_n) \\ w(p_{-m} + \cdots + p_i) - w(p_{-m} + \cdots + p_{i-1}) \end{cases} \text{for} \quad \begin{matrix} 0 \le i \le n \\ -m \le i < 0 \end{matrix}, \quad (2)$$

according to the weighting functions

$$w^+(p) = p^\gamma / (p^\gamma + (1-p)^\gamma)^{1/\gamma}, w^-(p) = p^\delta / \left(p^\delta + (1-p)^\delta\right)^{1/\delta}. \quad (3)$$

As indicated by the parameters γ and δ the weighting function is different for gains and losses. The same is true for the value function:

$$v(x) = \begin{cases} x^\alpha \text{ if } x \ge 0 \\ -\lambda(-x)^\alpha \text{ if } x < 0 \end{cases}, \quad (4)$$

An individual then assigns the prospect in (1) the following value:

$$\sum_{i=-m}^{n} \pi_i v(x_i). \quad (5)$$

Put into words, an individual analyses each pair of outcome and probability according to (2)-(4) and values the prospect as a sum of these parts. I now turn to discuss essential components of prospect theory, namely probability weighting, reference dependence, diminishing sensitivity and loss aversion.

2.2.2. Components of prospect theory and related concepts

Probability weighting

Probability weighting refers to the transformation of probabilities into decision weights. These weights are a representation of the perceived impact an outcome has on the evaluation of the prospect. As a consequence decision weights do not equal the objective probability of an outcome. An interesting case in probability weighting is individuals' preference for lotteries and insurance. Kahneman and Tversky (1979) describe two gambles that illustrate these preferences. In the first gamble, the options are a gain of 5000 with a 0.001 chance or a gain of 5 with certainty. In the second gamble, the options are a loss of 5000 with a 0.001 chance or a loss of 5 with certainty. Notice that the expected value of the options in both gambles is identical. The majority of respondents prefer the very unlikely gain of 5000 in the first gamble, while choosing the certain loss of 5 in the second gamble. These choices show that individuals overweight very small probabilities, causing them to prefer lottery-like outcomes in the domain of gains and insurance in the domain of losses. The weighting function that transforms probabilities into decision weights in prospect theory accounts for this by its inverse s-shaped slope. Figure 1 plots the weighting functions for γ=0.61 and δ=0.69, as estimated by Tversky and Kahnemann (1992). In the original version the probabilities used as input to the weighting function are exact, individual probabilities. As Quiggin (1982) shows, the application of decision weights to exact probabilities can lead to violations of dominance. He suggests using the entire cumulative distribution of probabilities instead. In cumulative prospect theory, the formulation of decision weights is in accordance with this approach. For any positive outcome, the decision weight is the difference between the probabilities of "the outcome is at least as good as x_i" and "the outcome is strictly better than x_i", the weighting function applied to both. For any negative outcome, the decision weight is the difference between the probabilities derived from "the outcome is at least as bad as x_i" and "the outcome is strictly worse than x_i", again, the weighting function applied to both. These types of decision weights capture the contribution of a single outcome in context of the entire probability distribution. Consequently the overweighting of small probabilities carries over into an overweighting of the tails of any distribution in cumulative prospect theory.

Reference dependence

Reference dependence implies that individuals evaluate the attractiveness of an outcome based on gains and losses in relation to a reference point. In expected utility there is no

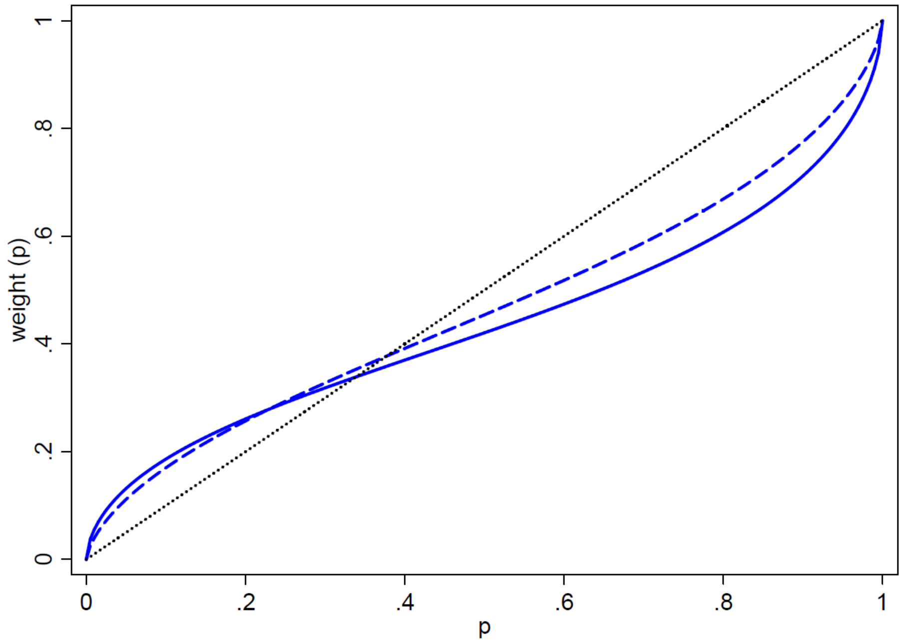

Figure 1: The weighting function (Based on Tversky and Kahnemann (1992), p.313).

Graphic illustration of the probability weighting function in prospect theory for γ=0.61 (solid line) and δ=0.69 (dashed line).

difference to whether the outcome of a gamble represents a relative increase or decrease in wealth to the individual. Only the final wealth is considered. Intuitively though, the same final wealth can feel different depending on how it was reached, through a gain or a loss. In accordance with this rationale Kahneman and Tversky (1979) find experimental evidence that individuals' choices are dependent on their reference point. It is not clear though, what constitutes such a point. Köszegi, B. and Rabin, M. (2006) for example assume that the reference point is the individual's recently held expectations about an outcome.

Diminishing sensitivity

The principle of diminishing sensitivity determines the shape of the value function. The value function in prospect theory is concave for gains and convex for losses, as seen in figure 2. This shape is inspired by the proposition that the marginal value of gains and losses decrease with their magnitude. Kahneman and Tversky (1979) motivate this proposition through following problems posed to individuals in their experiments. The options in the first problem are a gain of 6000 with probability 0.25 or gains of 4000 and 2000, each with probability 0.25. The majority of respondents choose option two, implying that the value of the separate gains is larger than the value of the single gain, although both options offer a reward of 6000 in total. Since apparently $v(6000) < v(4000) + v(2000)$, the marginal value of smaller gains must be larger and the function for gains convex. The second problem has the same structure, only in the domain of losses, and in turn supports the convexity of the value function for losses. Tversky and Kahnemann (1992) estimate a median value of 0.88 for α in equation (4), in accordance with moderate levels of diminishing sensitivity.

Loss aversion

Loss aversion is the tendency of individuals to perceive losses stronger than gains of the same size. Therefore the value function for losses is steeper than for gains. This produces a kink at the origin. It illustrates how people are very reluctant to accept gambles with equal probabilities and exactly opposite outcomes. For example, a coin toss with a payment of 50 or -50 seems unattractive, because the individual is much more sensitive to the potential loss. In the value function in (4) the parameter λ determines the degree of loss aversion. Tversky and Kahnemann (1992) estimate a median value of 2.25 for λ, finding evidence of strong loss aversion in individuals' decisions under risk. Loss aversion is a widely accepted concept not only in neuroscience research[3], but also specifically in the finance literature. Benartzi and Thaler (1995), for example, use investor's loss aversion as a part of their explanation of the equity premium puzzle.

Narrow framing

One further concept in the analysis of choices under risk, not originally integrated in prospect theory, is sometimes applied together with prospect theory: narrow framing. Tversky and Kahnemann (1981) first describe the effect of "decision framing" on the preferences of test subjects. Decision framing refers to how individuals conceive a prospect i.e. how they perceive the choices they are presented with. The prefer-

[3]See Tom et al., 2007, for a study on the neural basis of loss aversion.

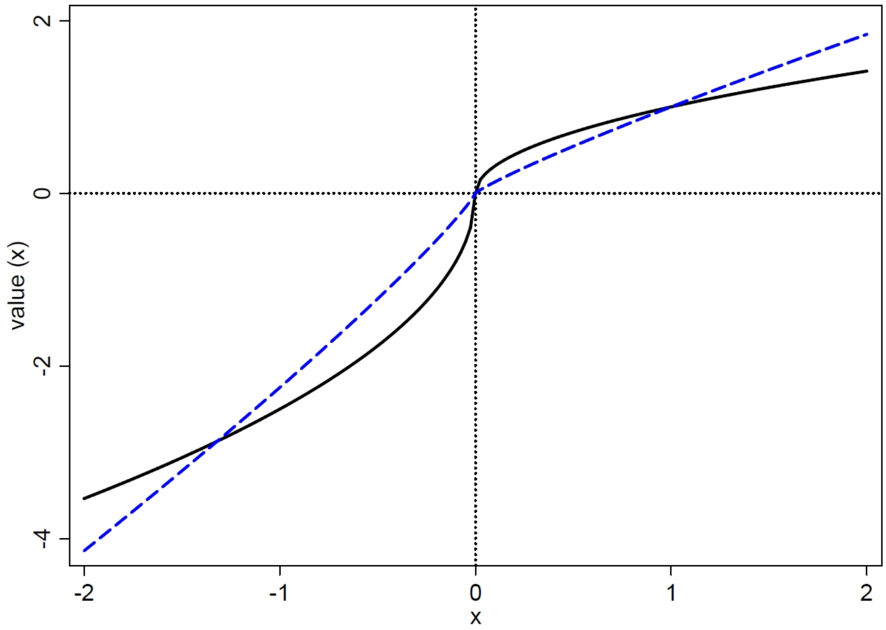

Figure 2: The value function (Based on Tversky and Kahnemann (1992), p.311).

Graphic illustration of the value function in Prospect Theory for α=0.5, λ=2.5 (solid line) and α=0.88, λ=2.25 (dashed line).The blue dashed line is the parameter specification used in the main analysis of this paper.

ences observed depend on the framing of the choice. Narrow framing is a specific way of framing choices: Each choice is viewed separately from other choices that might have to be made. This means that the current choice is isolated from other choices. Rabin, M. and Weizsäcker, G. (2009) replicate the experimental results of Tversky and Kahnemann (1981) and find in a survey that 89 percent of respondents engaged in narrow framing. The concept of narrow framing is also applicable in the stock market. Some investors might not track the performance of their overall portfolio, but rather the gains and losses of individual stock in the portfolio. Kumar and Lim (2008) use trade clustering as a proxy for narrow framing and show that investors engaging in narrow framing exhibit stronger disposition effects[4]. If investors track gains and losses on the individual stock level, they should indeed be more prone to the disposition effect, because they are risk-averse (risk-seeking) for gains (losses) and there is no offset over the entire portfolio. An intuitive reason why people could engage in narrow framing in the stock market is the complexity of the uncertainty involved. Managing a portfolio requires decision making with the correlations between stocks in mind. This task is cognitively demanding and it is therefore easier to consider only individual stocks.

2.3. Representation

2.3.1. Gains and losses in the stock market

Before applying the evaluation methods in prospect theory, investors first need to form a representation of the outcomes associated with a state of the world. In the context of stock market, it would be interesting to know how investors form a mental representation of a stock. What information are they using as the basis for their decision making? Prospect theory is defined over gains and losses. What is the reference point for measuring gains and losses? In the following section, I briefly discuss important research related to answering those questions.

The question of how to define a reference point in reference dependent models of preference, like prospect theory, is not limited to behavioural finance. There is an ongoing discussion in the related behavioural economics research, focussing on two main concepts for reference points. The first proposed reference point is the status quo. An investor in the stock market should care about returns on the wealth she invested. The status quo then is a return of zero over the considered period, since it reflects the initial wealth. Benartzi and Thaler (1995) for example use a return of zero as the reference point and assume initial wealth is reset to the most recent wealth for the subsequent evaluation period. Köszegi, B. and Rabin, M. (2006) propose a different definition of the reference point. In their model, the reference point is the recently held expectations about an outcome. Imagine an investor who predicts a positive stock market trend in the near future. She would expect an increase in stock prices and a positive return on her investment. Her reference point based on her expectations would be the increasing market return. In a setting where no specific market trend is expected by the investor, she could reasonably form her expectations based

[4]The disposition effect is the tendency to sell winners early and hold on to losers.

on the risk free rate[5]. Although the expectations-based definition of the reference point makes intuitive sense, there is one important assumption: individuals correctly predict the environment they face. Even with a short horizon it is questionable whether investors can form correct expectations of the market return or the risk-free rate and also predict their own reaction to them. In their experiments, Baillon et al. (2015) find support for the status quo being the preferred reference point. However, in an empirical study by Lin and Meng (2015), investors in the U.S. stock market have reference points above the risk-free rate, aligning closest with their portfolios past returns. Since there is no conclusive evidence on either side of the argument, three different reference points form part in my framework: zero or raw returns, returns in excess of the risk-free rate and returns in excess of the market. I do not make a specific distinction whether they are based on status quo or expectations. Returns in excess of the market are the benchmark case, because investors could reasonably compare their performance with the overall performance of the stock market.

2.3.2. Information sources

With the reference points in mind, what source are investors using to determine returns? Information sources probably vary greatly over time and investor groups. In the preinternet era investment booklets could be a common guide. In fact, Alexander et al. (1998) present survey results from US mutual fund shareholders showing that investment booklets are the most frequently used source of information. Other sources in their survey include financial publications, family and friends or brokers. Barberis et al. (2016) point to the Value Line Investment Survey as an example of an investment prospectus. Each stock page in this prospectus shows a chart of past prices and a future price target on top, fundamental variables over time in the middle and a written analysis at the bottom[6]. The price chart is the only visual representation of data and immediately catches the reader's eyes. It is not unreasonable to assume that the average individual investor looking at this kind of prospectus would be greatly influenced by the price chart in their decision making. Further, the focus of these charts on past price developments could suggest that past prices contain important information on future prices. Therefore I assume that investors use past prices to form their expectations about the future. Specifically investors are assumed to evaluate the attractiveness of a stock based on past return distributions.

But the stock return information from the prospectus is coded into the price charts. What investor biases could alter the information during the transition from price charts to returns? Nolte and Schneider (2018) test how the shape of price paths influences investment decisions. Although the underlying return distribution of each price path presented to their test subjects is equal, they find that that some price paths are preferred to others. It seems that the framing of price paths affects the perceived attractiveness of the underlying return distribution. It is also becoming clear that, in the mind of an investor, asking for prices and asking for returns is not essentially the same. Confronting subjects with return instead of price charts can lead to lower expectations (see Glaser et al. (2019)). Glaser et al. (2007) show that subjects expect trend continuation if asked for future price changes, while reverting to mean of the series if asked for future price levels. This evidence suggests that investors do no always end up with accurate stock returns, when gathering information from price charts. As consequence using stock returns observed in the marketplace as a depiction of the returns investors have in mind might not be correct. This bias is not only inherent to my framework, but also to other studies applying prospect theory to the stock market (see e.g. Barberis et al. (2016))[7]. The extent of this bias however depends on the assumption that price charts are the sole information source and on the fraction of unsophisticated investors in the marketplace that cannot accurately discern the difference between prices and returns.

About half of my sample period belongs to the internet era. Surely, this has an impact on the information sources used by investors. First, the channel through which many investors trade has moved to brokerage firms that offer online trading. This means that they receive a smaller amount of professional advice from their brokers, but in turn can avoid higher trading fees. Other new sources of information in the internet era are online investment communities or free financial websites like Yahoo Finance. Regarding the question whether one single display method (price charts versus return charts) is preferred, Glaser et al. (2019) observe no clear pattern for the different sources. All of the online investment information sources have in common that they offer investors much more data on stocks than in previous decades. More available information should enhance investors performance, but only if they can use it efficiently for their investment decisions. Keller and Staelin (1987) find that the decision effectiveness of people decreases with the amount of information. Information overload also causes subjects to opt for simple investment alternatives in an experiment reported by Agnew and Szykman (2005). When searching for information online investors have access to much more frequent data. This frequent feedback on stocks could actually hinder performance, as investors focus narrowly on very recent data and lose track of long term performance over multiple time periods (see Lurie and Swaminathan (2009)).

2.3.3. Prospect theory value of return distributions

It seems hard to accurately discern which information, and at what frequency, is used by each individual investor.

[5]In theory, even the status quo could be the reference point based on expectations, if the investor expects to retain initial wealth.

[6]See Appendix C, p. XII, where a sample stock page of the Value Line Investment Survey is presented.

[7]In addition Benartzi and Thaler (1995) as well as Menkhoff and Schmeling (2006) assume that investors evaluate stocks using past return distributions. Although they do not further specify how investors would get information on past returns, there might be a possibility that the same bias is inherent to their studies

In face of the uncomplete understanding on what gains and losses are in the stock market, I make following assumptions: (1) investors evaluate a stock based on its past return distribution, (2) investors track the stock's performance with monthly frequency and (3) they consider returns from the past five years in the benchmark case. The assumption of monthly returns is based on the idea that investors could think that they can see trends emerging at this frequency, which influences their perception of a stock. I use the same lookback period for the return distributions as Barberis et al. (2016) to keep comparability high. However, I also present the results of the main analysis for other periods in Appendix D. The prospect theory value is constructed as follows. For each stock, the past sixty monthly returns are sorted from most negative to most positive. The resulting return distribution is equally weighted. As a consequence recent returns have the same impact on the decision of an investor as the earliest returns in the distribution:

$$(r_{-m}, 1/60; r_{-m+1}, 1/60; \ldots; r_{-1}, 1/60; r_0, \\ 1/60; r_1, 1/60; \ldots; r_n, 1/60) \tag{6}$$

For convenience, the form and subscripts of equation (6) are equal to those of equation (1). It is easy to see that an equally weighted distribution results in the same probability for every outcome. The prospect theory value of this distribution then is the application of equations (2) until (5), keeping in mind the fixed probability. Rearranged in to one equation, investors determine the prospect theory value according to:

$$TK = \sum_{i=-m}^{-1} v(r_i)\left[w^{-}(i+m+1/60) - w^{-}(i+m/60)\right] + \\ \sum_{i=1}^{n} v(r_i)\left[w^{+}(n-i+1/60) - w^{+}(n-i/60)\right] \tag{7}$$

The parameters for the shape of the value and probability weighting functions are taken from Tversky and Kahnemann (1992). Other authors confirm the plausibility of their estimation (see e.g. Abdellaui (2000) or Bruhin et al. (2010)).

3. Model of asset prices

3.1. The model by Barberis et al. (2016)

The starting point in the model stock market is the mean-variance framework as proposed by classical Portfolio Theory. There are two types of assets; risky assets J with $j \in \{1, \ldots, J\}$ and the risk-free asset r_f. A risky asset j earns return r_j and is distributed around the mean μ_j with standard deviation σ_j. Assets are traded by two types of investors. Rational investors 'arbitrageurs' buy and sell stock maximizing the Sharpe-Ratio of their portfolio. Put another way, they invest in the tangency portfolio as proposed by Portfolio Theory. This portfolio has return r_t, and the J x 1 vector w_t determines the weight of the risky assets J. On the other hand, the individual investors exhibit prospect theory preferences. While these investors also start by building the efficient tangency portfolio, they chose to adjust their holdings positively for stocks with high prospect theory value and negatively for stocks with low prospect theory value. The asset weights in their portfolio w_p are given by:

$$w_p = w_t + k w_{TK} \tag{8}$$

for some $k > 0$, with w_{TK} as the vector of weights in high (low) prospect theory stocks. The weight of the j^{th} stock in the $J \times 1$ vector of w_{TK}, w_{TK}^j is:

$$w_{TK}^j = TK_j - \overline{TK} \tag{9}$$

The variable TK refers to the prospect theory value of a stock's return distribution as in equation (7). In the benchmark case, the return distributions represent sixty monthly returns in excess of the market. The individual investors in this model do not sort their portfolio on absolute TK values of stock, but rather on higher than average values. Accordingly, $\overline{TK}$ is the average TK value of all stocks in J, formally $\overline{TK} = \sum_{j=1}^{J} TK_j / J$.

The fraction of rational investors in the stock market is π. Conversely, $1-\pi$ is the fraction of irrational investors that adjust the tangency portfolio for stocks with high (low) TK value. Later in this chapter I will argue that π decreases during bubble periods. Given the fractions of the two types of investors in the population, the market portfolio w_m is:

$$\begin{aligned} w_m &= \pi w_t + (1-\pi)(w_t + k w_{TK}) \\ &= w_t + \eta w_{TK} \end{aligned} \tag{10}$$

where $\eta = (1-\pi)k$.

As stated above, the model of Barberis et al. (2016) starts with a mean-variance framework. Specifically, the basis is the matrix solution for the case of multiple risky assets and a risk-free asset. Guided by this solution, they develop a model of asset prices fitting the specifications in equations (8) through (10). In the resulting model, the mean return of assets μ_j is given by:

$$\begin{aligned} \frac{\mu_j - r_f}{\mu_m - r_f} &= \beta_j - \frac{\eta w_{TK}^j s_j^2}{\sigma_m^2 (1 - \eta \beta_{TK})} \\ &= \beta_j - \frac{\eta \left(TK_j - \overline{TK}\right) s_j^2}{\sigma_m^2 (1 - \eta \beta_{TK})} \end{aligned} \tag{11}$$

I present a commented proof of their model in the Appendix A. The β-factors are the product of regression of the respective excess returns on the market excess returns. The variable s_j^2 is the covariance of the residuals from the regression of j's excess returns. In equation (11) the mean excess return equals β_j of stock j, adjusted negatively for a term which is increasing in the prospect theory value TK_j of the stock. This leads to the original prediction that stocks with high prospect theory value earn low subsequent returns.

3.2. Model extension and hypothesis

For the main part of my analysis, I test how this prediction changes during bubble periods in the stock market. I extend the model of Barberis et al. (2016) by partly reinterpreting the variable η appearing in equation (11):

$$\eta = (1-\pi)k \tag{12}$$

The term for the fraction of irrational investors $(1-\pi)$ remains unaltered in its interpretation. The factor k however subsumes the extent to which rational investors face limits to arbitrage. I argue that both of these factors increase during bubbles. In the following paragraph I turn to the reinterpretation in more detail and show the effect of the proposed increase in $(1-\pi)$ and k on returns in equation (11).

As indicated in equation (8) individual investors split their holdings between the tangency portfolio with weight w_t and prospect theory portfolio with weight w_{TK}. The vector of weights w_{TK} is scaled by some factor $k > 0$. There is no further definition of this scaling factor. It subsumes all effects that increase or decrease the weight in the prospect theory portfolio. Limits to arbitrage for rational investors could be a factor that increases the weight in the prospect theory portfolio. During normal market conditions any mispricing introduced by individual investors should be swiftly corrected by contrary arbitrage positions of rational investors. Because of arbitrage the influence of individual investor's preferences on prices and the weight of the prospect theory portfolio in general are kept at low. If the limits to arbitrage rise during bubbles, rational investors have less interest to correct mispricing which gives the individual investor's preferences space to manifest themselves. So my argument is that arbitrage provides a boundary to the power of individual investors over stock prices. In summary, an increase in the limits to arbitrage increases k which in turn leads to a stronger weight in the prospect theory portfolio[8]. The product of the fraction of individual investors $(1-\pi)$ and the factor k determines the value of η, as in equation (12). Because k scales the fraction of individual investors by the extent of limits to arbitrage, η represents something akin to the effective participation of individual investors in the stock market. If k increases while keeping $(1-\pi)$ constant, the impact of their preferences on prices still increases: the effective participation of individual investors is larger. During bubble periods, both the fraction of individual investors and the limits to arbitrage should increase. Therefore the effective participation of individual investors should increase as well. The evidence supporting this proposition is presented in chapter four. First however it is necessary to take a closer look at the impact an increase in η has on the model in equation (11), which also leads up to my main hypothesis. If η increases the numerator in the fraction on the right hand side increases, while the denominator becomes smaller. The fraction as a whole becomes larger, which increases the deviation from β_j. Put differently; an increase in either k or $(1-\pi)$ i.e. the effective participation of individual investors η increases the mispricing and the deviation from rational prices based on the risk factor β_j. The formal proof of this intuition is the partial derivative of equation (11) with regards to η:

$$\frac{\partial EX}{\partial \eta} = -\frac{\left(TK_j - TK\right)s_j^2}{\sigma_m^2 (1-\eta\beta_{TK})^2} \tag{13}$$

where EX is an abbreviation for the mean excess return. The proof for equation (13) is in Appendix B. Under the assumption that the stock is high prospect theory value $TK_j > \overline{TK}$, an increase in η decreases excess returns. If $TK_j < \overline{TK}$ then the the excess return is rising in η. In the original prediction, stocks with high (low) prospect theory values earn low (high) subsequent returns. An increase in the effective participation rate of individual investors seems to work in the same direction. If η increases the effects should just be larger. My main hypothesis is that stocks with high (low) prospect theory values earn lower (higher) subsequent returns during bubble periods compared to non-bubble periods. Formally, the model representing this prediction is:

$$\frac{\mu_j - r_f}{\mu_m - r_f} = \beta_j - \frac{\eta_{non-bubble}\left(TK_j - \overline{TK}\right)s_j^2}{\sigma_m^2 (1-\eta_{non-bubble}\beta_{TK})} \tag{14}$$

$$\frac{\mu_j - r_f}{\mu_m - r_f} = \beta_j - \frac{\eta_{\text{bubble}}\left(TK_j - \overline{TK}\right)s_j^2}{\sigma_m^2 (1-\eta_{\text{bubble}}\beta_{TK})} \tag{15}$$

where $\eta_{\text{bubble}} > \eta_{\text{non-bubble}}$. This prediction is based on the proposition that the limits to arbitrage and/or the fraction of individual investors in the stock market increase. In the following chapter, I review the evidence for these propositions.

4. Review of the evidence

4.1. Limits to arbitrage

4.1.1. Arbitrage

Under the assumption of efficient markets all public information is incorporated into stock prices. In addition rational investors do not face constraints and can trade freely. If individual investors entered this market and drove stock price away from the fundamental value, rational investors immediately try to profit and correct prices i.e. conduct arbitrage. Since there are no constraints on trading, arbitrage is risk-free and requires no capital and any mispricing after price-shocks should quickly dissipate. What if the assumption on no constraints to trading is lifted? Rational investor now might not have the same incentive to correct mispricing, because they are facing some risks. The existence of risk associated with arbitrage means that rational investors face limits to arbitrage, which make persistent mispricing of

[8] Although I relate the factor k to the limit to arbitrage, it is a variable that subsumes any factor that increases the weight in the prospect theory portfolio. Because a high level of mispricing is one of the defining characteristics of bubbles, I come to the conclusion that any hindrance to correct this mispricing should lead to a higher weight in the prospect theory portfolio.

stocks possible. Barberis and Thaler (2003) identify three main risks and costs associated with taking arbitrage positions that could impose limits to arbitrage: fundamental risk, noise trader risk and implementation costs.

4.1.2. Risks associated with arbitrage

Fundamental risk

A risk factor facing the arbitrageur is changes in fundamental value of stocks. If a rational investor determines that a stock is undervalued, she would try to profit from the mispricing and buy the stock at a low price. Imagine that the fundamental value of the cash flows suddenly decreases, because unexpected negative company information arrives. The fundamental value now has dropped below the buying price and the arbitrageur has incurred a loss. Although shorting is a possibility, it might be hard to find a good substitute. Gromb and Vayanos (2010) explain fundamental risk in two parts, connected to cross-asset arbitrage and intertemporal arbitrage. Cross-asset arbitrage is rational investors exploiting price discrepancies between two different assets at one point in time, and intertemporal arbitrage is exploitation of price differences for one asset at two points in time. They find the impact of demand shocks on prices not only depends on the ability to hedge risk but also the risk aversion of arbitrageurs. Therefore fundamental risk could play an important role in the decision making and investment process of arbitrageurs.

Noise trader risk

Noise traders are unsophisticated investors that react to noise not related to actual cash flow information. Arbitrageurs face the risk that noise traders drive prices further away from fundamental value for a longer period. As Shleifer and Vishny (1997) point out, many rational investors do not invest their own money but rather their clients' money. The client may have short time horizons and demand positive returns. With this delegated portfolio management, clients could force the arbitrageur to liquidate positions early. If the arbitrageur has to liquidate before noise trader's beliefs bounce back to original levels, they incur a loss[9]. These problems do not only occur with delegation from clients, but also from other parties such as lenders, that make arbitrage not only risky but costly.

Implementation costs

Since arbitrage often involves shorting stocks, it is important to consider which costs are associated with such a transaction. In order to sell stocks short, an arbitrageur has to borrow the stock by giving some collateral, in most cases cash. Gromb and Vayanos (2010) point out that a low interest rate earned on the collateral compared to the market is costly for the arbitrageur. Further, the desired stock has to be shortable in the first place. D'Avolio (2002) analyses the market for borrowing stock and finds that most of the stocks in CRSP are shortable. Those that aren't can be found in the category of high illiquid stocks with low market capitalization. In addition, he shows that high loan fees seem to be rare on average, but high for some stocks. Duffie et al. (2002) develop a model where borrowers and lenders in the stock market bargain over the terms of their agreement. This leads to friction cost if the lender has problems finding lendable shares. Costs for conducting arbitrage transaction can also come from insufficient capital. Arbitrageurs do not invest their own money, so they either have to lend it directly or attract investor capital. Gromb and Vayanos (2002) show that financially constrained arbitrageurs are sometimes not able to hold their arbitrage positions. If the difference between the arbitrageur's position and the price increases, some lenders might first require higher margins and then terminate the loan, if possible. This is similar to the case of Shleifer and Vishny (1997), where investors pull out capital when they see a negative arbitrage position.

4.1.3. Measures for the limits to arbitrage

Proxies for the limits to arbitrage

Because the limits to arbitrage are difficult to measure by a single variable, researchers have tried to find adequate approximations instead[10]. Mashruwala et al. (2006) rely on idiosyncratic risk as a measure of the risk an arbitrage position entails. Their argument is that only a part of a stock's volatility can be hedged, specifically the part that is known to the arbitrageur, which is assumed to be the volatility explainable by the standard market model. The other part of the stock's volatility, coming from the residuals in the standard market model regression, is an approximation for the absence of close substitutes in the hedging process. This absence is a limit to arbitrage and a risk that the arbitrageur faces. Brav et al. (2009) also use idiosyncratic risk as a proxy for limits to arbitrage, but take the residual variance from a four-factor model, instead of the standard one-factor model. Size is also considered a factor related to the limits of arbitrage. Because the stocks of firms with large market capitalization tend to be easier to trade, borrow or short, size is an inverse estimate for the limits to arbitrage. In addition to size, Ali et al. (2003) use several measures of transaction costs to study the limits to arbitrage. They pick stock price as a measure of direct transaction costs, because it is inversely related to the bid-ask spread. Similar to Mashruwala et al. (2006), they employ the dollar trading volume as an indirect measure of transaction cost. A higher dollar trading volume suggests that an arbitrage position could be executed without delays and for low cost. Stoll (2000) also finds that the ask-bid price spread is negatively related to measures of trading activity, such as the average daily dollar trading volume. Theoretically there should be fewer frictions when implementing arbitrage strategies for stocks with high trading volume.

[9]See De Long et al., 1990, p.705.

[10]See Chu et al., 2017, p.5-6, for an overview of different approximations of limits to arbitrage.

Therefore the dollar trading volume is an inverse proxy for the limits to arbitrage.

Proxies for short-sale constraints

An arbitrageur has to have the ability to sell stock short, if she wants to hedge against adverse effects to her arbitrage position. As discussed in the paragraph on implementation costs, frictions in the search process for shortable stocks can be a limit to arbitrage. Chen et al. (2002) measure short-sale constraints indirectly by looking at the breadth of ownership, the number of investors with long positions in a stock. This is based on an argument by Miller (1977) that if short-sale constraints exist, stock prices are anchored around the valuation of optimists and pessimist do not have any impact on prices. Therefore a low breadth of ownership, a sign that mainly optimists trade a stock, should be an approximation of short-sale constraints. Nagel (2005) employs another measure related to the ownership, institutional ownership. He argues that borrowing stock for shorting is easier for stocks with high institutional ownership, related to the findings by D'Avolio (2002). Therefore high institutional ownership should indicate a lower degree of short-sale constraints. Similarly, Hirshleifer et al. (2011) use institutional holdings as proxy for availability of loanable shares. If the loanable shares are scarce, short-selling is difficult. Finding loanable shares should be easier for stock with high institutional holdings. In a second strand of the literature, short interest, the fraction of shares sold short to shares outstanding, is a constraint on the arbitrageur side. While the proxies using institutional holdings variables focus on the supply side, short interest can be seen as the demand side of the short-sale constraint issue (Asquith et al. (2005)). In this framework, a high short interest leads to more arbitrageurs competing for the shortable shares, increasing cost and building a short-sale constraint.

4.1.4. Limits to arbitrage and implementation costs during bubbles

For the main part of my analysis it is important to think about how the extent of limits to arbitrage could change during bubble-periods compared to non-bubble periods. I make a three-fold argument based on synchronization issues (Abreu and Brunnermeier (2003)), resale options (Scheinkman and Xiong (2003)) and implementation costs (D'Avolio (2002)). Abreu and Brunnermeier (2003) present a model of how bubbles in asset prices can form and persist due to synchronization issues among arbitrageurs. It is a description of how rational investor behaviour could result in a failure to correct mispricing. In the model, assets prices are assumed to deviate from fundamental value at some point. One after another, the arbitrageurs become aware of the mispricing. They have prior knowledge of how bubbles evolve. Therefore they want to profit from the mispricing as long as the bubble grows and exit right before it collapses. The bursting of the bubble occurs when a sufficient amount of arbitrageurs decides to sell off the positions that profit from the mispricing. But they have no knowledge when other arbitrageurs became aware of the bubble or when other will try to cancel out their respective positions. Arbitrageurs have to ty to time the market, because only a fraction of arbitrageurs can exit before the bubble collapses. The incentive to correct any mispricing in asset prices is therefore delayed until the point when a sufficient amount of arbitrageurs think the bubble will burst soon. This is a strong argument for the increased limits to arbitrage during bubble periods. Griffin et al. (2011) support the model of delayed arbitrage by rational investors during the dotcom-bubble. They find that institutional investors pull capital in the months before the bubble bursts, while individual investors increase their exposure. It seems that arbitrageurs drive the collapse of the bubble and delay the correction of mispricing. Scheinkman and Xiong (2003) offer a different approach to explain the continued overvaluation of assets during bubbles. In their model an investor not only buys an asset, but also the option for resale to others. The investors are not fully rational and exhibit overconfidence. Every investor is confident that her view of fundamental value is correct. This variability in beliefs leads to a situation where the resale option for the asset can be exercised with other overconfident investors, which see higher fundamental value in the asset. Therefore investors buy assets with prices in excess of their own evaluation of fundamental value. The resulting mispricing is a consequence of the irrational behaviour of investors. If the overconfident investors make up a significant part of the market, the limits to arbitrage rise. Another requirement for successful arbitrage is the ability to sell short. D'Avolio (2002) studies the characteristics of the market for lending and borrowing stock i.e. the processes underlying short selling. In the sample, most stocks are shortable for relatively low fees. Only few stocks demand very high loan fees. However, short-selling becomes more difficult when investor opinions on the value of a stock diverge. Imagine a short seller with a stock valuation lower than the average investor in the market. Since other investors are more optimistic, the lender could cancel the loan of the short seller and receive a higher fee. The short seller can now either close out the position by buying back the stock or chose to pay a higher fee. In both situations the short seller faces more costs associated with her position. It is reasonable to assume that, during bubble periods, some investors overvalue more than other, similar to the overconfidence model of Scheinkman and Xiong (2003). Therefore the short selling cost should rise during bubble periods, increasing the limits of arbitrage. In summary, any of the three explanations makes for a theoretical argument to why limits to arbitrage should rise during bubbles. Since evidence for the delayed arbitrage model by Abreu and Brunnermeier (2003) comes from the time period of the tech bubble and this accounts for the majority of bubble months I detect in my sample later, it is the most appealing explanation. However, the delayed arbitrage model cannot be measured with the proxies for limits to arbitrage or short-sale constraints above. Analysing the change in trading behaviour of arbitrageurs during the bubble like Griffin et al. (2011) seems to be the only way to test the delayed arbitrage model. The explanation based on Scheinkman and Xiong (2003) runs into similar issues. While it makes intuitive sense that overconfident investors continuously exercise resale options, it is hard to test. Explanations based on investor behaviour in general run into this issue. In these models, investors do not necessarily face higher limits to arbitrage, because they consciously decide to deviate from the own fundamental valuation of stocks. Only later, when the price pressure of this type of investor is high, arbitrageurs face higher limits to arbitrage.

4.2. Individual investor's participation during bubbles

The second proposition related to my hypothesis is that the fraction of irrational investors might increase during bubble periods compared to non-bubble periods. How could irrational investors be drawn in to participating in the stock market? Maybe their decision to invest in stocks depends on attention-grabbing news media coverage. If ever more positive information of returns to be earned arrives, more irrational investors might consider participating. Barber and Odean (2008) find that individual investors buy stocks that grab their attention, because they are faced with a search problem. There are many stocks they could buy, but they can only choose between those that they notice. Consequently they buy stocks that recently grabbed their attention. Institutional i.e. rational investors on the other hand do not confront a search problem as large as individual investors. They can narrow their search for stocks by demanding certain criteria (for example sector or certain multiples). Therefore institutional investors do not buy attention grabbing stocks to the same extent. During bubble periods individual investors receive continuous positive information about a set of stocks. They are incentivised to buy stocks driving the bubble and contribute to the overvaluation. Effectively, they have a greater influence on prices than during non-bubble periods. The evidence on higher participation of irrational investors during bubbles is scarce. Although Brennan (2004) observes an increase in individual investors owning stock during the time of the tech bubble, the reason was more the democratization of equity markets in the U.S. and not attention-grabbing stock information. Since one of the bubbles my test detects is the tech bubble, I might have to accept the notion that this one time increase in individual investors in the U.S. could drive results and not something inherent to all bubbles. However this would only be true if the results indicated a long term change even after the tech bubble. As an argument, it does not seem unreasonable that more individual investors will think about buying stocks during bubbles, because they are fed information about high returns.

In order to test my hypothesis and the propositions made, it is necessary to detect and time-stamp bubbles in my sample. In the following chapter I therefore discuss what constitutes a bubble and how to measure it, followed by presenting the econometric test for bubbles I apply to my sample.

5. Bubbles

5.1. A short history of bubbles

When hearing about bubbles, a sharp rise in prices coupled with investor euphoria and a subsequent bounce back of prices to pre-bubble levels comes to mind. Although bubbles seem like a new phenomenon, many date them back until the tulip mania that occurred in the Netherlands leading up to 1637. There, prices for rare species of tulips increased rapidly until, in some cases, they crashed to 10 percent of their maximum[11]. Two other commonly identified bubbles before the start of the 20th century are the Mississippi bubble and the South Sea Bubble. Both shared similar characteristics, they occurred when private enterprises sought to refinance government debt. The Mississippi bubble was closely connected to the French Compagnie des Indes, which successively took over large projects like collecting taxes or running trade monopolies outside Europe, financed by the issuances of shares. In the beginning stages of the Compagnie des Indes shares sold at 500 livres. From the start of the expansion if its operations in April 1719 to the height of the bubble share prices rose to 10,000 livres in February 1720 until falling sharply after May 1720. The South Sea Bubble happened during the same time period, but the English South Sea company had the refinancing of government debt as its only goal. The acquisition of government debt was again finance through share issues. From January until July 1720 the share prices rose from close to 100 pounds to 900 pounds, before crashing down in September. Although both of these developments are often seen as bubbles, Garber (1990) argues that there were rational explanations for the respective price declines[12]. The next period of explosive prices probably is more well-known today: the bubble leading up to the great market crash in 1929. White (1990) attributes the bull markets in the 1920s to the stable growth after the post-war recession and large companies' trend towards equity financing of new plants and equipment. In addition, commercial banks started to make investment for individuals easier through investment trusts. While stock prices and dividends commoved during most of the 1920s, prices started to run away from 1928 onwards up to the crash[13]. For the following decades there is now consensus whether bubbles existed in the financial markets. One of these shorter and more contested bubbles is the one before Black Monday in October 1987. Some argue that it was computer high frequency trades that caused the crash but there is no broadly accepted version of why a bubble could have existed. The most recently recorded bubble in stock prices is the tech bubble in the late 1990s. Driven by the promise of the internet economy, investors heavily bought in to stocks of technology firms, although the fundamentals did not justify the high valuations. Shiller (2000) describes how investors often willingly ignored concerning facts and

[11]See Garber, 1990, p.37.

[12]See Garber, 1990, for an overview of the three bubbles and alternative explanations.

[13]See White, 1990, Figure 3, p.73.

were blinded by positive news coverage. Ofek and Richardson (2002) show that the internet sector was highly overvalued, where the valuations of a quarter of the firms suggest a required excess return on capital of almost 50 percent over a 10 year period[14].

5.2. Detecting bubbles in asset prices

Because of the enormous impact bubbles can have on the economy and the livelihoods of many people, economists have tried to develop methods for testing explosiveness in asset prices. Gürkaynak, R. S. (2008) provides a survey of the literature on econometric methods. These methods rely on the rational bubble model, where the price is not only based on cash flow streams from dividends and fundamentals, but also has a bubble component. Investors incorporate this component into their valuation and accept the corresponding prices. The variance bound test by Shiller (1981) is based on a model of dividend cash flows. With perfect knowledge of the future dividend streams, the asset price would reflect the present value of the resulting cash flows. Since investors don't know the future, the price reflects the present value of expected dividends plus the forecast error. Consequently, the variance of this price together with the variance of the forecast error should be equal to the variance of prices with perfect knowledge. If they exceed this variance bound, a possible reason is the existence of a bubble in prices. There are several issues with this method. First, it should be noted that it was originally not supposed to be a specific test for bubbles, but rather market efficiency. However, Marsh and Merton (1986) conclude that the variance bound test does not even test for market rationality, when changing the assumptions underlying the stochastic process for dividends. In addition it can only be applied as an ex post test, because the calculation of the variance bound requires information on prices formed on perfect knowledge. West (1987) develops another test that is supposed to indirectly prove the existence of bubbles. The starting point is a model where asset prices are determined by the discounted dividend streams. The basic idea is to compare a set of parameters from different regressions, one run on an actual time-series and the other one run on simulated data. The first set of parameters is obtained from a regression of the stock price on lagged dividends. The second set comes from a pair of equations, determining the dividend process and discount rate. Because both regressions estimate the same model, there is no bubble if the parameters are equal. If there is a bubble the underlying model is slightly different from before. Now asset prices are determined by discounted dividends and a bubble component. The parameter estimates from the regression with actual data are then biased. However, the second set of parameters obtained by simulation is still consistent. The test by West (1987) compares the two sets of parameters. If they are different after controlling for sampling errors, this is indirect evidence of a bubble. Problems for this method arise in the specification of the stochastic process for dividends. Both of the test presented until now rely on ruling out any other option than bubbles, instead of testing for them directly. Diba and Grossman (1988) design a more direct test of bubbles and lay the foundation for the test by Phillips et al. (2015a,b) that I use to detect bubbles. Prices in their model are determined by dividends and unobservable fundamentals. Essentially, they conclude that if no bubble exists, the processes for dividends and prices should be stationary after the equal amount of differencing. The explosive process for bubbles in their model breaks this relationship. Therefore a bubble exists if the same amount of differencing that makes the process for dividends stationary, does not make prices stationary, which can be tested with unit-root tests. Evans (1991) doubts the power of such tests. He shows that the specification of the explosive bubble process is problematic: for periodically collapsing bubbles, unit root tests have low power of detecting an explosive process. Diba and Grossman (1988) assume that bubbles cannot restart after bursting and therefore their test would fail to detect explosive prices with periodically collapsing bubbles. Although the idea behind the test is promising, reliably testing for bubbles requires a different method, where unit root tests do not lose detection power. This is the issue that Phillips et al. (2015a,b) address with their method that I present in the next section.

5.3. The Phillips et al. (2015a) approach for bubble detection

5.3.1. Rational bubble model

In a recent series of papers Phillips et al. (2011) and Phillips et al. (2015a,b) develop a new approach based on unit root testing that aims to correct for theses fallacies. In order to understand hoy unit root tests can help to identify explosiveness in asset prices, it is useful to start with a rational bubble model, similar to those discussed in the previous section:

$$P_t = \sum_{i=0}^{\infty} \left(\frac{1}{1+r_f} \right)^i \mathbb{E}_t \left(D_{t+i} + U_{t+i} \right) + B_t \qquad (16)$$

where P_t is the after-dividend price of the asset, D_t is the payoff from dividends, U_t represents the unobservable fundamentals and $\mathbb{E}_t$ is an expectation operator. Accordingly the price of an asset depends on two components. A fundamental component that is based on expected cash flows from dividends D_t and fundamentals U_t, dubbed P_t^f here, and a bubble component B_t. This results in a simplified form of equation (16):

$$P_t = P_t^f + B_t \qquad (17)$$

The stochastic process underlying the bubble component has a submartingale property[15]:

$$\mathbb{E}_t \left(B_{t+1} \right) = \left(1 + r_f \right) B_t \qquad (18)$$

[14] See Ofek and Richardson, 2002, p.270.

[15] See Phillips et al., 2011, Figure 2, p.204, for an illustration of this kind of process.

Therefore the expected value for the bubble component in $t+1$ is an increasing function of the bubble component in t. This is in accordance with the self-reinforcing price expectations during bubbles. Under the assumption that $B_t = 0$ asset prices are driven by the characteristics of D_t and U_t. If now the stochastic processes for dividends and unobservable fundamentals are at most integrated of order one, then the fundamental component P_t^f in general is at most integrated of order one, then the fundamental component P_t^f in general is at most integrated of order one. In this case asset prices follow a random walk. If there is a bubble in asset prices i.e. $(B_t \geq 0)$, then the explosive process described in equation (18) would dominate the behaviour of asset prices. Therefore it is possible to infer the existence of a bubble by testing whether prices follow a random walk or an explosive autoregressive process. One disadvantage of testing prices is the possibility that the detected explosiveness is driven by the fundamental component rather than a bubble. Using the price-dividend ratio controls for a false positive interpretation of test results, assuming that the unobservable fundamentals are still at most driven by a I(1) process. Under the null hypothesis asset prices follow the process in equation (19):

$$y_t = dT^{-\eta} + \theta y_{t-1} + \varepsilon_t, \varepsilon_t \sim^{iid} \left(0, \sigma^2\right), \theta = 1 \qquad (19)$$

where d is a constant, T is the sample size and η is a control for the size of the intercept as the sample size approaches infinity. Phillips et al. (2015a) focus on the case of $\eta \geq 1/2$, for which the martingale component dominates the drift and y_t is close to a pure random walk. Phillips et al. (2014) report on the sensitivity of right-tailed unit root testing to specification of the null hypothesis. Instead of testing the null for stationarity against a random walk with intercept or trend, a random walk is the null against the explosive alternative. This corresponds to analysing the right tail of the distribution in unit root testing. To test for a unit root, Phillips et al. (2015a) apply Augmented Dickey-Fuller (ADF hereafter) tests on different subsamples using rolling regressions with sample start r_1 and end r_2 using following form:

$$y_t = \hat{\alpha}_{r_1,r_2} + \hat{\beta}_{r_1,r_2} y_{t-1} + \sum_{i=1}^{k} \hat{\psi}^i_{r_1,r_2} \Delta y_{t-i} + \hat{\varepsilon}_t \qquad (20)$$

where k is the number of lags and the ADF statistic is the t-statistic of the β-coefficient. The recursive approach is chosen based on the critique by Evans (1991) that unit root tests incorrectly detect stationary behaviour in a model with periodically collapsing bubbles.

5.3.2. SADF and GSADF test

The supremum augmented Dickey-Fuller test (SADF hereafter) is the first method that Phillips et al. (2011) develop, which applies unit root tests to an expanding sample sequence. The test statistic is obtained by running the regression in equation (20) on this sequence. Define the total sample as running on an interval of $[0, 1]$, r_0 as the minimum window size for the regression and r_1 and r_2 as the starting point and endpoint of the subsample, respectively. The first regression runs on the minimum window size r_0. Leaving the starting point r_1 fixed, the sample size increases by one observation at a time until the total sample size is reached. The supremum value of the resulting ADF test statistic sequence is the SADF test:

$$SADF\,(r_0) = \sup_{r_2[r_0,1]} ADF_0^{r_2} \qquad (21)$$

Imagine for example a series of 100 observations and a fixed minimum window size $r_0 = 0.1$. The first ADF test would then run a regression on observations 1 through 10, the second on observations 1 through 11 etc. until the last regression included all 100 observations. Figure 3 illustrates this procedure. Several papers explore this recursive approach for other unit root tests, such as the Bhargava (1986) statistic, the Busetti and Taylor (2004) statistic, the Kim (2000) statistic or the Chow-type Dickey-Fuller statistic[16].

The extension of this test is the general supremum augmented Dickey-Fuller test (GSADF hereafter) by Phillips et al. (2015a). This test is based on the same concept as the SADF test, but the number of subsamples used is much larger. Instead of only varying the endpoint r_2, the starting point r_1 is flexible as well. This double-recursive procedure is meant to make the detection of periodically collapsing bubbles more precise. Figure 4 illustrates the formation of subsamples the test is performed on. The GSADF test then is the supremum value of the resulting ADF test statistic sequence:

$$GSADF\,(r_0) = \sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} ADF_{r_1}^{r_2} \qquad (22)$$

The recommendation for the minimum window size r_0 is dependent on the size of the full sample T according to the rule $r_0 = 0.01 + 1.8\sqrt{T}$. On the one hand this ensures that the initial window is small enough to capture explosive price behaviour early in the sample if T is large. On the other hand the initial window is big enough for adequate estimation if T is small.

5.3.3. Time stamping bubbles

For empirical application it is important to know not only whether bubbles exist in a sample, but also when they originate and terminate. Since real-time stamping of bubbles is an important concern in practice, Phillips et al. (2015a) use a backward-looking implementation of the SADF test, called the backward sup augmented Dickey-Fuller test (BSADF hereafter). The endpoint r_2 is held fixed, while the starting point r_1 expands backwards towards the beginning of the full sample. The BASDF is then again the supremum value

[16]See Homm and Breilung, 2012, for a comparative analysis of these test statistics.

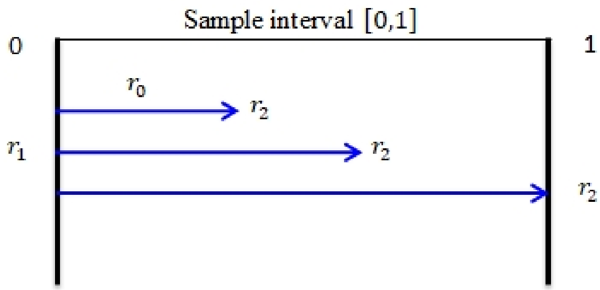

Figure 3: The SADF Test (Based on Phillips et al., 2015a, p.1049).

Set $r_1 = 0$ and $r_2 \in [r_0, 1]$. Use initial window $[0, r_2]$ and expand r_2.

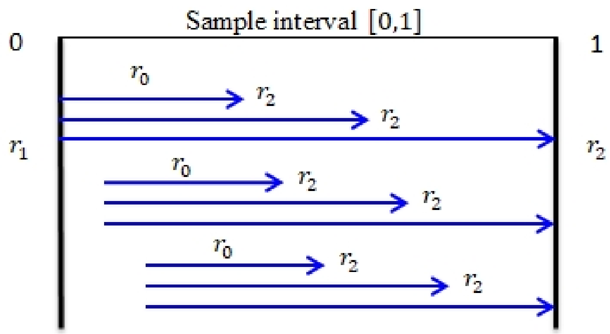

Figure 4: The GSADF Test (Based on Phillips et al., 2015a, p.1049).

Set $r_1 = [0, r_2 - r_0]$ and $r_2 \in [r_0, 1]$. Use moving window $[r_1, r_2]$.

of the resulting statistics sequence:

$$BSADF_{r_2}(r_0) = \sup_{r_1[0,r_2-r_0]} \left\{ ADF_{r_1}^{r_2} \right\}. \tag{23}$$

Similar to the double-recursive sampling technique of the GSADF test, varying the endpoint r_2 results in a statistic sequence of BSADF tests. This sequence consists of one BSADF statistic for every observation of the original sample provided the minimum window size restriction is fulfilled. Working from the first chronological test statistic to the last, bubble periods are defined according to the following equations:

$$\hat{r}_e = \inf_{r_2 \in [r_0,1]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\} \text{ and} \tag{24}$$

$$\hat{r}_f = \inf_{r_2 \in [\hat{r}_e + \delta \log(T)/T, 1]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\} \tag{25}$$

where $scv_{r_2}^{\beta_T}$ is the critical value for the ADF statistic based on Tr_2 observations. Here it is important to notice that a critical value is needed for every month in the sample. The origination of a bubble $\hat{r}_e$ is then the first chronological observation that passes the respective critical value. In order to avoid spurious explosive behaviour to be identified as a bubble a minimum duration based on sample size is required. After origination the explosive behaviour should persist at least for $\delta \log(T)$ observations. This rule is somewhat arbitrary, since δ is a parameter that can be chosen freely according to the frequency of the data and assumed minimum bubble duration. For example, if the data frequency is monthly, sample size $T = 480$ and bubbles should last a half a year at minimum, then $\delta \approx 2.2$. The end of a bubble $\hat{r}_f$ is the first observation after $\hat{r}_e + \delta \log(T)$ that falls below the critical value. This procedure can be used to time-stamp multiple bubbles that periodically expand and collapse, not only in stock prices, but also in many other assets traded on exchanges.

5.3.4. Other applications

Since its inception the BSADF test has been applied to detect bubbles in several assets. Cheung et al. (2015) search for bubbles in bitcoin prices, which seem to be driven by speculative behaviour. They find several smaller bubbles and one large bubble, collapsing the Mt. Gox exchange in 2014. Caspi et al. (2018) set out to detect explosiveness in oil prices. They test the nominal oil price against prior levels and the general price level in the U.S. and find periods of explosiveness in the 1970s and 1980s, as well as before the financial crisis in 2008. Escobari and Jafarinejad (2016) detect bubbles in real estate investment trust and are able to capture, among others, the reflection of the housing bubble before 2008. Hu and Oxley (2017) search for explosiveness in currency exchange rates and conclude that there were no bubbles in the

G10 countries during 1990 and 2014. In summary, this evidence suggests that the approach by Phillips et al. (2015a) can be readily applied to detect bubbles in a broad range of settings. I now turn to the application of this test to the U.S. stock markets.

5.4. Bubbles in the S&P 500

5.4.1. Data

The tests for the existence of bubbles and their time-stamping are performed on the S&P 500. This index covers approximately 80% of all available market capitalization of American equities[17]. The data for the S&P 500 is taken from Robert Shiller's website at Yale[18]. It consists of the monthly time-series of the index price, dividend and earnings. Additionally a consumer price index published by the U.S. Bureau of Labour Statistics allows for the conversion to real values. The data sampled stretch over 480 monthly observations from January 1978 up to December 2017. As indicated by the "rational bubble" model in section 5.3.1, using the price-dividend ratio controls for false positive interpretation of results. To obtain the price dividend ratio, I first compute the percentage dividend yield, which is real dividend divided by real price. The price dividend ratio then is the reciprocal value of the percentage dividend yield. The tests for bubbles run test regressions on an index of this price dividend ratio, starting with value 100 in January 1978.

5.4.2. SADF test of the S&P 500

Before applying the unit root test to the S&P 500 data, I obtain critical values under the null hypothesis by simulating the stochastic process specified in equation (19). The parameters controlling the size of the intercept are set to one i.e. $d = \eta = 1$. As discussed earlier, the data-generating process is then close to a pure random walk. The Monte-Carlo simulations for the finite sample critical values run over 480 observations with 1,000 replications. The SADF test statistic is stored for every replication. The basis for the 90%, 95% and 99% right-tailed critical values is the statistics sequence resulting from all replications. As suggested by Phillips et al. (2015a), the lag order is small and fixed, at $k = 1$, and the minimum window size for the recursive rolling regressions in the SADF test is 44 observations following their rule. I do not calculate the GSADF test statistic because the increased number of subsamples makes simulations of critical values much more extensive computationally. The alternative hypothesis of explosive behaviour is tested by running an SADF test on the S&P 500 data. Table 1 shows the finite sample critical values and the SADF test results. With a test statistic of 2.428 the SADF test surpasses its critical values at all significance levels. There is strong evidence for the existence of bubble(s) in the indexed price-dividend ratio of the S&P 500. These results are similar to the results reported by aforementioned authors.

5.4.3. Time stamping of bubbles in the S&P 500

After establishing that one or more bubbles exist in the S&P 500, I turn to the BSADF test (equations (23) to (25)) to determine the exact number of bubbles and the time stamping of each bubble. First critical values for the BSADF tests are obtained by Monte-Carlo simulations with 1,000 replications. The calculation is more complex than for finite sample critical values, since the simulations run on each subsample in the sequence Tr_2. In the end, there is a critical value for every observation; each based on the number of months passed until that point in the sample. Just as for the SADF, the minimum window size is 44 observations and the order of lags is fixed at $k = 1$. I then compare this sequence of simulated critical values to the BSADF test statistic sequence for the price-dividend ratio of the S&P500. In accordance with the crossing-time equations in (24) and (25), the starting point of a bubble is the first chronological observation in the sequence, where the BSADF test statistic passes its corresponding 95% critical value. The assumption is that a bubble has to last at least six months in order to be considered as such. Therefore if the observation six months after the test statistic first passed the critical value remains above the critical value, this time period in the S&P 500 is considered a bubble. Any observation that drops below its critical value afterwards is the end of the bubble. The arbitrary definition of a six month minimum bubble period (or any other chosen duration) seems to be one of the biggest issues for correctly specifying the BSADF test in empirical applications. What if, for example month five and seven after the start of the bubble surpass their respective critical values, but month six does not? The dating algorithm in equations (24) and (25) would not capture this time period as a bubble. Although I did not run into this problem, I attempt to use averaged values over a time span in order to avoid short dips (e.g. month six in the example) in the BSADF test statistic sequence. However, the definition of average values inevitably becomes arbitrary as well: it is not clear how many months should be used for averaging. The conclusion is that economic intuition has to be applied. It is unreasonable to assume that the bubble suddenly stops for one or two months. Ex post it is therefore best to fill in theses gaps, if they are just short breaks. If one wants to apply the BSADF test as a real time detector for bubbles, this is a bigger issue since it is not clear whether the drop is the end of the bubble or just a short break. Despite these problems, the time-stamping of bubbles seems to be surprisingly accurate. Figure 5 shows the BSADF test statistic sequence plotted against the 95% critical value sequence for the entire sample. The grey areas are the bubble periods identified by the BSADF test. They coincide with two bubbles corresponding to the tech bubble in the late 20th century and the lead-up to the crash of Black Monday in 1987. In fact, the starting point of the bubble is 1995m12, exactly one month later as reported by Phillips et al. (2015a). The ending point of the tech bubble indicated by my tests is 2000m9, which is eleven months earlier. Due to the characteristics of the time stamping algorithm described earlier, there are two short

[17] See the S&P 500 website, https://eu.spindices.com/indices/equity/sp-500.

[18] See "U.S. Stock Markets 1871-Present and CAPE Ratio", www.econ.yale.edu/~shiller/data.htm.

Table 1: The SADF test with critical values for the S&P 500 (Based on Phillips et al., 2015a, p.1066).

The table shows the results of running the SADF test specified in equation (21) on the indexed price-dividend ratio of the S&P 500. The sample stretches over 480 monthly observations, from January 1978 to December 2017. The finite sample critical values are obtained via Monte-Carlo simulation with 1,000 replications.

	Finite sample critical values		
SADF	90%	95%	99%
2.428	1.227	1.568	1.998

breaks between explosive periods that are assumed to be part of the tech bubble. In addition there are two visible spikes of the BSADF tests statistic. The first spike indicates explosive prices leading up to the market crash in October starting in 1987m2. This is the second and much shorter bubble I detect in the S&P 500. The second spike coincides with the financial crisis and passes the critical value sequence from 2008m10 to 2009m3. The detection of a crisis period as explosive is an issue that Phillips and Shi (2018) try to address by implementing a wild bootstrap for the critical values. Phillips et al. (2015a) again find similar periods of explosive prices, but the start and end are different. From their online appendix[19], it becomes clear that they use a variable lag depending on Bayesian information criteria, while I use a fixed lag. In my own informal tests I find that the specification of the lags in the augmented Dickey-Fuller tests influences results. Therefore the differences could be due to using fixed versus variable lags. Figure 6 shows the real price and price-dividend ratio of the S&P500 for my sample as well as the bubble periods indicated earlier. The BSADF test runs regression of the time-series of the indexed price-dividend ratio (the dotted line in Figure 6). As discussed when presenting the rational bubble model (equation (16)), this is necessary for the correct specification of bubbles. The detected bubble periods coincide well with price increases in the real price index of the S&P500, but only when dividends are not keeping up with this increase. For example, the BSADF test does not detect a bubble in the period from 2016 onwards, although the price of the S&P500 reaches record highs. This is because the price dividend ratio remains more or less constant. Of course the BSADF test is always limited to the specific timer series analysed. The detection of bubbles, say in housing prices, requires a separate implementation of the test. In summary, the bubble periods I detect coincide with the tech bubble and the lead up to Black Monday. They run from December 1995 to September 2000 and from February 1987 to August 1987, respectively. The strategy for the empirical analysis is to split the sample into two different sub samples. The first sample contains all months that were not stamped as a bubble and the second sample contains all months that were stamped as a bubble. This allows me to analyse the changes in the relationships between returns, the prospect theory value and other important control variables. It is therefore not only a test of my hypothesis that stocks with high prospect theory value earn lower returns during bubbles, but also gives an indication how the predictive power of other variables varies with market conditions.

[19]See https://sites.google.com/site/shupingshi/GSADF.zip?attredirects=0.

6. Prospect theory and stock returns during bubbles

6.1. Empirical approach

Barberis et al. (2016) show that stock with high (low) prospect theory value earn low (high) subsequent returns. Earlier, I made the prediction that this relationship should be more pronounced during bubble periods compared to non-bubble periods. In order to these two predictions in my sample, I rely on several tests. The first tests are portfolio sorts with a low high prospect theory value strategy that should earn abnormal returns if the predictions hold. These include checks for robustness, where the parameters for constructing the prospect theory value or the underlying sample are modified. In order to test the main predictions in the cross-section of returns, I use the Fama-Macbeth approach. This lets me test while controlling for known return predictors. The plausibility of the rising limits to arbitrage argument is evaluated using the same method, but the regressions include an interaction term of prospect theory value with different approximations for the limits to arbitrage, as presented in section 4.1.3. In the last section I look at the underlying components of prospect theory likely to cause the observed relationship with returns. The empirical analysis closely follows Barberis et al. (2016).

6.2. Data

6.2.1. Databases

The data for the empirical analysis is provided by Wharton Research Data Services. The CRSP/WRDS Merge database is the source for monthly security information of all companies listed on US stock exchanges, namely security identifiers, exchange codes and prices. The fundamental information on these companies is obtained from the Compustat database. Besides balance sheet items it also contains information related to security markets like common shares outstanding, all sampled with yearly frequency. The Fama and French database is the basis for the monthly market returns and risk-free rate. The sample for these three datasets starts in January 1978 and ends in December 2017, corresponding to the time period used for the time-stamping of bubbles.

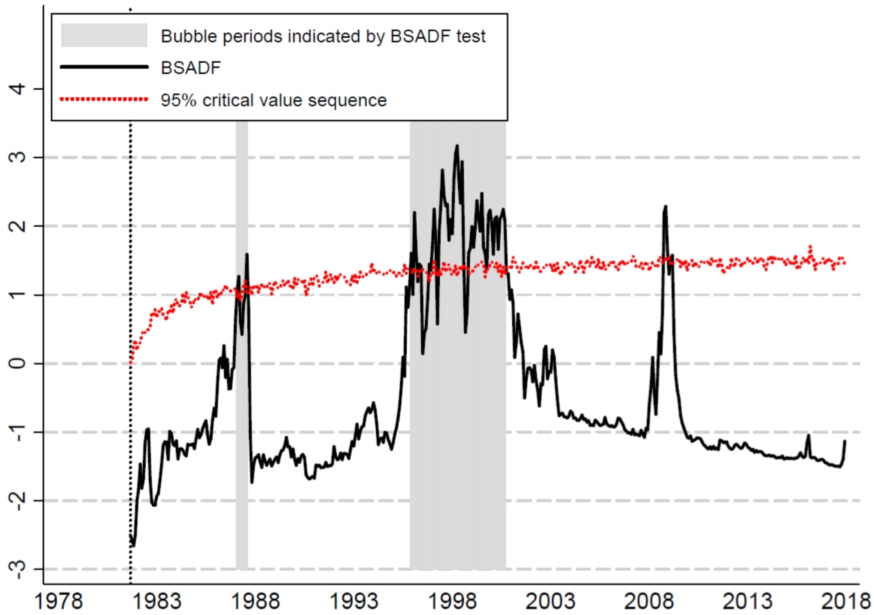

Figure 5: The BSADF test for bubbles in the S&P 500 (Based on Phillips et al., 2015a, p.1066).

The graph shows the results of the BSADF test procedure. The bold line represents the test statistic for the indexed price-dividend ratio of the S&P 500, the red-dotted line are the 95% critical values obtained via Monte-Carlo simulations with 1,000 replications for each point in time. The sample of the S&P 500 starts in Janaury 1978 and ends in December 2017. The dotted black line represents the end of the minimum window for estimation.

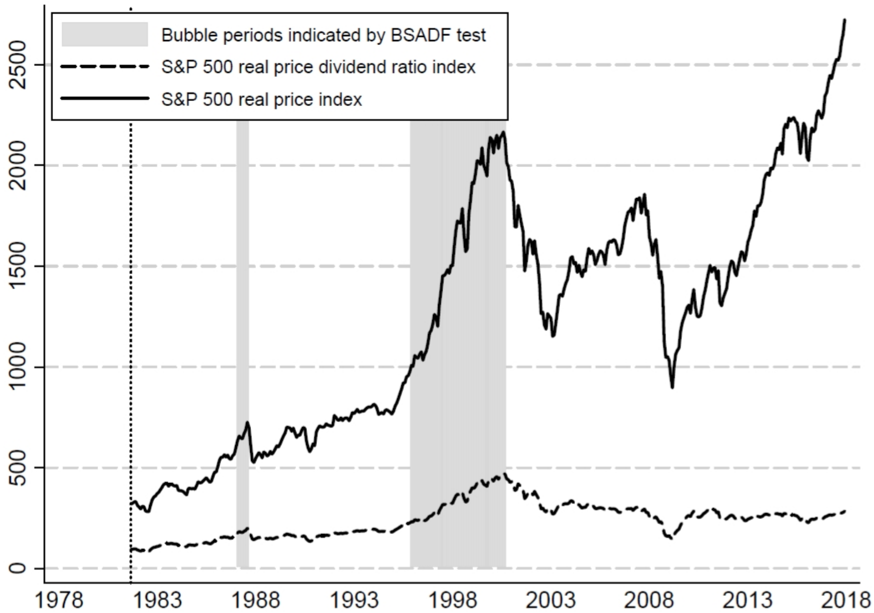

Figure 6: Bubbles in the S&P 500 (Based on Phillips et al., 2015a, p.1066).

The graph shows the results of the BSADF test procedure. The bold line is the price index of the S&P 500, adjusted for inflation. The dashed line is the index of the price-dividend ratio fo the S&P 500, again with prices adjusted for inflation. The sample of the S&P 500 starts in Janaury 1978 and ends in December 2017. The dotted black line represents the end of the minimum window for estimation.

In addition, I need daily data on prices and trading volume, which comes from CRSP/WRDS Merged daily. The daily data sample starts in January 1983 and ends in December 2017, as available. After merging the four datasets, I exclude stocks from smaller exchanges because of missing price information. The stocks in the final dataset are listed on NYSE, NYSE Amex, NASDAQ-NMS, NASDAQ OMX BX or NYSE Arca. Finally, stocks are required to have a history long enough as to make the calculation of the prospect theory value possible i.e. a history of at least 60 months.

6.2.2. Data Summary

Table 2 summarizes the variable for prospect theory value, dubbed TK, and the control variables I use in the empirical analysis. In addition, Table 3 shows the correlations of all variable. TK is the prospect theory value of a stock's return distribution, computed following equation (7). In the benchmark case, the return distribution consists of the past 60 monthly returns. Size is the log of the market value of equity in month t-1, obtained by multiplying the common shares outstanding at the end of year t-1 with the end of month price. Beta is a stock's beta as in Fama and French (1992). Each June of a given year stocks are sorted into 100 portfolios based on size and pre beta, a stock's market beta over the last six years. Then the return of each portfolio in the following year is used in a regression on the market return in the current and prior month. Beta is the sum of the coefficients from this regression. Bm is the log book-to-market ratio computed as the difference of the logs of book value of equity and market value of equity. The book value of equity is the shareholder's equity minus the preferred stock value. Following Daniel and Titman (2006), shareholder's equity is approximated by stockholder's equity. If the stockholder's equity is missing, I replace it with common equity plus preferred stock par value or assets minus liabilities plus book minority interest, as available. The preferred stock value is redemption, liquidation or carrying value in that order. After subtracting the stockholder's equity with the preferred stock value, book deferred taxes are added to get the final book value of equity. Mom is the control for the momentum effect reported by Jegadeesh and Titman (1993), the cumulative return from month t-12 to t-2. Illiq is Amihud (2002) measure for illiquidity, the average ratio of absolute return to dollar trading volume in month t-1. Further controls include rev, for the short term reversal, an ltrev, for the long term reversal. Rev is a stock's return in t-1, while ltrev is the cumulative return from t-60 to t-13. Max is the maximum daily return in month t-1, while Min is the negative of the lowest return in month t-1. Skew is the sample skewness of returns over the past five years. Because probability weighting is an important factor for the prospect theory value, skewness should be included in the analysis (see Barberis and Huang (2008)). The computation for the variables starts in January 1983 to match the minimum window required by the time-stamping of the bubbles and stock requirements mentioned earlier. On average, the prospect theory value of a stock's return distribution is negative, which is in accordance with the result in Barberis et al. (2016). Menkhoff and Schmeling (2006) also find that the prospect theory value of simulated return distributions tends to be negative. Table 3 presents the correlations between TK and the control variables employed. Beta is negatively correlated to the prospect theory value of a stock. Probability weighting could be an intuitive explanation for the positive relation between the prospect theory value of a stock and the sample skewness. A skewed return distribution should be preferred, since tails are assigned higher weights in prospect theory. TK is also positively correlated with the controls for past return anomalies, namely momentum, short term reversal and long term reversal. The correlation with ltrev is the strongest out of all control variables. It is also the variable that is constructed with the return information closest to TK. The reason behind the positive correlation with mom and rev should be similar, although the correlation coefficient on rev seems particularly large considering it represents only one month of shared return information. Finally, TK is positively related to size, the log market value of equity, and negatively related to the log book-to-market ratio Bm. Looking at correlations is only a first rough approximation of the relationship between the variables, therefore I now turn to the first step laid out in the empirical approach, portfolio analysis of stocks sorted on the prospect theory values.

6.3. Portfolio sorts on prospect theory value

Sorting stocks into portfolios on prospect theory values make it possible to exploit the prediction that high prospect theory value stocks should earn low returns. The portfolio with the highest TK values should earn the lowest returns and the portfolio with the lowest TK values should earn the highest returns. If the prediction is correct, going short and long in these portfolios, respectively, should generate positive excess returns. This approach of building long-short portfolios is popular in the finance literature, especially as a first assessment of the power of anomalies. For example, momentum (Jegadeesh and Titman (1993)) is exploited by forming such a portfolio, only with past winners and losers. It is also a well-known strategy for asset pricing tests (e.g. Fama and French (1996)), although Berk (2000) points out that sorting into many groups could weaken otherwise strong explanatory power of pricing models. I begin building the TK portfolios by sorting stocks into deciles based on their prospect theory values starting in January 1983 and ending in December 2017. For each of the ten portfolios formed on TK, the average equal weighted return over the next month is computed. The returns in the benchmark case are in excess of the market. The long-short portfolio takes a long position in the low TK portfolio (decile one) and a short position in the high TK portfolio (decile ten). This portfolio should generate positive excess returns, if the prediction is accurate. Table 4 shows the results of the portfolio sorts on prospect theory value. Decile one offers the highest average returns while decile ten exhibit's the lowest returns. Consequently, the low high TK portfolio earns positive returns. The excess returns to this strategy are highly significant. Overall, all decile portfolios are significant at least at the 10% level. The significance of the low TK and high TK decile is most important, since the return on the portfolio rely on both. While I use equal weighted returns, Barberis et al. (2016) also look at value weighted returns. They present similar results and show that the low-high TK portfolio even generates excess returns with value-weighting. Further, they show that the predictive power of prospect theory values for returns lasts for several months, although the level of excess returns declines very fast after the first month. My second prediction

Table 2: Summary statistics.

The table shows mean, standard deviation and median for each of the variables used in the empirical analysis over the period starting in January 1983 and ending December 2017. TK is the prospect theory value of a stock's return calculated using a distribution of the past 60 returns. Size is the logarithm of the market value of equity. Beta is a stock's beta based on 100 portfolio sorted on size and pre-beta as in Fama and French (1992). Bm is the log book-to-market ratio, calculated as the difference of the logarithm of the book value of equity and the logarithm of the market value of equity. Mom is the momentum of a stock; the cumulative return from month t-12 to t-2. Illiq is the illiquidity measure of Amihud (2002), scaled by 10^5. Rev is the return in month t-1. Ltrev is the cumulative return of a stock from month t-60 to t-13. Max is the highest daily return in month t-1, while Min is the negative of the lowest daily return. Skew is the sample skewness of returns over the past five years.

Variables	Mean	SD	Median
TK	-0.078	0.032	-0.075
Size	5.902	2.115	5.759
Beta	1.058	0.297	1.046
Bm	-0.575	0.821	-0.500
Mom	0.099	0.394	0.081
Illiq	0.122	0.610	0.002
Rev	0.010	0.123	0.004
Ltrev	0.412	0.753	0.344
Max	0.065	0.593	0.042
Min	0.052	0.052	0.038
Skew	0.0534	0.893	0.146

Table 3: Correlation matrix.

The table shows the correlation between TK and the control variables presented in Table 2. The sample starts in January 1983 and ends in December 2017.

	TK	Beta	Size	Bm	Mom	Illiq	Rev	Ltrev	Skew	Max	Min
TK	1.00										
Beta	-0.22	1.00									
Size	0.35	0.01	1.00								
Bm	-0.15	-0.07	-0.32	1.00							
Mom	0.21	0.05	0.08	-0.09	1.00						
Illiq	-0.26	-0.01	-0.33	0.16	-0.10	1.00					
Rev	0.06	0.02	0.03	0.03	-0.02	-0.01	1.00				
Ltrev	0.44	0.16	0.15	-0.34	-0.08	-0.14	-0.02	1.00			
Skew	0.21	0.19	-0.27	0.10	0.24	0.09	0.10	0.22	1.00		
Max	-0.06	0.03	-0.03	0.01	-0.02	0.07	0.07	-0.02	0.02	1.00	
Min	-0.36	0.17	-0.25	0.02	-0.05	0.31	-0.31	0.00	0.05	0.08	1.00

was that high (low) prospect theory value stocks should earn lower (higher) returns during bubble periods compared to non-bubble periods. Therefore I do the same portfolio sorting exercise comparing the date-stamped bubble periods to the non-bubble periods. Table 5 presents the results of this test.

The results for non-bubble periods indicate no large changes of the returns to the low high TK strategy. The first decile still has the highest average excess return and the tenth the lowest average excess return. Returns in the first decile are higher, while the magnitude does not change in the tenth decile. Overall, the low-high TK portfolio earns less compared to the test using the full sample. During bubble periods the returns on the same portfolio are on average more than a percent higher. This is in accordance with my main hypothesis. However, most of the decile portfolio are not significantly different from zero anymore. The increase in returns mainly comes from decile one: while returns in decile ten are almost on the same level, returns in decile one increase compared to non-bubble periods. This could indicate that some investors during bubble periods only invest in stock with high prospect theory value, causing low prospect theory value stocks to become relatively undervalued and earn higher returns. Generally the returns on the lower half of TK decile portfolios are higher during bubble periods. In order to confirm the results from the portfolio sorts, I first conduct robustness checks employing different specifications related to the construction of the prospect theory value. Table 6 presents the returns on the low-high TK portfolio for these tests, for both bubble and non-bubble periods. First, the construction window of the return distributions used for the computation of prospect theory values changes from

Table 4: Portfolio sorts on TK.

The table reports the average monthly excess return of portfolios formed on the prospect theory value. Stocks are assigned to deciles based on their TK value each month, ranging from low TK (decile one) to high TK (decile ten). The sample starts in January 1983 and ends in December 2017. The numbers in parentheses show the t-statistic.

Decile	Excess return	
1	0.0184	(5.45)
2	0.0137	(4.93)
3	0.0121	(4.88)
4	0.0097	(4.31)
5	0.0097	(4.50)
6	0.0079	(3.86)
7	0.0074	(3.77)
8	0.0062	(3.48)
9	0.0058	(3.51)
10	0.0036	(2.04)
Low-high TK	0.0148	(5.97)
N	420	

Table 5: Portfolio sorts during bubbles.

The table reports the average monthly excess return of portfolios formed on the prospect theory value. Stocks are assigned to deciles based on their TK value each month, ranging from low TK (decile one) to high TK (decile ten). Bubble periods were date-stamped following PSY (2015). The sample starts in January 1983 and ends in December 2017.

Decile	Excess return			
	Non-Bubble		Bubble	
1	0.0168	(4.50)	0.0268	(3.53)
2	0.0129	(4.18)	0.0181	(2.89)
3	0.0114	(4.13)	0.0155	(2.97)
4	0.095	(3.79)	0.0110	(2.15)
5	0.0094	(3.93)	0.0117	(2.29)
6	0.0077	(3.40)	0.0089	(1.91)
7	0.0074	(3.36)	0.0076	(1.78)
8	0.0065	(3.24)	0.0044	(1.29)
9	0.0062	(3.29)	0.0041	(1.22)
10	0.0039	(2.02)	0.0021	(0.46)
Low-high TK	0.0130	(4.73)	0.0247	(4.43)
N	355		65	

the benchmark of 60 months to 48 and 36 months. For bubble periods I expect lower returns to the low-high TK the longer the construction window, because longer equal weighted return distributions cannot capture the explosive nature of prices. During non-bubble periods the returns to the low-high TK portfolio are slightly higher compared to the construction window of 60 months and are still highly significant. Like for the benchmark case, the returns on the portfolio are higher during bubbles. Taking the results from Table 5 into consideration, the relationship between longer construction windows and lower returns on the low-high TK portfolio during bubbles seems to hold as predicted. The second technical robustness check for the computation of the prospect theory values deals with the returns used. Instead of excess returns over the market, I now use returns in excess over the risk-free rate and raw returns. Using returns in excess of the risk-free rate leads to a similar decrease in returns on the low-high TK portfolio as using raw returns. The impact of different specifications of returns is generally small. Dividing the sample into two sub-periods generates more interesting results. Since the bubble periods are already a sub periods, it is not reasonable to analyse them together with non-bubble periods for this test. The returns on the portfolio are lower for the period before the new millennium. In both

Table 6: Robustness tests (Based on Barberis et al., 2016, p.3088).

The table reports the average returns of a low-high prospect theory value portfolio, formed each month by going long (short) in stocks in the lowest (highest) TK decile. In the panel construction window the TK values are computed with different window sizes for the return distributions. The panel returns tests the robustness for raw returns and returns in excess of the risk-free rate. Time period splits the sample into two sub periods. In the last panel, I exclude stocks that fall under 5 Dollars in the month before the computation of TK values.

		Low-high TK			
		Non-Bubble		Bubble	
Construction window	48 months	0.0117	(4.26)	0.0250	(4.44)
	36 months	0.0112	(4.07)	0.0255	(4.57)
Returns	Raw returns	0.0131	(4.11)	0.0258	(4.12)
	Returns in excess of the risk-free rate	0.0133	(4.20)	0.0260	(4.13)
Time periods	1983-2000	0.0103	(2.92)		
	2001-2017	0.0150	(3.74)		
Exclude low price stock	Price >= 5$	0.0109	(6.43)	0.0153	(3.17)

periods, the returns are lower than the average in Table 4 because I excluded bubble periods for this robustness test. The next robustness test shows that low price stocks could be part of the explanation for the return on the portfolio. The magnitude of returns drops after excluding stocks that fall below a price of five dollars in the months before computation of the prospect theory value. Low priced stocks tend to be more illiquid and less subject to the forces to arbitrage. They are also more risky and could have a more skewed distribution. Both of these characteristics attract investors with prospect theory preferences; therefore low price stocks should have higher prospect theory values. Individual investors in general seem to prefer lotter like stocks. Kumar (2009) analyses the propensity to gamble that different types of investors exhibit. He also finds that investors consider a stock a lottery especially if it has a low price. In summary, the exclusion of low price stocks should have a negative effect on the returns of the low-high TK portfolio. Since the returns on the portfolio are still high and significant after excluding low price stocks, prospect theory preferences seem to influence prices in general and not only for a small part of the stock market. All robustness tests confirm the initial observation that there are significant abnormal returns to a low-high TK portfolio during non-bubble periods. In addition, the returns to the portfolio are larger during bubbles, but suffer from lower significance. How do the returns to the portfolio develop during the tech bubble? Figure 7 illustrates the returns on the low-high TK portfolio for every month date-stamped as a bubble, from December 1995 up to September 2000. It is noticeable that the returns to the portfolio are extreme for most of the tech bubble.

There is one month in 1997 when they drop below minus eight percent, while they are especially positive in months from 1998 onwards. This fits earlier results and the magnitude of returns seems to coincide with the progress of the bubble itself, which hits its highest levels in the course of 1998[20]. In the analysis for the portfolio sorts during bubbles in Table 5, I concluded that the returns on the low-high TK portfolio seem to be driven mainly by the positive returns of the lower TK deciles. These returns also seem to be highly volatile. The high average returns seem to come from fewer extreme values rather than a constant trend. Also the low significance in the high TK deciles could be a problem for the implementation of the strategy during bubbles. Portfolio sorts are not appropriate to test a relationship between prospect theory value and returns. This leads up to the question whether the prospect theory value of a stock in fact has predictive power for returns in the cross section.

6.4. Prospect theory and stock returns in the cross section

In this section, I test my two predictions using the Fama-MacBeth approach. They develop a method for testing the relationship between risk and return using three steps. In the first step, they run a time-series regression to obtain beta coefficient estimates. They then use these estimates to a run cross-sectional regression each month. The coefficients and standard errors are then obtained as the time-series average. This approach is frequently applied in asset pricing research. For my analysis only the last two steps are necessary, since I am not trying to measure the relationship between return and risk and therefore do not need market beta-regression coefficients. The Fama-MacBeth approach is very useful, because it allows capturing the cross-sectional variation of factors. I can therefore test my predictions while controlling for variables known to explain returns. Starting in January 1983 and ending in December 2017, I run a regression of the prospect theory value TK and my control variables mentioned in Table 2 on the percentage return in that month. The coefficients

[20] See Figure 5, where the BSADF test statistic reaches its highest level in 1998.

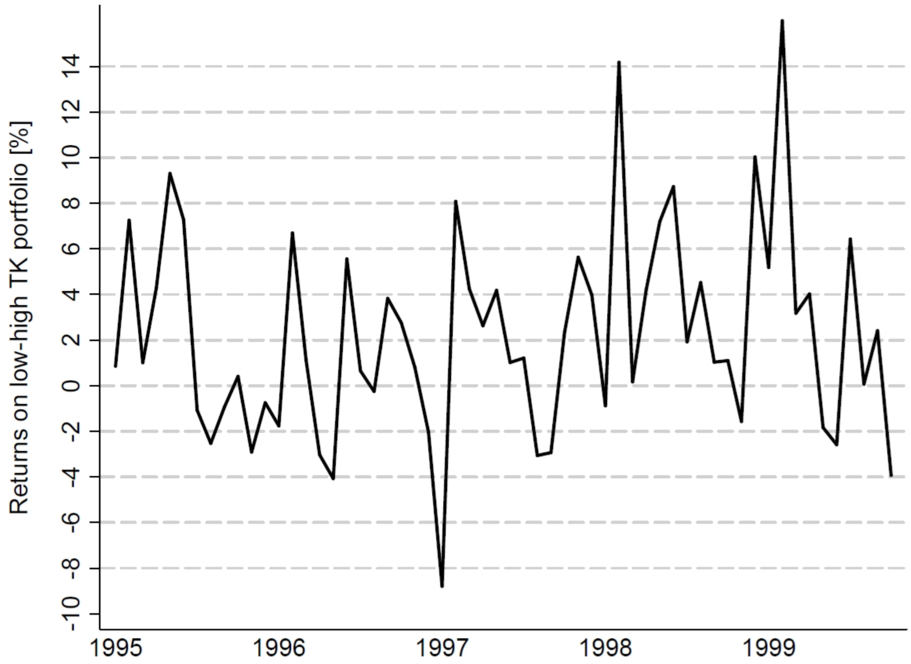

Figure 7: The low-high TK portfolio during bubbles.

The figure shows the returns of a low-high prospect theory value portfolio, formed each month by going long (short) in stocks in the lowest (highest) TK decile. The sample runs from 1995m12 up to 2000m9, which falls into bubble periods detected applying a method by Phillips et al. (2015b). The returns are monthly percentage returns.

and standard errors reported are the time-series average of these monthly regressions. Panel A of Table 7 reports the results of this exercise for the months not stamped as a bubble using the method by Phillips et al. (2015a), while Panel B reports the results for the months stamped as a bubble period using the same method. In each panel, I run the regressions with five different specifications. Model (1) is the first test of my predictions using only TK as an independent variable. Model (2) brings in controls for market beta, size, book-to-market ratio, momentum effect and illiquidity. Model (3) and Model (4) include the control for short term reversal, rev, and for long term reversal, ltrev, together with the maximum and minimum daily return variable, respectively. I do this to highlight the effect especially rev has on the predictive power of TK. Model (5) includes the control for skewness of past returns. The results in Panel A support the original prediction by Barberis et al. (2016). The coefficient on TK is negative, indicating that an increase in TK measured at the beginning of the month leads to a decrease in returns. This effect is not only sizeable, but also economically significant. After including the controls for size, beta, the book-to-market ratio and the momentum effect the coefficient on TK remains significant and loses only little in magnitude. I also find that including the control for the short term reversal effect, Rev, has a strong impact on the magnitude and significance of TK. Comparing the coefficients of TK in Model (2) and Model (3) shows that the introduction of rev leads to a decrease in magnitude with the t-statistic shrinking. Why should the predictive power of TK for returns be affected most by the return in month t-1? Barberis et al. (2016) argue that investors get the most recent monthly return from a different source than returns that are further in the past. If this is correct, investors might attribute the previous month's return more weight than the other returns in the distribution. In that case, a high return in the previous month would overshadow possible low returns in the past and investors would be incentivised to buy the stock, causing it to become overvalued and earn low returns. This would also be an alternative explanation of the short term reversal effect. The most negative daily return in month t-1 has a similar and strong negative effect on subsequent returns. The strong effect of the control variable rev could lead to the conclusion that TK's predictive power is due to reversal, but Barberis et al. (2016) find that leaving out month t-1 in the construction of TK has no large impact. Skewness should influence the results to some degree, because probability weighting favours extreme outcomes with low probability. Therefore a skewed distribution with a heavy weigh on the right tail should be preferred by investors with prospect theory preferences. As expected the inclusion of the skewness control in model (5) leads to a slight decrease of the significance on TK. I now turn to the results for the Fama-MacBeth regressions during bubbles in panel B. This is the test of my main prediction that high (low) prospect theory values stocks earn lower (higher) returns in bubble periods compared to non-bubble periods. The results suggest that, in the cross-section, the prospect theory value does not have more predictive power for returns during bubbles. The significance of TK is very similar to non-bubble

Table 7: Fama-MacBeth regressions (Based on Barberis et al., 2016, p.3090).

The Table shows the results of the Fama-MacBeth approach. In each cross-sectional regression, percentage return is the dependent variable. TK is a stock's prospect theory value, measured at the beginning of the month using a distribution consisting of 60 monthly returns (see equation (7)). Beta is a stock's market beta, computed with the returns of 100 portfolios formed on size and pre-beta in June of each year, following Fama and French (1992). Size is the log equity market value at month t-1; the product of common shares outstanding in December of year t-1 and monthly stock price. Bm is the log book-to-market ratio calculated as the difference of the log book value and the log equity market value in December of year t-1. Book value is computed following Daniel and Titman (2006). Mom is a stock's cumulative return from month t-12 until t-1. Illiq is the illiquidity measure of Amihud (2002), scaled by 105. Rev is a stock's return in month t-1. Ltrev is a stock's cumulative return from month t-60 until t-13. Max is the highest daily return in month t-1, while Min is the negative of the lowest daily return. Skew is the sample skewness of monthly returns over the past five years. TK, Rev, Mom, Max and Min are scaled by factor 100. The entire sample starts in January 1983 and ends in December 2017. Panel A shows the results of the Fama-MacBeth regressions for the months (N=355) not stamped as a bubble using a method by Phillips et al. (2015a), while Panel B shows the results for months (N=65) stamped as a bubble using the same method. The numbers in brackets show the t-statistics.

Panel	A. Non-Bubble					B. Bubble				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
TK	-0.121***	-0.082***	-0.057**	-0.080***	-0.092***	-0.251***	-0.226***	-0.192***	-0.190**	-0.243**
	(-4.68)	(-4.49)	(-3.31)	(-3.68)	(-3.45)	(-4.67)	(-5.51)	(-4.77)	(-3.36)	(-2.94)
Beta		0.247	0.306	0.516*	0.411		0.675	0.763	0.758	0.573
		(0.95)	(1.20)	(2.15)	(1.58)		(1.24)	(1.46)	(1.90)	(1.45)
Size		-0.115***	-0.113***	-0.134***	-0.116***		-0.126	-0.113	-0.128	-0.073
		(-4.44)	(-4.39)	(-4.81)	(-4.07)		(-1.25)	(-1.15)	(-1.37)	(-0.76)
Bm		0.189***	0.204***	0.170***	0.160***		0.194	0.208	0.170	0.169
		(3.83)	(4.17)	(3.99)	(3.60)		(1.17)	(1.28)	(1.44)	(1.50)
Mom		0.002	0.001	0.002	-0.001		0.007	0.006	0.007	0.005
		(1.42)	(0.64)	(1.04)	(0.57)		(1.82)	(1.58)	(1.68)	(1.05)
Illiq		0.226***	0.232**	0.339***	0.344***		0.217	0.254	0.279	0.264
		(2.85)	(2.95)	(4.09)	(3.33)		(1.34)	(1.57)	(1.57)	(1.32)
Rev			-0.049***	-0.056***	-0.058***			-0.054***	-0.066***	-0.073***
			(-13.90)	(-14.17)	(-12.62)			(-5.84)	(-6.60)	(-6.38)
Ltrev				-0.075	-0.071				-0.049	0.128
				(-1.01)	(-0.84)				(-0.23)	(0.49)
Max				0.007	0.015**				0.017	0.016
				(1.94)	(2.90)				(1.48)	(1.12)
Min				-0.052***	-0.053***				-0.045**	-0.049**
				(-8.30)	(-6.76)				(-3.79)	(-3.74)
Skew					0.117*					0.187
					(2.56)					(1.56)
R-squared	0.015	0.044	0.050	0.057	0.062	0.012	0.049	0.056	0.065	0.073

t-statistics in parentheses *p<0.05, **p<0.01, ***p<0.001

periods when not controlling for other return predictors in model (1). In some models it is actually greater than during non-bubble periods, but for the most sophisticated specification TK exhibits slightly weaker predictive power for returns. This leads to the conclusion that the predictive power of TK is roughly equal during bubble and non-bubble periods. There is no strong evidence that my prediction holds on average. Another interesting finding is that the coefficients of TK are greater in magnitude during bubbles. But comparing the size of coefficients between the two different sub samples is not a statistical test, even if the independent variables are standardized. In general it is important to note that a precise interpretation of the predictive power of variables in a Fama-MacBeth regression is only valid under the assumption of linear relationship. In my case the variables are scaled differently so the interpretation has to be restricted to the t-statistics. Nonetheless I can compare the effects of the same variables between the different panels. Bubble periods are extreme market periods characterised by overvaluation, so the predictive power for some variables should be expected to change. While the book-to-market ratio, for example, has significant positive predictive power for returns during non-bubble periods in panel A, it has no power in panel B. The book-to-market ratio is a measurement of the valuation of a firm's assets. A higher book-to-market ratio signals undervaluation. During non-bubble periods an undervaluation should lead to a future increase in prices and higher returns, which is suggested by panel A. The book-to-market ratio could lose its predictive power for returns during bubbles, because on average it could be more difficult for investors to determine the fundamental value of stocks. Just like in panel A, the short term reversal still seems to play a role during bubbles. Investors continue to buy stocks that performed well in the last month, causing them to be overvalued and earn low returns. Although the control for the momentum effect remains to have no significant influence on returns, it seems to be somewhat more robust. These results make intuitive sense for the market conditions during bubbles, where investors are searching for the best stocks to profit from, but it also describes the buying behaviour during non-bubble periods. It is important to notice that the predictive power of all control variables is weaker during bubble periods. Given this fact, the predictive power of TK seems rather high. Maybe the mispricing during bubbles makes it harder for any variables to predict returns consistently. For TK it seems that both standard errors and coefficients are larger on average. Taking the results from the portfolio sort and Figure 7 into account, the relationship between TK and returns might be stronger during bubbles, but only for a smaller set of stocks or particular time and not on average. This is entirely speculative though. In summary, I do not find support for the prediction that high prospect theory value stocks earn lower returns during bubbles, but can confirm the original prediction of its general effect. I the next section, I try to answer part of the question which stocks are favoured by prospect theory, namely if the limits to arbitrage have an impact on TK's relationship with returns. This could also shed some light on whether the limits to arbitrage of a stock have more influence on prices during bubbles.

6.5. Prospect theory value, limits to arbitrage and individual investors

Generally, stocks with high limits to arbitrage tend to be owned by less sophisticated investors. Since it is more costly for arbitrageurs to employ strategies that correct the mispricing, the prices of these kinds of stocks should be affected by the behaviour of individual investors to a greater extent. The prospect theory value is assumed to be a reflection of individual investor's behaviour. Therefore the relationship between high TK values and low returns should be stronger for stocks exhibiting higher limits to arbitrage. The limits to arbitrage are hard to measure, but in section 4.1.3, I presented several approximations. To test whether the effect of TK depends on the proxies, a Fama-MacBeth regression with interaction terms is employed. Each regression includes TK, the previous control variables and adds an interaction term of TK with the proxy for limits to arbitrage. The first proxy for the limits to arbitrage is Mashruwala et al. (2006) arbrisk variable. It captures the idiosyncratic volatility of a stock's past returns. Arbitrage often requires being able to hedge against contrary outcomes. The basic idea for the arbrisk variable is that an arbitrageur can only find substitutes for hedging of the risk explained by the standard market model. She is then exposed to the risk that she cannot hedge, the excess volatility from her model. The variance of residuals of a standard market model regression, or of models with more factors, therefore is a proxy for arbitrage risk. Arbrisk therefore is the residual variance from a standard market model regression of a stock's return on the market return from month t-36 to month t-1. Although calculating idiosyncratic volatility with daily returns would probably be more accurate, I do not have the resources to execute the required number of regressions in a large dataset. The second proxy for limits to arbitrage is the dollar trading volume, an indicator for transaction costs. Stoll (2000) finds that the ask-bid price spread is negatively related to measures of trading activity, such as the average daily dollar trading volume. Theoretically there should be fewer frictions when implementing arbitrage strategies for stock with high trading volume. Following Mashruwala et al. (2006), the measure of transaction costs tvol is the average daily dollar trading over the past year. Size is also an approximation of transaction costs and the limits to arbitrage, since stocks with high market capitalization should be easier to loan or short. The last proxy for the limits to arbitrage is the illiquidity measure of Amihud (2002), which is the ratio of absolute daily stock return to trading volume. It measures the marginal impact of one dollar trading volume on prices. More illiquid stock should be more difficult to incorporate into arbitrage. Table 8 presents the results of the regression using the interaction terms of TK with the proxies for limits to arbitrage. For non-bubble periods, the results confirm the prediction that stocks which are more likely to have higher limits to arbitrage also exhibit a stronger predictive power of TK for returns. The interaction terms all show the expected signs on coefficients. The interaction terms with the measures for arbitrage risk and illiquidity both have significant negative effect on returns, because they measure frictions for arbitrage strategies. The terms with the measures for dollar trading volume and size have a significant positive effect, because they make it easier for arbitrageurs to trade against mispricing introduced by individual investors. During bubbles periods there seems to be no reinforcing effect of the limits to arbitrage to the effect of TK. The interaction terms with arbitrage risk, trading volume, size and illiquidity are not significant anymore. These don't seem to be limits to arbitrage influencing the relationship between prospect theory and returns. Are the limits to arbitrage even rising during bubbles? The arbitrage risk is a highly significant during non-bubble periods, but does this limit to arbitrage rise during bubble periods? I test for difference between means for bubble and non-bubble periods. Indeed I find that the arbitrage risk is significantly higher during bubble periods. This suggests that although the limits to arbitrage are rising, they might not affect the relationship between prospect theory and returns. Of course this is just a narrow view of a much broader problem in research. There are many other limits to arbitrage that would need to be considered. Instead of directly measurable limits to arbitrage, there are also models explaining mispricing by investor behaviour. In the model of Abreu and Brunnermeier (2003) arbitrage is delayed, because arbitrageurs try to profit from

Table 8: Prospect theory value and proxies for limits to arbitrage (Based on Barberis et al., 2016, p. 3096).

The table reports the result of Fama-MacBeth regressions with the independent variable percentage return. The entire table with all control variables can be found in Appendix E. Arbrisk is the residual variance from a standard market model regression on market returns from month t-36 to t-1. Tvol is the average daily dollar trading value over year t-1. Size is the log equity market value at month t-1; the product of common shares outstanding in December of year t-1 and monthly stock price. Illiq is the illiquidity measure of Amihud (2002), scaled by 105. TK is scaled by 100. Panel A shows the results of the Fama-MacBeth regressions for the months (N=355) not stamped as a bubble using a method by Phillips et al. (2015a), while Panel B shows the results for months (N=65) stamped as a bubble using the same method. The numbers in brackets show the t-statistics.

Panel	A. Non-Bubble				B. Bubble			
	(1) Arbrisk	(2) Tvol	(3) Size	(4) Illiq	(1) Arbrisk	(2) Tvol	(3) Size	(4) Illiq
TK	0.140***	-0.310***	-0.213***	-0.068**	0.084	-0.169	-0.195	-0.184**
	(8.17)	(-5.41)	(-5.77)	(-3.16)	(1.87)	(-1.20)	(-1.86)	(-3.26)
TKarbrisk	-0.022**				-0.032			
	(-3.07)				(-1.52)			
TKtvol		0.017***				-0.001		
		(4.34)				(-0.11)		
TKsize			0.023***				0.001	
			(5.06)				(0.08)	
TKilliq				-0.115***				-0.047
				(-4.85)				(-1.58)

t-statistics in parentheses *p<0.05, **p<0.01, ***p<0.001

the bubble. Once the majority of arbitrageurs enter in a selling effort, the initial mispricing is corrected as the bubble collapses. In other words, arbitrageurs could knowingly ignore the mispricing and would not necessarily care about the measurable limits to arbitrage during bubbles. While delayed arbitrage is a possible cause for mispricing and bubbles, it is hard to test empirically. Only investor level data shines some light on the issue. For example, Griffin et al. (2011) find that institutional investors i.e. arbitrageurs were more heavily invested in technology stocks during the tech bubble than individual investors. It is also a coordinated selling effort of institutions in mid-March of 2000 that drives the collapse of the bubble and not individual investors. Individual investors should have more influence on prices during bubbles, because arbitrageurs have more interest in riding it than correcting the mispricing. Other proposition to why the tech bubble built up is the increase number of individual investors. Brennan, for example, argues for an increase in individual investors during the 1990s due to changes in pension plan policy, from benefit pension plans to contribution plans. Benefit pension plans were managed by the employer and protected against failure, while contribution plans put the responsibility for the investment in the hands of the employee. Contribution plans also began to lean more towards equity investments. In addition, households in the 1990s followed the same trend, increasing the share of equities in their portfolio from 13.2% in 1990 to 32.5% in 1999[21]. Although the discussion on limits to arbitrage and individual investors during bubbles shows that overall individual investor's preferences should be able to manifest themselves to a greater degree, this seems to have no impact on the prospect theory variable. How does the mechanism of the prospect theory value work at a more fundamental level? Maybe the mechanism offers some insights on the issue. In the next section, I therefore analyse the effect of the components of prospect theory on returns individually.

6.6. On the relevance of prospect theory components

In order to better understand what drives the results, I test the importance of the three components of prospect theory reflected in the equations of the value function and the probability weighting function. Each of the components reflects certain behavioural patterns individuals exhibited in experiments. By analysing the components it could be possible to infer which of these patterns is especially important for the predictive power of the prospect theory value. Since I focus on the comparison of bubble and non-bubble periods, the following analysis can also show whether the relative importance of the components undergoes significant changes through time. Hence the analysis could give insights into the evolution of behavioural patterns of investors, de-

[21] See Brennan, 2004, p.6.

pending on the situation they are facing in the stock markets. The empirical test is based on the Fama-MacBeth approach used in Table 7. Specifically it is Model 5 of Table 7, which includes all control variables. But instead of using all the parameters provided by Tversky and Kahnemann (1992) to construct the variable TK, parameters are chosen such that only one specific component of prospect theory is active. This means that either probability weighting, loss aversion or diminishing sensitivity is active. In order to study the predictive power of TK constructed with a focus on probability weighting, the parameters for loss aversion and diminishing sensitivity are set to one. So TK in the column labelled PW (probability weighting) is computed with the parameters λ=1;α=1;γ=0.61;δ=0.69 . The same rationale applies to the other columns in Table 7. TK in the column labelled LA (loss aversion) corresponds to applying the parameters λ=2.25;α=1;γ=1;δ=1 and TK in the column labelled DS (diminishing sensitivity) to the parameters λ=1;α=0.88;γ=1;δ=1. In contrast to Table 5, the independent variables in Table 7 are normalized to have a mean of zero and a standard deviation of one. Barberis et al. (2016) present a very similar analysis, but also look at interactions of the three components. The results I get for this extension do not provide much additional information. Therefore I don't present them here and avoid making the table more complex than needed, especially since there are two sections of results comparing bubble to non-bubble periods. The first three columns in Table 9 show the result for non-bubble periods. It seems that loss aversion by itself explains a great deal of the effect of prospect theory. Probability weighting and diminishing sensitivity on the other hand become insignificant predictors after including all control variables. Especially the control for the long term reversal has a great impact on their t statistics and on the sign of the coefficients. Before including ltrev, probability weighting is actually a significant predictor of returns on its own, with a negative sign on the coefficient. Maybe the turning of the sign is due to the high correlation between ltrev and TK constructed with probability weighting and diminishing sensitivity. However, I do not find any evidence for collinearity after testing for the VIF measure. Taking all this into account, I cautiously conclude that the positive sign on the coefficients should not be assigned too much weight, since they are negative for any model excluding ltrev. In general probability weighting by itself is a more significant predictor of returns than diminishing sensitivity. This suggests that the skewness of a return distribution should also be somewhat attractive to investors with prospect theory preferences. If this is true, then the combination of probability weighting and loss aversion in the TK variable should result in the most predictive power for returns. I indeed find that this TK variable has the most predictive power for returns, even higher than the original specification. The respective results are located in Appendix F. During bubbles the relative importance of the prospect theory components does not seem to change. Loss aversion remains to be the most reliable predictor of returns by itself. Overall, the effects are less significant, which reflects the result form Table 7 to some extent. If loss aversion is important during both bubble and non-bubble periods, what does this mean for the return distribution investors acting according to prospect theory prefer? Under loss aversion a loss has a greater weight than a gain of similar size. Therefore investors should prefer return distributions with low standard deviation. In summary then, the optimal distribution of returns should first and foremost have low standard deviation and combine this with some skewness. These results are just guesses, as a comparison of t statistics across the columns is no conclusive test of the matter. Since loss aversion seems to play such a big role, it is possible that the effect on returns is not unique to prospect theory. After all, expected utility also has a concave value function, indicating that individuals are loss averse. With this alternative setup, the prediction would be that high expected utility stocks earn low subsequent returns. In the next section I test this conjecture using the expected power utility formula from Barberis et al. (2016).

6.7. Prospect theory versus expected utility

The test for expected utility will consist of the Fama-MacBeth regressions in Table 7, but this time I compare the effect of the TK variable with the effect of the variable EU, the power expected utility:

$$EU = (1/60)\sum_{i=-m}^{n}(1+r_i)^{1-\theta}/1-\theta, \tag{26}$$

where the parameter theta is initially set at 5, indicating a high degree of loss aversion. EU is constructed using the same return distributions as the TK variable, where each return is weighted with equal probability. Of course it is stated in terms of final wealth; therefore it is $1+r_i$ instead of just r_i. This means that the loss aversion has the most impact in the region near-1. There, the curve for expected utility is especially steep. In addition meaning of subscripts m and n is in accordance with equation (6), m is the lowest ranked return and n is the highest ranked return in the distribution. Table 10 presents the results of this exercise. Similar to the prospect theory value, I find a negative realationship between the expected utility and returns. This would mean that the effect observed in Table 7 is not unique to prospect theory. A possible cause for this has been discussed earlier. Both models for decision making under risk share the incorporation of loss aversion. Maybe loss aversion is the dominating force behind the results and it does not matter whether returns are measured over final wealth or gains and losses. This result also would stand in contra to Barberis et al. (2016) that find no relationship between returns an expected utility. In fact, the regressions leading up to Table 7 show that the relationship between expected utility and returns is spurious. In the model that only tests expected utility and returns, the relationship is positive. This positive relationship holds until the control for the long term reversal is included. Again one could cautiously conclude that the elevated correlation between EU and ltrev has a confounding impact on the regression. Therefore the interpretation of the result form Table 10 is speculative. To quickly test what constitutes a high

Table 9: Relative importance of prospect theory components (Based on Barberis et al., 2016, p.3101).

The table presents the results of Fama-MacBeth regressions on the monthly percentage return as dependent variable. The independent control variables are the same as in Table 7, but normalized to have mean zero and standard deviation one. In each column, TK is computed with only one active parameter for probability weighting (PW), loss aversion (LA) and diminishing sensitivity. The sample runs from January 1983 until December 2017, but is split into bubble (N=355) and non-bubble (N=65) periods. T-statistics are in parentheses.

	Non-Bubble			Bubble		
	PW	LA	DS	PW	LA	DS
TK	0.071	-0.336***	0.164	0.179	-0.813**	0.091
	(1.17)	(-3.34)	(0.52)	(1.00)	(-2.76)	(0.11)
Beta	0.202*	0.128	0.204*	0.337*	0.180	0.351*
	(2.49)	(1.89)	(2.55)	(2.33)	(1.76)	(2.33)
Size	-0.320***	-0.225***	-0.341***	-0.320	-0.103	-0.375*
	(-5.64)	(-3.82)	(-5.99)	(-1.79)	(-0.56)	(-2.02)
Bm	0.110***	0.135***	0.113**	0.031	0.124	0.042
	(3.04)	(3.83)	(3.14)	(0.27)	(1.33)	(0.37)
Mom	-0.040	0.077	-0.089	-0.012	0.286	0.033
	(-0.53)	(1.28)	(-0.66)	(-0.07)	(1.68)	(0.08)
Illiq	0.209***	0.202***	0.210***	0.189	0.160	0.193
	(4.23)	(4.01)	(4.22)	(1.80)	(1.45)	(1.86)
Rev	-0.709***	-0.688***	-0.727***	-0.872**	-0.806**	-0.861**
	(-13.36)	(-14.34)	(-12.46)	(-7.00)	(-6.78)	(-4.59)
Ltrev	-0.241***	-0.006	-0.342	-0.509**	-0.106	-0.457
	(-3.48)	(-0.09)	(-1.27)	(-2.86)	(0.49)	(-0.65)
Max	0.518*	0.364	0.531*	1.252	0.589	1.360
	(2.13)	(1.69)	(2.21)	(1.78)	(1.03)	(1.87)
Min	-0.225***	-0.275***	-0.224***	-0.131	-0.235**	-0.135
	(-6.11)	(-8.73)	(-6.11)	(-1.77)	(-4.07)	(-1.79)

t-statistics in parentheses *p<0.05, **p<0.01, ***p<0.001

expected utility stock, I sort stocks into deciles each month and calculate the expected value for the different control variables. Then I analyse the time series average of these statistics, with a special focus on the return measures. The average returns in the high expected utility decile are lower than those of the first decile. In addition, the long term reversal is decreasing in the expected utility decile. Therefore I would expect a positive relationship between expected utility if anything, as the counterpart to the effect of prospect theory value. I don't present the resulting table here, because the point was to quickly back test which relationship between expected utility and returns could make sense. However, the corresponding table can be found in Appendix G. In any case, the prospect theory has higher predictive power over returns and the estimation process is more robust. In summary, I do not find definitive proof that expected utility has a similar effect on returns like prospect theory. Generally, the prospect theory value seems to reflect investor behaviour in the U.S. stock markets more accurately. Then, the following question arises from these results: Which return characteristics do investors with prospect theory preferences like the most? Previous results suggest that low standard deviation and skewness should be preferred. In order to get a broader picture regarding the return characteristics, the next section presents the results of a decile sort on prospect theory value, similar to the one for expected utility.

6.8. Decile sorts on prospect theory value

Each month I sort stocks into deciles based on their prospect theory value. For each decile the average value for several of the control variables is computed. The sample periods starts in January 1983 and ends in December

Table 10: Prospect theory versus expected utility.

The Table reports the results of Fama-MacBeth regressions with percentage returns as the dependent variable. The control variables are the same as in Table 7. EU is a stock's expected utility based on its past return distribution of 60 months. The sample runs from January 1983 until December 2017, but is split into bubble (N=355) and non-bubble (N=65) periods. T-statistics are in parentheses.

	Non-Bubble		Bubble	
	(1)	(2)	(1)	(2)
TK	-0.092*** (-3.45)		-0.243** (-2.94)	
EU		-0.057** (-3.07))		-0.147** (2.69)
Beta	0.411 (1.58)	0.543* (2.23)	0.573 (1.45)	0.834* (2.04)
Size	-0.116*** (-4.07)	-0.119*** (-4.42)	-0.073 (-0.76)	-0.082 (-1.00)
Bm	0.160*** (3.60)	0.157*** (3.71)	0.169 (1.50)	0.127 (1.09)
Mom	-0.001 (-0.57)	-0.003 (1.31)	0.005 (1.05)	-0.004 (-0.95)
Illiq	0.344*** (3.33)	0.323*** (3.95)	0.264 (1.32)	0.249 (1.38)
Rev	-0.058*** (-12.62)	-0.060*** (-13.72)	-0.073*** (-6.38)	-0.075*** (-7.76)
Ltrev	-0.071 (-0.84)	-0.443*** (-5.04)	0.128 (0.49)	-0.950*** (-4.35)
Max	0.015** (2.90)	0.007 (1.84)	0.016 (1.12)	0.012 (1.21)
Min	-0.053*** (-6.76)	-0.050*** (-8.08)	-0.049*** (-3.74)	-0.039** (-3.31)
Skew	0.117* (2.56)	-0.005 (-0.18)	0.187 (1.56)	-0.027 (0.44)

t-statistics in parentheses *p<0.05, **p<0.01, ***p<0.001

2017. The reported values are the averages of the time series consisting of the average value per variable and month. In addition to sourcing from the control variables introduced earlier, the standard deviation is also included. Sd is computed as the standard deviation of monthly returns over the past three years. Including sd makes it possible to test whether loss aversion indeed plays a role when individual investors pick stocks. Table 11 presents the results of this exercise. First, prospect theory values are negative on average, even in the high TK decile. The prospect theory value of a stock is also increasing in size and decreasing in the book-to-market ratio. This is a first indication that large and less volatile stocks are preferred by individual investors with prospect theory preferences. In accordance with this conjecture, the measure for illiquidity is decreases monotonically form decile one to decile ten. The regressions from Table 7 have shown that past return measures have a large impact on the predictive power of the prospect theory value. Table 11 makes the same point in a different way. Rev and ltrev increase monotonically from decile one through decile

Table 11: Prospect theory value portfolios (Based on Barberis et al., 2016, p.3100).

The Table reports the time series average of the monthly mean statistic of variables for deciles formed on TK. The description of variables can be found in Table 2. Sd is the standard deviation of monthly returns over the past five years. The sample starts in January 1983 and ends in December 2017.

Deciles	TK	Beta	Size	Bm	Rev	Illiq	Ltrev	Skew	Sd
1	-0.138	1.200	4.317	-0.285	-0.007	0.567	-0.403	0.166	0.184
2	-0.107	1.159	5.313	-0.445	0.003	0.211	0.126	-0.044	0.153
3	-0.093	1.120	5.789	-0.519	0.006	0.127	0.283	-0.166	0.136
4	-0.084	1.089	5.984	-0.562	0.008	0.093	0.384	-0.183	0.124
5	-0.076	1.055	6.107	-0.581	0.009	0.068	0.437	-0.145	0.114
6	-0.068	1.029	6.152	-0.589	0.010	0.055	0.468	-0.066	0.105
7	-0.062	1.001	6.218	-0.617	0.012	0.042	0.501	0.045	0.097
8	-0.054	0.986	6.317	-0.625	0.013	0.033	0.553	0.178	0.090
9	-0.046	0.961	6.487	-0.656	0.015	0.023	0.628	0.318	0.084
10	-0.032	0.989	6.772	-0.763	0.023	0.019	0.925	0.646	0.087
Total	-0.076	1.060	5.945	-0.564	0.010	0.124	0.390	0.075	0.117

ten. At first sight this could lead to the conclusion that TK really measures the effects of these reversals. Although these measures take significant bites out of the predictive power of TK, its negative relationship with returns remains significant. The results from Table 9 suggest that skewness should be preferred by investors with prospect theory preferences. As expected, skewness generally increases from decile one to decile ten. But the relationship is less straightforward than assumed. Apparently, returns in the first decile are skewed more than returns up to decile seven. This could be connected to the fact that the standard deviation in the first decile is at its highest point over all deciles. Returns exhibiting a high standard deviation are more likely to be skewed as well. The skewness of returns particularly increases from decile seven on, with the sharpest rise from decile nine to ten. The standard deviation also increases from decile nine to ten, while it decrease from decile one to nine. Again, this could be due to the relationship between skewness and standard deviation. However, the standard deviation remains below its average value. Stocks in the high prospect theory value deciles exhibit the characteristics expected earlier. They have a high return, high skewness and low standard deviation. Investors with prospect theory preferences should indeed like high returns combined with a low standard deviation, because they are loss averse. Loss aversion causes these investors to “feel” losses more heavily than gains, so they tilt towards low standard deviation. On the other hand investors with prospect theory preferences like skewness because of probability weighting. They overweight returns that are less likely to occur and tend to buy “stock lotteries”; high reward and low probability. Together with the results from previous tests, this leads to the conclusion that investors evaluating stock using prospect theory look for a certain mix of both worlds. High prospect theory value stocks are those which optimally combine skewness, volatility and high returns.

7. Conclusion and summary of results

There are several questions related to recent research I try to address with the analysis in this paper. The first goal was to test the model by Barberis et al. (2016). The fundamental assumption in this framework is that individual investors evaluate stocks based on the prospect theory value of their past return distributions. Individual investors add stocks with a high prospect theory value to their portfolio and sell those with low prospect theory value. Therefore stocks with high (low) prospect theory value become overvalued (undervalued) and earn low (high) subsequent returns. I find support for this prediction in the cross section of U.S. stock markets during the period of January 1983 to December 2017. This extends the evidence of the impact the preferences of individual investors have on prices in the stock market. Prospect theory seems to capture the behaviour and choices of individual investors to a certain degree. Based on the formal model of Barberis et al. (2016), I made the additional prediction that the relationship between prospect theory value and returns should be more pronounced during bubbles. This prediction was based on a two-fold argument: that the limits to arbitrage rise during bubbles and the number of individual investors increases. The limits to arbitrage are assumed to rise, because rational investors delay arbitrage and arbitrage becomes more costly. The increase in individual investors on the other hand comes from the attention grabbing nature of stocks during bubbles. Testing the prediction during bubbles is a way to find out whether the predictive power of the prospect theory value varies through time. It is also related to the question how investor behaviour changes depending on the market conditions. In order to test the prediction, I had to define what constitutes a bubble and apply a method to time stamp the start and collapse of the bubbles. The BSADF test developed by Phillips et al. (2015a) is tool for the detection of bubbles, which has been proven to work in many different markets. It focuses on the nature of the stochastic pro-cesses underlying the price development. The application of this method has led to the conclusion that two bubbles existed in the price dividend ratio of the S&P 500 during the time period of January 1983 to December 2017. The first corresponds to the lead up to Black Monday in October and spans from February 1987 to August 1987. The second coincides with the tech bubble during the late stages of the past century, ranging from December 1995 to September 2000. The implementation of the BSADF test in the context of the time series of the S&P 500 supports the reliability of the approach. The time stamping of the start and end of the detected bubbles seems to be consistent across different studies, if the underlying time series used is similar. But I also find some issues for the specification of the test, especially concerning the necessary assumption of the minimum bubble duration, where the application of economic intuition becomes necessary. In the main empirical test I find no support for the prediction that high prospect theory value stocks earn lower subsequent returns during bubbles compared to non-bubble periods. In contrast, the returns to a strategy that buys stocks with low prospect theory value and sells stock with high prospect theory value are higher during bubbles. However both tests suggest that the significance could be a problem and the relationship between prospect theory value and returns is, on average, not stronger during bubbles. The changes in significance are not large enough to justify an argument for an increase or decrease in the predictive power of the prospect theory value. The main analysis also sheds some light on the predictive power of other commonly used variables in financial research. Variables drawing their power from a measurement of over- or undervaluation suffer noticeably during bubbles. The same goes for predictors that are based on past performance, although they mostly remain significant. The short term reversal especially has shown a very robust relationship with returns. In light of the findings, the almost constant effect of prospect theory on returns shows that behavioural explanations of asset prices can offer valuable insights. Being able to time stamp bubbles in the stock markets might make it possible to better understand the forces determining prices. This is an interesting avenue for future research in asset pricing. The analysis on the accuracy of the limits to arbitrage argument has also proven insightful. The first and maybe most important question is how to measure the limits to arbitrage. Most of the research is focussed on approximations for costs or risks involved with arbitrage. But they might not be able to explain many risks that arbitrageurs face in the stock market. The delayed arbitrage model by Abreu and Brunnermeier (2003) is just one example for this. By applying several proxies for the limits to arbitrage I was able to show that stocks with higher limits to arbitrage are more prone to the effect of the prospect theory value. Future research in the measurement of limits to arbitrage could be helpful, although it is already quite extensive. Looking into each of the individual components of prospect theory has shown that loss aversion and probability weighting together predict returns with the highest significance. Because of loss aversion, investors with prospect theory preferences like stocks that have a low standard deviation. They also tend to prefer stocks with high skewness, because low probability outcomes receive more weight in the return distribution due to probability weighting. In general, investors seem to prefer a combination of low standard deviation and high skewness under prospect theory. There are several open questions with regards to prospect theory that future research could address. One important aspect is the definition of gains and losses. The modelling of gains and losses based on recent expectations is a concept that, with further testing, could prove valuable to applications of prospect theory. Empirically testing the reference points of investors could prove difficult though, since the necessary data has to be available on the investor level. Further research on the information sources could provide equally helpful guidelines in this matter. In the short review under section 2.3.2 several interesting questions arise. Where do investors get information on stock prices and returns? How does the presentation form influence the perceived attractiveness of a stock? An overview that connects the findings of the related research would be very helpful for the empirical application of prospect theory. It is also possible to consider other models of individual's choices under risk, as the section on expected utility has shown. The assumption that investors evaluate stocks based on a stock's past return distribution could maybe be lifted in such models. It is possible, that these models have even greater explanatory power for asset prices. After all, one of the most interesting findings of my analysis is that behavioural models like prospect theory might be more consistent predictors of stock prices over time than measures of valuation or risk-return relationships.

References

Abdellaui, M. Parameter-free elicitation of utility and probability weighting functions. *Management Science*, 46(11):1497–1512, 2000.

Abreu, D. and Brunnermeier, M. K. Bubbles and crashes. *Econometrica*, 71 (1):173–204, 2003.

Agnew, J. R. and Szykman, L. R. Asset allocation and information overload: The influence of information display, asset choice, and investor experience. *The Journal of Behavioral Finance*, 6(2):57–70, 2005.

Alexander, G. J., Jones, J. D., and Nigro, P. J. Mutual fund shareholders: Characteristics, investor knowledge, and sources of information. *Financial Services Review*, 7(4):301–316, 1998.

Ali, A., Hwang, L. S., and Trombley, M. A. Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics*, 69(2):355–373, 2003.

Amihud, Y. Illiquidity and stock returns cross-section and time-series effects. *Journal of Financial Markets*, 5(1):31–56, 2002.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299, 2006.

Asquith, P., Pathak, P. A., and Ritter, J. R. Short interest, institutional ownership, and stock returns. *Journal of Financial Economics*, 78(2):243–276, 2005.

Baillon, A., Bleichrodt, H., and Spinu, V. Searching for the reference point. Working paper., 2015.

Barber, B. M. and Odean, T. All that glitters: The effect of attention and news on the buying behaviour of in-dividual and institutional investors. *The Review of Financial Studies*, 2008.

Barberis, N. and Huang, M. Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, 98(5): 2066–2100, 2008.

Barberis, N. and Thaler, R. A survey of behavioral finance. *Handbook of the Economics of Finance*, 1:1053–1128, 2003.

Barberis, N., Mukherjee, A., and Wang, B. Prospect theory and stock returns: An empirical test. *The Review of Financial Studies*, 29(11):3068–3107, 2016.

Benartzi, S. and Thaler, R. H. Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, 110(1):73–92, 1995.

Berk, J. B. Sorting out sorts. *The Journal of Finance*, 55(1):407–427, 2000.

Bhargava, A. On the theory of testing for unit roots in observed time series. *The Review of Economic Studies*, 53(3):369–384, 1986.

Brav, A., Heaton, J. B., and Li, S. The limits of the limits of arbitrage. *Review of Finance*, 14(1):157–187, 2009.

Brennan, M. J. How did it happen. *Economic Notes*, 33(1):3–22, 2004.

Bruhin, A., Fehr-Duda, H., and Epper, T. Risk and rationality. Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4):1375–1412, 2010.

Busetti, R. and Taylor, A. R. Tests of stationarity against a change in persistence. *Journal of Econometrics*, 123(1):33–66, 2004.

Caspi, I., Katzke, N., and Gupta, R. Date stamping historical periods of oil price explosivity. 1876–2014. *Energy Economics*, 70:582–587, 2018.

Chen, J., Hong, H., and Stein, J. C. Breadth of ownership and stock returns. *Journal of Financial Economics*, 66(2-3):171–205, 2002.

Cheung, A., Roca, E., and Su, J. J. Crypto-currency bubbles: an application of the Phillips–Shi–Yu (2013) method-ology on Mt. Gox bitcoin prices. *Applied Economics*, 47(23):2348–2358, 2015.

Chu, Y., Hirshleifer, D., and Ma, L. The causal effect of limits to arbitrage on asset pricing anomalies (No. w24144). National Bureau of Economic Research, 2017.

Daniel, K. and Titman, S. Market reactions to tangible and intangible information. *The Journal of Finance*, 61(4):1605–1643, 2006.

D'Avolio, G. The market for borrowing stock. *Journal of Financial Economics*, 66(2-3):271–306, 2002.

De Bondt, W. F. and Thaler, R. Does the stock market overreact? *The Journal of Finance*, 40(3):793–805, 1985.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. Noise trader risk in financial markets. *Journal of Political Economy*, 98(4):703–738, 1990.

Diba, B. T. and Grossman, H. I. Explosive rational bubbles in stock prices. *American Economic Review*, 78(3):520–530, 1988.

Duffie, D., Garleanu, N., and Pedersen, L. H. Securities lending, shorting, and pricing. *Journal of Financial Economics*, 66(2-3):307–339, 2002.

Escobari, D. and Jafarinejad, M. Date stamping bubbles in real estate investment trusts. *The Quarterly Review of Economics and Finance*, 60:224–230, 2016.

Evans, G. W. Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review*, 81(4):922–930, 1991.

Fama, E. F. and French, K. R. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.

Fama, E. F. and French, K. R. Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, 51(1):55–84, 1996.

Fama, E. F. and MacBeth, J. D. Risk, return, and equilibrium empirical tests. *Journal of Political Economy*, 81(3):607–636, 1973.

Garber, P. M. Famous first bubbles. *Journal of Economic Perspectives*, 4(2): 35–54, 1990.

Glaser, M., Langer, T., Reynders, J., and Weber, M. Framing effects in stock market forecasts. The difference between asking for prices and asking for returns. *Review of Finance*, 11(2):325–357, 2007.

Glaser, M., Iliewa, Z., and Weber, M. Thinking about prices versus thinking about returns in financial markets. *The Journal of Finance*, 74(6):2997–3039, 2019.

Griffin, J. M., Harris, J. H., Shu, T., and Topaloglu, S. Who drove and burst the tech bubble. *The Journal of Finance*, 66(4):1251–1290, 2011.

Gromb, D. and Vayanos, D. Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics*, 66(2-3): 361–407, 2002.

Gromb, D. and Vayanos, D. Limits of arbitrage. *Annual Review of Financial Economics*, 2(1):251–275, 2010.

Gürkaynak, R. S. Econometric tests of asset price bubbles taking stock. *Journal of Economic Surveys*, 22(1):166–186, 2008.

Hirshleifer, D., Teoh, S. H., and Yu, J. J. Short arbitrage, return asymmetry, and the accrual anomaly. *The Review of Financial Studies*, 24(7):2429–2461, 2011.

Homm, U. and Breilung, J. Testing for speculative bubbles in stock markets a comparison of alternative methods. *Journal of Financial Econometrics*, 10(1):198–231, 2012.

Hu, Y. and Oxley, L. Are there bubbles in exchange rates? Some new evidence from G10 and emerging market economies. *Economic Modelling*, 64:419–442, 2017.

Jegadeesh, N. and Titman, S. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1): 65–91, 1993.

Kahneman, D. and Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–292, 1979.

Keller, K. L. and Staelin, R. Effects of quality and quantity of information on decision effectiveness. *Journal of Consumer Research*, 14(2):200–213, 1987.

Kim, J. Y. Detection of change in persistence of a linear time series. *Journal of Econometrics*, 95(1):97–116, 2000.

Köszegi, B. and Rabin, M. A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165, 2006.

Kumar, A. Who gambles in the stock market? *The Journal of Finance*, 64(4): 1889–1933, 2009.

Kumar, A. and Lim, S. S. How do decision frames influence the stock investment choices of individual investors. *Management Science*, 54(6): 1052–1064, 2008.

Lehmann, B. N. Fads, martingales, and market efficiency. *The Quarterly Journal of Economics*, 105(1):1–28, 1990.

Lin, W. and Meng, J. The Status Quo or Expectation. Estimating Investors' Reference Points on Stock Returns. Working Paper, 2015.

Lurie, N. H. and Swaminathan, J. M. Is timely information always better. the effect of feedback frequency on decision making. *Organizational Behavior and Human Decision Processes*, 108(2):315–329, 2009.

Markowitz, H. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

Marsh, T. A. and Merton, R. C. Dividend variability and variance bounds tests for the rationality of stock market prices. *American Economic Review*, 76 (3):483–498, 1986.

Mashruwala, C., Rajgopal, S., and Shevlin, T. Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs. *Journal of Accounting and Economics*, 42(1-2):3–33, 2006.

Menkhoff, L. and Schmeling, M. A prospect-theoretical interpretation of momentum returns. *Economic Letters*, 93(3):360–366, 2006.

Miller, E. M. Risk, uncertainty, and divergence of opinion. *The Journal of Finance*, 32(4):1151–1168, 1977.

Nagel, S. Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics*, 78(2):277–309, 2005.

Neumann, M., J.v. annd Oskar. *Theory of Games and Economic Behavior*. Princeton, NJ, Princeton University Press, 1944.

Nolte, S. and Schneider, J. C. How price path characteristics shape investment behavior. *Journal of Economic Behavior & Organization*, 154(33–59), 2018.

Novy-Marx, R. The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28, 2013.

Ofek, E. and Richardson, M. The valuation and market rationality of internet stock prices. *Oxford Review of Economic Policy*, 18(3):265–287, 2002.

Phillips, P. C. and Shi, S. Real time monitoring of asset markets bubbles and crises. *Cowles Foundation Discussion Paper (no.2152)*, 2018.

Phillips, P. C., Wu, Y., and Yu, J. Explosive behavior in the 1990s Nasdaq. When did exuberance escalate asset values. *International Economic Review*, 52(1):201–226, 2011.

Phillips, P. C., Shi, S., and Yu, J. Specification sensitivity in right-tailed unit root testing for explosive behaviour. *Oxford Bulletin of Economics and Statistics*, 76(3):315–333, 2014.

Phillips, P. C., Shi, S., and Yu, J. Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review*, 56(4):1043–1078, 2015a.

Phillips, P. C., Shi, S., and Yu, J. Testing for multiple bubbles: Limit theory of real-time detectors. *International Economic Review*, 56(4):1079–1134, 2015b.

Quiggin, J. A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343, 1982.

Rabin, M. and Weizsäcker, G. Narrow bracketing and dominated choices. *American Economic Review*, 99(4):1508–43, 2009.

Scheinkman, J. A. and Xiong, W. Overconfidence and speculative bubbles. *Journal of political Economy*, 111(6):1183–1220, 2003.

Shiller, R. C. Irrational exuberance. *Philosophy & Public Policy Quarterly*, 20 (1):18–23, 2000.

Shiller, R. J. The use of volatility measures in assessing market efficiency. *The Journal of Finance*, 36(2):291–304, 1981.

Shleifer, A. and Vishny, R. W. The limits of arbitrage. *The Journal of Finance*, 52(1):35–55, 1997.

Stoll, H. R. Presidential address: friction. *The Journal of Finance*, 55(4): 1479–1514, 2000.

Tom, S. M., Fox, C. R., Trepel, C., and Poldrack, R. A. The neural basis of loss aversion in decision-making under risk. *Science*, 315(5811):515–518, 2007.

Tversky, A. and Kahnemann, D. The framing of decisions and the psychology of choice. *Science*, 211(4481):453–458, 1981.

Tversky, A. and Kahnemann, D. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992.

West, K. D. A specification test for speculative bubbles. *The Quarterly Journal of Economics*, 102(3):553–580, 1987.

White, E. N. The stock market boom and crash of 1929 revisited. *Journal of Economic Perspectives*, 4(2):67–83, 1990.