ESSAYS IN EMPIRICAL FINANCE

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von

TILMAN H. DRERUP

aus DÜREN

Bonn 2015

Dekan:	Prof. Dr. Rainer Hüttemann
Erstreferent:	Prof. Dr. Erik Theissen
Zweitreferent:	Prof. Dr. Hans-Martin von Gaudecker

Tag der mündlichen Prüfung:	11.12.2014

To my parents.

Acknowledgements

This thesis has benefitted from comments and support by many individuals. Foremost, I want to express deep gratitude to my supervisor Erik Theissen. The times at his chair and in his lectures have kindled my interest in economics and finance. I thank him for his guidance, patience, and constructive comments throughout the years.

I owe a lot to my second advisor Hans-Martin von Gaudecker who has greatly shaped the way I think and approach research. Working with him and Benjamin Enke on the third chapter of this thesis has been both fun and a challenge. I also thank Rainer Haselmann and Hendrik Hakenes whose comments have improved the second chapter of this thesis.

I am grateful to the Bonn Graduate School of Economics (BGSE), in particular to the tireless efforts of Urs Schweizer, Silke Kinzig, and Pamela Mertens, for providing a great research environment and financial support.

My fellow Ph.D. students each deserve special mention. However, since you know who you are and hopefully remember the times we shared, let this be my thank-you to you all. Special thanks also goes to the BGSE football group. Weekly meetings on the pitch were one of the highlights of my time in Bonn. Thank you, Matthias Wibral, for organising these events.

Moreover, I thank my (human and canine) colleagues at IAME for many fruitful discussions and good times. I thank Holger Gerhardt for help in matters of style and entertaining discussions of all things Cupertino.

I thank E., K., K., and S. for keeping me smiling and G., B., and F. for continuously lifting me up. Moreover, I am thankful to J. for showing me that creation is sometimes an act of sheer will.

To Inga van den Bongard: Thank you for thoroughly reading my work. Your valuable comments and your enduring support during the good and the bad times contributed immensely to this thesis.

I am greatly indebted to my family for providing unconditional support. Thank you for accepting “I don’t know.” as a satisfactory answer to “When?”.

Contents

List of Figures

List of Tables

Introduction

The three chapters of this thesis contain contributions to distinct branches of the finance literature—corporate finance, accounting, and household finance. Although the chapters address questions with little overlap, they share one common element. All chapters emphasise the paramount role of expectations in financial markets.

Theoretical arguments, both rational and irrational (e.g., Fama, 1970; Shiller, 1981), suggest that investors' expectations are central to the formation of stock prices. Because expectations themselves are rarely directly observable, much research in empirical finance uses changes in stock prices to infer changes in expectations. The first two chapters of this thesis build on the appealing intuition of this approach. In both chapters, I inspect the behaviour of stock prices around events that prior literature considers value-relevant. In neither chapter, however, I treat changes in prices as ends in themselves. Rather, to understand why investors changed their expectations of firm value, both chapters attempt to simultaneously identify changes in fundamentally relevant information. While the first chapter takes a purely empirical, exploratory approach to this issue, the second chapter embeds the same rationale in a stylised theoretical model that allows me to put more structure on the empirical analysis.

The third chapter of this thesis analyses the role of expectations in investors' decisions to invest in the stock market. Expectations, in particular those about risk and return, are assigned a major part in models of stock market participation. Patterns in empirically elicited expectations, however, suggest that the latter are measured with error (e.g., Manski, 2004; Hurd, 2009). In the paper that underlies this chapter, Benjamin Enke, Hans-Martin von Gaudecker, and I argue that the magnitude of this measurement error can be used to uncover heterogeneity in households' choice behaviour. We empirically assess this hypothesis using data on expectations and stock market participation specifically collected for this purpose as part of a representative survey.

The following pages describe each chapter and its relation to the respective literature in more detail.

Chapter 1.[1] The first chapter is a contribution to the literature on shareholder activism by hedge funds. Evidence in Anglo-Saxon countries indicates that activist hedge funds create value by successfully affecting their target firms (e.g., Clifford, 2008; Brav et al., 2008; Klein and Zur, 2009). The empirical literature on hedge fund activism in Germany, however, is comparably scarce—even though the issue has been part of intense public and political debate. Since Anglo-Saxon and German institutional environments differ substantially, in particular in terms of ownership concentration and shareholder rights, it is questionable whether activism by hedge funds as minority shareholders can be equally successful in Germany.

In this chapter, I analyse whether and how hedge funds affect the firms they invest in in the context of the German capital market. Thereby, I extend both the general literature on shareholder activism by hedge funds and the literature for Germany in particular, specifically Bessler, Drobetz, and Holler (2013) and Mietzner and Schweizer (2014), in a number of ways.

Both Bessler, Drobetz, and Holler (2013) and Mietzner and Schweizer (2014) put their focus on the analysis of long-term changes in stock market valuation. To thoroughly understand the effects and effectiveness of hedge funds' investments, however, one should also explore whether potential changes in firm valuation are accompanied by changes in firm fundamentals and policies. To this end, I analyse how central variables an activist hedge fund might attempt to affect within a firm, e.g., payout policy, capital structure, or management turnover, change in the two years following a fund's investment. Moreover, since the analysis of long-term valuation effects tends to be sensitive to methodology, I probe the robustness of the two studies' results under alternative methods of risk adjustment, namely by estimating calendar-time regressions in the spirit of Carhart (1997) and calculating buy-and-hold abnormal returns akin to Daniel et al. (1997). Finally, by working with hand-collected hedge fund investments between 1999 and 2010, I employ a larger and more comprehensive sample than both studies. Also, existing evidence predominantly exists for periods of extended up-markets. Since my sample contains a number of events following the advent of the financial crisis, my results help assessing whether the effects of activism persist in less favorable market conditions.

The chapter's empirical analyses begin with a characterisation of hedge funds' target selection and investment patterns. Resembling previous evidence, I show that activist hedge funds purchase more votes and commit larger amounts to their investments than their passive peers. While they tend to invest in firms with relatively

[1] This chapter is based on Drerup (2014a).

weak incumbent shareholders (relative to other CDAX firms), the stakes they purchase are still on average very small in comparison to the size of the largest blockholder in their target. Notably, I find that activist hedge funds' targets are not particularly exceptional along central firm characteristics. In particular, they are neither exceptionally profitable nor distressed—both in absolute terms and in relation to other firms in CDAX.

Next, I analyse investors' reactions to hedge fund engagements. I show that market participants only respond significantly when the investing fund is known to be an activist. Average positive abnormal returns of more than 4% in activists' targets are accompanied by abnormally high levels of trading volume. While the immediate return response is consistent with expectected value creation, the strong correlation of 0.4 between abnormal returns and abnormal volumes suggests that price pressure associated with intense buying may be an alternative explanation. A cross-sectional analysis finds that none of the variables suggested to affect the immediate reaction (e.g., the size of the initial stake or the measure for illiquidity by Amihud, 2002) is significantly associated with it, thus favouring neither explanation over the other.

In the main part of this chapter, I turn to the long-term effects of activism on fundamentals and valuation. If hedge fund activism generates shareholder value, then the initially positive abnormal returns should be followed by non-negative returns in the long run. I document, however, that long-term abnormal stock returns following the initial investments are negative and reverse the initially positive return response. This results in a combined effect that is indistinguishable from zero. At the same time, I find that only management turnover seems to change significantly and increase in response to the funds' investments, while other central corporate variables like profitability or capital structure seem unaffected.

In combination, the results suggest that hedge fund activism in Germany is fundamentally ineffective. I interpret the patterns in returns and trading volume as evidence that investors buy attention-grabbing stocks as suggested in Barber and Odean (2008). Anecdotal evidence indicates that hedge funds cause and even proactively enforce public reactions to their investments. Calls for changes in management in particular may make for interesting news and direct market participants' attention towards a stock. As investors respond to such news and flock into a fund's target, they temporarily drive up its price. When the target firm eventually leaves the spotlight and investors' excess demand dies down, the firm's valuation returns to its initial level.

Chapter 2.[2] The second chapter is a contribution to the literature on the value of analyst research. Existing research (e.g., Womack, 1996 or more recently Bradley et al., 2014) finds that revisions of analysts' recommendations are associated with significant movements in stock prices, typically increasing following upgrades and decreasing following downgrades. These cross-sectional patterns suggest that investors use analysts' opinions to revise their own expectations of a firm's prospects. In this chapter, I attempt to assess what information investors respond to when reacting to recommendation revisions. For this purpose, I compare patterns in investors' responses to recommendations—a proxy for the information investors extract—to patterns in how recommendations relate to firms' earnings processes—a proxy for the information recommendations contain.

Recommendation revisions following an earnings release provide a unique setting to compare investors' and analysts' earnings interpretation. As analysts revise their pre-earnings recommendations following an earnings release, they indicate whether they agree with investors' response to the earnings release. Employing a simple model built on this premise, I devise two distinct empirical strategies to assess the informational content of recommendation revisions. The first strategy is to use observable data to model investors' expectation of future earnings. If analysts' earnings interpretation is on average superior to investors', then recommendation revisions should improve this estimate in a specific, predictable way. A caveat with this strategy, however, is that it relies heavily on the appropriateness of the model imposed to proxy investors' expectations. The second strategy is based on the idea that investors' responses to recommendation revisions can be used as a proxy for the extracted information. Thus, because changes in expected earnings are value-relevant, changes in investors' estimates of future earnings should cause patterns in returns around revisions that are in line with the information the revisions contain about future earnings.

I test these predictions using a large sample of recommendation revisions by analysts on record with IBES. Supporting the notion that recommendation revisions are informative about firms' earnings processes, I find that recommendation revisions are associated with differences in the earnings process. More concretely and as predicted by the model, the association between current earnings surprises and future earnings is exceptionally strong, i.e., more positive than average, when upgrades follow positive surprises or downgrades follow negative surprises. In contrast, it is exceptionally weak, i.e., less positive than average, when downgrades follow positive

[2] This chapter is based on Drerup (2014b).

surprises or upgrades follow negative surprises. Cross-sectional patterns in the return response to recommendation revisions suggest that investors value this source of information. Supporting the model's predictions, I find that abnormal returns are positively associated with past surprises when revisions indicate a stronger-than-average association between past surprises and future earnings, and they are negatively associated when revisions indicate the opposite. Both sets of results persist in a number of sensitivity checks. They are robust to, among other things, different ways of measuring expectations, different models of return, or inclusion of control variables. In further regressions I document that recommendations by more experienced analysts and analysts from larger brokerage firms provide stronger signals, indicating enhanced skill at interpreting earnings. Cross-sectional differences in investors' responses to revisions suggest that they are also aware of these between-analyst differences.

The chapter's main contribution to the extensive literature on the value of analyst research (e.g., Stickel, 1991; Womack, 1996; Francis and Soffer, 1997; Ivković and Jegadeesh, 2004; Asquith, Mikhail, and Au, 2005; Chen, Cheng, and Lo, 2010; Livnat and Zhang, 2012; Bradley et al., 2014) is to show that patterns in how revisions in analysts' recommendations relate to fundamental, value-relevant information—future earnings—parallel the patterns in how investors respond to their releases. It extends prior evidence that investors revise their interpretation of past earnings using corroborating firm information like subsequent earnings releases (Freeman and Tse, 1989), dividend changes (Koch and Sun, 2004), or insider transactions (Veenman, 2012). In particular, it lends credence to findings in Mendenhall (1991) and Park and Pincus (2000), whose results suggest that patterns in returns around changes in consensus forecasts and recommendations indicate a reinterpretation of past earnings. In addition, my cross-sectional results on analyst heterogeneity provide an explanation why the accuracy of analyst forecasts is larger for more experienced analysts or analysts from larger brokerage firms (Mikhail, Walther, and Willis, 1997, 2003; Clement, 1999).

Chapter 3.[3] In the paper that constitutes the third chapter of this thesis, Benjamin Enke, Hans-Martin von Gaudecker, and I show that measurement error in subjective expectations data can be put to productive use in understanding the adequacy of economic models of portfolio choice. Stock market expectations are typically considered central determinants of a household's decision to participate in the stock market. However, many empirical studies suggest that measurement error

[3] This chapter is based on Drerup, Enke, and von Gaudecker (2014).

permeates subjective beliefs data. For example, large fractions of elicited subjective expectations do not obey the laws of probability (Manski, 2004; Hurd, Rooij, and Winter, 2011).

Standard techniques of using corrected estimates instead of the misreported values (Wansbeek and Meijer, 2000; Schennach, 2013) may lead to improved estimates in contexts like past income or consumption data, where measurement error can arise due to imperfect recall (Hoderlein and Winter, 2010) or incongruent variable definitions, for example. In the context of measurement error in subjective expectations data, however, analysts may be chasing an elusive target: Many people may simple not hold well-formed beliefs about a given phenomenon. For example, Bruine de Bruin et al. (2000) and Bruine de Bruin and Carman (2012) interpret the prevalence of 50-50 responses in expectations surveys in exactly this way. This suggests that key structural parameters of economic models might not be present in the form envisioned by the econometrician (Stiglitz, 2002; Rust, 2014). If this is the case, using corrected estimates will not lead to improved fit of choice models because the corrected estimates are not closer to forming the basis of decisions. Our main idea is to instead employ the extent of measurement error to uncover heterogeneity in choice behaviour. Put differently, we argue that the magnitude of measurement error in stated beliefs will inform us about the extent to which an economic model constitutes an adequate description of the process underlying an individual's decision to invest in stocks.

To motivate our empirical analysis, we write down a simple economic model that relates the decision to participate in the stock market to expectations about risk and return, risk preferences, and transaction costs. We argue that individuals who employ alternative decision rationales, e.g., those who rely on others' advice or follow rules of thumb, need not maintain particularly meaningful beliefs about the future evolution of the stock market. In consequence, we hypothesise that data collected for these individuals will be characterised by two features. First, their stated beliefs will be prone to measurement error. Second, the sensitivity of their stockholdings to changes in model primitives will be comparably low. To empirically evaluate this hypothesis, we collect data on households' expectations, risk preferences, and financial portfolios as part of the LISS (Longitudinal Internet Studies for the Social Sciences) panel.

We then estimate a Klein and Vella (2009) semiparametric double index model. In the first index of this model we include the primitives of our theoretical model of stock market participation. In the second index we include quantitative and qualitative indicators of measurement error. We allow both indices to interact in a fully nonparametric fashion to obtain predicted probabilities of stock market participation. Consistent with our hypothesis, we demonstrate that changes in primitives of the economic model induce large variation in stock market participation only at low levels of the measurement error index. If measurement error is high, however, stockholdings respond much less sensitively to variation in the index containing beliefs and preferences. We show that the results hold up in several different specifications. In addition, we demonstrate that our modelling approach remains useful when analysing less detailed data. In particular, we show that restricting ourselves to variables that are commonly available or inexpensive to collect, i.e., simpler measures of expectations and purely qualitative measurement error proxies, yields a similar overall pattern. Still, as one would expect, the differences along the measurement error distribution are less pronounced.

1 Long-Term Effects of Hedge Fund Activism in Germany

1.1 Introduction

Hedge fund activism in Germany has been an issue of intense public and political debate. Publicly fought contests for control of firms like TUI AG or Deutsche Börse AG have led to allegations that hedge funds myopically pillage German corporations at the cost of their long-term prospects. The empirical literature on shareholder activism in Germany, however, is comparably scarce. While evidence in Anglo-Saxon countries indicates that hedge funds are in fact creators of shareholder value,[1] the German institutional environment, in particular the high level of ownership concentration and relatively weak shareholder rights, presents a very different and perhaps more challenging setting for activism by minority shareholders to work.

This paper attempts to answer the questions whether and how hedge funds affect German corporations and thereby extends the existing literature on this topic, primarily Mietzner and Schweizer (2014) and Bessler, Drobetz, and Holler (2013), in a number of ways. Foremost, both Mietzner and Schweizer (2014) and Bessler, Drobetz, and Holler (2013) put their focus on the analysis of long-term changes in stock market valuation following hedge funds' investments. However, a full characterisation of how hedge funds affect their target companies should also explore whether potential changes in firm valuation are accompanied by changes in firm fundamentals and policies. To fill this gap,[2] I explore the unexpected one- and two-year changes in central variables an activist hedge fund might attempt to affect within a firm, e.g., payout policy, capital structure, or management turnover. In addition, the findings in Bessler et al. and Mietzner and Schweizer are ambiguous as to whether hedge funds generate value in the long run. Bessler, Drobetz, and Holler

---

[1] A longer literature review on the theoretical background and existing empirical evidence on hedge fund activism is relegated to Section 1.2

[2] The analyses in Mietzner, Schweizer, and Tyrell (2011) resemble those in this paper. I will discuss this point in the next section.

(2013) find that less aggressive hedge fund investments are associated with short- and long-term increases in shareholder value, while more aggressive, activist investments are associated with initially positive abnormal stock returns and eventual reversals. Mietzner and Schweizer (2014) focus on activist hedge funds and also document long-term reversals of initially positive abnormal returns. Given the importance of risk-adjustment in long-term event studies, I probe the robustness of the two studies' results by estimating calendar-time regressions in the spirit of Carhart (1997) and calculating buy-and-hold abnormal returns akin to Daniel et al. (1997). Finally, my study builds on a hand-collected sample of hedge fund shareholdings that covers the period between 1999 and 2010, whereas the samples in both Bessler, Drobetz, and Holler (2013) and Mietzner and Schweizer (2014) stop short before the advent of the financial crisis. This extended timeframe lets me work with a larger sample than both studies and extends the evidence concerning activism to a period that includes times of emphasised down-markets. Prior evidence in US studies mainly comes from activist investments during long up-markets. Including the recent extended downturn thus helps in assessing whether hedge funds' abilities to create shareholder value persist in less favorable market conditions.

My analyses begin with a characterisation of hedge funds' target selection and investment patterns. Similar to previous work, I document that activist hedge funds purchase more votes and commit larger amounts to their investments than their passive peers. However, with an average fraction of 6.4% of the voting rights in their target companies, the purchased blocks are small in comparison to the average size of the largest blockholder in their target firms whose share is almost 30% of the votes. A comparison of targets to the remaining firms in Germany's CDAX reveals that these 30% are still comparably small relative to the votes owned by the average largest blockholder in CDAX, suggesting that hedge funds invest into firms with relatively—but not absolutely—weak incumbent shareholders. Beyond this result, activist hedge funds' targets are not particularly exceptional along central firm characteristics. Notably, activist hedge fund targets are neither exceptionally profitable nor distressed—both in absolute terms and in relation to other firms in CDAX.

The main analyses evaluate how hedge funds affect their target companies. I start with an analysis of the market's immediate reaction to hedge fund engagements. To this end, I use standard event study methodology and estimate abnormal returns around the initial publication of hedge funds' shareholdings. I find that there is only a significant market reaction when the investing fund is known to follow an activist strategy. The 20 days leading to the publication of an activist fund's investment are associated with abnormal returns of 4.22% in the target firms. This increase is accompanied by abnormally high levels of trading volume. The correlation between abnormal returns and abnormal volumes is 0.4, indicating that firms with stronger stock market return responses also have larger abnormal trading volumes.

Both the anticipation of value creation as well as price pressure due to intense buying provide possible explanations for these patterns. A cross-sectional analysis of abnormal returns, however, favors neither of the explanations over the other as none of the variables suggested to affect the reaction (e.g., the size of the initial stake or the measure of illiquidity by Amihud, 2002) is significantly associated with the immediate response.

I next turn to the long-term effects of activism. If hedge fund activism generates shareholder value, then the initially positive abnormal returns should be followed by non-negative returns in the long run. Moreover, fundamental changes to corporate policies and performance induced by hedge fund activism should also be detectable in the dynamics of firm characteristics in the years following the funds' investments. This is, however, not what I find. In fact, I document that long-term abnormal stock returns following the initial investments are negative and reverse the initially positive return response, resulting in a combined effect that is indistinguishable from zero. Moreover, when looking at changes in the targets' values of central corporate variables like profitability, capital structure, or management turnover, I find that only management turnover seems to change significantly and increase in response to the funds' investments.

Together, the results provide an ambiguous picture. On the one hand, the short-term response suggests that market participants attribute value to activist hedge fund shareholders. On the other hand, they seem to reverse this assessment in the long run. Moreover, the lack of any observable changes in fundamental performance suggests that they are right to do so. I conjecture that there are two possible (and mutually not exclusive) explanations for these results, both of which hint at a short episode of market inefficiency. Irrespective of which of them accounts for the results, both have in common that hedge fund activism in Germany seems to be fundamentally ineffective.

The first explanation is that market participants initially overestimate the degree to which hedge funds can overcome the existing blockholders' resistance to meaningfully affect a firm. As a consequence, they react enthusiastically to the presence of a hedge fund. However, as market participants realise that hedge fund activism is not associated with the expected improvements in corporate fundamentals, they reverse their initially positive response. This explanation can account for the initially positive response, the long-term reversal, and the lack of changes in firm fundamentals. Still, it is unclear why market participants would misjudge the implications of hedge funds' investments to begin with.

A second interpretation of the results is that the pronounced initial return and volume responses indicate that investors buy attention-grabbing stocks as suggested in Barber and Odean (2008). Hedge funds are known to stir and even proactively enforce public reactions to their investments. Calls for changes in management in particular may direct market participants' attention towards a stock. It is conceivable that investors respond to such news by flocking into the funds' targets, thereby temporarily driving up their prices. As target firms eventually leave the spotlight, investors' excess demand dies down and firms' valuations return to their initial levels. This explanation can also explain the initial response, the eventual reversal, and the lack of changes in fundamentals. In addition, it finds support in the strong association between initial abnormal returns and abnormal volumes.

The rest of this paper is structured as follows. Section 1.2 describes the previous literature on hedge fund activism and the institutional background. Section 1.3 presents the data collection procedure. Section 1.4 summarises the sample's characteristics and compares passive and activist targets and investment characteristics. Section 1.5 presents the short- and long-run analyses of the funds' investments. Section 1.6 concludes.

1.2 Previous Literature

The idea that large blockholders can overcome the problems that result from the separation of ownership and control (Shleifer and Vishny, 1986; Grossman and Hart, 1980) has intuitive appeal, yet surveys of previous empirical literature like Karpoff (2001) document that shareholder activism has little if any positive effect on target performance or value.

The effectiveness of activism by institutional investors is presumably hampered by a variety of reasons: collective attempts suffer from the incentive to free-ride on the monitors' efforts (Black, 1990); business ties between investor and target prevent confrontational voting on the investor's side (Davis and Kim, 2007); the preference to sell positions rather than engage in insecure outcomes (Parrino, Sias, and Starks, 2003); a lack of sufficiently credible threatening power for classic institutional investors to be recognised as influential monitors by either management or market (Clifford, 2008); political (Romano, 1993) and regulatory (Black, 1990) barriers; insufficient monetary incentives for fund managers to bear the personal costs of activism (Rock, 1991); inadequate managerial expertise to alter central firm policies (Lipton and Rosenblum, 1991). Moreover, managers in a state of activist siege may devote considerable time and resources to assure their position and in turn inadvertently impair corporate performance (Karpoff, 2001; Bebchuk and Cohen, 2005).

The literature lists a number reasons why one class of market participants, hedge funds, may still be able to do well in a role as shareholder activists.[3] For one, because hedge fund managers participate in the performance of their own fund both through performance-linked pay and their own investment in the fund, they have strong incentives to maximise returns. While institutional investors like mutual and pension funds also disburse boni to their managers, they do so to a lesser degree than hedge funds and are limited by regulation (Brav et al., 2008). Accordingly, hedge fund managers may be more incentivised to increase firm value by affecting a firm's management than other institutional investors. Secondly, hedge funds can be more credible and thus successful as monitors due to their organizational form. For instance, they are free from investment limitations and liquidity constraints other investors can be subject to. And in contrast to most institutional investors, hedge funds can circumvent the problem of capital redemption by locking up capital or keeping it in a side pocket for illiquid investments (Aragon, 2007), thus guaranteeing a sufficient investment horizon to affect a firm. Thirdly, contrary to pension or mutual funds, hedge funds are free from political pressure and other conflicts of interest that prevent them from solely focusing on their own and their investors' benefits (Brav et al., 2008). And absent any conflicting business ties to the target firm, hedge funds are able to adopt a fully confrontational demeanor towards management. Fourthly, the strategic maneuverability that comes with freedom from regulation allows for an optimal mix of value-maximising techniques. For example, hedge funds sometimes engage in what is referred to as empty voting or hidden ownership (Hu and Black, 2007). They often either lend votes or have readily available access to further votes and thereby increased effective or potential share in a company's control. Also, anecdotal evidence suggests that hedge funds regularly engage in concerted behaviour. Even though this is legally controversial, especially if the funds circumvent regulatory obligations for collective ownership, it can exert significant pressure on management. In addition, while it is hard for institutional investors to combine activism with an overall diversification strategy (Kahan and Rock, 2007), hedge funds are free from diversification restrictions (Brav et al., 2008) and can therefore take very concentrated positions in target companies. Thus, in case a

3 See among others Clifford (2008), Klein and Zur (2009), Boyson and Mooradian (2011), Brav et al. (2008), Kahan and Rock (2007), or Bratton (2007).

management does not acquit to a fund's demands, the fund can threaten to buy out the company (Clifford, 2008). Finally, due to their alleged aptitude towards activist influence, hedge funds can invest strategic and ex ante by identifying promising targets early, whereas other institutional investors tend to act incidental and ex post by reacting to possible insufficiencies if they occur (Kahan and Rock, 2007, p. 1022). Activism by hedge funds is not merely a coerced reaction but rather a conscious approach.

Recent literature explores the consequences of hedge fund activism around the world. In one of the most extensive studies, Brav et al. (2008) document that hedge funds often succeed in provoking a variety of changes in corporate structure and performance in the US. For example, they find both initial and long-term positive valuation effects subsequent to a hedge fund's entry as well as changes in corporate performance and policy. Klein and Zur (2009) similarly detect significant and persistent increases in stock valuation at hedge funds' target firms. They show that the initial market reaction seems to correctly anticipate the outcome of following activist campaigns as stated in obligatory filings. However, Klein and Zur also find decreases in various measures of target profitability. Boyson and Mooradian (2011) find that hedge funds following activist strategies outperform their non-activist peers in terms of risk-adjusted performance. Clifford (2008) puts specific focus on the investing funds' stated objectives. He shows that the outcome and returns to hedge fund investments differ between activist and passive blockholdings and observes that funds seem to align their operational form with their applied strategy.

Even if hedge funds follow an activist strategy, it is unclear whether they can be successful in Germany. For one, German shareholder rights are comparably weak (La Porta et al., 2000), perhaps withholding from hedge funds the legal arsenal necessary to affect target firms' policies. Moreover, the presence of large incumbent shareholders is a central feature of the German market (Franks and Mayer, 2001; Andres, 2008). Stronger incumbent shareholders in controlling positions may present a considerable obstacle for shareholder activism to work.

The literature on hedge fund activism in Germany in general and its long-term consequences in particular is comparably scarce. In a study of 67 German transactions, Mietzner and Schweizer (2014) analyse the short- and long-term valuation effects. Even though their results of significantly positive announcement returns are consistent with hedge funds generating shareholder value, they observe that targeted firms underperform their benchmark, the CDAX, in the long run. They attribute this to the structure of the German corporate governance system that prevents funds from exerting control. In an attempt to explain their seemingly contradictory observations, Mietzner and Schweizer (2014) hypothesise that initially positive and ultimately negative returns may reflect a market misjudgment of hedge funds' intentions and capabilities to reduce agency costs. Bessler, Drobetz, and Holler (2013) show that hedge funds target poorly governed firms. Similar to Mietzner and Schweizer, they detect significant positive market reactions to hedge fund shareholdings. In addition, they find that only non-aggressive hedge fund investments are followed by future stock market outperformance. However, they also show that the initially positive returns to aggressive forms of activism are followed by a reversal in abnormal returns thereafter.

Achleitner, Betzer, and Gider (2010) compare the German targets of hedge funds and private equity funds. They suggest that targeting choices of hedge funds indicate a short-term orientation, but since they focus on the targeting choice only, they concede that "[a] more comprehensive assessment therefore necessitates analysis of the long-term consequences" (Achleitner, Betzer, and Gider, 2010, p. 826). Weber and Zimmermann (2013) provide evidence suggesting that hedge fund investments in Germany are associated with both an information-driven effect around their publication and a volume-driven effect around the actual transaction.

My paper is very close to Mietzner, Schweizer, and Tyrell (2011), who also look at the development of long-term fundamentals following hedge fund investments in German firms. I only became aware of their work after having finished the initial draft of this paper.[4] In consequence, some of the analyses resemble those in their work. There are, however, a few notable differences. First, while Mietzner, Schweizer, and Tyrell (2011) focus on hedge fund activists and their industry peers (as well as private equity firms and their peers), I compare the effects of hedge funds both with and without an activist agenda. Second, my analyses include a substantially larger number of events (142 vs. 78 activist events) and cover the period following the advent of the financial crisis. This allows me to both check the robustness of their findings and provide an assessment of whether the ability of hedge funds to create shareholder value in the long run is contingent upon market environment. Third, I use a differences-in-differences design to compare how hedge fund targets develop over time to how their industry peers develop. Mietzner, Schweizer, and Tyrell (2011) instead compare the levels of target firms' fundamentals to the levels of their peers' fundamentals at different points in time. In order to make a (statistical) statement on how hedge funds affect their target firms, it is necessary to benchmark the

---

[4] The first draft of this paper was finished in 2009. The sample period of their paper ends in the same year.

changes in the fundamentals of target firms against a proxy for expected changes, e.g., the changes in the fundamentals of industry peers. To this end, a differences-in-differences design is more appropriate than the analysis of differences-in-levels, which only allows statements on how targets and peers differ at a particular point in time. Finally, my conclusions are different. In contrast to Mietzner, Schweizer, and Tyrell (2011), I conclude that there are —with the exception of a change in management turnover— no discernible effects of hedge funds' investments on firm fundamentals.

1.3 Dataset

To construct my sample, I analyse all firms listed in the CDAX market segment between 1999 and July 2010. I work with several sources to gather all relevant hedge fund shareholdings over the sample period. My point of departure is the filings database provided by BaFin, the German supervisory organization for financial services, that contains mandatory filings of significant shareholdings according to article 21 of the German Securities Trading Act ('Wertpapierhandelsgesetz WpHG').[5] To compensate for the possibility that some funds do not submit their holdings to BaFin, I further search all disclosures provided by dgap.de, a website focusing on corporate news reporting. In addition, I go through all filings reported on corporate websites, where I hand-collect shareholdings from annual documents as well as voting rights announcements not found in the previous two sources.

The next step in my analysis is the classification of which shareholder constitutes a hedge fund. Absent an official definition of the term hedge fund, I use industry- and fund-websites to classify investors. To this end, I individually check the identity of all investors owning significant stakes in each company. In cases where the manager of a hedge fund is listed as a shareholder, I assume that he represents the fund. Some companies run both private equity and hedge fund subsidiaries. Lacking a definite means of drawing a distinction, I leave out these ambiguous cases. For each hedge fund shareholder found with this procedure, I then search *Bing*, *Yahoo!*, and *Google* as well as news on Bloomberg for the combination of the fund's name and the terms *'wphg'*, *'voting rights'*, *'holding'*, or *'shareholdings'* as well as the respective German translations to identify additional investments by the same fund. Beyond mandatory disclosures, I also collect shareholdings that become publicly

[5] By the 2010 version of article 21 of WpHG it has to be filed with BaFin once voting rights cross the thresholds of 3, 5, 10, 15, 20, 25, 30, 50, or 75% either way not later than 4 days subsequent to the crossing.

known through press releases or alternative and publicly accessible channels. Most of these cases constitute a hedge fund's open critique of corporate policy.

To enhance the accuracy of event dates, i.e., the dates where hedge fund shareholdings become publicly known, I search for the conjunction of firm and fund name using the above sources. For every firm-fund combination I finally pick the earliest date of publication as the effective event date. Later increases or decreases of a fund's holdings are not included as separate events, yet I record the maximal stake during a funds' investment period.

This yields an interim sample of 353 potential observations. From these, I exclude various types of events (75 in total). Following Brav et al. (2008), I drop all cases where the target company is subject to an announced merger and cases where the hedge fund enters by a debt-equity swap. I further exclude events where the fund holds shares prior to the IPO, the company is not publicly traded, or the fund quits within the first 30 days of investment. I do explicitly not exclude co-investments into a company where another fund is or was present. If available, I also collect the date of the funds' exits which is usually the date where it drops below the lowest threshold with filing requirement.

Next, I distinguish between events with activist background and those where the investment indicates a passive position. I classify events as of an activist nature if one of the following holds true for the fund in question:

1. The fund describes itself as taking activist stances towards firm management on its website;
2. Publications concerning the hedge fund industry or the item of news surrounding the launch of the fund describe it as a shareholder activist;
3. I contact the fund to inquire and activism is named as the applied strategy;
4. I observe cases of overt activism by the fund in the past or upon entry into the company. I use a news search in Bloomberg and the above mentioned search engines to look for news in this regard. A public call for a higher dividend, for example, leads me to classify a fund as an activist;
5. Press releases surrounding the event indicate that the fund is taking a possibly confrontational stance towards management.

I classify all remaining funds as passive, and I classify each of the sample's hedge funds as either always active or always passive. Admittedly, this classification is imperfect. However, it should come reasonably close to the actually applied strategy.

In contrast to the studies in the US, the absence of filings of purpose prevents a clear-cut dissection of strategies in Germany. Importantly, this strategy of identifying activists is more likely to misclassify activists as passive funds than passive funds as activists. In consequence, if it biases the sample of activists, then it will most likely bias it to stronger forms of activism rather than the opposite.

Several of the subsequent analyses compare the collected sample of firms to companies that are identified to be of similar type. These matching firms serve as the benchmark to explore abnormal changes in firm policies and fundamentals in the years following the hedge fund investments. Moreover, they provide each firm's individual benchmark in the analysis of long-run buy-and-hold returns.[6] The procedure I use to assign matching firms follows Brav et al. (2008). In the first step, I choose all CDAX firms that fall into the same industry as each target. Then, I select the firms among these whose size in terms of sales is not larger than 125% or smaller than 75% of the target company's pre-event fiscal year value. In the resulting list, the company with the closest book-to-market ratio is selected. If step two or three fails, the closest in sales is picked. I drop financial and real estate firms in all analyses with a focus on accounting variables. For the event studies, I select the respective Deutsche Börse sector index as the benchmark for all financial and real estate firms to minimise the interference of the financial crisis. All event-study results are robust to their exclusion.

1.4 Investment Characteristics and Target Selection

The final sample consists of 278 events across 170 firms. 53 of the 111 different hedge fund families investing are activists, resulting in 142 activist blockholdings. Tables 1.1 and 1.2 provide a breakdown of all investments in terms of industry and temporal distributions. No specific industrial sector seems to be strongly over- or underrepresented. Still, there is a distinctive temporal pattern. Germany has seen a steady increase in hedge fund investments until 2007, yet the number of new as well as concurring investments dropped substantially following the onset of the financial crisis. This development is remarkably similar to the change in assets under management in the hedge fund industry in general and coincides with the overall stock market conditions. Some of the target firms are subject to only one significant blockholding over the entire sample period. However, a large number of firms (≈59%) see various contemporaneous as well as subsequent investments, with one

6 Owing to the small number of firms in CDAX, I assign only one firm as each target's benchmark instead of a matching portfolio as in Daniel et al. (1997).

firm hosting 9 different hedge funds between 2003 and 2010.

Table 1.1. Industry distribution

Sector	Sample		CDAX	
Automobiles & Parts	10	3.60%	21	3.08%
Banks	7	2.52%	7	1.03%
Basic Resources	2	0.72%	11	1.61%
Chemicals	7	2.52%	17	2.49%
Construction & Materials	5	1.80%	24	3.52%
Financial Services	17	6.12%	54	7.92%
Food & Beverage	2	0.72%	18	2.64%
Healthcare	21	7.55%	49	7.18%
Industrial Goods & Services	73	26.26%	121	17.74%
Insurance	0	0.00%	14	2.05%
Media	12	4.32%	45	6.60%
Oil & Gas	5	1.80%	19	2.79%
Personal & Household Goods	11	3.96%	54	7.92%
Real Estate	15	5.40%	31	4.55%
Retail	20	7.19%	31	4.55%
Technology	54	19.42%	140	20.53%
Telecommunications	8	2.88%	4	0.59%
Travel & Leisure	9	3.24%	10	1.47%
Utilities	0	0.00%	12	1.76%
Sum	278	100%	682	100%

Table 1.1 reports the industry distribution of events compared to the distribution in CDAX in 2007.

Table 1.2. Temporal distribution

Year	Activist Blocks	Passive Blocks
2001	1	1
2002	4	0
2003	1	0
2004	14	5
2005	18	8
2006	21	14
2007	46	68
2008	22	25
2009	12	9
2010	3	6

Table 1.2 shows the distribution of events over time. The sample contains no events prior to 2001.

Table 1.3. Stakes, invested capital, and holding periods

	Activist Blocks				Passive Blocks			
	Percentage Ownership		Initial Investment (in million €)	Realised Holding Period	Percentage Ownership		Initial Investment (in million €)	Realised Holding Period
	Initial	Maximum			Initial	Maximum		
Average	6.40%	8.30%	59.60	467	4.38%	5.47%	25.96	461
Minimum	1.00%	1.00%	0.54	42	3.00%	3.01%	0.12	39
P(10)	3.04%	3.05%	1.73	70	3.03%	3.04%	1.52	113
P(25)	3.21%	3.45%	4.40	154	3.11%	3.19%	3.45	192
P(50)	5.02%	5.31%	14.56	344	3.63%	5.01%	10.08	365
P(75)	6.84%	10.30%	45.64	663	5.10%	5.58%	28.70	613
P(90)	12.65%	18.67%	141.64	1060	6.91%	9.46%	77.16	884
Maximum	29.57%	29.57%	873.31	1865	10.13%	25.32%	264.97	1526
				19 still invested				24 still invested

Table 1.3 reports details about the sizes and holding periods for the hedge funds' investments, separetely for activist blocks and passive blocks. I report the distributions of the initially purchased voting rights, the maximum reported percentage during the investment, and the approximate €-value of the initial purchase in millions. The initial investments are calculated using the targets' market value 30 days prior to the entry date. I also provide an estimate of the realised holding periods as the number of days between exceeding the lowest threshold with reporting requirement and falling below it again. In addition, I report the number of funds still invested as of July 2010, the end of the sample period.

Table 1.3 presents summary statistics for activist and passive blocks in terms of voting rights acquired, approximate Euro value of the investment, and its holding period. All information used in the estimates stems from either the news item that publicised the shareholding or the obligatory filings released by the fund initially and over the invested period. Similar to Clifford (2008), I find that both initial and maximal percentage ownership are significantly larger for activist blocks (Wilcoxon p-value < 0.01). Whereas activists acquire a mean initial stake of 6.40%, passive funds purchase on average 4.38% of voting rights with their initial investment. The same pattern characterises the respective maximum stakes over the investment of both types of funds. While activists average maximum stake in their targets is 8.30%, the average maximum for passive blocks is 5.47%. The median initial capital committed to the purchase of an activist block is 14.56 million Euros and 10.08 million Euros for a passive one. The blocks' sizes are nevertheless highly variable, with the smallest block at a mere 120,000 Euros and the largest at 873 million.

Two dominant patterns in Table 1.3 should be highlighted: For one, activists purchase larger shares of their target companies and invest more capital than passively investing hedge funds. This may reflect that activism has to pay off to the activist, and for this to happen a necessarily large amount has to be invested on which returns can be realised. Spending time and effort to actively engage management for a fund equipped with possibly billions of capital may hardly be worthwhile if the stake in question is comparably small. And secondly, the stakes activist hedge funds purchase are in the majority of cases small relative to the strongest incumbent shareholder who on average holds about 30% of the votes. Anecdotal evidence has it that hedge funds influence corporate policy from a minority position and attempt to organise support by further shareholders when necessary.[7] Brav et al. (2008) suggest that this is a central feature that differentiates these activists from previous raiders who attempted to take over control of their targets. Median holding periods of about a year for both activist and passive blocks casts doubt on the public assertion that hedge funds are short-sighted investors as discussed in Kahan and Rock (2007). One activist fund even stays invested for a total of 1865 days or 5.1 years. Brav et al. (2008) observe similar holding periods, reporting a median period of 369 days across all events. 43 or 15.5% of investments are still held at the sample period's end, though none of them is held for excessively long.

Both time horizon as well as percentage held and capital investment need to

---

[7] After hedge funds got involved with Deutsche Börse AG and ousted its CEO Werner Seifert, Seifert labeled this behaviour *dictatorship of minority shareholders* in his book "Invasion of the locusts" (2007), written by him and Hans-Joachim Voith.

be interpreted with caution. The holding period leaves out all time before initially exceeding and after finally falling below the lowest mandatory threshold. Furthermore, percentages of voting rights may largely underestimate a funds' actual share in the target as they do not involve derivative instruments, debt positions, or alternative sources of additional control. At the same time, hedge funds need not report changes of voting rights in between reporting thresholds. This necessarily results in a somewhat inaccurate reporting. In case activist funds tend to report the largest rather than the average or lowest value after crossing a threshold, the difference between activist and passive blocks might be overstated.

Next, I turn to the question how hedge funds select their targets. I analyse the targeting choice with a bivariate probit regression.[8] The dependent variable in both equations is binary, taking on the value 1 if a hedge fund purchased a stake in a particular month and firm, and 0 otherwise. One of the equations includes all activist investments, the other the passive ones. The regressions include all firms that have been listed in the CDAX between 2001 and 2010. As control variables I add dummies for the presence of both activist and passive funds, the logarithm of the market value of equity, market-to-book, the ratio of cash holdings to total assets, sales growth, leverage, return on assets, dividend yield, and free float, all of which are lagged by one month. In addition, I include year dummies to allow for unobserved temporal effects triggering hedge fund entry. I winsorise the accounting variables at the 1 percent level.

Table 1.4 presents the results. All funds seem to exhibit a tendency to cluster in target firms. Both of the coefficients indicating the presence of either an activist or passive hedge fund in the previous month have positive and significant loadings, regardless of the fund investing. For example, the presence of an activist hedge fund increases the monthly probability of being targeted by another one for an otherwise average firm by 0.40%, the presence of a passive hedge fund increases it by 0.16%. One explanation for this clustering is that specific unobservable target characteristics trigger entry of hedge funds. However, the comparably larger effect for activist funds is also suggestive of the interpretation of activism as a joint effort between several hedge funds. Market value is positively related to the probability of being targeted by both activist and passive funds, in line with the idea that hedge fund

[8] Since I restrict each hedge fund to either always invest as an activist or as a passive blockholder when classifying funds, I abstain from modelling the target selection sequentially. In other words, I treat being an activist or a passive blockholder as a fixed hedge fund characteristic and not as a choice each fund makes upon investment. This setup aligns with anecdotal evidence that some hedge funds (e.g., Wyser-Pratte Management Inc.) are notorious for their tendencies to get involved as shareholder activists once they are invested.

Table 1.4. Targeting choice, bivariate probit

	Activist Blocks			Passive Blocks		
	Δ Marg.	Marg.	Coef.	Δ Marg.	Marg.	Coef.
Activist Presence		0.4039%	0.5099***		0.0137%	0.2513*
			(4.38)			(1.88)
Passive Presence		0.1635%	0.2918***		0.0234%	0.3550***
			(2.82)			(3.49)
Market Value	0.0624%	0.0237%	0.6516***	0.0039%	0.0015%	0.0455***
			(4.24)			(3.33)
Market/Book	-0.0048%	-0.0032%	-0.0088	-0.0000%	-0.0000%	-0.0000
			(-0.76)			(-0.00)
Cash/Assets	0.0032%	0.0162%	0.0445	0.0027%	0.0122%	0.3598*
			(0.20)			(1.73)
Sales Growth	0.0127%	0.0183%	0.0501	0.0013%	0.0018%	0.0534
			(1.27)			(1.59)
Leverage	0.0173%	0.0557%	0.1529	-0.0002%	-0.0006%	-0.0188
			(1.19)			(-0.13)
Return on Assets	-0.0017%	-0.0092%	-0.0253	0.0035%	0.0158%	0.4669***
			(-0.12)			(2.81)
Dividend Yield	-0.0173%	-0.0075%	-0.0206	-0.015%	-0.0007%	-0.0199
			(-0.94)			(-1.39)
Float	0.0593%	0.1621%	0.4452***	0.0056%	0.0147%	0.4349***
			(4.30)			(3.71)
Constant			-4.0012***			-3.5829***
			(-17.65)			(-22.23)
					Observations	52,928
					ρ	0.2567
					p-Value	0.0157

Table 1.4 summarises the results of a bivariate probit regression. The dependent variables in both equations are binary, taking the value of 1 in case a hedge fund reported holdings in the particular month and firm, 0 otherwise. One equation contains all activist reported holdings, the other those of passive funds. Included in the analyses are all firms with necessary data on all covariates that have been listed in the CDAX between 2001 and 2010. All covariates are lagged by one month. Activist Presence and Passive Presence are dummy variables indicating whether an activist or passive hedge fund, respectively, were present in the month preceding the month in question. Market Value is the logarithm of the market value of equity lagged by one month. Market/Book is the market value of equity lagged by one month divided by the book value of equity from the previous fiscal year. Cash/Assets is the ratio of cash and equivalents to total assets. Sales Growth is the percentage change in sales from two fiscal years ago to the last fiscal year. Leverage is the ratio of lagged debt to lagged assets. Return on Assets is last year's EBITDA divided by total assets two years ago. Dividend Yield is the dividend yield lagged by one month. Float is the percentage of shares in free float in the previous month. For all covariates, I report coefficients and their z-statistics, marginal probabilities (Marg.) at the means as well as the change in marginal probabilities induced by a one standard deviation increase from the averages (Δ Marg.), holding all remaining variables at their means. I also include a set of year dummies. *, **, and *** mark significance levels of 10, 5, and 1 percent, respectively. I winsor accounting variables at the 1 percent level. ρ is the correlation between the respective equations' residuals.

targets need to be sufficiently large to be worthwhile an investment. Alternatively stated, many firms may be too small to legitimate investments in the intended size. Also, as small firms' businesses are often complicated by personal ties and firm-specific management knowledge, they may be hard to influence and understand. Both funds' investments are more probable if larger percentages of shares are considered as free float. This likely reflects that hedge funds avoid targets with strong incumbent shareholders. In addition, if all remaining ownership is dispersed, even a small blockholder can effectively be the largest and supposedly most influential among all shareholders. The results also suggest that passive hedge funds seek firms that operate a profitable business, equipped with a solid amount of cash, and realising relatively high returns on assets. Contrary to that, none of the variables aimed at capturing corporate performance is related to the probability of becoming an activist target. Specifically, the results do neither indicate that activists' targets are relatively distressed nor that they are particularly profitable relative to other firms in CDAX.

Table 1.5 provides summary statistics that describe the target firms in absolute terms. The most notable result in this table is that the median target of both activist and passive hedge funds is profitable both in terms of return on equity and return on assets. The results of probit analysis and summary statistics render it questionable that activist hedge funds intend to make a short-term profit by pillaging profitable companies as is sometimes suggested publicly, specifically since the companies they actually target are not particularly profitable, underleveraged, or cash-laden to begin with. However, the results also do not support the opposite view that hedge funds invest in firms in financial distress.

1.5 Effects of Hedge Fund Investments

Short-Run Response. I use standard event-study methodology to analyse the short-run impact of hedge funds' investments. Defining the event date as the day the hedge fund's shareholdings become public knowledge, I estimate normal returns from day -110 to -21 relative to the event using a market model with the CDAX as the market return.[9] I then calculate cumulative abnormal returns (CAR) as the sum of daily abnormal returns over different event periods and evaluate the significance of their differences from zero using the tests of both Patell (1976) and Boehmer, Musumeci, and Poulsen (1991). The latter accounts for event-induced

[9] Different estimation windows produce the same results.

Table 1.5. Target characteristics and differences in target selection

	Activist Blocks		Passive Blocks		Difference between	
	Average	Median	Average	Median	Averages	Medians
Total Assets (in million €)	2,431.93	276.68	848.58	149.94	1,583.34***	126.75**
					(3.09)	(2.47)
Market Value (in million €)	1,007.97	299.22	686.19	221.89	321.78	77.34
					(1.63)	(0.15)
Return on Equity	-6.14%	7.53%	6.37%	11.71%	-12.51%***	-4.18%***
					(-2.50)	(-2.60)
Return on Assets	8.62%	10.84%	11.03%	12.60%	-2.41%	-1.76%*
					(-1.59)	(-1.95)
Return on Lagged Assets	12.96%	11.49%	16.88%	15.73%	-3.91%*	-4.24%***
					(-1.66)	(-3.09)
Market/Book	2.24	1.91	2.84	2.30	-0.61***	-0.39***
					(-2.75)	(-2.91)
Leverage	0.3582	0.3200	0.2882	0.2445	0.0699*	0.0755**
					(1.88)	(2.05)
Dividend Yield	1.32%	0%	1.41%	0%	-0.18%	0%
					(-0.76)	(-0.55)
Cash/Assets	0.2420	0.0953	0.3167	0.1435	-0.0748	-0.0482
					(-1.30)	(-1.52)
Free Cash Flow/Sales	0.0505	0.0583	0.0742	0.0826	-0.0238	-0.0243**
					(-1.45)	(-2.16)

Table 1.5 reports summary statistics as well as a comparison between activist and passive targets. Total Assets is the sum of the book values of debt and equity. All remaining variables are defined as in Table 1.4. I present means and medians for both activist and passive targets. In addition, I report differences between the means and the medians alongside the t-statistics of unpaired t-tests and the z-statistics of Wilcoxon signed rank tests. *, **, and *** mark significance levels of 10, 5, and 1 percent, respectively.

variance. Since hedge fund investments and the subsequent days are often characterised by highly volatile returns, this seems particularly appropriate.

Table 1.6 presents cumulative abnormal returns over different timeframes for activist and passive blocks. Figure 1.1 plots them and adds the cumulative abnormal volume.[10] While there is no discernible pattern in the response to passive events, market participants react strongly to the presence of activist hedge funds. On the event day itself, the activists' targets earn a statistically significant mean abnormal return of 1.27%, significantly ($p < 0.05$) larger than the 0.14% CAR associated with

[10] As a measure of volume I employ the Euro value of shares traded. I define abnormal volume for each firm-day as the percentage excess volume over the median volume estimated over trading days -100 to -21 prior to the event. The graph plots the median volume across all firms for each day relative to the event. By using medians rather than means I attempt to circumvent the high degree of skewness in volume data. With very small stocks, I observe volumes during the event period of up to 17 times the average volume in the estimation period. Using medians allows incorporating this information, yet reduces its high leverage.

Table 1.6. Immediate market reaction

		Activist Blocks			Passive Blocks		
τ_{start}	τ_{end}	CAR	Patell	Boehmer	CAR	Patell	Boehmer
0	1	1.27%	7.09***	4.62***	0.14%	0.11	0.07
-1	1	1.22%	4.04***	3.20***	-0.07%	-0.61	-0.50
-20	0	4.22%	4.22***	3.31***	0.83%	0.17	0.13
-10	0	3.36%	4.54***	3.33***	1.31%	1.15	0.82
-20	20	3.47%	2.94***	2.37**	-0.44%	-0.31	-0.24
1	10	0.90%	1.72*	1.96**	-1.39%	-1.27	-1.03
1	20	-0.75%	-0.11	-0.11	-1.27%	-0.62	-0.59

Table 1.6 summarises the short-term event study's results. Normal returns are calculated using a market model with a 90 day estimation period from day -110 to -21 and the CDAX as the market index. After dropping events with insufficient return history, there are 137 activist events and 125 passive events. I present the results for various windows surrounding the initial event. The initial event is defined as the first day the investment of the hedge fund becomes publicly known. *, **, and *** mark significance levels of 10, 5, and 1 percent. I report values of both Patell's (1976) test and Boehmer's (1991) test that accounts for event-induced variance.

passive blocks. The lack of significant reactions to the investments of hedge funds following a passive agenda suggests that market participants do not consider their investments value-relevant information on average. The largest part of abnormal returns around activist events seems to materialise well in advance of the publication of the investments with a significant run-up CAR of 4.22% in the 20 days preceding day 0. Abnormal volume for activist events increases during the run-up period, peaks on the event day with a median abnormal volume of 54.59%, and gradually moves downward afterwards. One interpretation of the significantly positive returns and volumes around activist events is that investors buy the target's stock because they anticipate that hedge fund activists create shareholder value. Brav et al. (2008) remark that fund managers at times hint other investors at their upcoming engagement, which may explain why a large fraction of the return response precedes the actual event. An alternative interpretation why prices increase alongside large abnormal volumes is that buying pressure exerted by either the fund itself or other market participants drives up the stock price. The strong correlation of 0.39 between CARs and abnormal trading volumes for activist events ($\rho = -0.08$ for passive events) is consistent with both interpretations.

In sum, the market reacts strongly to the presence of hedge funds, yet only if the funds have a reputation for applying activist strategies. This positive reaction is common to previous studies on hedge fund activism, including the studies of Bessler, Drobetz, and Holler (2013) and Mietzner and Schweizer (2014).

Figure 1.1. Abnormal returns and volumes around activist and passive events

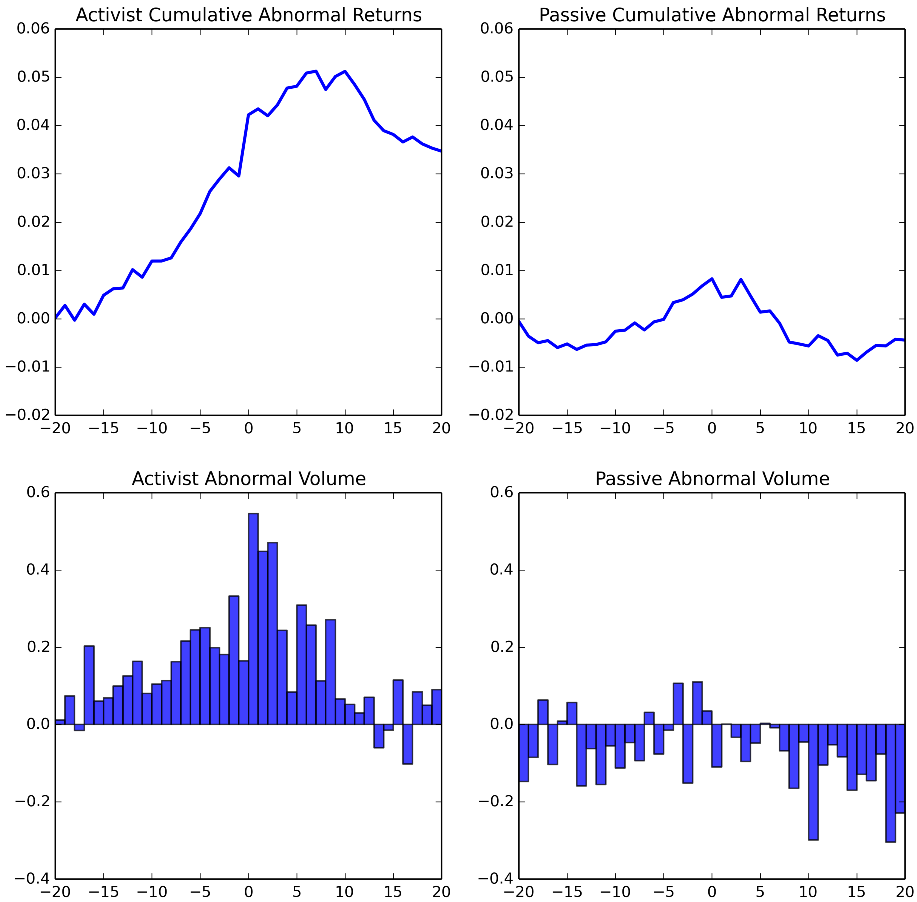


Figure 1.1 plots average cumulative abnormal returns and median daily abnormal volumes for activist and passive events. Cumulative abnormal returns are the sum of daily abnormal returns. The latter are calculated as the difference between actual returns and the predictions of a market model with parameters estimated over days -110 to -21 relative to the event. A firm's daily abnormal volume is defined as the percentage excess volume over the median volume estimated from -100 to -21 trading days prior to the event, using the Euro value of shares traded as the measure of trading volume. The upper two graphs plot the average cumulative abnormal return for each day for activist (left) and passive events (right). The lower two graphs plot the daily abnormal volumes for activist (left) and passive events (right).

Short-Run Cross-Section. To get a better understanding of the driving forces behind individual CARs, I next regress the CARs for the window with the most pronounced average reaction (day -20 to 0) on a number of variables: the natural logarithm of sales, the return on equity, the market-to-book ratio, the size of the initial stake, the Amihud illiquidity measure,[11] and the free float portion of shares. Rather

[11] The Amihud (2002) measure of illiquidity is calculated as $1/80\ \sum_{t=-100}^{-21} abs(R_{i,t})/V_{i,t}$, where $abs(R_{i,t})$ and $V_{i,t}$ are the stock's absolute return and Euro trading volume on the specific day relative to the event date.

than including a set of covariates that proxy for potential causes of agency problems,[12] I include the return on equity as a measure of profitability and the market-to-book ratio as a valuation measure to proxy for their consequences.

The results are presented in Table 1.7. I report separate regressions for active and passive blocks. Not a single coefficient is significant in either of the regressions. When I consider different timeframes (e.g., abnormal returns on day 0 only or cumulative abnormal returns from day -20 to +20), some coefficients at times become significant, yet there is no consistent picture across regressions that would allow to speak of a robust effect.[13] Of course, it is possible that the regressions simply fail to capture characteristics that drive the cross-sectional differences in abnormal returns. I relegate an interpretation of this non-result to a little later and next explore the long-term stock response and look for changes in central corporate fundamentals.

Long-Run Response. I follow two distinct approaches to evaluate the long-term stock performance of hedge fund targets. The first approach estimates buy-and-hold returns in excess of either the CDAX market index or the characteristic-matched benchmark firm as in Daniel et al. (1997).[14] The CDAX index serves as the first benchmark, even though it is very likely that an index this broad provides an insufficient adjustment as Barber and Lyon (1997) show that empirical rejection rates of tests based on reference portfolios like the CDAX exceed theoretical ones. Still, using the CDAX is a nearby method of calculating an investor-oriented assessment of the hedge funds' impact and makes the results comparable to some of the results in the studies of Bessler, Drobetz, and Holler (2013) and Mietzner and Schweizer (2014). For all buy-and-hold abnormal returns, significance is evaluated using the skewness-adjusted t-statistic suggested by Lyon, Barber, and Tsai (1999). This statistic builds on the statistic by Johnson (1978) that captures skewness in returns, supplemented by a bootstrapping procedure as recommended by Sutton (1993).[15] I drop events with overlapping return windows to avoid misspecified test statistics as documented in Lyon, Barber, and Tsai (1999).

---

[12] Bessler, Drobetz, and Holler (2013), for example, include dummies for different types of incumbent shareholders.

[13] The untabulated cross-section of abnormal volumes provides the same, completely insignificant results.

[14] I only use size, market-to-book, and industry classification to find matching firms. There are too few firms left within the set of candidate matches to also select firms with similar past stock performance after matching on the first three characteristics.

[15] In an earlier version of this paper I also employed the covariance matrix estimator suggested in Jegadeesh and Karceski (2009). The results are essentially the same.

Table 1.7. Cross-section of cumulative abnormal returns

	CAR (-20 to 0)	
	Activist Blocks	Passive Blocks
Sales	-0.0089	0.0035
	(-1.12)	(0.51)
Return on Equity	-0.0002	-0.0001
	(-0.90)	(-0.11)
Market/Book	-0.0022	-0.0075
	(-0.19)	(-0.82)
Initial Stake	-0.0024	0.0139
	(-0.92)	(1.41)
Amihud (× 1000)	-7.5600	16.4142
	(-1.54)	(1.56)
Free Float	0.0007	0.0008
	(1.09)	(1.13)
Constant	01346	-0.1350
	(1.32)	(-0.93)
Observations	98	97
R^2	4.32%	7.94%

Table 1.7 reports the results of cross-sectional analyses for both all activist and all passive events. Financial and real estate firms are excluded due to the incomparability of their accounting numbers. The dependent variables in both panels are the individual firms' cumulative abnormal returns (CARs) for days -20 to 0 relative to the event. Sales is the natural logarithm of the previous fiscal year's sales. Amihud is the measure for illiquidity developed in Amihud (2002) times 10^3, a higher value indicating less liquid stocks. The remaining variables are defined as beforehand. *, **, and *** mark significance levels of 10, 5, and 1 percent. I employ White's (1980) covariance estimator. Amihud's measure, Market/Book, and Return on Equity are all winsorised at 2.5%. t-values are in parentheses.

The second approach employs calendar-time portfolio regressions with the Carhart (1997) 4-factor model. In order to estimate the 4-factor model, I create German equivalents of the size, book-to-market, and momentum factors of Fama and French (1993) and Carhart (1997) from the universe of all CDAX-listed firms. All steps in the construction of the factors mimic those taken by Fama and French (1993) and Carhart (1997) with US data. To mitigate heteroskedasticity in calendar-time portfolio returns that is due to the changing number of constituents over time (Mitchell and Stafford, 2000), I also run calendar-time regressions using only observations where the portfolio includes at least 5 firms. In addition, this approach avoids giving excessive weights to periods with little events. The calendar-time regressions incorporate the findings of Artmann et al. (2012) that the Carhart model is superior to the Fama-French model in explaining the cross-section of German stock returns.

Table 1.8 reports results for the pre-event period (day -360 to -20) and for the post-event period (day +1 to +360). My focus on approximately 1 year after the event aligns with the observation concerning the funds' average holding periods. As a robustness check I also evaluate abnormal returns for the entire holding period of each investment. While this generates problems in the statistical analysis due to overlapping investments in the same firm, the general tendency of the results is unchanged.

In the pre-event period, neither buy-and-hold returns nor calendar-time alphas are significantly different from zero for activist or passive funds' investments. This suggests that target firms on average do not experience any abnormal stock market performance prior to the funds' investments and adds to the contradicting findings in US studies. Klein and Zur (2009) observe that targets of activist hedge funds significantly outperform prior to the purchase. In contrast, Clifford (2008) sees activists' targets significantly underperforming.

The results for the post-event period are insignificant for the subsample of passive blocks. However, the results provide some evidence that activists' targets significantly underperform in the year following the investment. Median buy-and-hold returns are significantly negative when the CDAX is the benchmark ($p \approx 0.05$). They are also negative with the control-firm approach, albeit insignificant. The factor model alpha is negative and significant ($p \approx 0.08$), but only when the restriction of having at least 5 portfolio constituents is applied. Still, the results in combination indicate that the initially positive reaction to the hedge funds' entry is subsequently followed by significantly negative returns.

In unreported analyses I split the sample into pre- (before August 2007) and post-financial crisis (after August 2007) subsamples. In both subsamples I find that activist portfolios earn negative alphas in the post-event period. However, the alphas are insignificant, possibly owing to the comparably small number of observations left in each subsample. These results support the findings of Mietzner and Schweizer (2014) who analyse events until 2007 and find that activist hedge funds' targets significantly underperform in the 250 days following the investment. Bessler, Drobetz, and Holler (2013) look at events until 2006 and provide qualitatively similar results for activist investments.

In summary, both the analysis of buy-and-hold abnormal returns as well as calendar-time portfolios suggest that hedge funds do not create shareholder value in the long run as documented in US studies (Clifford, 2008; Klein and Zur, 2009; Brav et al., 2008) but neither do they seem to destroy it. If anything, the targets' stock

Table 1.8. Long-term stock performance

Activist Blocks	Buy-and-Hold-Returns		4-Factor-Alpha	
	Market Adjusted	Size- and B/M-Adjusted	A)	B)
Pre-Event (Day -360 to -20)	-8.40%	2.87%	0.0002	0.0006
	(-1.24)	(0.96)	(0.62)	(1.63)
Post-Event (Day +1 to +360)	-18.06%**	-9.17%	-0.0007	-0.0005*
	(-1.96)	(-1.32)	(-1.49)	(-1.76)

Passive Blocks	Buy-and-Hold-Returns		4-Factor-Alpha	
	Market Adjusted	Size- and B/M-Adjusted	A)	B)
Pre-Event (Day -360 to -20)	-6.61%	3.19%	0.0006	0.0006
	(-0.73)	(0.41)	(1.64)	(1.64)
Post-Event (Day +1 to +360)	-14.36%	-11.30%	0.0002	0.0001
	(-0.45)	(-1.63)	(0.65)	(0.19)

Table 1.8 presents the analysis of the long-term stock performance for the period preceding the event (day -360 to -20) and the period following the event (day +1 to +360). I report median buy-and-hold returns in excess of the market index (CDAX) and in excess of the matched firms' returns (see text). The alphas are from 4-factor regressions employing the Carhart (1997) model, with (B) or without (A) the restriction that there are at least 5 constituents in the portfolio. The significance of buy-and-hold abnormal returns is assessed via the bootstrapped statistic suggested by Lyon, Barber, and Tsai (1999). *, **, and *** indicate statistical significance of 10, 5, and 1%, respectively. t-statistics are reported in parentheses.

performance seems to be average when considering the combination of the period surrounding the initial investment and the following year.

Long-Run Changes in Characteristics. If hedge funds affect corporate policies, then this might result in discernible changes in observable firm characteristics following their investments. To explore this conjecture, I compare several target firm characteristics over time, adjusted for the respective changes in matched firms.[16] I first calculate the changes in each characteristic between year 0 and 1 and between year 0 and 2 for a given firm, and then I subtract the respective contemporaneous change in the characteristics of the matched firm. The resulting difference provides a test whether the target firm has abnormally changed in the particular dimension. I restrict my attention to observations around the entry of the first activist to avoid

[16] I winsorise the benchmark-adjusted changes at 2.5% to reduce the influence of a number of far-outlying observations.

overlapping timeframes.

One would expect that successful activism is mirrored in some way in core corporate characteristics. For example, if firms suffer from corporate governance problems that impair firm performance, successful activism aimed at improving corporate performance should result in abnormal changes in return on equity or assets. To evaluate the impact of activism, I consider changes in the following firm characteristics: total assets, return on equity, return on (lagged) assets, the ratio of free cash flow to sales, the market-to-book ratio, leverage, dividend yield, the ratio of cash to assets, and board turnover.[17]

Table 1.9 provides the results. With the exception of a significant increase in board turnover, none of the fundamental characteristics changes significantly when compared to changes in matched firms.[18] The increase in board turnover is consistent with the findings in Klein and Zur (2009) that US hedge funds achieve their stated objective of altering a board's composition in more than 7 of 10 activist investments. However, the observation that targets do not improve in either of the other firm characteristics indicates that activism in Germany is ineffective in fundamentally affecting target firms. Of course, it may be the case that heterogeneity in activist attempts and goals prevents the identification of effects.[19] Nevertheless, the results are consistent with earlier literature on shareholder activism that finds little evidence of increases in operational performance (Del Guercio and Hawkins, 1999; Karpoff, 2001; Wahal, 1996) as well as the study by Klein and Zur (2009) on hedge fund activism in particular that even shows a deterioration of profitability.

The results run counter to Clifford's (2008) finding of significant increases in targets' returns on assets. The set of variables considered in this analysis is restricted to a small number of central variables and it did not include a number of potential alternative variables that relate to a firm's governance, for example. However, even if hedge funds did affect corporate fundamentals in some way, the long-term event study results suggest that such changes are either not recognised or valued by investors.

---

[17] I define board turnover as the number of changes in the composition of the board of executives divided by the number of seats in the year of the fund's entry.

[18] When interpreting the changes, I focus on medians and the results of Wilcoxon's (1945) signed rank test.

[19] Del Guercio and Hawkins (1999) observe heterogeneity in pension fund activists' strategies and outcomes. Similarly, Cronqvist and Fahlenbrach (2009) observe heterogeneity between different types of blockholders.

Table 1.9. Changes in activists' targets' characteristics

	+1 Year				+2 Years			
	Average Difference (1)	Average Adj. Difference (2)	Median Difference (3)	Median Adj. Difference (4)	Average Difference (5)	Average Adj. Difference (6)	Median Difference (7)	Median Adj. Difference (8)
Ln(Assets)	0.2054**	0.1726	0.0291**	-0.0283	0.2392**	0.0618	0.0822**	0.0500
	(0.04)	(0.12)	(0.02)	(0.98)	(0.02)	(0.59)	(0.02)	(0.76)
Return on Equity	-5.90%	-3.53%	-0.06%	1.56%	-5.19%	3.71%	-1.86%	-1.60%
	(0.37)	(0.61)	(0.35)	(0.76)	(0.48)	(0.69)	(0.47)	(0.89)
Return on Assets	1.85%	1.19%	0%	0.29%	-1.40%	1.40%	0.99%	0.63%
	(0.25)	(0.44)	(0.73)	(0.48)	(0.53)	(0.50)	(0.55)	(0.84)
Return on Lagged Assets	-2.83%	-5.37%**	-1.09%	-1.78%	-6.84%**	-5.09%	-1.79%	-0.75%
	(0.24)	(0.03)	(0.24)	(0.16)	(0.04)	(0.15)	(0.16)	(0.52)
Free Cash Flow/Sales	0.0152	0.0356	0.0030	-0.0136	-0.0250	0.0335	0.0045	0.0182
	(0.39)	(0.29)	(0.83)	(0.45)	(0.46)	(0.59)	(0.90)	(0.77)
Market/Book	0.0314	-0.3164	-0.0500	0.1100	-0.2749	0.1184	-0.3750**	-0.2600
	(0.90)	(0.41)	(0.59)	(0.88)	(0.30)	(0.71)	(0.04)	(0.68)
Leverage	-0.0097	-0.0572**	-0.0270	-0.0132	0.0185	-0.0250	-0.0054	0.0062
	(0.69)	(0.05)	(0.40)	(0.14)	(0.55)	(0.50)	(0.93)	(0.73)
Dividend Yield	0.54%**	0.53%	0%**	0%	0.91%**	0.88%*	0%**	0%
	(0.04)	(0.13)	(0.01)	(0.12)	(0.03)	(0.10)	(0.01)	(0.15)
Cash/Assets	-0.0855	-0.0578	0.0096	0.0048	-0.0863*	-0.0475	-0.0062	-0.0082
	(0.11)	(0.16)	(0.57)	(0.56)	(0.10)	(0.28)	(0.43)	(0.59)
Management Turnover	0.0312	0.1000	0	0.1389*	0.0671	0.1769**	0*	0**
	(0.49)	(0.10)	(0.57)	(0.06)	(0.23)	(0.03)	(0.09)	(0.03)

Table 1.9 reports changes in various measures of profitability and further characteristics in the first and second year following the entry of the first activist hedge fund into a company. I separately present unadjusted differences ((1), (3), (5), and (7)) and differences adjusted for the change in the respective matched firm's characteristics ((2),(4),(6), and (8)). Mean differences are given in (1), (2), (5), and (6), including the p-values of a simple t-test in parentheses. Median differences are given in (3), (4), (7), and (8), including the p-values of a Wilcoxon (1945) signed rank test in parentheses. The columns of interest are (2), (4), (6), and (8); they present the abnormal changes in target characteristics. Management Turnover is the number of changes in the board of executives divided by the board's size in the year of the fund's entry. The remaining variables are defined as in the preceding tables. *, **, and *** mark significance levels of 10, 5, and 1 percent, respectively.

Interpretation. In summary, activist hedge fund investments are associated with positive short-term stock returns and abnormal trading volumes, negative long-term stock returns, and with no changes in central corporate variables in the years following the funds' engagement apart from an increase in board turnover.

But how do the significantly positive immediate abnormal stock returns fit into this picture? Though hedge funds seem to affect corporate control structures, neither the long-term stock returns nor the variables typically considered value-relevant justify the immediate positive reactions. One possible explanation is that this pattern reflects overoptimism on investors' side. As pointed out by Brav et al. (2008), shareholders build some belief regarding the expected gains from activism upon seeing a hedge fund enter a firm. This requires —among other things— an estimate of the strength of the incumbent shareholders' resistance. These incumbent shareholders own on average approximately 29% of all voting shares (~30% for passive blocks) in the month prior to entry. To the extent that hedge fund activism targets corporate inefficiencies that benefit the shareholders in control, agreement between the fund and the dominant incumbent becomes unlikely. Owing to the comparably weak rights of minority shareholders in Germany, this likely puts the incumbent shareholders in a strong defensive position. The fiercely fought public debates around hedge fund activism at Deutsche Börse, Kuka, TUI, and several other firms present anecdotal evidence for such disagreements. Still, it raises the question as to why investors would initially form biased beliefs when responding to hedge fund investments.

I favour the explanation that the immediate market response to hedge fund investments reflects attention-induced buying as suggested in Barber and Odean (2008). Activist hedge fund investments are typically accompanied by plenty of media coverage as well as press releases by target firms and the investing funds. This is likely to raise attention among potential shareholders and might induce them to flock into the target stock. It is thus conceivable that as hedge funds push particular stocks into the spotlight, these stocks are equally pushed up in price — yet not because hedge funds are believed to generate value but because they generate substantial attention among prospective investors. This would explain high abnormal volumes, and initially positive as well as subsequently negative abnormal returns. It would also explain why none of the covariates in the cross-sectional regressions of short-run CARs is significant.

1.6 Conclusion

This paper explores the effects of hedge fund investments in Germany based on a sample of events covering the period between 1999 and 2010. The results I portray indicate that—all-in-all—hedge fund activism in German firms is ineffective. I neither find that target firms improve (or change at all, for that matter) in terms of corporate fundamentals, nor do I find a permanent change in firm valuations, though hedge funds seem to affect the composition of the targets' boards of management. Interestingly, market participants seem to respond inefficiently to the initial publication of an activist's investment. Though activist hedge fund investments are accompanied by substantial and significantly positive abnormal returns as well as high abnormal trading volumes in the days around the publication of the investment, long-run abnormal returns in the post-event period are negative. Over the entire period, i.e., from shortly before the announcement of investment to one year later, abnormal returns are insignificant. I interpret the combination of these results as evidence of attention-based buying in the spirit of Barber and Odean (2008). As hedge funds invest into firms, they generate news surrounding the companies they target. This raises the attention of investors and temporarily pushes up the targets' prices. Eventually, prices revert to the levels before the investment.

A question left unanswered in this paper is why hedge fund activism is ineffective. One may conjecture that intensely concentrated ownership structures in combination with rather weak shareholder rights pose a considerable challenge for shareholder activism. It is conceivable that changes to a firm's management as indicated by the increase in board turnover are the only concession that large shareholders are willing to make when confronted with hedge fund activists. This, however, may not be enough for the latter to affect the target firm in terms of fundamentals and valuation. I leave a more thorough discussion of this question to further research.

2 Recommendation Revisions and Differences in the Interpretation of Earnings

2.1 Introduction

Recommendations to buy, hold, or sell specific stocks are one of the main outlets of sell-side analyst research. Existing literature (e.g., Womack, 1996 or more recently Bradley et al., 2014) finds that recommendation revisions are associated with significant movements in stock prices, typically increasing following upgrades and decreasing following downgrades. Evidence of this type, i.e., cross-sectional patterns in returns around the release of recommendations, suggests that investors attribute value to analysts' opinions. What is less clear is what information investors respond to when reacting to recommendation revisions. In this paper, I attempt to shed light on this issue by contrasting patterns in investors' responses to recommendations—a proxy for the information investors extract—to patterns in how recommendations relate to firms' earnings processes—a proxy for the information recommendations contain.

Post-earnings revisions of pre-earnings recommendations provide a unique setting to compare investors' and analysts' earnings interpretation. When revising their recommendations, analysts indicate that the difference between their own estimate of a firm's value and the current market price has changed (e.g., Francis and Soffer, 1997). An upgrade from a hold to a buy recommendation, for example, reveals that a firm's market value has fallen below the value the analyst considers appropriate. The central idea behind this paper is that recommendation revisions following an earnings release indicate whether analysts agree with investors' earnings interpretation. Employing a simple model, I build on this premise and devise two distinct empirical strategies to identify whether analysts have an edge over investors in the interpretation of earnings.

The first strategy aims to approximate the process by which investors build expectations of future earnings. If analysts' interpretation of earnings surprises is on average superior to investors', then recommendation revisions should capture variation in earnings that is not predicted by the process underlying investors' expectations. Moreover, they should do so in a specific, predictable way. While confirmatory revisions, i.e., upgrades following positive surprises or downgrades following negative surprises, should indicate exceptionally permanent earnings surprises, contradictory revisions, i.e., downgrades following positive surprises or upgrades following negative surprises, should indicate exceptionally transitory earnings surprises. A caveat with these predictions, however, is that empirically testing them necessitates imposing a model to proxy for investors' expectations of future earnings.

The second strategy attempts to overcome this problem. If investors value analysts' assessment of their own initial response to the earnings release, then their response to recommendation revisions should exhibit predictable patterns. Specifically, the association between stock returns around revisions and past earnings surprises should vary with the type of the recommendation. While returns should be positively associated with the preceding earnings surprise if a revision is confirmatory, the association should be negative if it is contradictory. Put differently, the idea behind the second strategy is to explore whether changes in investors' perception of earnings surprises induced by recommendation revisions are in line with what the model predicts them to be if investors consider analysts' skill at interpreting earnings superior to their own.

I test these predictions using a large sample of recommendation revisions released by analysts on record with IBES in the period between 1994 and 2013. Supporting the idea that revised recommendations contain information about past earnings surprises, I find that the association between future earnings and past earnings surprises varies as predicted by the model. It is exceptionally strong, i.e., more positive than average, when earnings are followed by confirmatory revisions, and it is exceptionally weak, i.e., less positive than average, when they are followed by contradictory revisions. These results are robust to different ways of measuring earnings surprises and they hold while accounting for firm-, industry-, size-, or time-specific differences in how current earnings surprises relate to future earnings.

To understand whether investors value this source of information, I analyse cross-sectional variation in returns around recommendation revisions. In support of the model's prediction, I find that abnormal returns are positively associated with past surprises around confirmatory revisions, and they are negatively associated around contradictory revisions. These results hold while accounting for a host of covariates known to drive investors' responses to recommendations, including the level of the recommendation and the change in recommendation levels, and they are insensitive to different ways of measuring abnormal returns. This suggest that investors are aware of the information that recommendation revisions reveal about past earnings surprises, and they incorporate this information to reassess their estimate of firm value.

In further regressions I explore differences in analyst skill. I document that recommendations by more experienced analysts and analysts from larger brokerage firms provide stronger signals, suggesting enhanced skill at assessing whether an earnings surprise is likely to reflect a persistent change in earnings. I then show that cross-sectional differences in investors' responses to revisions indicate awareness of these between-analyst differences. Return responses are more positively associated with the earnings surprise around confirmatory revisions and more negatively associated with the earnings surprise around contradictory revisions if either of them are released by more experienced analysts or analysts working with larger brokerage firms.

This paper relates to the literature along several lines. First, it is a contribution to the extensive literature on the value of analyst research (e.g., Stickel, 1991; Womack, 1996; Francis and Soffer, 1997; Ivković and Jegadeesh, 2004; Asquith, Mikhail, and Au, 2005; Chen, Cheng, and Lo, 2010; Livnat and Zhang, 2012; Bradley et al., 2014). The extant literature's primary focus has been the cross-sectional analysis of stock market reactions to analyst publications. The main contribution of this paper is to show that patterns in how revisions in analysts' recommendations relate to fundamental, value-relevant information—future earnings—parallel the patterns in how investors respond to their releases.

Second, there is some evidence that investors revise their interpretation of past earnings using corroborating firm information like subsequent earnings releases (Freeman and Tse, 1989), dividend changes (Koch and Sun, 2004), or insider transactions (Veenman, 2012). However, little research has explored whether investors use non-firm releases to the same end. Particularly close to my paper are Mendenhall (1991) and Park and Pincus (2000). Findings in both papers suggest that patterns in returns around changes in consensus forecasts and recommendations indicate a reinterpretation of past earnings. Their results, however, remain suggestive in that they do not relate the patterns in returns to those in actual earnings. My paper extends and lends credence to their findings by providing the fundamental analogue to the patterns they document in investors' responses.

Third, previous literature finds that the accuracy of analyst forecasts is larger for more experienced analysts or analysts from larger brokerage firms (Mikhail, Walther, and Willis, 1997, 2003; Clement, 1999). My results provide an explanation for this pattern in that they show the same differences in analysts' abilities to extrapolate future earnings from current earnings surprises. In other words, my results indicate that one reason for the observed differences in analysts' forecasting accuracy is that more experienced analysts and analysts from larger brokerage firms are better at deducing what current earnings surprises imply for future earnings.

The paper proceeds with the development of hypotheses in Section 2.2, where I use a simple model to motivate my empirical analyses. Section 2.3 describes the data collection procedure and defines the main variables. Section 2.4 presents the results. In Section 2.5, I explore cross-sectional heterogeneity in analysts' abilities and the associated differences in investors' responses. I present robustness checks and additional analyses in Section 2.6 and conclude in Section 2.7.

2.2 Empirical Predictions

I develop my empirical predictions using a stylised model that builds on three premises: (1) The implications of current earnings surprises for future earnings are uncertain. (2) Market participants and analysts sometimes differ in their interpretation of earnings surprises, and (3), analysts revise recommendations when their assessment of relative firm valuation changes, i.e., when the difference between their own estimate of firm value and the current market price changes.

Based on the above premises, I derive two distinct empirical approaches to compare analysts' and investors' skill in the interpretation of earnings. The first approach attempts to approximate investors' estimate of the association between a given period's earnings surprise and the following period's earnings. The model makes directional predictions how addition of recommendation revisions should improve this estimate—if analysts' interpretation of earnings is superior to investors'. The second approach looks at investors' responses to the publication of recommendation revisions instead and attempts to identify analyst superiority by exploring whether differences in the response to recommendations suggest a reinterpretation of previous earnings surprises.

Let q and $q+1$ indicate two consecutive periods and let E_q and E_{q+1} be the associated actual values of earnings per share (henceforth just earnings). Denote investors' (INV) and analysts' (AN) expectations of earnings in q just before observing the earnings announcement in q by $\mathbb{E}_{pre-q}[E_q]^{INV}$ and $\mathbb{E}_{pre-q}[E_q]^{AN}$. Similarly, let $\mathbb{E}_{pre-q}[E_{q+1}]^{INV}$ and $\mathbb{E}_{pre-q}[E_{q+1}]^{AN}$ be their expectations of earnings in $q+1$ before observing earnings in q.

Assume earnings surprises contain both permanent and transitory components. While permanent components reflect shifts in the expected level of all future earnings, transitory components have no consequences beyond the current period.[1] Let p^{INV} and p^{AN} denote investors' and analysts' estimate of the fraction of an earnings surprise that is permanent and let $1-p^{INV}$ and $1-p^{AN}$ be the fractions they believe to be transitory.

Following the announcement of period q earnings, investors' and analysts' expectation of earnings in period $q+1$ can be written as the sum of their priors and their expectation of the permanent fraction of the earnings surprise:

$$\mathbb{E}_{post-q}[E_{q+1}]^{INV} = \mathbb{E}_{pre-q}[E_{q+1}]^{INV} + p^{INV}(E_q - \mathbb{E}_{pre-q}[E_q]^{INV}) \tag{2.1}$$

and

$$\mathbb{E}_{post-q}[E_{q+1}]^{AN} = \mathbb{E}_{pre-q}[E_{q+1}]^{AN} + p^{AN}(E_q - \mathbb{E}_{pre-q}[E_q]^{AN}). \tag{2.2}$$

Thus, both analysts and investors alter their estimates of future earnings by the fraction of the current surprise they consider permanent.

To link this change in expectations to market valuations, I use earnings response coefficients (e.g., Freeman and Tse, 1989). Intuitively, earnings response coefficients or ERCs describe how a (1 unit-) change in unexpected earnings maps into stock returns. Let ERC_p and ERC_t be the ERCs for permanent and transitory components of earnings surprises and assume that both analysts and investors apply the same ERCs. Note that because permanent changes in the level of earnings have more pronounced implications for firm value, $ERC_p > ERC_t$.

The initial stock market response to the earnings announcement in period q (R_q) can be written as the sum of ERC-weighted earnings surprise components:

$$R_q = (p^{INV} ERC_p + (1-p^{INV})ERC_t)(E_q - \mathbb{E}_{pre-q}[E_q]^{INV}). \tag{2.3}$$

Because investors' opinions determine market prices, this response includes p^{INV} and $\mathbb{E}_{pre-q}[E_q]^{INV}$ rather than p^{AN} and $\mathbb{E}_{pre-q}[E_q]^{AN}$. In consequence, the change in value observed around the earnings announcement, R_q, does not need to coincide with the change in value analysts deem appropriate. To understand how this affects an analyst's assessment of relative firm value, what investors can learn about

[1] This dichotomy serves to keep matters as simple as possible. The essential assumption necessary for all of the following arguments is that earnings surprises contain different components, some of which are longer-lived and thus more relevant for firm value than others.

p^{AN} from recommendation revisions, and why this may help in contrasting analysts' and investors' skill in the interpretation of earnings, it is necessary to take a step back and discuss why analysts revise recommendations.

Francis and Soffer (1997, p.193) view "stock recommendations as expressions of analysts' beliefs about share values relative to their market prices". The release of a "sell", for example, suggests that an analyst believes a firm is currently trading above its fundamental value. As an analyst revises a recommendation and upgrades, reiterates, or downgrades a stock, the analyst reveals whether her assessment of a firm's relative valuation has changed since she last released a recommendation. Thereby, a revised recommendation indicates whether an analyst agrees with how investors' responded to news released following the previously released recommendation. This distinguishes recommendation revisions not only from earnings forecasts and changes in earnings forecasts but also from recommendation levels.[2]

Note that even though analysts and investors may disagree on their estimates for earnings in period $q+1$ prior to period q ($\mathbb{E}_{pre-q}[E_{q+1}]^{AN}$ and $\mathbb{E}_{pre-q}[E_{q+1}]^{INV}$), there can only be changes in an analyst's assessment of relative firm value if either p^{AN} differs from p^{INV} or $\mathbb{E}_{pre-q}[E_q]^{AN}$ differs from $\mathbb{E}_{pre-q}[E_q]^{INV}$, because only the earnings surprise and its composition change analysts' resp. investors' estimate of firm value. To simplify the argument, I equate $\mathbb{E}_{pre-q}[E_q]^{INV}$ and $\mathbb{E}_{pre-q}[E_q]^{AN}$ as well as $\mathbb{E}_{pre-q}[E_{q+1}]^{INV}$ and $\mathbb{E}_{pre-q}[E_{q+1}]^{AN}$ in the following and assume that analysts and investors agree on their expectation of E_q and E_{q+1} just before the announcement in q, henceforth denoted by $\mathbb{E}_{pre-q}[E_q]$ and $\mathbb{E}_{pre-q}[E_{q+1}]$. I thus assume that—at least for one and two periods ahead—analyst estimates provide a reasonable proxy for investors' expectations of earnings and I abstract from situa-

---

[2] To see this, first consider a change in an analyst's forecast following an earnings announcement. Without knowing investors' post-announcement estimate of future earnings, it is unclear whether this new forecast is higher, lower, or equal to investors' new expectation. Thus, revised forecasts on their own only reveal whether the analyst changed her expectation of future earnings, but not whether she disagrees with the market's earnings interpretation. Similarly, without contrasting recommendations to their previous level, they provide an ambiguous signal about the analysts' assessment of the market's response to earnings. To see this, consider a "hold" recommendation following an earnings announcement. If the previous recommendation was a "buy", the new recommendation suggests investors' response to the earnings release was too positive. If the previous recommendation was a "sell", however, the new recommendation suggests the opposite, i.e., that investors responded too negatively to the release. Thus, while recommendation levels can indicate disagreement between analysts and market participants about the level of a firm's valuation, they cannot without further information indicate disagreement concerning changes in firm valuation.

tions in which investors and analysts disagree on the surprise itself.[3,4]

How will analysts revise their recommendations after observing earnings surprises and the market's response to earnings? I will now go through the possible cases.[5]

If an analyst's belief about the permanent component of the earnings surprise coincides with investors' ($p^{AN} = p^{INV}$), the immediate response in (2.3) aligns with the analyst's assessment and her estimate of relative firm value stays constant. Thus, the analyst will reiterate her old recommendation. This changes when analyst and investors disagree on the permanent fraction of the surprise.

There are two situations in which the immediate response increases an analyst's assessment of relative firm value: Either if the analyst perceives investors' response not positive enough following better-than-expected earnings or too negative following worse-than-expected earnings. In both cases, the analyst will want to release an upgrade following the earnings announcement. In other words, the analyst will use an upgrade to communicate that a firm has become more undervalued / less overvalued over the earnings release:

$$Upgrade_q \text{ if} \begin{cases} E_q > \mathbb{E}_{pre-q}[E_q] \text{ and } p^{AN} > p^{INV} \text{ or} \\ E_q < \mathbb{E}_{pre-q}[E_q] \text{ and } p^{AN} < p^{INV}. \end{cases} \tag{2.4}$$

In contrast, every time an analyst thinks that the response to a negative earnings surprise is not negative enough or the response to a positive surprise is is too positive, she will use a downgrade to communicate that a stock has become more overvalued / less undervalued:

$$Downgrade_q \text{ if} \begin{cases} E_q > \mathbb{E}_{pre-q}[E_q] \text{ and } p^{AN} < p^{INV} \text{ or} \\ E_q < \mathbb{E}_{pre-q}[E_q] \text{ and } p^{AN} > p^{INV}. \end{cases} \tag{2.5}$$

---

3 There is an extensive literature that debates whether analyst forecasts or time-series models provide better surrogates for investors' earnings expectations. See Kothari (2001) for a review. I present results for a time-series model of earnings expectations in Section 2.6.

4 This does not imply that analysts and investors agree on firm value prior to earnings. For one, differences between the analysts' estimate of firm value and investors' may also derive from differences in expectations of earnings beyond period $q+1$. Also, in line with the basic idea of the model, analysts' estimate of firm value may deviate from investors' in case they consider previous responses to earnings inadequate.

5 I focus on the distinction between upgrades, downgrades, and reiterations, and I do not make a distinction between upgrades and downgrades that change the level of recommendation by one notch only and those that change it by several notches.

Two variables, $Confirm_q$ and $Contradict_q$, capture these arguments and characterise what recommendation revisions reveal about an analysts' assessment of the stock market's immediate response to earnings:

$$Confirm_q = 1 \text{ if} \begin{cases} E_q > \mathbb{E}_{pre-q}[E_q] \text{ and } Upgrade_q \\ E_q < \mathbb{E}_{pre-q}[E_q] \text{ and } Downgrade_q \end{cases}, \text{else } 0, \tag{2.6}$$

and

$$Contradict_q = 1 \text{ if} \begin{cases} E_q < \mathbb{E}_{pre-q}[E_q] \text{ and } Upgrade_q \\ E_q > \mathbb{E}_{pre-q}[E_q] \text{ and } Downgrade_q \end{cases}, \text{else } 0. \tag{2.7}$$

The labelling of the variables is borrowed from Freeman and Tse (1989) and indicates that a recommendation revision with the same sign as the preceding earnings surprise confirms the latter, while a revision with the opposite sign contradicts it.

I now turn to the derivation of empirical predictions to identify analysts' superiority in the interpretation of earnings. Assume investors' best estimate of p^{INV} can be modelled by observable characteristics $\mathbf{X}_q$. Then, one can write their posterior estimate of next period's earnings as

$$\mathbb{E}_{post-q}[E_{q+1}]^{INV} = \mathbb{E}_{pre-q}[E_{q+1}] + \delta'\mathbf{X}_q(E_q - \mathbb{E}_{pre-q}[E_q]). \tag{2.8}$$

Essentially, this equation states that investors use observable characteristics to estimate how current surprises map into future earnings. Assume an analyst can extract this guess from investors' earnings response and releases recommendations according to (2.4) and (2.5). Relying on the definitions of $Confirm_q$ and $Contradict_q$, her best estimate of next period's earnings can be written as

$$\begin{aligned} \mathbb{E}_{post-q}[E_{q+1}]^{AN} =& \mathbb{E}_{post-q}[E_{q+1}]^{INV} \\ &+ Confirm_q \delta_1 (E_q - \mathbb{E}_{pre-q}[E_q]) + Contradict_q \delta_2 (E_q - \mathbb{E}_{pre-q}[E_q]), \end{aligned} \tag{2.9}$$

where δ_1 and δ_2 indicate by how much the analyst's estimate of p differs from investors'. The comparison of actual earnings to this posterior is the basis for Hypothesis 1:

H1 If analysts' earnings interpretation is on average superior to investors', then current earnings surprises should be positively (negatively) associated with unexpected future earnings following confirmatory (contradictory) revisions.

Hypothesis 1 can be written in the form of a regression equation:

$$\begin{aligned} E_{q+1} - \mathbb{E}_{pre-q}[E_{q+1}] = \alpha &+ \beta_1(E_q - \mathbb{E}_{pre-q}[E_q]) + \boldsymbol{\beta_2}'\mathbf{X}_q + \boldsymbol{\beta_3}'\mathbf{X}_q(E_q - \mathbb{E}_{pre-q}[E_q]) \\ &+ \gamma_1 Confirm_q(E_q - \mathbb{E}_{pre-q}[E_q]) \\ &+ \gamma_2 Contradict_q(E_q - \mathbb{E}_{pre-q}[E_q]) + \epsilon_{q+1}. \end{aligned} \tag{2.10}$$

If analysts' skill in interpreting earnings surprises exceeds investors', γ_1 in (2.10) should be positive and γ_2 should be negative. For (2.10) to identify whether analysts' earnings interpretation is superior to investors', the included variables need to contain a sufficient approximation of investors' expectations regarding next period's earnings. To this end, the regression includes a vector of controls, $\mathbf{X}_q$, as well as its interaction with the earnings surprise. The interaction is intended to approximate investors' estimate of p^{INV}. The levels of $\mathbf{X}_q$ allow for the characteristics to also have an effect on the level of next period's earnings expectation.

The problem with (2.10) is that it jointly tests whether analysts have skill at interpreting earnings and whether the choice of controls $\mathbf{X}_q$ adequately describes investors' earnings expectation. Unless the controls are adequate, the coefficients γ_1 and γ_2 might pick up something not included in the model but part of investors' information set. In the empirical section I will approach this issue by including a large number of controls in $\mathbf{X}_q$ to allow for variation of expectations in time-, firm-, and industry-specific factors.

An alternative way to identify whether analysts' earnings interpretation is superior to investors' is to look at market responses to recommendation revisions. Consider the return around a recommendation revision following an earnings release, R_{REV}. This return can be written as

$$R_{REV} = (p^{NEW} - p^{INV})(ERC_p - ERC_t)(E_q - \mathbb{E}_{pre-q}[E_q]) + \beta'\mathbf{X}_{REV}. \tag{2.11}$$

While part of the response is due to factors unrelated to past earnings, $\mathbf{X}_{REV}$, (2.11) includes the product of the past earnings surprise, $E_q - \mathbb{E}_{pre-q}[E_q]$, and the difference between investors' revised expectation of p after observing the revision, p^{NEW}, and their initial estimate, p^{INV}. Because $ERC_p > ERC_t$, the sign of the association between the return around the revision and the past surprise depends on the sign of $p^{NEW} - p^{INV}$. In consequence, this sign can be used to infer how investors alter their estimate of the permanent fraction of the surprise, p^{INV}, upon observing an analyst's recommendation revision. The previous arguments suggest three possible situations.

If analysts release confirmatory revisions, they indicate that $p^{AN} > p^{INV}$. Investors who attribute value to this signal should adjust their estimate p^{INV} to some value $p^{NEW} > p^{INV}$,[6] and returns around confirmatory revisions should be positively associated with the preceding earnings surprise. In case analysts release a contradictory revision, they signal that $p^{AN} < p^{INV}$. Thus, investors who believe that an analyst's interpretation is valuable should adjust their estimate of p to some value $p^{NEW} < p^{INV}$. In consequence, returns should be negatively associated with the past earnings surprise around contradictory revisions. Finally, if analysts reiterate their previous recommendation, they indicate that the market's estimate of p coincides with their own. In this situation, returns should not be associated with past earnings.

Following this argument and using the previous definitions, I can rewrite (2.11) as a regression equation:

$$\begin{aligned} R_{REV} = \alpha + \beta' \mathbf{X}_{REV} + \gamma_1 Confirm_q (E_q - \mathbb{E}_{pre-q}[E_q]) \\ + \gamma_2 Contradict_q (E_q - \mathbb{E}_{pre-q}[E_q]) + \epsilon_{REV}. \end{aligned} \tag{2.12}$$

If revisions conform to the logic outlined previously and if investors believe that analysts' interpretation of earnings is superior to their own, then γ_1 should be positive, whereas γ_2 should be negative.

H2 If investors consider analysts' earnings interpretation superior to their own, then abnormal returns should be positively (negatively) related to past earnings surprises around confirmatory (contradictory) recommendation revisions.

Equation (2.12) identifies (investor-assessed) analyst superiority in interpreting earnings in case the controls in $\mathbf{X}_{REV}$ adequately control for alternative drivers' of investors' responses to recommendation revisions.

Taken together, I propose two hypotheses aimed at testing for analyst superiority in the interpretation of earnings. Both attempt to identify skill relative to investors. Testing of the first hypothesis relies on the assumption that equation (2.10) is successful at capturing investors' information set concerning the relation between current surprises and future earnings. Essentially, the underlying idea is to treat investors like an econometrician who attempts to model the mapping between current earnings surprises and future earnings and see whether the addition of analyst

[6] How strongly investors' posterior moves into the direction of p^{AN} depends—among other things—on how strong a signal a revision is. I will return to this point in Section 2.5

information improves upon her hypothesised model in the way predicted by theory. Testing the second hypothesis instead requires that specification (2.12) is successful at modelling the determinants of (abnormal) stock returns around recommendations revisions.

2.3 Sample and Central Variables

2.3.1 Sample

I obtain data for this study from 2 sources. I download data on earnings, earnings forecasts, and recommendations from the Institutional Brokers' Estimate System (IBES), and I download data on stocks, indices, and industry affiliations from the Center for Research in Security Prices (CRSP). My sample covers the period between January 1994 and December 2013.

Because the focus of this study lies on recommendation revisions released between two adjacent quarterly earnings announcements that revise older recommendations released prior to the earlier announcement, I impose the following data requirements:

1. Actual values for earnings per share and earnings announcement dates for period q and $q+1$ are available on the IBES tapes;

2. Data on CRSP are available around the recommendation date and for at least 30 days prior to the period q announcement;

3. The recommendation has a non-zero analyst identifier on the IBES tapes;

4. The old recommendation is released prior to the period q earnings announcement; The new recommendation is released between the announcements in q and $q+1$;

5. There is at least 1 trading day between the release of the new recommendation and both the previous and subsequent earnings release;

6. There are at least 3 analyst forecasts for earnings in period q and period $q+1$ in the 120 calendar days preceding the announcement for period q;

7. The earnings surprise in period q is not 0.

Data requirements (1) to (3) are necessary to construct the variables employed in the empirical analyses. Criterion (4) restricts the sample to post-earnings revisions of recommendations released prior to the earnings release. (5) is a precaution to ensure that revisions are released strictly after the period q announcement and strictly before the period $q+1$ announcement. (6) ensures that the measures of earnings expectation I employ are sufficiently accurate. I impose (7) because there is no new value-relevant information that might later be reconsidered in situations where earnings are as expected.[7]

From all observations that fulfil these requirements, I drop firms with stock prices below 10 dollars or market values below 100 million dollars. The market value requirement ensures that my results are not driven by small firms. I impose the price requirement because I standardise the earnings variables by the stock price. Setting a minimal value of 10 dollars prevents the occurrence of extreme values. I also require that there are at most 180 calendar days between the two earnings announcements and thus drop cases with an unusually large period of time between releases for consecutive quarters.

My final sample comprises $85,011$ pairs of sequential earnings announcements with $187,288$ intermittent recommendation revisions.

2.3.2 Variables and Descriptive Statistics

Testing Hypotheses 1 and 2 requires expectations of the current and the next period's earnings before the announcement in period q, $\mathbb{E}_{pre-q}[E_q]$ and $\mathbb{E}_{pre-q}[E_{q+1}]$. As a surrogate for the unobservable market expectations I employ the consensus analyst forecasts in IBES. To construct the consensus, I download the IBES Detail tape with all individual analyst forecasts for earnings per share over the sample period. I then calculate $\mathbb{E}_{pre-q}[E_q]$ as the mean of all most recent analyst forecasts over the 120 days preceding the earnings announcement in q. Analogously, I calculate expectations for period $q+1$ earnings before period q, $\mathbb{E}_{pre-q}[E_{q+1}]$, as the mean of all most recent analyst forecasts for earnings in $q+1$ over the 120 days preceding the earnings release of q. Thus, both the expectations for earnings in q and $q+1$ are calculated in the days preceding the period q earnings release.

Using actual earnings for E_q and E_{q+1} from the IBES Detail tape and the split-adjusted stock price (P_q) 30 days prior to the announcement in q, I then define

$$SUE_q = \frac{E_q - \mathbb{E}_{pre-q}[E_q]}{P_q} \tag{2.13}$$

[7] Koch and Sun (2004) impose a similar restriction when exploring whether dividend announcements convey information about previously released earnings. They eliminate observations where seasonal earnings changes are zero.

and

$$FSUE_{q+1} = \frac{E_{q+1} - \mathbb{E}_{pre-q}[E_{q+1}]}{P_q}. \tag{2.14}$$

SUE (standardised unexpected earnings) quantifies the period q earnings surprise. $FSUE$ is the difference between period $q+1$ earnings and the expectation for period $q+1$ earnings that prevailed prior to observing the announcement in period q. I employ the price to deflate both variables to make earnings comparable across firms and mitigate heteroskedasticity in later regressions.[8] In all my analyses, I winsorise SUE and $FSUE$ at the lower and upper 1 percent tails of their respective distributions to reduce the impact of outlying observations.

From IBES I also download all recommendations released over the sample period. IBES records recommendations as a numerical score from 1 ("strong buy") to 5 ("sell"). I use these ranks to quantify the content of a recommendation revision as in Jegadeesh et al. (2004) or Loh and Stulz (2011). To this end, I first invert the numerical rank (5 = "strong buy", 1 = "sell") and then calculate the recommendation change ($\triangle Rec$) as the difference between the new recommendation rank and the last preceding recommendation rank. $\triangle Rec$ thus ranges from -4 to +4. While positive values correspond to upgrades, negative values indicate downgrades.

I then define two binary variables to characterise whether a given recommendation revision on some day t confirms or contradicts the last preceding earnings surprise (in quarter q):

$$Confirm_t = 1 \text{ if} \begin{cases} SUE_q > 0 \text{ and } \triangle Rec_t > 0 \\ SUE_q < 0 \text{ and } \triangle Rec_t < 0 \end{cases}, \text{else } 0; \tag{2.15}$$

$$Contradict_t = 1 \text{ if} \begin{cases} SUE_q < 0 \text{ and } \triangle Rec_t > 0 \\ SUE_q > 0 \text{ and } \triangle Rec_t < 0 \end{cases}, \text{else } 0. \tag{2.16}$$

Thus, $Confirm_t$ is 1 when the revision on day t is an upgrade and the last preceding earnings surprise is positive or when the revision is a downgrade and the last preceding earning surprise is negative. $Contradict_t$ is 1 if a downgrade follows a positive surprise or an upgrade follows a negative surprise.

Table 2.1 shows that 36.7 percent of the sample's revisions are confirmatory and

---

8 Using the standard deviation of actual earnings over the 20 quarters preceding the earnings release in q instead of the price does not affect the results.

Table 2.1. Summary statistics

Panel A: Earnings announcements and recommendation revisions

	N	SD	P5	Mean	P50	P95
SUE_q	85,011	0.004	-0.005	0.001	0.000	0.006
$FSUE_q$	85,011	0.008	-0.014	-0.001	0.000	0.008
$\triangle Rec_t$	187,288	1.263	-2	-0.115	0	2
$Confirm_t = 1$	68,693					
$Contradict_t = 1$	71,544					
$Reiteration_t = 1$	47,051					
AR_t	187,288	0.072	-0.120	-0.005	-0.001	0.094

Panel B: Joint distribution of $\triangle Rec_t$ and SUE_q

	Upgrade ($\triangle Rec_t > 0$)	Reiteration ($\triangle Rec_t = 0$)	Downgrade ($\triangle Rec_t < 0$)
Better-than-expected earnings	43,151	32,780	52,022
($SUE_q > 0$)	33.7%	25.6%	40.7%
Worse-than-expected earnings	19,522	14,271	25,542
($SUE_q < 0$)	32.9%	24.1%	43.0%

Panel A presents summary statistics for the sample's earnings announcements and recommendation revisions. Panel B presents the distribution of revision types conditional on the sign of the preceding earnings surprise.
SUE_q is the earnings surprise in quarter q, measured as the difference between actual earnings for q and the consensus forecast on the IBES Detail tape in the 120 days preceding the earnings release in q, divided by the stock price 30 days prior to the release. $FSUE_{q+1}$ is the difference between period $q+1$ earnings and the IBES consensus for earnings in $q+1$ in the 120 days preceding the announcement for period q, divided by the stock price 30 days prior to the release in q. $\triangle Rec_t$ is the change in IBES recommendation ranks associated with a recommendation revision on day t. $Confirm_t$ is 1 if $\triangle Rec_t$ has the same sign as the last preceding earnings surprise, SUE_q, and 0 otherwise. $Contradict_t$ is 1 if $\triangle Rec_t$ and the last preceding surprise have opposite signs, and 0 otherwise. $Reiteration_t$ is 1 if $\triangle Rec_t$ is 0. AR_t is a firm's buy-and-hold return in excess of the value-weighted CRSP market index from day -1 to $+1$ around a recommendation revision on day t. SUE_q, $FSUE_{q+1}$, and AR_t are winsorised at the lower and upper 1% tails of their respective distributions.

have the same sign as the preceding earnings surprise, 38.2 percent are contradictory with the opposite sign, and 25.1 percent reiterate the old recommendation. Panel B presents the distribution of recommendation revisions conditional on the sign of the preceding earnings surprise. The distributions differ only slightly between positive and negative surprises. The Spearman correlation (not tabulated) between $\triangle Rec_t$ and SUE_q is very close to zero (0.018). Similar to findings in prior literature (e.g., Loh and Stulz, 2011), downgrades in general are more prevalent than upgrades and reiterations.

Hypothesis 2 requires estimates of the market's response to the release of recommendation revisions. I measure buy-and-hold excess returns as the difference between a firm's buy-and-hold return and the buy-and-hold return of the CRSP value-weighted market index from day -1 to +1 around the revision date (t):

$$AR_{firm,t} = \prod_{\tau=t-1}^{t+1}(1+R_{firm,\tau}) - \prod_{\tau=t-1}^{t+1}(1+R_{crsp,\tau}). \tag{2.17}$$

I winsorise abnormal returns at the lower and upper 1 percent tails of their distribution.[9]

2.4 Empirical Results

2.4.1 Hypothesis 1

Hypothesis 1 builds on the idea that recommendation revisions reveal differences in the interpretation of earnings surprises between analysts and investors. It predicts that—if analysts' earnings interpretation is superior to investors'—there will be an (unexpectedly) positive association between current earnings surprises and future earnings following confirmatory revisions and an (unexpectedly) negative association following contradictory revisions. I estimate the following pooled OLS regression as the analogue of equation (2.10) to evaluate these predictions:[10]

$$\begin{aligned} FSUE_{q+1} = \alpha + \beta_{\mathbf{1}}'\mathbf{X}_q + \beta_{\mathbf{2}}'\mathbf{X}_q SUE_q + \gamma_1 Confirm_q + \gamma_2 Contradict_q + \gamma_3 SUE_q \\ + \gamma_4 Confirm_q SUE_q + \gamma_5 Contradict_q SUE_q + \epsilon_{q+1}. \end{aligned} \tag{2.18}$$

Hypothesis 1 predicts that γ_4 should be positive and γ_5 should be negative. I estimate (2.18) with the sample's $85{,}011$ earnings announcements as the unit of observation. (2.18) includes earnings releases followed by reiterations. Therefore, they serve as the base group against which I judge differences in the association between current surprises and future earnings. Because many earnings announcements in q are followed by several recommendation revisions before $q+1$, I set $Confirm_q$ and $Contradict_q$ to 1 if at least one of the revisions between the announcements in period q and $q+1$ is either confirmatory or contradictory. In case there is both a confirmatory and a contradictory revision, both variables are 1, and if there are only

9 Section 2.6.2 presents results for alternative approaches to estimating the immediate response.

10 For notational convenience, I omit a firm index in all equations.

reiterations, both variables are 0.[11] I cluster standard errors by firm and year-quarter of the earnings announcement.

The vector of control variables ($\mathbf{X}_q$) includes dummies for each quarter of the sample period, dummies for each of the 48 Fama-French industries, and 10 size dummies, which I construct by ranking firms into quarterly market value deciles. (2.18) also includes SUE_q as well as the interaction between SUE_q and $\mathbf{X}_q$. Owing to the large number of controls in many of the following regressions, I restrict the tabulation of coefficients to those of the variables of interest. The inclusion of industry dummies as well as their interactions with SUE_q alone would, for example, add $47 * 2 = 94$ additional coefficients to the tables.[12]

Table 2.2 presents the results with the main variables printed in bold. Model 1 in the first column estimates (2.18) with the period q earnings surprise as the only explanatory variable. Its coefficient is 0.709 and highly significant (p-value < 0.01), suggesting that the average fraction of earnings surprises that is permanent is about 70.9 percent. Model 2 adds $Confirm_q$, $Contradict_q$, and the interactions of SUE_q with $Confirm_q$ and $Contradict_q$. As predicted by Hypothesis 1, the coefficient of the interaction between period q earnings surprises and confirmatory revisions is positive and highly significant ($\gamma_4 = 0.298$, p-value < 0.01), whereas the coefficient of the interaction with contradictory revisions is negative and highly significant ($\gamma_5 = -0.106$, p-value < 0.01). When I add the controls in Model 3, the coefficients become slightly smaller in absolute magnitude, though both still support Hypothesis 1 ($\gamma_4 = 0.265$, p-value < 0.01, and $\gamma_5 = -0.067$, p-value ≈ 0.06).

To put the coefficients into perspective and get an understanding of their economic magnitudes, it is helpful to consider γ_4 and γ_5 in relation to γ_3 in Model 2 as relative measures of differences in how current surprises relate to future earnings between announcements followed by different types of recommendation revisions.[13] When earnings surprises are followed by a reiteration, about 60 percent of the current surprise (γ_3) maps into future earnings. This fraction increases to about 90 percent ($\gamma_3 + \gamma_4$) when the earnings release is followed by a confirmation, but it drops to about 50 percent ($\gamma_3 + \gamma_5$) when it is followed by a contradiction.

Taken together, revised recommendations contain information that may help investors in reinterpreting past earnings surprises. Next, I turn to the analysis of in-

[11] Estimation of (2.18) without observations where both $Confirm_q$ and $Contradict_q$ equal 1 does not qualitatively affect the results.

[12] Tables including all coefficients are available upon request.

[13] I use the coefficients from Model 2 because Model 3 does not allow for this comparison due to the inclusion of interactions with SUE_q.

Table 2.2. Recommendation revisions, past earnings surprises, and future earnings

	$FSUE_{q+1}$		
	(1)	(2)	(3)
SUE_q	0.709***	0.599***	0.475***
	(0.032)	(0.042)	(0.140)
$Confirm_q$		-0.000***	-0.001***
		(0.000)	(0.000)
$Contradict_q$		-0.000**	-0.000***
		(0.000)	(0.000)
$\mathbf{Confirm_q SUE_q}$		0.298***	0.265***
		(0.040)	(0.035)
$\mathbf{Contradict_q SUE_q}$		-0.106***	-0.067*
		(0.039)	(0.036)
$Constant$	-0.001***	-0.001***	-0.002***
	(0.000)	(0.000)	(0.000)
Controls	No	No	Yes
R-squared	0.121	0.128	0.186
Observations	85,011	85,011	85,011

Table 2.2 reports results of pooled OLS regressions that estimate equation (2.18) to evaluate whether recommendation revisions contain information about the mapping of past earnings surprises into future earnings. The left-hand variable in all columns is $FSUE_{q+1}$, the difference between period $q+1$ earnings and the IBES consensus for earnings in $q+1$ in the 120 days preceding the announcement for period q, divided by the stock price 30 days prior to the release in q.
The units of observation are individual earnings announcements. SUE_q is the earnings surprise in quarter q. It is defined as the difference between actual earnings for q and the consensus forecast on the IBES Detail tape in the 120 days preceding the earnings release in q, divided by the stock price 30 days prior to the release. $Confirm_q$ is 1 if at least one of the recommendation revisions between quarter q and $q+1$ has the same sign as earnings surprise in period q, SUE_q, and 0 otherwise. $Contradict_q$ is 1 if at least one of the recommendation revisions between quarter q and $q+1$ has the opposite sign of the quarter q surprise, and 0 otherwise. Controls in (3) include dummies for 48 Fama-French industries, dummies for each quarter of the sample period, and dummies for a firm's size decile in q. All controls are interacted with SUE_q. To save space, I do not report the coefficients of the control variables. They are available upon request. Standard errors are clustered by firm and by year and quarter of the earnings announcement. ***, **, and * denote significance levels of 0.01, 0.05, and 0.1.

vestors' responses to recommendation revisions to assess whether this is information that investors extract from recommendations.

2.4.2 Hypothesis 2

Hypothesis 2 predicts that investors use recommendation revisions to update their interpretation of past earnings surprises. It predicts a positive association between earnings surprises and abnormal returns around confirmatory revisions and a negative association around contradictory revisions. To test these predictions, I estimate the following version of equation (2.12) using the 177,792 firm-days with recommendation revisions following one of the sample's earnings surprises:

$$\begin{aligned} AR_t = \alpha + \beta'\mathbf{X}_t + \gamma_1 Confirm_t + \gamma_2 Contradict_t + \gamma_3 SUE_q + \\ + \gamma_4 Confirm_t SUE_q + \gamma_5 Contradict_t SUE_q + \epsilon_t. \end{aligned} \tag{2.19}$$

(2.19) regresses abnormal returns surrounding a recommendation revision on day t on the last preceding earnings surprise interacted with dummies to indicate whether the revision on day t is confirmatory or contradictory. While individual earnings surprises can enter the estimation through different firm-day observations, each firm-day with at least one recommendation revision is included once at most. As some of the sample's recommendation revisions are made concurrently with other recommendations for the same firm, I pool all recommendations released for a firm on the same day into one observation. For these observations $Confirm_t$ and $Contradict_t$ are 1 in case at least one analyst confirmed or contradicted the preceding earnings surprise on this day, and 0 otherwise.

The focus in (2.19) is on γ_4 and γ_5. Hypothesis 2 suggests that as investors increase their initial estimate of the fraction of an earnings surprise that is permanent around confirmatory revisions, returns move into the direction of the earnings surprise ($\gamma_4 > 0$), and they decrease this estimate around contradictory revisions, resulting in a negative association between past surprise and abnormal returns ($\gamma_5 < 0$).

Recommendation revisions are likely to bring more to the market than only new information about past earnings surprises. I thus add a set of controls ($\mathbf{X}_t$) to (2.19). The controls include the change in recommendation ranks ($\triangle Rec_t$) since Womack (1996) shows that stock prices respond strongly to changes in recommendations. I calculate $\triangle Rec_t$ as the average recommendation change across all analysts that release a revision for the same firm on the same day t. Prior literature finds that recommendation revisions tend to have larger effects in worse information environments (e.g., Stickel, 1985). $\mathbf{X}_t$ therefore includes dummies for quarterly firm size deciles and their interaction with $\triangle Rec_t$. It is possible that the response to recommendation revisions depends both on the level of the recommendation as well as the recommendation change (e.g., Jegadeesh et al., 2004; Barber, Lehavy, and Trueman, 2010; Stickel, 1995). I add indicator variables for all 5 IBES recommendation ranks and their interaction with $\triangle Rec_t$ as additional controls.[14] Finally, I include dummy variables for each quarter of the sample period and the 48 Fama-French industries as well as their interactions with $\triangle Rec_t$ to allow the impact of revisions to vary in time and across industries.

Table 2.3. Returns around recommendation revisions and past earnings surprises

	AR_t
	(1)
SUE_q	0.078
	(0.079)
$Confirm_t$	-0.003***
	(0.000)
$Contradict_t$	-0.002***
	(0.001)
$\mathbf{Confirm_t SUE_q}$	0.805***
	(0.115)
$\mathbf{Contradict_t SUE_q}$	-0.538***
	(0.121)
$Constant$	0.007
	(0.009)
Controls	Yes
R-squared	0.117
Observations	177,792

Table 2.3 reports results of pooled OLS regressions that estimate equation (2.19) to assess whether investors reinterpret past earnings surprises using recommendation revisions. The left-hand variable is AR_t, a firm's buy-and-hold return in excess of the value-weighted CRSP market index from day -1 to $+1$ around the date of a recommendation revision (t).
The units of observation are firm-days with recommendation revisions. SUE_q is the last earnings surprise preceding the recommendation revision. $Confirm_t$ is 1 if at least one of the recommendation revisions on day t has the same sign as the last preceding earnings surprise, SUE_q, and 0 otherwise. $Contradict_t$ is 1 if at least one of the recommendation revisions on day t and the last preceding surprise have opposite signs, and 0 otherwise. Controls include the average change in IBES recommendation ranks associated with the recommendation revisions on t, size decile dummies, indicator variables for all 5 IBES recommendation ranks (rounded to the nearest integer in case I average over several recommendations on the same day), dummies for each quarter of the sample period, and dummies for 48 Fama-French industries. All controls are interacted with the change in recommendation ranks. For brevity, I again omit the coefficients of the control variables. They are available upon request. Standard errors are clustered by firm and by calendar date of the recommendation revision. ***, **, and * denote significance levels of 0.01, 0.05, and 0.1.

[14] Since recommendation ranks are averaged across all recommendations on a given day, I round them to the nearest integer before generating rank-dummies.

Table 2.3 presents the results. I again restrict the presentation to the central coefficients.[15] In line with the predictions of Hypothesis 2, abnormal returns are positively associated with past surprises around confirmatory revisions ($\gamma_4 = 0.805$) and negatively associated around contradictory revisions ($\gamma_5 = -0.538$). Both coefficients are highly significant (p-values < 0.01).

The magnitudes of the coefficients imply an economically relevant adjustment in investors' interpretation of past earnings. To see this, consider the predicted effects of confirmatory and contradictory revisions on abnormal returns at the 1 percent and 99 percent quantiles of SUE_q (−0.019 and 0.015). For stocks with exceptionally positive surprises, the coefficient estimates predict an abnormal return of 1.21 percent around confirmatory revisions and an abnormal return of −0.81 percent around contradictory revisions. In contrast, for stocks with exceptionally negative earnings surprises, the coefficients imply an abnormal return of −1.53 percent around confirmatory revisions and an abnormal return of 1.02 percent around contradictory revisions.

The results suggest that investors extract information about past earnings surprises from recommendation revisions and use this information to alter their own interpretation as predicted by Hypothesis 2. This reaffirms the findings concerning Hypothesis 1 and indicates that part of what revisions reveal about past earnings surprises reflects genuinely new information.

[15] Tables including all coefficients are avaialable upon request.

2.5 Analyst Heterogeneity

2.5.1 Empirical Predictions and Variable Construction

My central analyses show that analysts' recommendation revisions are *on average* informative about the permanence of past earnings surprises (Hypothesis 1), and that investors *on average* respond to revisions in a way that suggests they recognise this association (Hypothesis 2). I now examine whether there is heterogeneity between analysts in interpreting earnings surprises, and whether investors' responses to revisions indicate that they are aware of this heterogeneity.

Prior work (e.g., Mikhail, Walther, and Willis, 1997; Clement, 1999) finds that more experienced analysts and analysts from larger brokerage firms provide more accurate forecasts of future earnings. I contend that one reason for differences in analysts' ability to forecast earnings is that they vary in their ability to extrapolate future earnings from current earnings surprises.

To measure differences in analyst ability to interpret earnings, I compare the strength of the association between recommendation revisions, earnings surprises, and future earnings across analysts. If more experienced analysts and analysts from larger brokerage firms are more able to extrapolate from current earnings surprises, then their recommendation revisions should more accurately signal differences in future earnings. This suggests Hypothesis 3:

H3 The positive (negative) association between future earnings and past earnings surprises around confirmatory (contradictory) revisions is more positive (negative) as analyst experience or the size of the employing brokerage firm increases.

If investors recognise that analysts vary in their ability to interpret earnings surprises as suggested in Hypothesis 3, then they should respond more strongly to revisions by analysts whose recommendations indicate larger differences in the association between future earnings and past surprises. This provides the basis for Hypothesis 4:

H4 The positive (negative) association between past earnings surprises and abnormal returns around confirmatory (contradictory) revisions is more positive (negative) as analyst experience or the size of the employing brokerage firm increases.

Both hypotheses can be motivated in a slightly different way: Assume analysts signal their perception of differences in the components of earnings surprises via recommendation revisions, but their ability to do so is imperfect. In consequence, in some situations they will falsely suggest that earnings surprises contain an exceptionally large or small permanent component. The rationale behind Hypotheses 3 is that the probability of this happening decreases as analysts' experience or the size of the employing brokerage firm increases. As a result, the revisions of more experienced analysts and analysts from larger brokerage firms are more strongly associated with actual differences in how current surprises relate to future earnings. Hypothesis 4 suggests that investors recognise which analysts are more likely to correctly interpret earnings.

To construct operational measures of analyst experience and brokerage size, I use variants of the approaches in Mikhail, Walther, and Willis (1997) and Clement (1999). I then follow Clement and Tse (2005) and scale the resulting variables to allow for a simple interpretation of their coefficients.

I construct analyst experience by first counting the number of past quarters on the IBES Detail tape with at least one forecast by an analyst a prior to every quarter q ($PastQuarters_{a,q}$). Then, I calculate the maximum and minimum values of $PastQuarters_{a,q}$ across all analysts within a quarter, $max[PastQuarters_q]$ and $min[PastQuarters_q]$, and define $Exp_{a,q}$ as

$$Exp_{a,q} = \frac{PastQuarters_{a,q} - min[PastQuarters_q]}{max[PastQuarters_q] - min[PastQuarters_q]} - 0.5. \qquad (2.20)$$

$Exp_{a,q}$ ranks all analysts within a quarter by the number of past quarters they have actively released forecasts in. While the most experienced analysts within a quarter are assigned a value of 0.5, the least experienced analysts are assigned −0.5.

The variable used to describe the size of the employing brokerage firm, $BSize_{a,q}$, is constructed in a similar way. First, I use all recommendations in the sample to count the number of distinct analysts releasing recommendations for each brokerage firm b within a quarter q ($Analysts_{b,q}$). Next, I calculate the maximum and minimum values of $Analysts_{b,q}$ across all brokers within a quarter, $max[Analysts_q]$ and $min[Analysts_q]$. For each analyst a working for a brokerage firm b in a quarter q, I then define

$$BSize_{a,q} = \frac{Analysts_{b,q} - min[Analysts_q]}{max[Analysts_q] - min[Analysts_q]} - 0.5. \qquad (2.21)$$

$BSize_{a,q}$ measures the size of the brokerage firm an analyst a is working with. Analysts from the largest brokerage firms are assigned a value of 0.5, whereas analysts from the smallest brokerage firms are assigned −0.5.

2.5.2 Regressions and Results

To test Hypothesis 3, I estimate an augmented version of the specification used to test Hypothesis 1. The specification allows the effects of the interactions between confirmatory and contradictory revisions and the earnings surprise to vary in analysts' experience and brokerage size. I construct two variants of both variables that each distinguish between the ability of the analysts releasing confirmatory revisions and the analysts releasing contradictory revisions. Exp_q^+ is defined as the maximum of $Exp_{a,q}$ of all analysts releasing confirmatory revisions (for a particular firm) following the announcement in q, and Exp_q^- is the maximum of $Exp_{a,q}$ of the analysts releasing contradictory revisions. Analogously, I define $BSize_q^+$ and $BSize_q^-$ as the maxima of $BSize_{a,q}$ for all analysts releasing confirmatory and contradictory revisions, respectively.

Using these definitions, I then estimate the following regression:[16]

$$\begin{aligned}FSUE_{q+1} = \alpha &+ \beta_1' \mathbf{X_q} + \beta_2' \mathbf{X_q} SUE_q + \\ &+ \gamma_1 SUE_q + \gamma_2 Confirm_q + \gamma_3 Contradict_q + \gamma_4 Exp_q + \gamma_5 BSize_q \\ &+ \gamma_6 Confirm_q SUE_q + \gamma_7 Contradict_q SUE_q \\ &+ \gamma_8 Exp_q SUE_q + \gamma_9 BSize_q SUE_q \\ &+ \gamma_{10} Confirm_q Exp_q^+ + \gamma_{11} Contradict_q Exp_q^- \\ &+ \gamma_{12} Confirm_q BSize_q^+ + \gamma_{13} Contradict_q BSize_q^- \\ &+ \gamma_{14} Confirm_q Exp_q^+ SUE_q + \gamma_{15} Contradict_q Exp_q^- SUE_q \\ &+ \gamma_{16} Confirm_q BSize_q^+ SUE_q + \gamma_{17} Contradict_q BSize_q^- SUE_q + \epsilon_{q+1}.\end{aligned} \tag{2.22}$$

The regression again uses earnings announcements as the unit of observation. The variables of interest in (2.22) are $\gamma_{14}, \gamma_{15}, \gamma_{16}$, and γ_{17}. Because Exp and $BSize$ are standardised, their coefficients indicate the difference in the baseline effects (γ_6 and γ_7) as one moves from the lowest to the highest levels of the variables, i.e., from -0.5 to $+0.5$. Hypothesis 3 predicts that revisions by more experienced analysts and analysts from larger brokerage firms are more informative than those by less experienced analysts or analysts from smaller brokerage firms. Thus, it predicts that γ_{14} and γ_{16} are positive, whereas γ_{15} and γ_{17} are negative. The regression includes the same set of controls I used when testing Hypothesis 1.

[16] (2.22) also includes Exp_q and $BSize_q$, the maximal experience / broker size of all analysts revising recommendations between q and $q+1$ and their interactions with SUE_q.

Table 2.4. Analyst heterogeneity and the informational content of recommendation revisions

	$FSUE_{q+1}$		
	(1)	(2)	(3)
SUE_q	0.609***	0.577***	0.580***
	(0.129)	(0.131)	(0.131)
$Confirm_q$	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
$Contradict_q$	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)
$\mathbf{Confirm_q SUE_q}$	0.271***	0.364***	0.365***
	(0.039)	(0.043)	(0.044)
$\mathbf{Contradict_q SUE_q}$	-0.082**	-0.069	-0.077*
	(0.038)	(0.045)	(0.046)
Exp_q		-0.000	0.000
		(0.000)	(0.000)
$Exp_q SUE_q$		-0.156	-0.186
		(0.113)	(0.117)
$Confirm_q Exp_q^+$		-0.001***	-0.001***
		(0.000)	(0.000)
$Contradict_q Exp_q^-$		-0.001***	-0.001***
		(0.000)	(0.000)
$\mathbf{Confirm_q Exp_q^+ SUE_q}$		0.488***	0.496***
		(0.119)	(0.121)
$\mathbf{Contradict_q Exp_q^- SUE_q}$		-0.089	-0.018
		(0.125)	(0.124)
$BSize_q$	-0.000**		-0.000***
	(0.000)		(0.000)
$BSize_q SUE_q$	0.050		0.082
	(0.076)		(0.079)
$Confirm_q BSize_q^+$	-0.000**		-0.000
	(0.000)		(0.000)
$Contradict_q BSize_q^-$	-0.000		0.000
	(0.000)		(0.000)
$\mathbf{Confirm_q BSize_q^+ SUE_q}$	0.084		0.011
	(0.081)		(0.084)
$\mathbf{Contradict_q BSize_q^- SUE_q}$	-0.281***		-0.288***
	(0.073)		(0.071)
$Constant$	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)
Controls	Yes	Yes	Yes
R-squared	0.187	0.187	0.188
Observations	85,011	85,011	85,011

Continued

Table 2.4. Continued

Table 2.4 reports results of pooled OLS regressions that estimate equation (2.22) to evaluate whether the information that recommendation revisions contain about the mapping of past earnings surprises into future earnings varies between analysts. The left-hand variable in all columns is $FSUE_{q+1}$, the difference between period $q+1$ earnings and the IBES consensus for earnings in $q+1$ in the 120 days preceding the announcement for period q, divided by the stock price 30 days prior to the release in q.
Exp_q^+ and Exp_q^- are the maximum values of experience for analysts releasing confirmatory (Exp_q^+) and contradictory (Exp_q^-) revisions between the announcements of quarter q and $q+1$. Both variables are scaled to range between −0.5 and 0.5. Larger values correspond to more experienced analysts. An individual analyst's experience is defined as the number of quarters with forecasts by the same analyst prior to quarter q. $BSize_q^+$ and $BSize_q^-$ are the maximum values of brokerage size for analysts releasing confirmatory ($BSize_q^+$) and contradictory ($BSize_q^+$) revisions between the announcements of quarter q and $q+1$. Both variables are scaled to range between −0.5 and 0.5. Larger values correspond to larger brokerages. Brokerage size is measured as the number of different analysts releasing recommendations for the same broker in quarter q. The remaining variables and controls are defined as in Table 2.2. For brevity, I omit the coefficients of the controls. They are available upon request. Standard errors are clustered by firm and by year and quarter of the earnings announcement. ***, **, and * denote significance levels of 0.01, 0.05, and 0.1.

Table 2.4 presents the results including only analyst experience (Model 1), only brokerage size (Model 2), and both variables jointly (Model 3). Because the qualitative results are very similar across specifications, I discuss only Model 3. 2 of 4 coefficients are highly significant (p-values < 0.01) and have the predicted sign. The other two coefficients have the correct sign, albeit they are insignificant. The standardisation of the variables allows for an intuitive interpretation of the coefficients' magnitudes. The increase in the fraction of earnings surprises that maps into future earnings associated with confirmatory revisions varies from 12 percent ($\gamma_6 - 0.5\gamma_{14} = 0.12$) for novice analysts ($Exp_q^+ = -0.5$) to 61 percent ($\gamma_6 + 0.5\gamma_{14} = 0.61$) for very experienced analysts ($Exp_q^+ = 0.5$). The decrease in the fraction of earnings surprises that maps into future earnings for contradictory revisions by analysts from large brokerage firms ($BSize_q^- = 0.5$) is about 22 percent ($\gamma_7 + 0.5\gamma_{17} = -0.22$). The effect switches signs for analysts from very small brokerage firms ($BSize_q^- = -0.5$), whose contradictory revisions are associated with an increase of 7 percent ($\gamma_7 - 0.5\gamma_{17} = 0.07$) in the fraction of earnings surprises that maps into future earnings.[17]

In combination, the results thus suggests that there is pronounced heterogeneity in the information that revisions contain about the implications of past earnings surprises. Experienced analysts and analysts from large brokerage firms seem to exhibit substantially more skill at extrapolating future earnings from current surprises

[17] Since γ_7 is estimated with limited precision, the abilitiy to make quantitative statements about variation in the effect of contradictory revisions in general is limited.

than novice analysts or analysts from small brokerage firms.

Do investors recognise these differences as suggested in Hypothesis 4? To evaluate this, I use an augmented version of the specification used to test Hypothesis 2 that allows the effect of confirmatory and contradictory revisions on abnormal returns to vary in analyst experience and brokerage size. For this regression, I define $BSize_t^+$ and $BSize_t^-$ as the maximal values of $BSize_{a,q}$ over all analysts a releasing confirmatory and contradictory revisions (for a firm) on day t. In the same way, I construct Exp_t^+ and Exp_t^-. I then estimate the following specification:[18]

$$\begin{aligned}
AR_t = {} & \alpha + \beta'\mathbf{X}_t + \gamma_1 SUE_q + \gamma_2 Confirm_t + \gamma_3 Contradict_t + \gamma_4 Exp_t + \gamma_5 BSize_t \\
& + \gamma_6 Confirm_t SUE_q + \gamma_7 Contradict_t SUE_q + \gamma_8 Exp_t SUE_q + \gamma_9 BSize_t SUE_q \\
& + \gamma_{10} Confirm_t Exp_t^+ + \gamma_{11} Contradict_t Exp_t^- \\
& + \gamma_{12} Confirm_t BSize_t^+ + \gamma_{13} Contradict_t BSize_t^- \\
& + \gamma_{14} Confirm_t Exp_t^+ SUE_q + \gamma_{15} Contradict_t Exp_t^- SUE_q \\
& + \gamma_{16} Confirm_t BSize_t^+ SUE_q + \gamma_{17} Contradict_t BSize_t^- SUE_q + \epsilon_t. \qquad (2.23)
\end{aligned}$$

The focus in (2.23) is on coefficients γ_{14} to γ_{17}. Hypothesis 4 predicts that if investors are aware of the differences in analysts' abilities to extrapolate from past earnings surprises, then γ_{14} and γ_{16} will be positive and γ_{15} and γ_{17} will be negative.

Table 2.5 presents the results for Exp and $BSize$ separately (Model 1 and Model 2) and jointly (Model 3). All models include the same set of control variables employed in the testing of Hypothesis 2. Because the results on individual coefficients are largely insensitive to whether the variables are included separately or jointly, I discuss the results from Model 3. All coefficients have the sign predicted in Hypothesis 4 and all of them are significant at conventional levels. The positive association between abnormal returns and past surprises around confirmatory revisions is more positive and the negative association around contradictory revisions is more negative for both more experienced analysts and analysts from larger brokerage firms.

In summary, the results in this section suggest that analysts differ in their abilities to interpret earnings and investors incorporate these differences in their response to recommendation revisions.

[18] Similar to (2.22), (2.23) also includes Exp_t and $BSize_t$, the maximal experience / broker size of all analysts releasing recommendations on t and their interactions with SUE_q.

Table 2.5. Returns around recommendation revisions and analyst heterogeneity

	AR_t		
	(1)	(2)	(3)
SUE_q	0.069	0.066	0.060
	(0.080)	(0.108)	(0.110)
$Confirm_t$	-0.002***	-0.002***	-0.002***
	(0.001)	(0.001)	(0.001)
$Contradict_t$	-0.002***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)
$\mathbf{Confirm_t SUE_q}$	1.003***	1.115***	1.240***
	(0.122)	(0.161)	(0.165)
$\mathbf{Contradict_t SUE_q}$	-0.772***	-0.812***	-0.977***
	(0.125)	(0.158)	(0.163)
Exp_t		-0.000	-0.000
		(0.001)	(0.001)
$Exp_t SUE_q$		-0.021	-0.045
		(0.342)	(0.345)
$Confirm_q Exp_t^+$		-0.001	-0.002
		(0.002)	(0.002)
$Contradict_q Exp_t^-$		-0.006***	-0.005***
		(0.002)	(0.002)
$\mathbf{Confirm_t Exp_t^+ SUE_q}$		1.397***	1.156**
		(0.487)	(0.495)
$\mathbf{Contradict_t Exp_t^- SUE_q}$		-1.218**	-0.961*
		(0.506)	(0.505)
$BSize_t$	-0.001		-0.001
	(0.001)		(0.001)
$BSize_t SUE_q$	0.175		0.170
	(0.233)		(0.234)
$Confirm_q BSize_t^+$	0.004***		0.004***
	(0.001)		(0.001)
$Contradict_q BSize_t^-$	-0.003**		-0.002*
	(0.001)		(0.001)
$\mathbf{Confirm_t BSize_t^+ SUE_q}$	1.764***		1.670***
	(0.383)		(0.386)
$\mathbf{Contradict_t BSize_t^- SUE_q}$	-2.426***		-2.358***
	(0.345)		(0.344)
$Constant$	0.005**	0.005**	0.005**
	(0.002)	(0.002)	(0.002)
Controls	Yes	Yes	Yes
R-squared	0.119	0.118	0.119
Observations	177,792	177,792	177,792

Continued

Table 2.5. Continued

Table 2.5 reports results of pooled OLS regressions that estimate equation (2.23) to evaluate whether investors adjust their interpretation of past earnings surprises more strongly when responding to recommendation revisions of more experienced analysts or analysts from larger brokerages.
AR_t is a firm's buy-and-hold return in excess of the value-weighted CRSP market index from day -1 to $+1$ around the date of a recommendation revision (t). Exp_t^+ and Exp_t^- are the maximum values of experience for analysts releasing confirmatory (Exp_t^+) and contradictory (Exp_t^+) revisions on day t. Both variables are scaled to range between -0.5 and 0.5 with larger values corresponding to more experienced analysts. An individual analyst's experience is defined as the number of quarters with forecasts by the same analyst prior to quarter q. $BSize_t^+$ and $BSize_t^-$ are the maximum values of brokerage size for analysts releasing confirmatory ($BSize_t^+$) and contradictory ($BSize_t^-$) revisions released on day t. Both variables are scaled to range between -0.5 and 0.5 with larger values corresponding to larger brokerages. Brokerage size is measured as the number of different analysts releasing recommendations for the same broker in quarter q. The remaining variables are defined as in Table 2.3. For brevity, I omit the coefficients of the controls. They are available upon request. Standard errors are clustered by firm and by calendar date of the recommendation revision. ***, **, and * denote significance levels of 0.01, 0.05, and 0.1.

2.6 Robustness

2.6.1 Firm-Specific Heterogeneity in Earnings Persistence

My main analyses account for variation in the relation between earnings surprises and future earnings over time, between industries, and for firms of different sizes. However, they do not account for firm-specific variation. If analysts release confirmatory revisions in firms where surprises have on average larger permanent components and release contradictory revisions for firms where surprises have on average smaller permanent components, this effect might be captured in the interactions between $Confirm$ and $Contradict$ and the earnings surprise. To control for this possibility, I rerun the regressions in Table 2.4, but I replace $FSUE_{q+1}$ by $FSUE^e_{q+1}$, the residual of firm-level regressions of $FSUE_{q+1}$ on SUE_q. In other words, $FSUE^e_{q+1}$ is the share of $FSUE_{q+1}$ that cannot be explained by firm-specific variation in the association between surprises and future earnings.[19] Table 2.6 presents results in line with the previous findings. Thus, recommendation revisions contain information about the relation between current earnings surprises and future earnings even after accounting for firm-specific variation in this relation.

2.6.2 Alternative Models of Abnormal Return

To evaluate investors' responses to recommendation revisions, the main analyses use market-adjusted buy-and-hold returns surrounding the revision. Though the estimation of abnormal short-term returns is typically insensitive to the choice of normal return model, I re-estimate (2.23) using two alternative models of normal return. Table 2.7 presents results when I either use the return of the size-matched decile of CRSP stocks (Model 1) or the characteristic-based matching portfolio of Daniel et al. (1997) (Model 2) to model normal returns.[20] Again, the results confirm the previous findings.

The results are also insensitive to measuring the market's response to recommendation revisions using cumulative abnormal returns, $CAR_{firm,t} = \sum_{\tau=t-1}^{t+1}(R_{firm,\tau} - R_{crsp,\tau})$, instead of buy-and-hold abnormal returns.

[19] I require at least three observations per firm to run the firm-level regressions.

[20] The DGTW benchmarks are available via `http://www.rhsmith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm`.

Table 2.6. Heterogeneity in firms' earnings processes

	$FSUE^{e}_{q+1}$ (1)
SUE_q	-0.058
	(0.058)
$Confirm_q$	-0.001***
	(0.000)
$Contradict_q$	-0.000***
	(0.000)
$\mathbf{Confirm_q SUE_q}$	0.201***
	(0.032)
$\mathbf{Contradict_q SUE_q}$	-0.103***
	(0.039)
Exp_q	-0.000
	(0.000)
$Exp_q SUE_q$	-0.038
	(0.102)
$Confirm_q Exp^{+}_{q}$	-0.001***
	(0.000)
$Contradict_q Exp^{-}_{q}$	-0.001**
	(0.000)
$\mathbf{Confirm_q Exp^{+}_{q} SUE_q}$	0.358***
	(0.096)
$\mathbf{Contradict_q Exp^{-}_{q} SUE_q}$	-0.127
	(0.110)
$BSize_q$	-0.000
	(0.000)
$BSize_q SUE_q$	0.005
	(0.074)
$Confirm_q BSize^{+}_{q}$	-0.000
	(0.000)
$Contradict_q BSize^{-}_{q}$	-0.000
	(0.000)
$\mathbf{Confirm_q BSize^{+}_{q} SUE_q}$	0.052
	(0.085)
$\mathbf{Contradict_q BSize^{-}_{q} SUE_q}$	-0.174***
	(0.059)
$Constant$	-0.000
	(0.000)
Controls	Yes
R-squared	0.022
Observations	84,350

Table 2.6 revisits the analysis from Table 2.4. First, I estimate $FSUE_{q+1} = \alpha + \beta SUE_q + \epsilon_{q+1}$ for each firm and over the entire sample period. Then, I define $FSUE^{e}_{q+1} = FSUE_{q+1} - \hat{\alpha} - \hat{\beta} SUE_q$ and re-estimate equation (2.22).
The regression includes only firms with at least three observations of SUE_q. Otherwise, the specification is the same as the one in the third column of Table 2.4. Standard errors are clustered by firm and by year and quarter of the earnings announcement. I again omit coefficients of controls from the presentation due to their large number. They are available upon request. ***, **, and * denote significance levels of 0.01, 0.05, and 0.1.

Table 2.7. Alternative models of normal return

	AR_t^{size} (1)	AR_t^{dgtw} (2)
SUE_q	0.075	0.100
	(0.109)	(0.117)
$Confirm_t$	-0.002***	-0.002***
	(0.001)	(0.001)
$Contradict_t$	-0.002***	-0.002***
	(0.001)	(0.001)
$\mathbf{Confirm_t SUE_q}$	1.166***	1.058***
	(0.162)	(0.169)
$\mathbf{Contradict_t SUE_q}$	-0.981***	-0.890***
	(0.162)	(0.169)
Exp_t	0.001	0.000
	(0.001)	(0.001)
$Exp_t SUE_q$	0.037	0.021
	(0.338)	(0.389)
$Confirm_q Exp_t^+$	-0.002	-0.002
	(0.002)	(0.002)
$Contradict_q Exp_t^-$	-0.004**	-0.004**
	(0.002)	(0.002)
$\mathbf{Confirm_t Exp_t^+ SUE_q}$	0.953**	0.959*
	(0.486)	(0.551)
$\mathbf{Contradict_t Exp_t^- SUE_q}$	-0.974*	-0.996*
	(0.498)	(0.557)
$BSize_t$	-0.000	0.001
	(0.001)	(0.001)
$BSize_t SUE_q$	0.217	0.138
	(0.225)	(0.252)
$Confirm_q BSize_t^+$	0.003***	0.002*
	(0.001)	(0.001)
$Contradict_q BSize_t^-$	-0.002	-0.001
	(0.001)	(0.001)
$\mathbf{Confirm_t BSize_t^+ SUE_q}$	1.423***	1.382***
	(0.371)	(0.387)
$\mathbf{Contradict_t Bsize_t^- SUE_q}$	-2.246***	-2.170***
	(0.340)	(0.384)
$Constant$	0.002	0.004*
	(0.003)	(0.003)
Controls	Yes	Yes
R-squared	0.133	0.132
Observations	177,792	146,575

Table 2.7 reports results for regressions that vary the model of normal return used to calculate AR_t. The model in the first column defines abnormal returns as the firms' returns in excess of the return of CRSP-firms in the same size decile. The model in the second column uses the approach of Daniel et al. (1997) to define abnormal return as the firms' return in excess of the return of a portfolio matched on size, momentum, and book-to-market. Owing to a lack of portfolio assignment for some firms, the second column includes less observations than the first. All remaining variables are defined as in Table 2.5. Standard errors are clustered by firm and by calendar date of the recommendation revision. I omit coefficients of controls from the presentation due to their large number. They are available upon request. ***, **, and * denote significance levels of 0.01, 0.05, and 0.1.

2.6.3 Alternative Model of Earnings Expectations

A central aspect of all preceding analyses is the approximation of investors' earnings expectations by current and next period consensus forecasts on IBES. To probe whether the results are robust to this choice, I repeat my central analyses using a time-series model to calculate earnings expectations and surprises. To capture predictable variation in firms' earnings processes, I take a model suggested in Foster (1977) that allows for trend growth in earnings, seasonal dependence, and business cycle effects.[21]

For each stock on the IBES Actuals History file and each quarter q of the sample period, I first run the following time-series regression over quarters $q-21$ to $q-1$:[22]

$$E_q = \phi_0 + \phi_1 E_{q-4} + \phi_2 (E_{q-1} - E_{q-5}) + \epsilon_q. \tag{2.24}$$

I use the estimated parameters from (2.24) to predict earnings for period q, denoted $\mathbb{E}^{seasonal}_{pre-q}[E_q]$, and $q+1$, denoted $\mathbb{E}^{seasonal}_{pre-q}[E_{q+1}]$.[23] Employing these two variables as proxies for period q and $q+1$ expectations, I then define FE_q and FFE_{q+1} as the analogues of SUE_q and $FSUE_{q+1}$:

$$FE_q = (E_q - \mathbb{E}^{seasonal}_{pre-q}[E_q]) / sd(\epsilon) \tag{2.25}$$

and

$$FFE_{q+1} = (E_{q+1} - \mathbb{E}^{seasonal}_{pre-q}[E_{q+1}]) / sd(\epsilon). \tag{2.26}$$

I deflate both variables by the standard deviation of the residuals over the respective estimation period of (2.24), $sd(\epsilon)$, to mitigate heteroskedasticity and make them comparable across firms. Using FE_q to redefine which revisions are confirmatory

---

21 Much recent research like Battalio and Mendenhall (2011) uses a seasonal random walk to model earnings expecations. I opt for a more complex model to explicitely capture variation in earnings an investor might have deduced from historical patterns.

22 I require at least 10 observations in each regression.

23 To predict earnings for $q+1$, I use the parameters estimated up to (but not including) period q and the forecast for q instead of the actual value for q: $\mathbb{E}^{seasonal}_{pre-q}[E_{q+1}] = \hat{\phi}_0 + \hat{\phi}_1 E_{q-3} + \hat{\phi}_2(\mathbb{E}^{seasonal}_{pre-q}[E_q] - E_{q-4})$.

and which are contradictory, I then re-estimate the main regressions:[24]

$$
\begin{aligned}
FFE_{q+1} = \alpha + \beta_{\mathbf{1}}'\mathbf{X_q} + \beta_{\mathbf{2}}'\mathbf{X_q}FE_q + \\
&+ \gamma_1 FE_q + \gamma_2 Confirm_q + \gamma_3 Contradict_q + \gamma_4 Exp_q + \gamma_5 BSize_q \\
&+ \gamma_6 Confirm_q FE_q + \gamma_7 Contradict_q FE_q + \gamma_8 Exp_q FE_q + \gamma_9 BSize_q FE_q \\
&+ \gamma_{10} Confirm_q Exp_q^+ + \gamma_{11} Contradict_q Exp_q^- \\
&+ \gamma_{12} Confirm_q BSize_q^+ + \gamma_{13} Contradict_q BSize_q^- \\
&+ \gamma_{14} Confirm_q Exp_q^+ FE_q + \gamma_{15} Contradict_q Exp_q^- FE_q \\
&+ \gamma_{16} Confirm_q BSize_q^+ FE_q + \gamma_{17} Contradict_q BSize_q^- FE_q + \epsilon_{q+1}
\end{aligned} \tag{2.27}
$$

and

$$
\begin{aligned}
AR_t = \alpha + \beta'\mathbf{X}_t + \gamma_1 FE_q + \gamma_2 Confirm_t + \gamma_3 Contradict_t + \gamma_4 Exp_t + \gamma_5 BSize_t \\
&+ \gamma_6 Confirm_t FE_q + \gamma_7 Contradict_t FE_q + \gamma_8 Exp_t FE_q + \gamma_9 BSize_t FE_q \\
&+ \gamma_{10} Confirm_t Exp_t^+ + \gamma_{11} Contradict_t Exp_t^- \\
&+ \gamma_{12} Confirm_t BSize_t^+ + \gamma_{13} Contradict_t BSize_t^- \\
&+ \gamma_{14} Confirm_t Exp_t^+ FE_q + \gamma_{15} Contradict_t Exp_t^- FE_q \\
&+ \gamma_{16} Confirm_t BSize_t^+ FE_q + \gamma_{17} Contradict_t BSize_t^- FE_q + \epsilon_t.
\end{aligned} \tag{2.28}
$$

As before, I winsorise AR_t and FE_q at the upper and lower 1 percent tails of their respective distributions. Table 2.8 presents the results of these regressions with specification (2.27) in the first column and (2.28) in the second column. All of the coefficients have the predicted sign and with the exception of two, γ_{15} and γ_{17} in (2.27), they are significant at conventional levels. This supports the main findings and suggests that they persist when I use a statistical model instead of analyst forecasts to approximate investors' earnings expectations.

[24] I restrict this robustness check to the sample of the main analyses. Owing to insufficient quarterly observations of E_q, the eventual sample size in this robustness check is smaller than in the original analyses.

Table 2.8. Alternative model of expected earnings

	FFE_{q+1} (1)	AR_t (2)
FE_q	0.924***	-0.000
	(0.152)	(0.000)
$Confirm_{q,t}$	-0.243***	-0.005***
	(0.023)	(0.001)
$Contradict_{q,t}$	-0.162***	-0.003***
	(0.023)	(0.001)
$\mathbf{Confirm_{q,t}FE_q}$	0.191***	0.003***
	(0.019)	(0.000)
$\mathbf{Contradict_{q,t}FE_q}$	-0.055***	-0.003***
	(0.019)	(0.000)
$Exp_{q,t}$	0.422***	0.008***
	(0.079)	(0.001)
$Exp_{q,t}FE_q$	-0.081	-0.001
	(0.054)	(0.001)
$Confirm_{q,t}Exp^+_{q,t}$	-0.528***	-0.012***
	(0.077)	(0.002)
$Contradict_{q,t}Exp^-_{q,t}$	-0.399***	-0.009***
	(0.065)	(0.002)
$\mathbf{Confirm_{q,t}Exp^+_{q,t}FE_q}$	0.293***	0.003***
	(0.052)	(0.001)
$\mathbf{Contradict_{q,t}Exp^-_{q,t}FE_q}$	-0.014	-0.003***
	(0.053)	(0.001)
$BSize_{q,t}$	0.087*	0.005***
	(0.050)	(0.001)
$BSize_{q,t}FE_q$	0.004	-0.001
	(0.041)	(0.000)
$Confirm_{q,t}BSize^+_{q,t}$	-0.147***	-0.004***
	(0.051)	(0.001)
$Contradict_{q,t}BSize^-_{q,t}$	-0.058	-0.005***
	(0.051)	(0.001)
$\mathbf{Confirm_{q,t}BSize^+_{q,t}FE_q}$	0.072*	0.007***
	(0.039)	(0.001)
$\mathbf{Contradict_{q,t}BSize^-_{q,t}FE_q}$	-0.052	-0.006***
	(0.045)	(0.001)
$Constant$	0.263***	0.009***
	(0.102)	(0.003)
Controls	Yes	Yes
R-squared	0.426	0.119
Observations	76,695	162,864

The analyses in Table 2.8 replace earnings expectations based on analyst forecasts by predictions from a time-series model as described in Section (2.6.3). The left-hand variable in the first column is the next period's forecast error of this model. The left-hand variable in the second column is the abnormal return around the recommendation revision. Controls in the first column are the same as in Table 2.4. Controls in the second column are the same as in Table 2.5. All remaining variables are defined as beforehand. Standard errors are clustered by firm and by year and quarter of the earnings announcement in the first column and by firm and by calendar date of the recommendation revision in the second column. I omit coefficients of controls from the presentation due to their large number. They are available upon request. ***, **, and * denote significance levels of 0.01, 0.05, and 0.1.

2.7 Conclusion

This paper uses post-earnings revisions of pre-earnings recommendations to contrast analysts' and investors' skill in the interpretation of earnings surprises. Building on the predictions of a simple model, I show that patterns in how revisions in analysts' recommendations relate to future earnings parallel patterns in how investors respond to their releases. In addition, I show that cross-sectional differences in the response to the release of recommendation revisions align with cross-sectional differences in the information they contain about future earnings, each varying in characteristics that prior work associates with analyst skill.

In combination, the results indicate that recommendation revisions help investors in understanding the implications of current earnings surprises for future earnings and that investors correctly recognise whose recommendations contain on average more accurate information. Contrasting investors' and analysts' opinion via recommendation revisions provides a means of assessing how analysts contribute to the market's information environment. This approach lends itself to different types of news like mergers or dividend announcements, where comparing analysts' and investors' abilities may provide a worthwhile field for future research.

3 Measurement Error in Subjective Expectations and the Empirical Content of Economic Models

3.1 Introduction

Stock market expectations are among the most important primitives of economic models of portfolio choice, but measurement error is pervasive in subjective beliefs data. For example, many empirical studies have discarded large fractions of data because answers do not obey the laws of probability (Manski, 2004; Hurd, Rooij, and Winter, 2011). The semantics of measurement error in subjective expectations data, however, is potentially quite different from contexts in which measurement error is usually studied, like past income, savings, or consumption. In the latter cases, there is a precisely defined "true" value and measurement error arises because of imperfect recall (Hoderlein and Winter, 2010) or incongruent variable definitions.

In case of subjective expectations, however, analysts may be chasing an elusive target: It is not evident that all people hold well-formed beliefs about a given phenomenon. For example, the prevalence of 50-50 responses in expectations surveys has been interpreted in exactly this way by Bruine de Bruin et al. (2000) or Bruine de Bruin and Carman (2012). In consequence, key structural parameters of economic models might not be present in the form envisioned by the econometrician (Stiglitz, 2002; Rust, 2014). If this is the case, standard techniques of using corrected estimates instead of the misreported values (Wansbeek and Meijer, 2000; Schennach, 2013) will not lead to improved estimates of choice models because the corrected estimates do not form the basis of decisions, either.

In this paper, we take a very different route and use the extent of measurement error to uncover heterogeneity in choice behaviour. Put differently, we argue that the magnitude of measurement error in stated beliefs should provide insights into the extent to which an economic model constitutes an adequate description of an individual's portfolio choice behaviour.

To explore this channel and motivate our empirical strategy, Section 3.2.1 presents a simple economic model of stock market participation that clarifies the roles of expectations, preferences, and transaction costs. In Section 3.2.2, we argue that for a variety of alternative decision modes —trusting others' advice or following rules of thumb, among others— individuals need not hold particularly meaningful beliefs about the future evolution of the stock market. Consequently, the data for individuals who entertain such choice rules will be characterised by two features. First, their stated beliefs will be prone to measurement error. Second, the sensitivity of their stockholdings to changes in model primitives will be low. In order to empirically evaluate this hypothesis, we estimate a Klein and Vella (2009) semiparametric double index model. In this model, the first index contains the primitives of our theoretical model (such as beliefs and preferences), while the second index includes quantitative and qualitative indicators of measurement error. Both indices may interact in a fully nonparametric fashion to obtain predicted probabilities of stock market participation.

Section 3.3 describes the dataset that we collected specifically for this study. Section 3.4 presents the results of our empirical application. We demonstrate that changes in primitives of the economic model induce large variation in stock market participation at low levels of the measurement error index. If measurement error is high, the effect of changes in beliefs and preferences on stockholdings is much smaller. We perform a number of variations on this theme and show that the results hold up in several different specifications. We then demonstrate the usefulness of our modelling approach for the analysis of less detailed data by estimating a specification with variables that are commonly available or inexpensive to collect. In particular, we show that restricting ourselves to a simple measure of expectations and purely qualitative measurement error proxies yields a similar overall pattern. Though, as one would expect, the differences along the measurement error distribution are less pronounced. We discuss our results and conclude in Section 3.5.

3.2 Motivation and Empirical Strategy

We develop our econometric strategy in three steps. First, we characterise a household's portfolio problem by means of a simple choice model. We then explain in detail why we conjecture that the degree to which this model serves as an adequate description of the decision-making process varies across households and why we expect that variation in measurement error can be exploited to capture this adequacy. In the third step, we present our econometric strategy that implements these ideas.

3.2.1 A Simple Economic Model of Stock Market Participation

In our description of a household's portfolio choice problem we follow Campbell and Viceira (2002). We assume that the household maximises a power utility function defined over next period's expected financial wealth $E_t\left[W_{t+1}\right]$ by allocating fractions of period t wealth to one safe and one risky asset. If the household can neither short the risky asset nor leverage the position in it, the optimal risky asset share θ^{opt} solves:

$$\theta^{\text{opt}} = \arg\max_{\theta} \left\{ \frac{E_t\left[W_{t+1}(\theta)^{1-\gamma}\right]}{1-\gamma} \right\} \qquad \text{s.t.} \quad 0 \leq \theta \leq 1$$

Risk aversion and a household's beliefs about the returns of the two assets determine the optimal decision. Denote a household's expected return for the safe asset by μ_{t+1}^{safe} and assume that the household's expectations for the risky asset's return can be described by a log-normal distribution with mean μ_{t+1}^{risky} and standard deviation $\sigma_{t+1}^{\text{risky}}$. When returns are log-normally distributed, so is W_{t+1}. For a log-normal variable it holds that $\log E[X] = E\left[\log X\right] + {}^{1}\!/_{2}\operatorname{Var}\left[\log X\right]$. Thus, the maximisation problem can be rewritten as:

$$\theta^{\text{opt}} = \arg\max_{\theta} \left\{ (1-\gamma)E_t\left[w_{t+1}(\theta)\right] + \frac{1}{2}(1-\gamma)^2 \operatorname{Var}_t\left[w_{t+1}(\theta)\right] \right\} \qquad \text{s.t.} \quad 0 \leq \theta \leq 1$$

where lower case letters are logarithms. Using a first-order Taylor series approximation, next period's log wealth can be written as:

$$w_{t+1}(\theta) = w_t + (1-\theta)\mu_{t+1}^{\text{safe}} + \theta\mu_{t+1}^{\text{risky}} + \frac{1}{2}\theta(1-\theta)\left(\sigma_{t+1}^{\text{risky}}\right)^2$$

Substituting this into the expression for θ^{opt} and dividing by $1-\gamma$, we obtain for the maximand:

$$w_t + \theta\left(\mu_{t+1}^{\text{risky}} - \mu_{t+1}^{\text{safe}}\right) + \frac{1}{2}\theta\left(1-\gamma\theta\right)\left(\sigma_{t+1}^{\text{risky}}\right)^2$$

Solving the first-order condition of this problem for the optimal share θ^{opt} yields:

$$\theta^{\text{opt}} = \frac{\mu_{t+1}^{\text{risky}} - \mu_{t+1}^{\text{safe}} + \frac{1}{2}\left(\sigma_{t+1}^{\text{risky}}\right)^2}{\gamma\left(\sigma_{t+1}^{\text{risky}}\right)^2} \tag{3.1}$$

At plausible parameter values of γ, the optimal risky asset share will be positive when historical return data are used to estimate $\mu^{\text{safe}}, \mu^{\text{risky}}$, and σ^{risky} as proxies of households' expectations. However, studies on stock ownership find that a large fraction of the population does not participate in the stock market (e.g., Haliassos and Bertaut, 1995). Arguably the most prominent explanation for why households abstain from participation is the existence of (broadly defined) transaction costs (Vissing-Jørgensen, 2002). These transaction costs are likely to vary with household characteristics. If participation comes with fixed monetary costs, for example, wealthy households will be more likely to invest in risky assets, since for them the fixed costs are spread over larger investments. If information costs play an important role, transaction costs will be lower for numerate respondents who are quicker to grasp the basic functioning of the stock market. We assume that the variables affecting transaction costs can be modeled by observable household characteristics X^{ta}; denote the resulting transaction costs by $f\left(X^{\text{ta}}\right)$.

We now combine the optimal risky asset share (3.1), transaction costs, and random influences ε in a simple random utility model of stock market participation:

$$Y \equiv \mathbf{I}\{\theta > 0\} = \begin{cases} 1 & \text{if} \quad \theta^{\text{opt}}\left(\mu_{t+1}^{\text{risky}} - \mu_{t+1}^{\text{safe}}, \sigma_{t+1}^{\text{risky}}, \gamma\right) - f\left(X^{\text{ta}}\right) > \varepsilon \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

According to (3.2), the probability of participating in the stock market will depend on the mean and variance of beliefs over the risky asset, the expected risk-free rate, risk aversion, variables proxying transaction costs, and the stochastic properties of ε. If the latter was normally distributed, one could estimate (3.2) by means of a standard Probit model. Estimators that make minimal distributional assumptions but enable the researcher to recover marginal effects still require ε to either be homoskedastic or have a very particular form of heteroskedasticity (Klein and Vella, 2009). If our conjecture about a varying explanatory content of $\theta^{\text{opt}} - f\left(X^{\text{ta}}\right)$ is correct, this will be reflected in a form of heteroskedasticity that violates these assumptions. In particular, the variance of ε will vary with measurement error in a form that is unknown a priori.

3.2.2 Putting Measurement Error in Subjective Beliefs to Productive Use

The model combines effortful reasoning about future states of the world with personal risk tolerance to form a choice rule. While such behavior is at the heart of economic thinking, there are a number of reasons why the explanatory content of this model is likely to vary in the population. For example, almost half of the Dutch population report that they mostly rely on the advice of family, friends, or professionals when it comes to important financial decisions (Gaudecker, forthcoming). Individuals may take decisions intuitively (Kahneman, 2011; Binswanger and Salm, 2013) or follow simple rules of thumb like holding an equity share of 100 minus age (see, e.g., the discussion in Ameriks and Zeldes, 2004).

However, if some households base their investment decisions on such alternatives, then they have less incentives to maintain a meaningful, up-to-date, and/or reasonably stable belief about the evolution of the stock market. This should have at least two sets of consequences. First, different methods to elicit beliefs should lead to divergent reports, self-expressed confidence in one's estimates should decrease, and one should find tasks related to belief elicitation rather difficult: Measurement error in subjective beliefs will be high. Second, the marginal effects of changes in beliefs on portfolio choice behavior should be much smaller than for individuals whose choice behavior is well approximated by the economic model. Thus, the magnitude of measurement error in beliefs will be informative about economic quantities of interest.

There is a vast literature on measurement issues in subjective expectations of stock market developments. Manski (2004) and Hurd (2009) provide excellent overviews. First-order evidence for measurement error is provided by the facts that many answers to probabilistic survey questions violate basic laws of probability and that non-response tends to be concentrated among sub-groups who do not follow the development of the stock market (Hurd, 2009). In addition, consistent with our interpretation of measurement error as reflecting the absence of meaningful expectations, several authors argue that the prevalence of 50-50 responses to probability questions (Manski, 2004; Hurd, 2009; Kleinjans and Soest, 2014) reflects epistemic uncertainty rather than a genuine belief (Bruine de Bruin et al., 2000). This interpretation finds support in explicit follow-up questions (Hurd, 2009; Bruine de Bruin and Carman, 2012; Binswanger and Salm, 2013).

Similar patterns of imprecise measurements have been documented for risk preferences. Gaudecker, Soest, and Wengström (2011) and Choi et al. (2014) show that for respondents with high socio-economic status, sequences of lottery decisions are much more consistent with flexible parametric utility functions and the generalised axiom of revealed preferences, respectively. Put differently, risk preference parameters are much more precisely measured for these subgroups.

In sum, different pieces of evidence suggest that subjective stock market beliefs are measured with error and that such error-laden responses provide information about the meaningfulness of the underlying belief distribution. To the extent that the absence of meaningful beliefs is associated with the use of alternative choice rules, information on measurement error can be exploited to evaluate the explanatory content of the simple model of stock market participation discussed above.

3.2.3 Econometric Specification

In econometric terms, a consequence of measurement error is that ε in (3.2) will be heteroskedastic, i.e., its variance will increase in the amount of measurement error. Depending on the precise decision-making process, it may also have group-specific means different from zero. For example, the most prevalent advice by family and friends seems to be non-participation in the stock market (Gaudecker, forthcoming). For the group of individuals who follow this advice, participation rates will be low even if $\theta^{\text{opt}} - f\left(X^{\text{ta}}\right)$ takes on positive values on average. In order to capture these consequences, we require an econometric specification where the predictions of the choice model (3.2) interact with the extent of measurement error in a flexible way. The double index binary choice model of Klein and Vella (2009) is ideally suited for the structure of our problem. The model obtains an estimate of the probability of stock market participation by nonparametrically combining two linear indices.

We first aggregate $\mu^{\text{risky}}_{t+1} - \mu^{\text{safe}}_{t+1}, \sigma^{\text{risky}}_{t+1}, \gamma$, and X^{ta} into one vector X^{mod}; $X^{\text{mod}}\beta^{\text{mod}}$ approximates our choice model from 3.2.1.[1] We will refer to $X^{\text{mod}}\beta^{\text{mod}}$ as the economic model index in what follows. In a second vector X^{me}, we group quantitative and qualitative indicators of measurement error as well as covariates that we would expect to influence the "propensity to use economic reasoning"; the latter may overlap with covariates included in the economic model index to proxy transaction costs. Accordingly, we refer to $X^{\text{me}}\beta^{\text{me}}$ as the measurement error index. The Klein and Vella (2009) estimator models the relationship of both indices and risky asset holdings as:[2]

$$P\big(Y = 1 \mid X^{\text{mod}}\beta^{\text{mod}},\, X^{\text{me}}\beta^{\text{me}}\big) = h\big(X^{\text{mod}}\beta^{\text{mod}},\, X^{\text{me}}\beta^{\text{me}}\big) \tag{3.3}$$

This structure is directly related to (3.2) in that the measurement index further parameterises ε (i.e., the random component is systematic to some extent). The function $h(\cdot,\cdot)$ provides a nonparametric link mapping the indices for the economic model and measurement error into stock market participation probabilities.

To attain identification (up to location and scale) of the parameters β^{mod} and β^{me}, we require that at least one continuous variable per index is excluded from the

---

1 We also experimented with calculating (3.1) and including it alongside X^{ta}. This led to numerical difficulties as the covariance matrix of the two indices was near-singular for a wide range of parameter values. We attribute this to the lack of a quantitatively meaningful measure of γ (Rabin, 2000) and to a fat right tail of $\left(\sigma^{\text{risky}}_{t+1}\right)^2$. The latter is likely responsible for the numerical problems; it is also the reason why we use the standard deviation of beliefs instead of the variance.

2 Klein and Vella (2009) frame their discussion in terms of an estimator for a single-equation binary response model with dummy endogenous variable when no instruments are present. A first application that applies it directly to two indices is given in Maurer (2009).

other index. We normalise the coefficients on one of these variables per index to one. The resulting model satisfies the form in A5 of Klein and Vella (2009) without requiring reparameterisation. Under assumptions given in Klein and Vella (2009)—mainly smoothness of $h(\cdot,\cdot)$ and compact support of the covariates—the probability to participate in the stock market can be expressed as a function of the densities conditional on participation:

$$P\big(Y=1 \,\big|\, X^{\text{mod}}\beta^{\text{mod}},\, X^{\text{me}}\beta^{\text{me}}\big) = \frac{f_{Y=1}\big(X^{\text{mod}}\beta^{\text{mod}},\, X^{\text{me}}\beta^{\text{me}}\big)\cdot P\big(Y=1\big)}{f\big(X^{\text{mod}}\beta^{\text{mod}},\, X^{\text{me}}\beta^{\text{me}}\big)}, \tag{3.4}$$

where $f(\cdot)$ denotes the unconditional density of the bivariate index and $f_{Y=1}(\cdot)$ its density conditional on participation in the stock market. Kernel density estimators for these quantities are obtained under a multi-stage local smoothing procedure to achieve a sufficiently low order of the bias. Denoting the resulting estimator for (3.4) as $\hat{P}_i\left(\beta^{\text{mod}},\beta^{\text{me}}\right)$, we can write the semiparametric maximum likelihood estimator for $\beta^{\text{mod}},\beta^{\text{me}}$ as:

$$\big(\hat{\beta}^{\text{mod}}_{\text{ml}},\, \hat{\beta}^{\text{me}}_{\text{ml}}\big) = \underset{\beta^{\text{mod}},\,\beta^{\text{me}}}{\arg\max} \sum_{i=1}^{N} \hat{\tau}_i \left[Y_i \cdot \log \hat{P}_i\left(\beta^{\text{mod}},\beta^{\text{me}}\right) + (1-Y_i)\cdot \log\left(1-\hat{P}_i\left(\beta^{\text{mod}},\beta^{\text{me}}\right)\right)\right], \tag{3.5}$$

where $\hat{\tau}_i$ denotes a smooth trimming function ensuring that densities do not become too small (Klein and Spady, 1993). Klein and Vella (2009) show that $\big(\hat{\beta}^{\text{mod}}_{\text{ml}},\, \hat{\beta}^{\text{me}}_{\text{ml}}\big)$ converges at rate $\sqrt{N}$ to its true value. While the parameter values do not allow for a direct interpretation, various quantities of interest like average partial effects can be computed with little effort.

In sum, our empirical model allows for a flexible interplay between traditional economic parameters and measurement error proxies in generating choice behaviour. In particular, it will allow an analysis of how marginal changes in model parameters translate into stock market participation, and how this relationship varies across respondents with differential degrees of measurement error.

3.3 Data and Descriptive Statistics

Our data stem from the Dutch LISS study (Longitudinal Internet Studies for the Social Sciences), which regularly administers Internet surveys and experiments to a panel of households comprising a probability sample drawn from the population register kept by Statistics Netherlands.

Implementing our empirical strategy requires data on individual stock market participation, subjective beliefs and risk aversion, proxies for the degree of measurement error in individual responses, and a rich set of sociodemographic covariates. Only the latter are present in the LISS panel by default. In order to obtain measures for the main quantities of interest, we implemented a series of incentivised experiments and survey questions in August and September of 2013. We restricted our experiments to households with financial wealth in excess of 1,000€ to focus on respondents with substantial incentives to think about portfolio allocations. To increase turnout, we also included individuals who refused to answer questions about their exact amount of wealth. Within households, we selected the financial decision maker. In total, 2,125 individuals completed both survey waves. After dropping observations with missing data, we are left with a final sample of 2,072 observations.

3.3.1 Outcome Variable: Stock Market Participation

LISS routinely collects detailed data on respondents' financial background, including information on asset ownership. To ensure the relevance of elicited beliefs for current portfolio allocations, we asked respondents to update their information on asset holdings in August 2013. For this purpose, we asked them whether they had any type of bank or savings account and/or investments (stocks, bonds, funds, or options). Our outcome variable is a binary index that equals 1 if the respective respondent held any investments, and 0 otherwise. A quarter of the households in our sample holds risky assets (cf. Table 3.1). This is in the range of values reported for the Netherlands from other datasets and earlier periods (Alessie, Hochgürtel, and Soest, 2004; Rooij, Lusardi, and Alessie, 2011).

3.3.2 Variables Entering the Economic Model Index

Subjective Expectations. In August 2013, we asked respondents to describe their expectations about the one-year return of the Amsterdam Exchange Index (AEX). We employed a variation of the ball allocation procedure developed by Delavande and Rohwedder (2008), which was explicitly designed for usage in Internet experiments. For each individual, the procedure yields an 8-binned histogram for the expectation of the AEX's one-year return. Using the resulting 7 points on the cumulative distribution function, we follow Hurd, Rooij, and Winter (2011) and fit a log-normal distribution to obtain individual-level measures for μ_{t+1}^{risky} and $\sigma_{t+1}^{\text{risky}}$. Because our theoretical framework requires expected excess returns, we also asked respondents for a point estimate for the return of a one-year investment into a standard savings account as the most prevalent safe asset. Section A.1.1.1 of the Appendix contains detailed descriptions of both procedures.

Recent research in the experimental economics literature has shown that financial incentives induce more truthful reporting of beliefs in tasks like ours (see, for example, Palfrey and Wang, 2009; Gächter and Renner, 2010; Wang, 2011). In order to incentivise subjects, we employed the binarised scoring rule of Hossain and Okui (2013) which is incentive-compatible for a wide range of utility functions. As is common practice with large samples like ours, we randomly selected one in ten subjects for actual payment. The maximum earnings per selected subject were 100 € and average earnings equaled 39.66 € conditional on being selected for payment in September 2014.

We relegate a detailed presentation of summary statistics of the belief measures to Section A.1.1.1 of the Appendix and only discuss some notable features at this point. First, our data exhibit the same patterns found previously in the literature, i.e., male, richer, and better educated respondents tend to hold more optimistic expectations (e.g., Manski, 2004; Hurd, 2009; Hurd, Rooij, and Winter, 2011). Second,

Table 3.1. Descriptive statistics

	Statistic		Index	
	Mean	Std. Dev.	Model	Meas. Err.
Holds risky assets	0.25			
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	-1.18	8.10	×	
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	6.25	4.01	×	
Risk aversion	0.00	1.00	×	
Absolute difference between belief measures	11.20	13.57		×
Lack of confidence in AEX return estimate	0.54	0.23		×
Lack of confidence in sav. acc. return estimate	0.36	0.24		×
Experimental tasks difficult	0.49	0.33		×
Experimental tasks obscure	0.31	0.25		×
Financial wealth $\in$ (10000 €, 30000 €]	0.27		×	×
Financial wealth $\in$ (30000 €, ∞)	0.27		×	×
Financial wealth missing	0.17		×	×
Net income > 2500 €	0.46		×	×
Net income missing	0.07		×	×
High education	0.38		×	×
$30 < \text{Age} \leq 50$	0.30		×	×
$50 < \text{Age} \leq 65$	0.34		×	×
Age > 65	0.29		×	×

Sources: LISS panel and own calculations. Variables related to the confidence in return estimates, task difficulty, and task obscurity are scaled to range between 0 and 1. Risk aversion is the standardised average of 3 standardised risk aversion proxies. We omit standard deviations of binary variables. The number of observations is 2,072.

while our respondents expect a positive AEX return on average, their expectations are rather pessimistic relative to the AEX's historical return distribution. Third, our participants tend to place lower probabilities on extreme returns than what has historically been observed. In contrast, expectations for the return of savings accounts are high relative to the rates actually offered at the time of the survey and on average even exceed the expected return for the AEX. In our empirical analyses, we employ the difference between the expected mean return for the AEX and the expected return for the savings account as the empirical analogue of the expected excess return.

Risk Preferences. In September 2013, we elicited risk preferences by asking respondents to complete a variant of the "Preference Survey Module", which was developed in Falk et al. (2014) to measure economic preference parameters in large-scale surveys. As part of this survey module, which we describe in more detail in Section A.1.1.3 of the Appendix, respondents first provided a qualitative self-assessment of their willingness to take risks in general and in the financial domain. They then made choices in a series of hypothetical binary lottery tasks. In our main analysis, we employ the average of the three measures' standardised values.

Transaction Costs. We include several variables to empirically model the impact of transaction costs on stock market participation decisions. We focus on variables that proxy for variation in transaction costs in the form of either monetary or information costs. If monetary expenses of stock market participation are to some degree fixed—e.g., because banks charge a constant amount for setting up and keeping an investment account—then these costs will be less relevant for wealthy households. We therefore include net household income and financial wealth in the economic index to control for variation in the relevance of monetary transaction costs. If comprehension of the basic functioning of the stock market comes with information costs, then these costs will be lower for more numerate and cognitively able households. Both vary with educational attainment and age (McArdle, Smith, and Willis, 2011), which we include as further controls.

3.3.3 Variables Entering the Measurement Error Index

Several quantitative and qualitative measures serve to capture measurement error in individual responses. We employ variables for (i) the consistency with which participants report their expectations, (ii) their confidence in their own beliefs, and (iii) their self-assessment concerning both difficulty and clarity of our survey tasks. On top of such direct proxies, we also include the variables proxying transaction costs in the measurement error index. Indeed, it is difficult to argue for exclusion restrictions in one direction or another for education, income, financial wealth, or age.

In September 2013, one month after eliciting the distribution of beliefs, we asked the same set of respondents to provide a point estimate for the one-year return of the AEX. As a quantitative proxy for measurement error, we compute the absolute difference between the response to this question and the mean belief from the ball allocation task. We conjecture that large discrepancies between the two estimates are indicative of measurement error because respondents who do not have a stable set of beliefs or are incapable of articulating them meaningfully are likely to provide less consistent estimates.[3]

The first two qualitative proxies for measurement error relate to the confidence respondents have in their own estimates. Following the elicitation of the point estimates for the expected returns of the AEX and the savings account, we asked respondents to use a slider interface to express their confidence in their own belief on a scale from 0 to 10. We conjecture that respondents with little faith in their own estimates (e.g., because they did not put much cognitive effort into developing their prediction or did not have much of an interest in financial matters) provide error-ridden estimates. We invert responses to these questions so that higher values correspond to less confidence in ones estimates and scale them to the unit interval.

Both in August and September 2013, we asked subjects to use five-point scales to indicate how clear they found the task descriptions and how difficult they considered the belief elicitation itself. We expect that respondents without an elaborate belief distribution will find it hard to understand and to complete the tasks. For both questions, we aggregate the responses for August and September to create two further measurement error proxies.

The Appendix provides a more detailed description and further summary statistics of all measurement error proxies. The pairwise correlations between the individual proxies are all positive and range between .08 and .52. Notably, all of the proxies' correlations with sociodemographic variables show the same tendencies. For example, all measurement error proxies tend to be lower for highly educated households or households with higher net income, resembling previously-found patterns regarding inconsistent survey responses or item non-response (Manski, 2004; Hurd, 2009).

---

[3] We are not aware of changes in the economic environment between the two surveys that could have induced people to systematically and substantially revise their beliefs. Between August and September 2013, the AEX varied little with closing prices between 362.93 and 382.58.

3.4 Results

3.4.1 Main Specification

Table 3.2 presents parameter estimates for the coefficients of the main specification. In the economic model index, we normalise the coefficient on $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$ to 1, thus expressing the remainder of β^{mod} relative to subjective excess return expectations. In the measurement error index, we proceed in the same way with the coefficient on the absolute difference between the belief measures. As we will discuss in detail below, the link function $h(\cdot, \cdot)$ is (close to) monotonically increasing in the economic model index and monotonically decreasing in the measurement error index. This allows us to infer the direction of partial effects from the coefficient estimates.

The coefficients in both indices are estimated with reasonable precision; their signs and relative magnitudes are plausible given the aforementioned shape of the link function and the scaling of the variables (see Table 3.1). In particular, all variables with exclusion restrictions have the expected signs and most of them are significant. The economic model index increases in the level of the expected excess returns; it decreases in the standard deviation of returns and in risk aversion. The measurement error index increases in all of the 5 employed measurement error proxies.

Both indices vary significantly with a number of the common covariates. For example, financial wealth is positively related to the economic model index and negatively related to the measurement error index. This is consistent with wealthy households facing lower transaction costs, while at the same time having stronger incentives to form an opinion about stock market developments. Interestingly, education seems to mostly work through the measurement error index, but it has little impact on the economic model index.

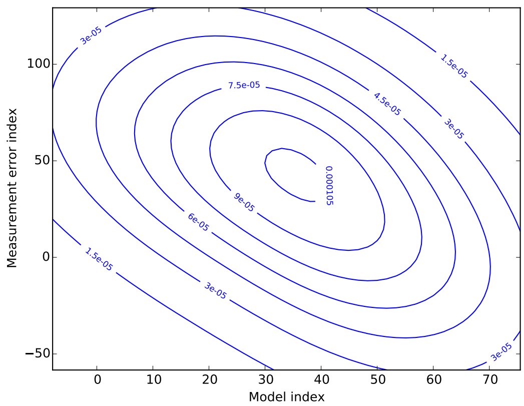
For presenting the results of semi- and nonparametric methods, it is particularly important to clarify the support of the data, which in our case refers to the two indices. Figure 3.1 shows a contour plot of the joint density of the estimated indices. We limit the area of Figure 3.1 and of all subsequent plots to the rectangle spanned by the 5% and 95% quantiles of the marginal distributions of both indices. With a correlation coefficient of -0.45, the indices are characterised by a pronounced negative correlation. Note that this negative correlation does not arise purely mechanically due to the previously noted influence of wealth on both indices—in a model that drops all variables common to both indices (described in the next section), we find the same pattern.

Table 3.2. Coefficient estimates for the economic model index and the measurement index

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	1.00	·	·	·
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.70	0.36	·	·
Risk aversion	-9.62	2.30	·	·
Absolute difference between belief measures	·	·	1.00	·
Lack of confidence in AEX return estimate	·	·	44.54	24.49
Lack of confidence in sav. acc. return estimate	·	·	17.09	21.43
Experimental tasks difficult	·	·	37.29	16.89
Experimental tasks obscure	·	·	30.58	18.99
Financial wealth ∈ (10000 €, 30000 €]	20.54	6.77	-23.17	26.11
Financial wealth ∈ (30000 €, ∞)	39.81	9.74	-77.90	41.40
Financial wealth missing	45.29	10.57	-12.78	29.30
Net income > 2500 €	8.85	3.73	22.45	11.73
Net income missing	-12.26	5.38	-5.90	13.13
High education	2.24	4.00	-47.84	19.50
30 < Age ≤ 50	20.34	7.87	34.20	21.52
50 < Age ≤ 65	16.42	6.51	5.32	15.21
Age > 65	4.64	6.26	-10.62	15.54

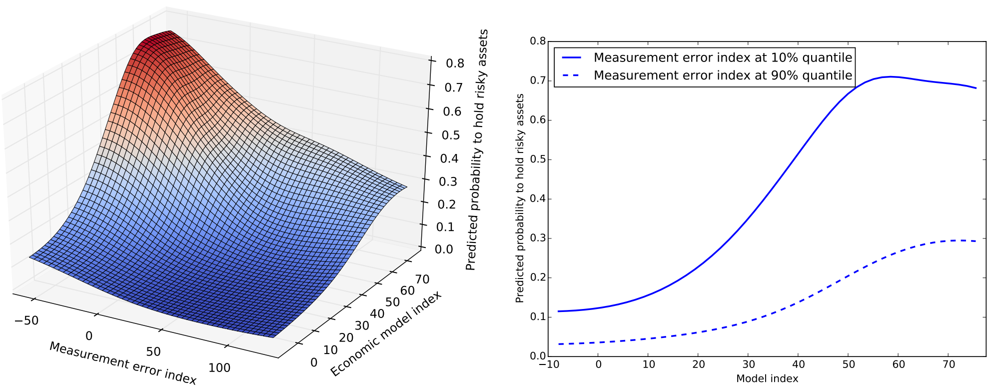
Sources: LISS panel and own calculations. The table shows coefficient estimates for the double index binary choice model of Klein and Vella (2009); see Section 3.2.3 for a detailed description. The dependent variable is a household's stock market participation decision, a binary variable equalling 1 in case the household reports holding any investments, and 0 otherwise. Columns 2 and 3 present estimates of the coefficients and standard errors for the variables contained in the economic model index. Columns 4 and 5 present estimates for the variables contained in the measurement error index.

Figure 3.1. Joint density of the two indices



Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model; see Section 3.2.3 for a detailed description.

Figure 3.2. Predicted probability to hold risky assets



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index (-40 and 115). Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

The left panel of Figure 3.2 plots the link function $h(\cdot,\cdot)$, i.e., the predicted probability of stock market participation, for varying levels of the economic model and measurement error indices. Three features of the plot stand out: First, predicted stock market participation rates vary substantially, ranging from single-digit values to more than 70%. Second, participation rates in general vary monotonically in both indices, increasing in the index for the economic model and decreasing in the measurement error index. Third and most importantly, the effects are highly non-linear and interact strongly. In particular, stock market participation is much more responsive to changes in the economic model ingredients at low levels of the measurement error index than at high levels.

To illustrate the last point more clearly, the second panel in Figure 3.2 extracts two slices from the first panel. The solid line shows the average response of stock market participation to variation in the model index at the 10% quantile of the measurement error index. There is a pronounced gradient in the middle region, causing predicted risky asset participation to rise from just over 10% to 70%. The dashed line plots the same relation for the 90% quantile of measurement error. Again, predicted stock market participation varies in the economic model index as expected, but to a much lesser extent. In particular, even for the highest levels of the economic model index, the predicted probability of participation does not rise above 30%. The discrepancy in shapes of the two lines highlights the importance of measurement error in understanding the relationship between the primitives of economic models and choices.

Table 3.3. Average partial effects

	Model	Meas. Err.	Combined
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	0.032	·	0.032
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.012	·	-0.012
Risk aversion	-0.041	·	-0.041
Absolute difference between belief measures	·	-0.020	-0.020
Lack of confidence in AEX return estimate	·	-0.015	-0.015
Lack of confidence in sav. acc. return estimate	·	-0.006	-0.006
Experimental tasks difficult	·	-0.019	-0.019
Experimental tasks obscure	·	-0.011	-0.011
Financial wealth ∈ (10000 €, 30000 €]	0.101	0.030	0.103
Financial wealth ∈ (30000 €, ∞)	0.215	0.153	0.369
Financial wealth missing	0.241	0.014	0.222
Net income > 2500 €	0.038	-0.034	0.003
Net income missing	-0.053	0.009	-0.045
High education	0.010	0.087	0.097
30 < Age ≤ 50	0.088	-0.054	0.034
50 < Age ≤ 65	0.073	-0.010	0.066
Age > 65	0.021	0.019	0.039

Sources: LISS panel and own calculations. The table presents average partial effects of the Klein and Vella (2009) model; see Section 3.2.3 for a detailed description. The effects are calculated for a change of 1 standard deviation in continuous variables. For binary variables, we calculate the effect by assigning individuals in the left-out category a value of 1.

We calculate average partial effects to quantify the dependence between individual covariates and stock market participation probabilities. In Table 3.3, we show how changes in covariates affect participation through either the economic or measurement error index. We also show the combined effect that operates through both indices simultaneously. To calculate average partial effects, we increase continuous variables by one standard deviation. For binary variables, we assign individuals in the left-out category a value of 1.

For the variables solely included in the economic model index, the average partial effects of expected excess return and risk aversion are somewhat larger than the effect of a change in the expected standard deviation of returns. An increase in the expected excess return by one standard deviation is associated with an increase of 3.2 percentage points in the probability to hold investments. Comparable increases in the expected standard deviation and risk aversion reduce the predicted participation rate by 1.2 and 4.1 percentage points, respectively. Increases in either of the measurement error proxies by one standard deviation reduce the propensity to participate by between 0.6 and 2 percentage points. If one thinks of the different indicators in terms of a factor structure (Section A.1.2 in the Appendix shows that all indicators are positively correlated), varying the underlying factor would likely yield effects comparable in magnitude to those of beliefs or risk aversion.

The effects of financial wealth tend to work through both indices, increasing the propensity to participate in the stock market through the economic model index as well as the measurement error index. In contrast, education seems to affect participation mainly through the measurement error index.

In sum, this section indicates that respondents' beliefs and risk attitudes are indeed predictive of economic choices. However, the extent to which this is the case varies strongly in the population. Hence, measurement error in the primitives of the economic model can be used to uncover heterogeneity in its explanatory power.

3.4.2 Robustness

To illustrate the robustness of our results to alternative specifications of both the economic model and the measurement error index, we now present an overview of a number of additional analyses. Section A.2 of the Appendix contains all tables, figures, and some additional information.

No Transaction Cost Proxies. Our main specification includes several covariates that proxy transaction costs. Some of them—financial wealth in particular—have strong effects on stock market participation through both the economic model index and the measurement error index. To investigate whether the predicted interactions between the economic model and measurement error are driven by these sociodemographics only, we estimate one specification without all of the corresponding proxies, i.e., we only include beliefs, risk preferences, and measurement error proxies. Except for lower predicted levels of stock market participation at high values of the model index, the overall results on $h(\cdot,\cdot)$ look very similar. Naturally, the partial effects change.

Mean Beliefs Only. In this specification, we restrict the model index to consist of expected excess returns only, thus endowing it with an interpretable scale. Section A.2.2 of the Appendix shows that the gist of our main results is present even in this stripped-down version. The relationship between beliefs and stock market participation is essentially flat at the 90th percentile of the measurement error index, while the probability to hold stocks doubles along the beliefs distribution at the 10th percentile of the measurement error index. This doubling is concentrated around expected excess returns of zero, whereas the relationship is flat at both extremes of the beliefs distributions. The pattern illustrates the usefulness of our semiparametric approach; typical parametric models such as Logit or Probit would yield the steepest gradient at the right tail of the support of the index instead of its center.

Additional Covariates. We also check the other extreme and employ a "kitchen-sink"-type approach, including binary variables for gender, having children, and being married in both indices along with the variables from our main specification. It turns out, however, that neither of these is significantly associated with the index of the economic model or the measurement error index nor does their inclusion affect the general tendency of our results.

Discarding Individuals with Missing Data on Financial Wealth. In our main specification, we included dummies for financial wealth terciles and for whether information on financial wealth was missing. Since wealth is among the strongest drivers of stock market participation in our model, it is possible that inclusion of respondents with missing information on portfolio value affects our results. To address this concern, we estimate our main specification only with respondents who provided all components of financial wealth. The results are very similar. In particular, the shape of $h(\cdot,\cdot)$ is virtually unchanged and the average partial effects of beliefs and preferences are almost identical to those in the main specification.

Alternative Belief Measure. We showed our main results using stated beliefs over the future development of the Amsterdam Exchange Index (AEX). While it is plausible that expectations over a composite index with high media exposure are a good proxy for "the" risky asset in our model, it is still conceivable that our results are biased due to this specific choice. We therefore elicited the same set of belief variables for the future stock return of Philips N.V., one of the largest publicly traded companies of the Netherlands. As one would expect for a single stock with additional idiosyncratic risk, average partial effects relating to the moments of the belief distribution are reduced by one third (mean) and one half (standard deviation), respectively. The general shape of the link function and all other results are essentially unchanged.

Disaggregated Risk Aversion Measures. By averaging over three distinct variables, we employed a particularly simple aggregation procedure for the risk aversion measure used in our main analysis. When including the three variables separately in the model index, aversion to risk in financial matters emerges as its most important component (Section A.2.6 of the Appendix). The remainder of our results is not affected.

Alternative Ways of Calculating the Moments of Belief Distributions. We arrived at our individual-level measures of μ_{t+1}^{AEX} and $\sigma_{t+1}^{\text{AEX}}$ by fitting log-normal distributions to respondents' stated cumulative distribution functions. We obtain very similar results when we estimate the moments assuming uniformly distributed expectations within bins (Section A.2.7 of the Appendix) or when we follow Bellemare, Bissonnette, and Kröger (2012) in approximating each respondent's distribution using a spline interpolation method (Section A.2.8).

3.4.3 Specification with Less Customised Data

Our analyses employ very detailed data on respondents' stock market expectations based on an incentivised online experiment. Our proxies for measurement error include a quantitative variable derived from repeated belief measurements and several qualitative indicators. In many surveys, asking for information this detailed is either impossible or impractical. We now evaluate the applicability of our empirical approach to situations with less customised data.

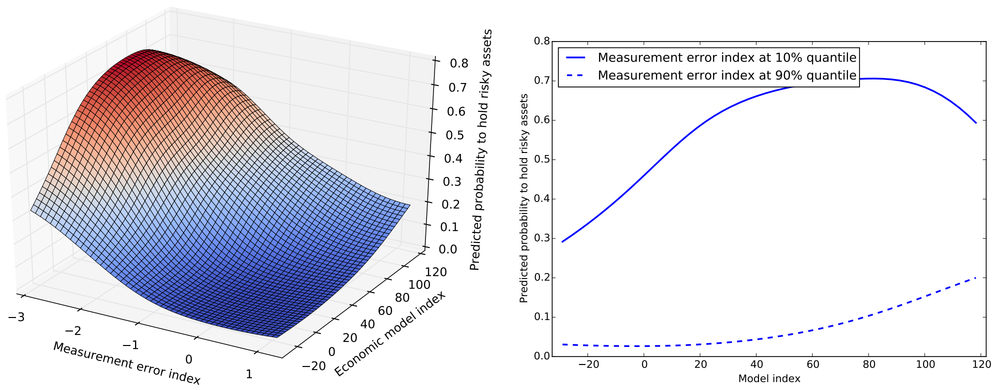
In the model index, we replace the mean of the log-normal belief distribution derived from the ball allocation task by individuals' point estimates. We drop the standard deviation of beliefs and use aversion towards risks in general instead of our composite variable (see Section A.1.1 of the Appendix for a detailed description of all measures). In the measurement error index, we only keep the answers to the qualitative questions which asked respondents about the difficulty and obscurity of our survey. We retain all sociodemographic covariates. We then re-run our main analyses using this limited set of variables.

Figure 3.3 illustrates that the main results for this model are broadly similar to those of our main specification.[4] As the left panel indicates, the predicted probability of holding risky assets strongly varies with both model indices. Importantly, we find strong variation in the gradient of the economic model even with these much coarser data: While the probability of investing in the stock market is sensitive to changes in the economic model index at low values of measurement error, the relationship is essentially flat for high levels of measurement error. The average partial effects in Table 3.4 again suggest that beliefs and willingness to take risks positively affect stock market participation, while the measurement error proxies decrease the probability to invest. All magnitudes are roughly similar to our main specification.

These results entail two consequences: On the one hand, they suggest that measurement error will also interfere with our understanding of stock market participation decisions when working with simple measures of beliefs and risk preferences. On the other hand, they suggest that our empirical approach to making productive use of measurement error of this kind does not seem to rely on very detailed data to work.

[4] Section A.3 of the Appendix provides the remaining results for this model.

Figure 3.3. Predicted probability to hold risky assets, specification with less customised data



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index. The estimation is based on a limited set of variables. Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

Table 3.4. Average partial effects, specification with less customised data

	Model	Meas. Err.	Combined
Subjective beliefs (direct question): Log expected excess return	0.029	·	0.029
Aversion to risks in general	-0.028	·	-0.028
Experimental tasks difficult	·	-0.034	-0.034
Experimental tasks obscure	·	-0.010	-0.010
Financial wealth ∈ (10000 €, 30000 €]	0.073	0.040	0.103
Financial wealth ∈ (30000 €, ∞)	0.049	0.355	0.400
Financial wealth missing	0.088	0.117	0.205
Net income > 2500 €	0.026	-0.011	0.014
Net income missing	-0.092	0.046	-0.053
High education	-0.001	0.117	0.116
30 < Age ≤ 50	0.091	-0.079	0.014
50 < Age ≤ 65	0.054	0.013	0.070
Age > 65	-0.032	0.059	0.024

Sources: LISS panel and own calculations. The table presents average partial effects of the Klein and Vella (2009) model with a limited number of variables. The effects are calculated for a change of 1 standard deviation in continuous variables. For binary variables, we calculate the effect of assigning individuals in the left-out category a value of 1.

3.5 Discussion and Conclusions

Attempts to measure subjective stock market expectations have dramatically increased over the last two decades. By and large, the results have been encouraging, but obvious signs of poor data quality remain for large fractions of the population regardless of particular survey devices (Manski, 2004; Hurd, 2009; Kleinjans and Soest, 2014). When these measures have been employed to predict portfolio choice behaviour (e.g., Hurd and Rohwedder, 2012; Hurd, Rooij, and Winter, 2011; Kézdi and Willis, 2011; Hudomiet, Kézdi, and Willis, 2011; Huck, Schmidt, and Weizsäcker, 2014), significant correlations in the expected direction have emerged. Nevertheless, it seems fair to say that these are not of the magnitude economists might have hoped for. In this paper, we have explored a mechanism that can explain both facts. We have argued that differences in the "propensity to use economic reasoning" may drive heterogeneity in measurement error and explain why the empirical content of portfolio choice models has been moderate on average.

While the idea of heterogeneous decision rules is certainly not new (e.g., Ameriks and Zeldes, 2004; Kahneman, 2011; Binswanger and Salm, 2013, among many others), we are the first to suggest that the magnitude of measurement error in subjective expectations data can be used to uncover such heterogeneity. To explore this link empirically, we have used a semiparametric double index model due to Klein and Vella (2009) on a dataset specifically collected for this purpose. Our results show that stock market participation reacts much more strongly to the primitives of an economic model (preferences, beliefs, and transaction costs) for low values of the measurement error index than for high values. This pattern obtains in a wide variety of specification choices, including a setting where we restrict ourselves to variables that are available in many datasets.

Two pieces of evidence lend further support to our interpretation of these patterns. First, if we were dealing with classical measurement error in beliefs, taking averages of multiple measurements with uncorrelated idiosyncratic variation should increase the predictive power of expectations. A simple exercise shows that such a pattern does not obtain in our data. We run OLS regressions of stock market participation on convex combinations of our two belief measures (the results are unchanged if we add controls). In Section A.4 of the Appendix, we show that the maximum R^2 is reached close to the point where all the weight is on the mean from the ball allocation task. In other words, the addition of the second measure hardly helps at all. Second, we found lower levels of stock market participation for high values of the measurement error index in all our specifications. This suggests that measurement error does not arise purely because of differential effort put into the subjective belief tasks. If some subjects gave random answers which were uncorrelated with portfolio allocations, participation rates should be the same on average. This suggests that the patterns we found do not merely reflect attenuation bias due to respondents' carelessness in providing subjective data.

Our method is applicable to a wide range of settings where subjective data is used. For example, we noted above that the precision of individual-level risk preference parameters obtained from experiments via revealed-preference paradigms varies tremendously in heterogeneous populations (Gaudecker, Soest, and Wengström, 2011; Choi et al., 2014). These findings strongly suggest that the degree to which meaningful structural parameters of economic models exist (Rust, 2014) varies across individuals. We have shown how the individual-level measurement error in structural parameters can be used when these parameters are employed to explain economically interesting outcomes. Doing so should help dampen the hostility of economists to subjective data (Manski, 2004) that has arisen largely because of perceived data quality. We have turned this argument around and shown that once there is direct information on measurement error at the individual level, it can be used to learn about the economic mechanism of interest.

A Appendix to Chapter 3

A.1 Extended Data Description

A.1.1 Variable Definitions and Descriptives

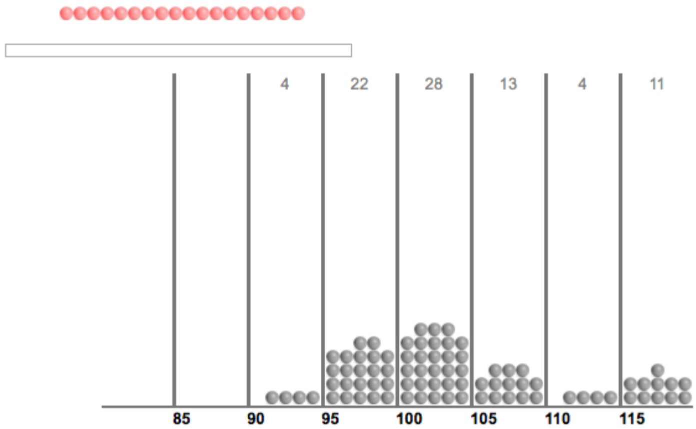
A.1.1.1 Subjective Expectations of Stock Market Returns

AEX return - Ball allocation task. In August 2013, we asked respondents to describe their expectations for the one-year return of the Amsterdam Exchange Index (AEX). To elicit the distribution of individual expectations, we employed a variation of the procedure presented in Delavande and Rohwedder (2008), which was explicitly developed for usage in Internet experiments and pays particular attention to the cognitive burden placed on heterogeneous subject pools. We asked respondents to imagine that they invested 100 € into an exchange traded AEX index fund today and to think about the likely value of this investment in one year. To aid respondents' thinking process and ensure comprehension of the task, the instructions clarified what an index fund is and provided an explicit formula for the value of the investment in one year (*value in a year = 100 € - 0.30 € (fees) + change in the AEX index*).

We then provided respondents with a visual interface that employed an iterative procedure to allow them to state their beliefs as accurately as possible (see Figure A.1). To familiarise subjects with the interface, we showed them an introductory video before asking for their beliefs about the stock market. The video used the example of expected annual rainy days in London to describe the intuition behind the ball allocation procedure and guided subjects through the controls of the interface.

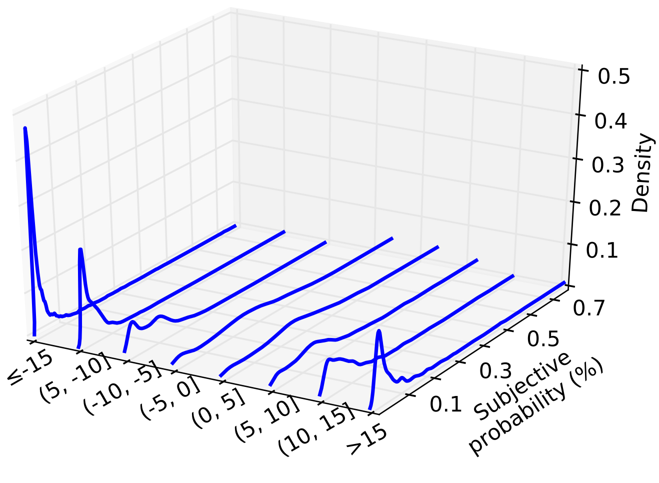
In the first step of the iterative procedure, the interface presented all possible values of the investment as two intervals, $[0,100]$ and $(100,\infty)$. We asked participants to use a slider to allocate 100 balls to indicate their relative confidence that the final value of the investment would fall into either of these intervals. We then split up the interval $(100,\infty)$ into $(100,105]$ and $(105,\infty)$, and we asked subjects to re-allocate the balls from the previous interval to this finer grid. This procedure continued successively until subjects had distributed all balls into 6 interior bins covering intervals of 5 € each and two exterior bins covering the intervals $[0,85]$ and $(115,\infty)$. Fig-

Figure A.1. Visual interface to elicit belief distribution (final step)



The figure shows the final step of the belief eliciation procedure. Respondents used the slider above to allocate 100 balls to the 8 bins below. The figure shows both the remaining balls and the number of balls assigned to each return interval in the previous steps.

Figure A.2. Distribution of probabilities within bins



Sources: LISS panel and own calculations. The picture shows Kernel density estimates of the distribution of probabilites for each of the 8 return intervals.

ure A.2 shows the resulting distribution of balls for each interval expressed in terms of expected returns. While the exterior bins contained only a small number of balls for the large majority of respondents, the distribution of balls in the interior bins was substantially more dispersed.

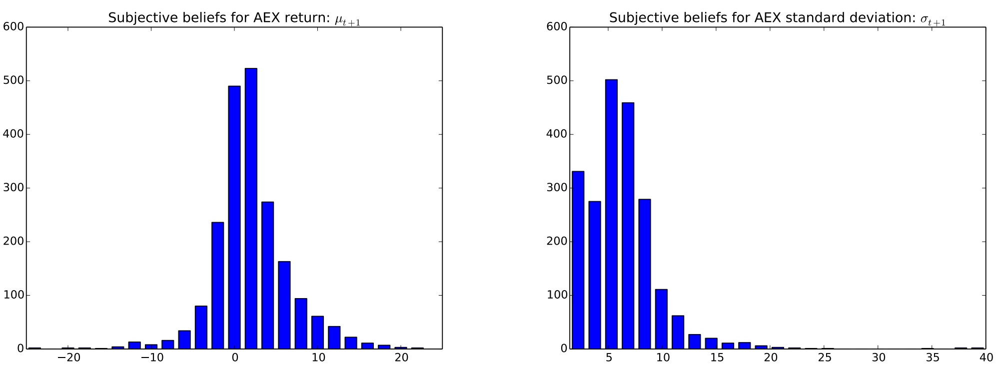
The iterative procedure provides an intuitively simple way of eliciting beliefs and the resulting distribution of balls lends itself to a straightforward interpretation as a histogram. One of its desirable properties is that it does not ask respondents for cumulative probabilities. In contrast, standard survey questions based on the elicitation of points on a cumulative probability distribution often yield logically inconsistent responses due to frequent monotonicity violations. This regularly forces researchers to discard large amounts of data, thereby potentially introducing severe selection effects into the empirical analyses (see, e.g., Manski, 2004; Hurd, Rooij, and Winter, 2011).

To obtain estimates of the mean and variance of individual belief distributions, we employ a procedure similar to Hurd, Rooij, and Winter (2011). We first cumulated the number of balls each respondent assigned to the bins to arrive at a discrete cumulative distribution function. We then used the 7 interior boundary points (b) and the associated values of the CDF (p) to minimise

$$\sum_{i=1}^{7}\left(p_i - \Phi\left(\frac{log(b_i/100)-\mu}{\sigma}\right)\right)^2$$

over μ and σ, our estimates of the mean and standard deviation of the repondent's belief distribution. On average, respondents expect a mean return of 2.01% and a standard deviation of 6.25%. Figure A.3 shows the distribution of estimated mean returns and the distribution of estimated standard deviations. As is evident from the two distributions, subjects have very heterogenuous expectations regarding both the expected return of the AEX as well as its expected standard deviation.

Figure A.3. Distribution of expected mean and standard deviation of returns



Sources: LISS panel and own calculations.

To financially incentivise the task, we used the binarised scoring rule of Hossain and Okui (2013). Subjects could either earn 100€ or 0€, depending on their stated beliefs, the actually realised value of a 100 € investment into the AEX after one year, and the outcome of a random draw. For each subject, we computed the sum of the squared deviations of the belief distribution from the actual value of a 100 € investment after 12 months, $\sum_{i=1}^{8}(b_i - 100 \times \mathbb{1}_i)^2$, where $\mathbb{1}_i$ equalled 1 if the realised value of the investment fell into bin i and 0 otherwise. We then drew a random number from $U[1, 20.000]$. If that random number turned out to be larger (smaller) than the sum of squared deviations, the participant received 100 (0) €.

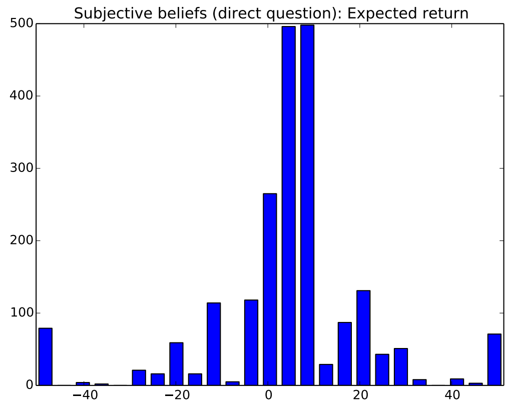
AEX return - One-shot estimate. In September 2013, we asked our full set of respondents for a second, this time non-incentivised, estimate of the one-year return of the AEX using a one-shot question similar to those commonly employed in large-scale surveys:

> *Please consider the Dutch stock market. The AEX index aggregates the stock prices of many of the largest Dutch companies. Now consider an investment fund tracking the AEX index, i.e., this investment exactly moves up and down with the AEX after subtracting rather small fees. If you invested 100 € in such a fund today, the amount of money you would have in a year from now will be:*
>
> $$\textit{value in a year} = 100\text{ €} - 0.30\text{ € } \textit{(fees)} + \textit{change in the AEX index}$$
>
> *What do you think will be this value in a year from now? Please type in your estimate (in Euros).*

Figure A.4 shows the distribution of expected returns implied by subjects' responses to this question. With an average expected return of 4.76%, subjects' point estimates are more optimistic than the mean estimates from the visual task. As is often the case in large-scale representative surveys, we observe a number of outliers in the unrestricted point estimates. Many of these are likely due to typing mistakes or lack of comprehension. Thus, before calculating returns, we winsorise the point estimates at the values of a 100 € investment into the AEX at the 2.5% and 97.5% percentiles of its historical return distribution (49.6 € and 151.3 €). This affected 99 responses.

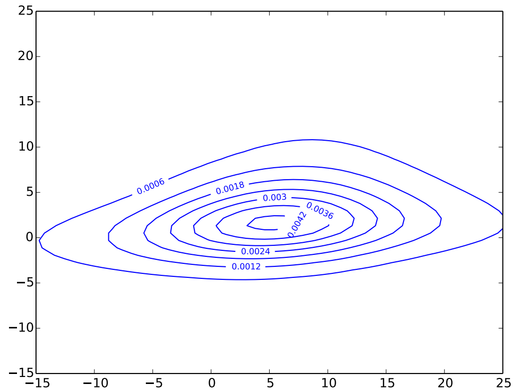
Joint distribution. Figure A.5 shows the joint distribution of the mean estimate from the visual task and the direct estimate from the one shot question. With standard deviations of 6.19% and 17.47%, respectively, the distribution of mean esti-

Figure A.4. Distribution of one-shot estimates for return of AEX



Sources: LISS panel and own calculations.

Figure A.5. Joint distribution of both average belief measures

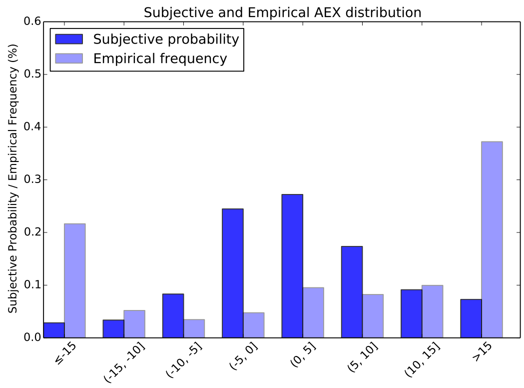


Sources: LISS panel and own calculations.

mates from the visual task is substantially less dispersed than the distribution of direct estimates.

Comparison to historical distribution of AEX returns. Figure A.6 plots the historical distribution of (inflation-adjusted) AEX returns alongside the average probabilities expected by our sample respondents. Respondents considered returns at both ends of the spectrum of the intervals we provided, i.e., in excess of +15% as well as below −15%, far less likely than what has historically been observed. For example, while our average repondent expects less than a 1 in 20 chance of observing returns below −15%, the historical probability of this happening exceeded 20%.

Figure A.6. Expected and historical distribution of AEX returns



Sources: LISS panel and own calculations.

Return to savings account - One-shot estimate. In August 2013, we asked respondents for an estimate of the return of a one-year investment into a standard savings account:

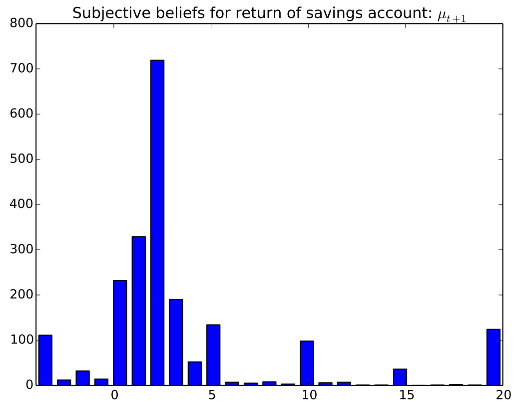
> *Suppose you invested 100 € into a standard savings account with a large Dutch bank. Then, in a year from now, the total amount of money you would have will be:*
>
> $$\textit{value in a year} = 100\text{ €} + \textit{interest payments}$$
>
> *What do you think will be this value in a year from now? Please type in your estimate (in Euros).*

To ensure comprehension of the question, the computer screen also contained a link with more detailed information and the example of a savings account with Rabobank (Rabo SpaarRekening). Figure A.7 shows the distribution of savings estimates. Somewhat surprisingly, subjects' average return estimate for the savings account is 3.35% and thus larger than their average estimate for the AEX in the visual task, though it is smaller than the average point estimate for the AEX. Similar to the one-shot AEX estimates, we winsorise point estimates for the savings account at the 5 and 95% percentiles of the sample distribution before calculating returns.

A.1.1.2 Measurement Error Proxies

Our rich data allow us to employ a number of different variables to proxy for measurement error. We use 5 proxies in total, 1 based on the consistency in stated beliefs,

Figure A.7. Distribution of one-shot estimates for savings account



Sources: LISS panel and own calculations.

2 based on subjects' confidence in their estimates, and 2 based on the subjects' perception of our survey.

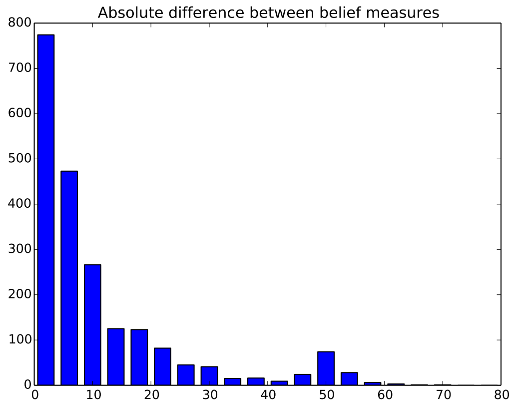
Consistency in beliefs. As discussed in Section A.1.1.1, we used the survey in September 2014 to ask our full set of respondents for a second estimate of the one-year return of the AEX. We use the absolute difference between the response to this question and the mean belief from the visual task as a quantitative proxy for measurement error. Figure A.8 shows a histogram of the absolute differences. On average, subjects' second estimate deviates from the mean estimate in the visual task by a considerable margin, 11.20 percentage points. This seems particularly large when compared to the average expected standard deviation of returns from the ball allocation task (6.25%). Note that these differences are not artifacts of the method we employ to estimate mean beliefs. Other methods, which we describe in Sections A.2.7 and A.2.8 of this appendix, yield very similar results.

Confidence in estimates. Following the elicitation of the point estimates for the expected returns of the AEX and the savings account, we asked respondents how certain they felt about their responses:

> *Please use the slider to indicate how certain you are that the value in a year will equal your estimate. 0 indicates "not certain at all" and 10 means "absolutely certain".*

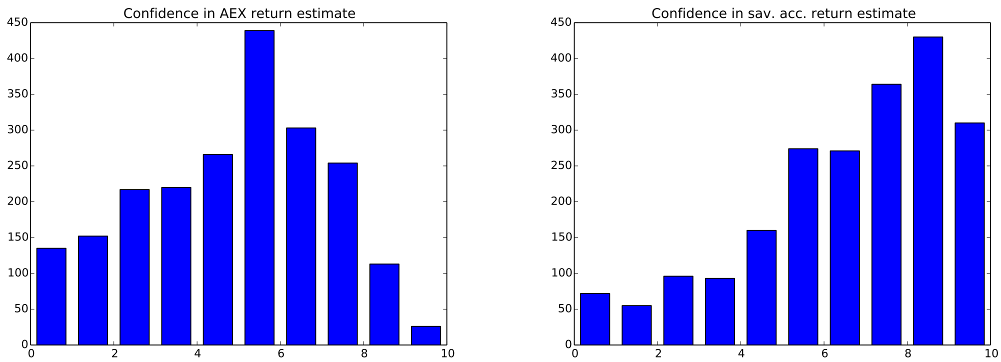
We conjecture that respondents with little confidence in their own estimates (e.g., because they know that they did not expend much cognitive effort into developing their predictions) provide estimates that are likely plagued by measurement

Figure A.8. Distribution of absolute differences between mean belief in visual task and point estimate



Sources: LISS panel and own calculations.

Figure A.9. Distribution of slider values for confidence in estimates

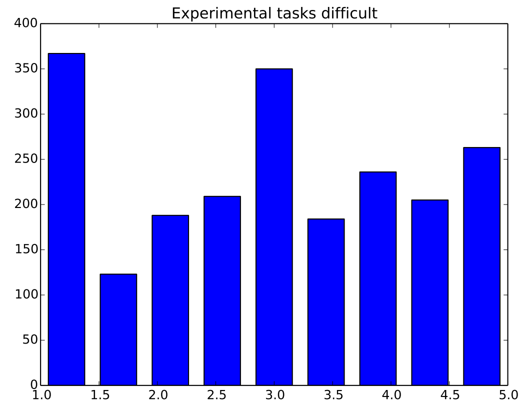


Sources: LISS panel and own calculations.

error and hence not very predictive of actual choices. Figure A.9 shows histograms for the answers to both questions. Respondents seem to be on average less confident in their estimates for the return of the AEX as compared to their estimates for the saving account. For the empirical analyses, we invert the responses so that larger values correspond to a lack of confidence and scale the resulting variables to range between 0 and 1.

Difficulty. Following the surveys in August 2013 and September 2013, we asked subjects to use five-point scales to indicate how difficult they considered the preceding belief elicitation task. We conjecture that answers by respondents who found it

Figure A.10. Distribution of assessments of difficulty



Sources: LISS panel and own calculations.

very hard to detail their stock market expectations are likely measured with error. Figure A.10 shows the distribution of the average of the responses in both surveys. Respondents vary greatly in their assessment of the tasks' difficulties. While some considered it simple, others seemed to find the task very demanding. We scale the average to range between 0 and 1 for our empirical analysis.

Clarity. In August 2013 and September 2013, we also asked subjects to use five-point scales to indicate how vague/obscure they found our questions. We expect that limited comprehension of the task on the side of respondents will introduce measurement error into the answers provided. Figure A.11 shows a histogram of the average response to this question in both surveys. For the empirical analysis, we also scale the average to range between 0 and 1.

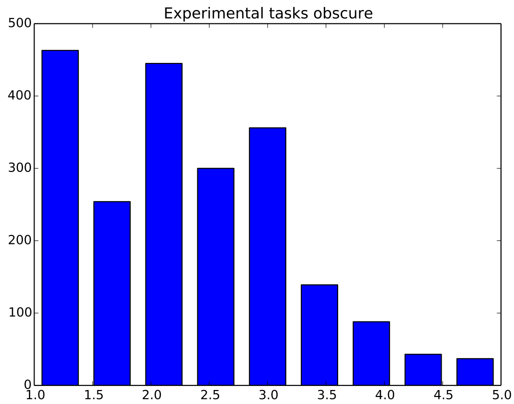
A.1.1.3 Risk Preferences

We use a composite variable to measure risk aversion. To construct this variable, we ask respondents two questions on their self-assessed willingness to take risks and we elicit one quantitative measure based on hypothetical lottery choices. In our empirical analyses, we use the average of the standardised values of all three measures to proxy for risk aversion, suitably coded so that larger values of individual variables as well as the composite variable correspond to larger values of risk aversion.

Risk questions. The subjective self-assessments directly ask for an individual's willingness to take risks, both in general terms and in financial matters:

"Different people have different opinions and characteristics. We are in-

Figure A.11. Distribution of assessments of obscurity

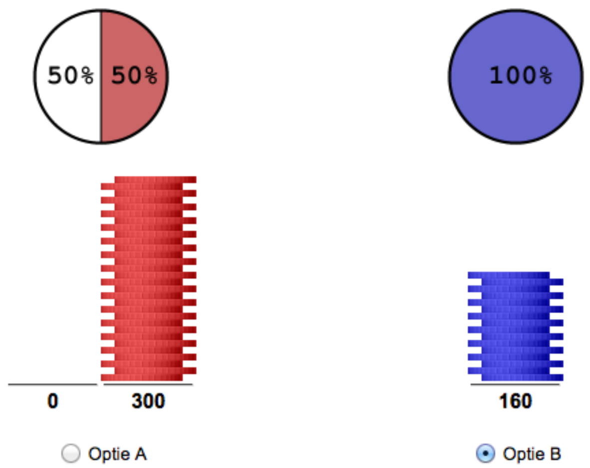


Sources: LISS panel and own calculations.

> *terested in how you describe yourself. In general, to what extent are you willing to take risks? You can answer this question by clicking somewhere on the slider (0-10).”*
>
> *“And, in general, to what extent are you willing to take risks in financial matters? You can answer this question by clicking somewhere on the slider (0-10).”*

Risk lottery. We derive a quantitative measure of risk aversion from a series of five interdependent hypothetical binary lottery choices, a format commonly referred to as the “staircase procedure”. In each of the questions, participants had to decide between a 50/50 lottery to win 300 € or nothing and a varying safe payment. The questions were interdependent in the sense that the choice of a lottery resulted in an increase of the safe amount being offered in the next question, while the choice of the safe payment resulted in a decrease of the safe amount in the next question. For instance, the fixed payment in the first question was 160 €. In case the respondent chose the lottery, the safe payment increased to 240 € in the second question. In case the respondent chose the safe payment, the next question’s fixed payment was reduced to 80 €. By adjusting the fixed payment according to previous choices, the questions allow for a relatively fine quantitative assessment of an individual’s attitudes towards risk. With 32 possible outcomes evenly spaced between 0 and 320 €, the procedure can in principle pin down a respondent’s certainty equivalent to a range of 10 euros. Because of the task’s abstract nature and our heterogeneous sub-

Figure A.12. Graphical illustration of hypothetical lottery choice



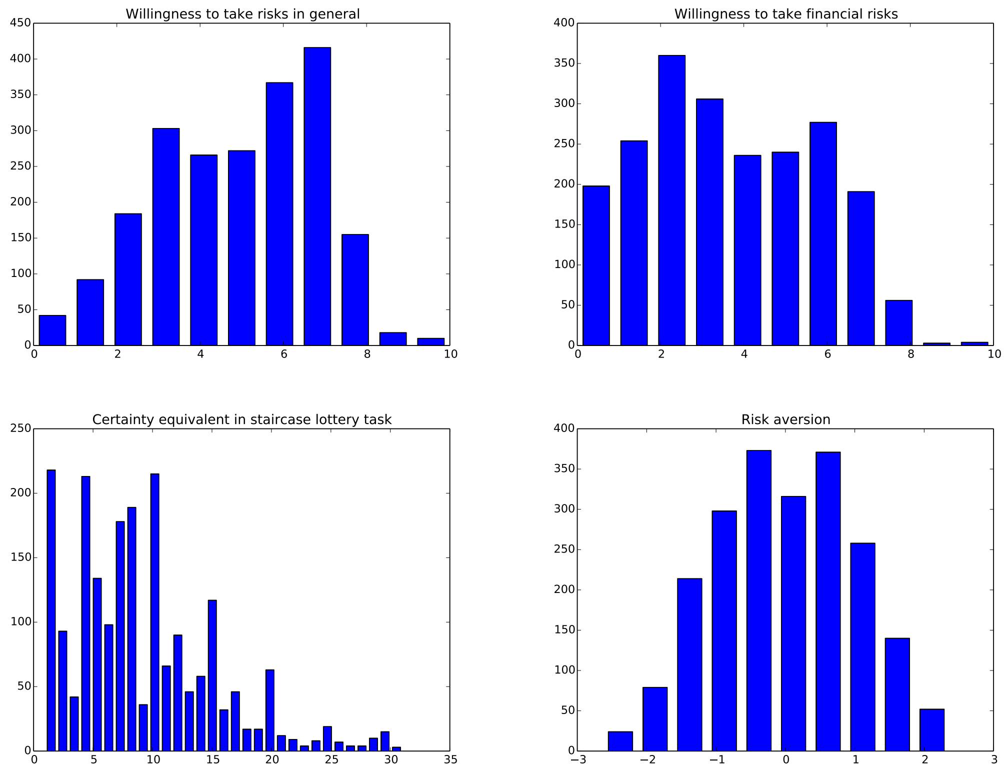
The figure shows the visual interface accompanying one of the lottery decisions.

ject pool, we accompanied each lottery decision with a visual representation of the current lottery to ensure comprehension (see Figure A.12).

The above variables resemble the variables developed for the "Preference Survey Module" in Falk et al. (2014) to measure economic preference parameters in large-scale surveys. Falk et al. (2014) use an experimental validation procedure to select behaviorally valid survey items to measure economic preferences. Dohmen et al. (2011) show that responses to our qualitative survey items correlate with many risky field choices, including stockholdings. Thus, even though the questions we asked were not financially incentivised, they are known to be behaviorally valid and were explicitly developed for the purpose of large-scale studies like ours.

In Figure A.13, we show histograms of the indiviual components as well the composite variable. There is substantial variation in the answers to all three questions. In the lottery task, most of our subjects end up with estimated certainty equivalents below 160 €, suggesting that the majority of our subjects is risk averse.

Figure A.13. Distribution of risk aversion components and aggregate variable



Sources: LISS panel and own calculations.

A.1.1.4 Transaction Cost Proxies / Sociodemographics

Portfolio value. LISS collects detailed information on the value of a respondent's financial assets. To calculate an estimate of the total value of a respondent's portfolio, we sum the amounts held as investments and and those in the bank, which we set to 0 in case the household reported negative values. LISS allows respondents to provide either continuous or interval statements for each category of assets. To calculate the overall portfolio value, we replace categorical answers by the midpoint of the respective interval. For example, we set an answer like "7.500 to 10.000 €" to 8.750 €. For all respondents, we use the most detailed level of information available. For investments, LISS asks both for the aggregate value of investments as well as for the value of the subcategories (stocks, funds, and other investments). We use the more detailed data if available, and we use the answer to the aggregate question otherwise.

Employing the resulting estimate of a respondent's portfolio value, we create categorical variables for each of the sample's portfolio value terciles. Some respondents prefer not to answer the questions concerning their financial situation, so we create one more binary variable for missing portfolio values.

Net household income. Using LISS's information, we create a binary variable for net household income in excess of 2.500 €, the median income of households providing an answer to the income question. We create a further dummy for households with missing values for income ($\approx$ 7% of the sample).

Education. LISS asks respondents for their highest educational degree. In our main estimation, we include a dummy variable for respondents who either report having a university degree or higher vocational education.

Age. Using LISS's data on birthyears, we create binary variables for several different age groups (31 to 50, 51 to 65, and for respondents older than 65).

A.1.2 Correlations

Table A.1 shows the correlation matrix for the main variables.

Table A.1. Correlation matrix

	Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	Risk aversion	Abs. diff. between belief measures	Lack of confidence in AEX return estimate	Lack of confidence in sav. acc. return estimate	Experimental tasks difficult	Experimental tasks obscure	Financial wealth ∈ (10000 €, 30000 €]	Financial wealth ∈ (30000 €, ∞)	Net income > 2500 €	High education	30 < Age ≤ 50	50 < Age ≤ 65	Age > 65
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	1	**-0.19**	**-0.11**	**-0.21**	**-0.16**	**-0.24**	**-0.17**	**-0.18**	-0.00	**0.18**	**0.13**	**0.15**	-0.02	**0.09**	-0.03
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	·	1	0.01	**0.25**	**0.09**	**0.09**	0.04	**0.08**	-0.03	**-0.08**	-0.04	**-0.06**	0.01	-0.04	0.00
Risk aversion	·	·	1	**0.06**	**0.33**	**0.22**	**0.20**	**0.13**	0.00	**-0.07**	**-0.12**	**-0.16**	**-0.09**	0.01	**0.14**
Abs. diff. between belief measures	·	·	·	1	**0.13**	**0.23**	**0.08**	**0.11**	-0.05	**-0.20**	**-0.14**	**-0.19**	-0.01	**-0.06**	0.05
Lack of confidence in AEX return estimate	·	·	·	·	1	**0.52**	**0.23**	**0.20**	-0.04	**-0.08**	**-0.13**	**-0.07**	0.00	-0.04	0.04
Lack of confidence in sav. acc. return estimate	·	·	·	·	·	1	**0.24**	**0.28**	**-0.06**	**-0.20**	**-0.20**	**-0.20**	**-0.07**	-0.04	**0.12**
Experimental tasks difficult	·	·	·	·	·	·	1	**0.48**	0.00	-0.04	**-0.13**	**-0.07**	**-0.07**	**-0.08**	**0.18**
Experimental tasks obscure	·	·	·	·	·	·	·	1	-0.01	**-0.10**	**-0.09**	**-0.10**	-0.05	**-0.07**	**0.13**
Financial wealth ∈ (10000 €, 30000 €]	·	·	·	·	·	·	·	·	1	**-0.37**	0.03	0.02	0.01	**-0.07**	0.05
Financial wealth ∈ (30000 €, ∞)	·	·	·	·	·	·	·	·	·	1	**0.20**	**0.17**	**-0.13**	**0.14**	0.05
Net income > 2500 €	·	·	·	·	·	·	·	·	·	·	1	**0.23**	**0.10**	0.03	**-0.10**
High education	·	·	·	·	·	·	·	·	·	·	·	1	0.05	-0.01	**-0.10**
30 < Age ≤ 50	·	·	·	·	·	·	·	·	·	·	·	·	1	**-0.47**	**-0.42**
50 < Age ≤ 65	·	·	·	·	·	·	·	·	·	·	·	·	·	1	**-0.46**
Age > 65	·	·	·	·	·	·	·	·	·	·	·	·	·	·	1

Significant correlations ($p < 0.01$) printed in bold.

A.1.3 Correlates of Beliefs

Table A.2 presents regressions of various measures of expectations on sociodemographic covariates. The dependent variable in column (1) is the mean belief from the ball allocation task, in column (2) it is the corresponding standard deviation, and column (3) employs the point estimate of the return of a savings account.

Table A.2. Beliefs and sociodemographics

	(1)	(2)	(3)
Constant	2.012***	6.824***	5.716***
	(0.506)	(0.352)	(0.608)
Financial wealth ∈ (10000 €, 30000 €]	0.027	-0.552***	-0.437
	(0.318)	(0.203)	(0.319)
Financial wealth ∈ (30000 €, ∞)	1.100***	-0.785***	-0.896***
	(0.334)	(0.216)	(0.288)
Financial wealth missing	-0.791**	-0.209	0.167
	(0.376)	(0.254)	(0.417)
Net income > 2500 €	0.482*	0.044	-0.357
	(0.254)	(0.161)	(0.249)
Net income missing	0.203	0.042	-1.299***
	(0.442)	(0.329)	(0.475)
High education	0.693***	-0.310**	-1.130***
	(0.238)	(0.155)	(0.218)
30 < Age ≤ 50	0.324	-0.028	-1.094*
	(0.475)	(0.337)	(0.625)
50 < Age ≤ 65	0.196	-0.347	-2.333***
	(0.486)	(0.343)	(0.596)
Age > 65	-0.626	-0.118	-1.766***
	(0.497)	(0.342)	(0.620)
Female	-1.413***	0.265*	1.251***
	(0.237)	(0.156)	(0.237)
Married	-0.043	-0.036	-0.563**
	(0.254)	(0.166)	(0.249)
Has children	0.251	-0.251	0.078
	(0.273)	(0.185)	(0.280)
Observations	2,108	2,108	2,125
Adj. R^2 (%)	5.3	1.2	6.6

The left-hand variable in column (1) is the mean return from the visual task. Column (2) contains the standard deviation of returns in the visual task. Column (3) includes the estimate for the return of the savings account as the left-hand variable. Robust standard errors in parentheses.

A.2 Robustness Checks

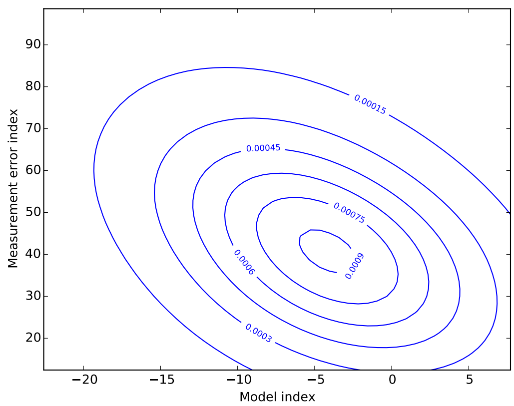
A.2.1 No Transaction Cost Proxies

Table A.3. Coefficient estimates for the economic model index and the measurement index, model without transaction cost proxies

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	1.00	·	·	·
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.75	0.24	·	·
Risk aversion	-4.56	1.00	·	·
Absolute difference between belief measures	·	·	1.00	·
Lack of confidence in AEX return estimate	·	·	-4.18	11.51
Lack of confidence in sav. acc. return estimate	·	·	55.99	20.65
Experimental tasks difficult	·	·	34.07	10.30
Experimental tasks obscure	·	·	10.07	10.67

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.2 in the main text. The model excludes all transaction cost proxies (financial wealth, net income, education, age).

Figure A.14. Joint density of the two indices, model without transaction cost proxies



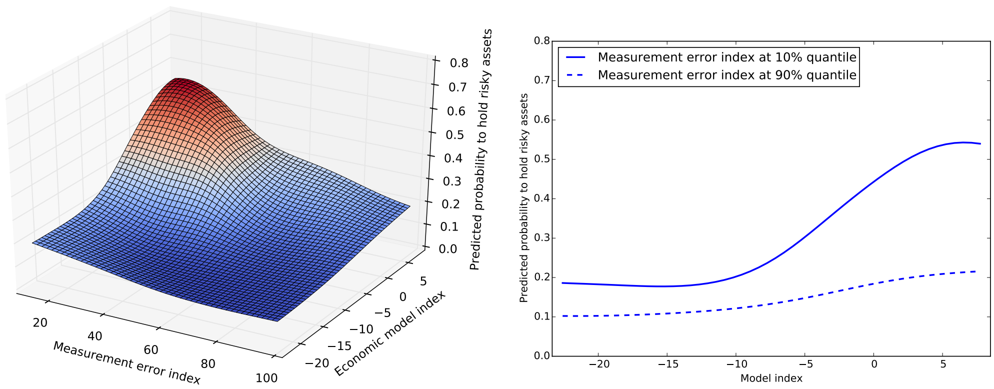
Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model.

Table A.4. Average partial effects, model without transaction cost proxies

	Model	Meas. Err.	Combined
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	0.065	·	0.065
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.031	·	-0.031
Risk aversion	-0.046	·	-0.046
Absolute difference between belief measures	·	-0.034	-0.034
Lack of confidence in AEX return estimate	·	0.002	0.002
Lack of confidence in sav. acc. return estimate	·	-0.035	-0.035
Experimental tasks difficult	·	-0.029	-0.029
Experimental tasks obscure	·	-0.006	-0.006

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.3 in the main text. The model excludes all transaction cost proxies (financial wealth, net income, education, age).

Figure A.15. Predicted probability to hold risky assets, model without transaction cost proxies



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index. Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

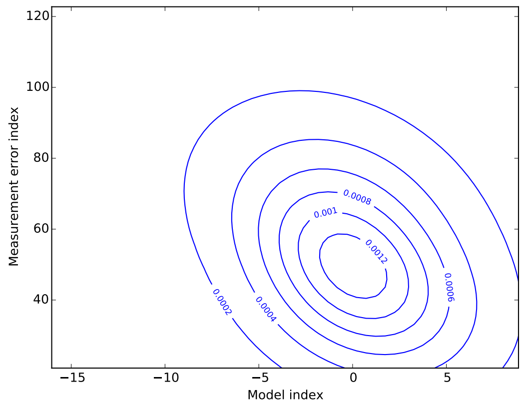
A.2.2 Mean Beliefs Only

Table A.5. Coefficient estimates for the economic model index and the measurement index, model with mean beliefs and measurement error proxies only

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	1.00	·	·	·
Absolute difference between belief measures	·	·	1.00	·
Lack of confidence in AEX return estimate	·	·	11.54	10.07
Lack of confidence in sav. acc. return estimate	·	·	80.10	25.04
Experimental tasks difficult	·	·	26.35	8.74
Experimental tasks obscure	·	·	12.40	10.75

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.2 in the main text. The model excludes the standard deviation of beliefs, risk preferences, and all transaction cost proxies (financial wealth, net income, education, age).

Figure A.16. Joint density of the two indices, model with mean beliefs and measurement error proxies only



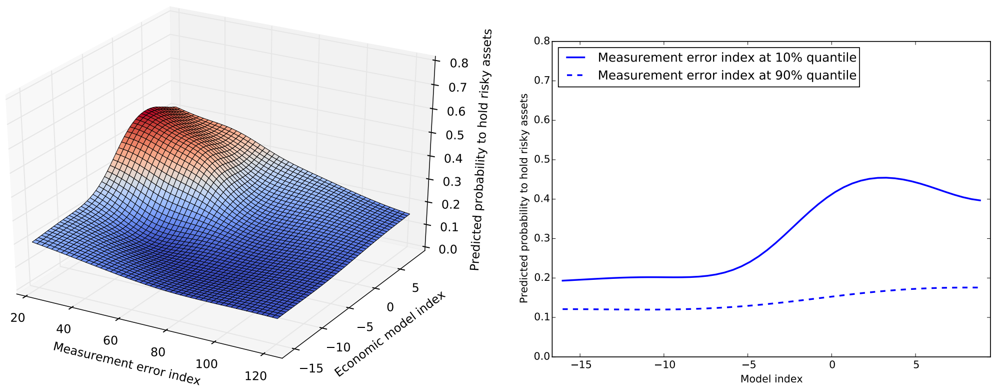
Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model.

Table A.6. Average partial effects, model with mean beliefs and measurement error proxies only

	Model	Meas. Err.	Combined
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	0.036	·	0.036
Absolute difference between belief measures	·	-0.036	-0.036
Lack of confidence in AEX return estimate	·	-0.007	-0.007
Lack of confidence in sav. acc. return estimate	·	-0.051	-0.051
Experimental tasks difficult	·	-0.023	-0.023
Experimental tasks obscure	·	-0.008	-0.008

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.3 in the main text. The model excludes the standard deviation of beliefs, risk preferences, and all transaction cost proxies (financial wealth, net income, education, age).

Figure A.17. Predicted probability to hold risky assets, model with mean beliefs and measurement error proxies only



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index. Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

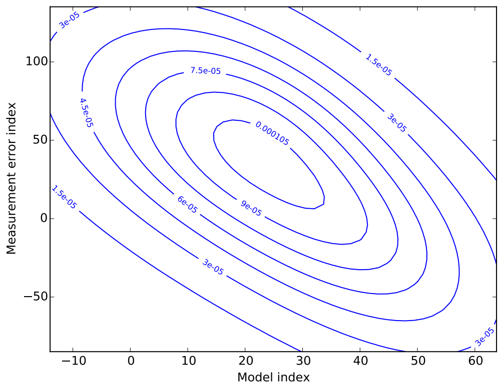
A.2.3 Additional Covariates

Table A.7. Coefficient estimates for the economic model index and the measurement index, model with additional covariates

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	1.00	·	·	·
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.82	0.28	·	·
Risk aversion	-7.98	2.07	·	·
Absolute difference between belief measures	·	·	1.00	·
Lack of confidence in AEX return estimate	·	·	51.16	26.69
Lack of confidence in sav. acc. return estimate	·	·	34.38	22.91
Experimental tasks difficult	·	·	51.04	18.85
Experimental tasks obscure	·	·	16.93	18.47
Financial wealth ∈ (10000 €, 30000 €]	20.20	6.17	-19.07	18.50
Financial wealth ∈ (30000 €, ∞)	42.29	9.71	-86.06	31.11
Financial wealth missing	30.35	7.82	-56.90	25.32
Net income > 2500 €	8.30	2.86	22.50	11.19
Net income missing	-6.92	4.13	-7.66	12.76
High education	2.32	3.30	-60.10	16.76
30 < Age ≤ 50	11.90	5.80	19.76	17.95
50 < Age ≤ 65	7.91	6.11	-21.13	15.98
Age > 65	0.69	6.09	-28.83	17.25
Female	-0.97	2.74	1.30	8.36
Married	-4.03	2.61	13.45	9.41
Has children	2.82	3.30	3.35	9.90

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.2 in the main text, except for the female, marriage, and having children dummies.

Figure A.18. Joint density of the two indices, model with additional covariates



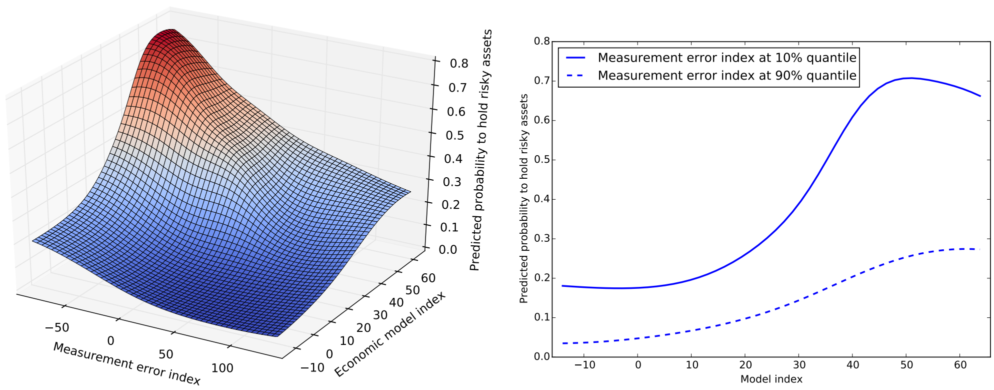
Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model.

Table A.8. Average partial effects, model with additional covariates

	Model	Meas. Err.	Combined
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	0.034	·	0.034
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.015	·	-0.015
Risk aversion	-0.038	·	-0.038
Absolute difference between belief measures	·	-0.015	-0.015
Lack of confidence in AEX return estimate	·	-0.012	-0.012
Lack of confidence in sav. acc. return estimate	·	-0.009	-0.009
Experimental tasks difficult	·	-0.018	-0.018
Experimental tasks obscure	·	-0.004	-0.004
Financial wealth ∈ (10000 €, 30000 €]	0.101	0.018	0.099
Financial wealth ∈ (30000 €, ∞)	0.245	0.119	0.369
Financial wealth missing	0.173	0.072	0.226
Net income > 2500 €	0.041	-0.024	0.017
Net income missing	-0.033	0.009	-0.026
High education	0.011	0.081	0.093
30 < Age ≤ 50	0.056	-0.022	0.027
50 < Age ≤ 65	0.039	0.027	0.063
Age > 65	0.004	0.036	0.033
Female	-0.004	-0.001	-0.006
Married	-0.018	-0.015	-0.034
Has children	0.013	-0.004	0.009

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.3 in the main text, except for the female, marriage, and having children dummies.

Figure A.19. Predicted probability to hold risky assets, model with additional covariates



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index. Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

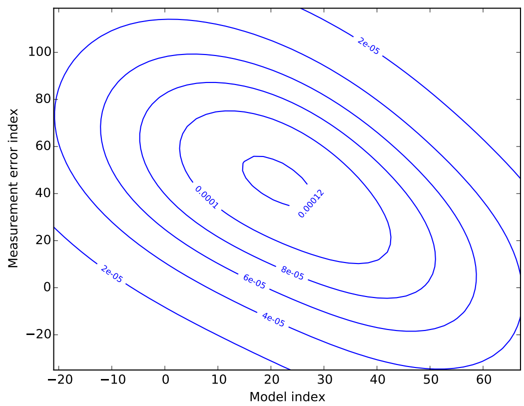
A.2.4 Discarding Individuals with Missing Data on Financial Wealth

Table A.9. Coefficient estimates for the economic model index and the measurement index, sample restricted to individuals with available information on financial wealth

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	1.00	·	·	·
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.94	0.42	·	·
Risk aversion	-10.72	2.90	·	·
Absolute difference between belief measures	·	·	1.00	·
Lack of confidence in AEX return estimate	·	·	47.76	28.36
Lack of confidence in sav. acc. return estimate	·	·	16.69	25.15
Experimental tasks difficult	·	·	38.28	19.24
Experimental tasks obscure	·	·	17.83	17.17
Financial wealth ∈ (10000 €, 30000 €]	24.20	7.52	-6.35	16.73
Financial wealth ∈ (30000 €, ∞)	48.56	13.36	-46.36	29.01
Net income > 2500 €	6.24	3.35	23.50	12.67
Net income missing	-10.15	7.24	-16.84	12.81
High education	-1.26	4.13	-42.15	17.40
30 < Age ≤ 50	12.08	7.80	24.16	17.82
50 < Age ≤ 65	2.81	7.38	-6.44	13.48
Age > 65	-4.40	7.10	-16.90	15.26

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.2 in the main text. The model excludes respondents with missing information on financial wealth.

Figure A.20. Joint density of the two indices, sample restricted to individuals with available information on financial wealth



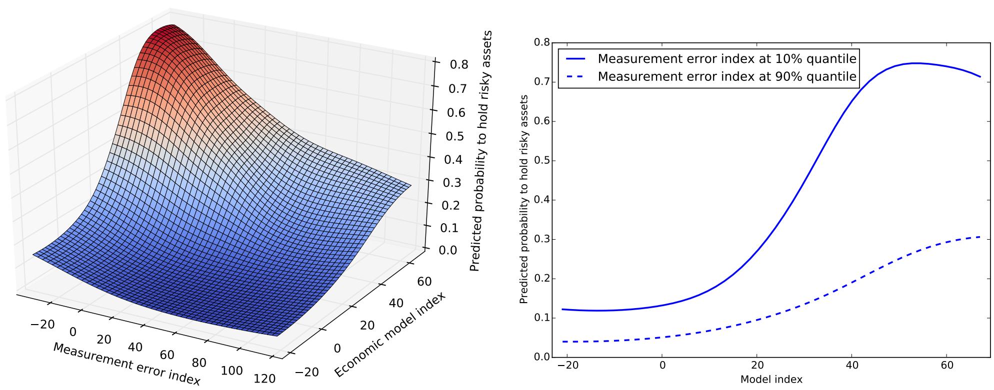
Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model.

Table A.10. Average partial effects, sample restricted to individuals with available information on financial wealth

	Model	Meas. Err.	Combined
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	0.031	·	0.031
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.016	·	-0.016
Risk aversion	-0.045	·	-0.045
Absolute difference between belief measures	·	-0.021	-0.021
Lack of confidence in AEX return estimate	·	-0.017	-0.017
Lack of confidence in sav. acc. return estimate	·	-0.006	-0.006
Experimental tasks difficult	·	-0.021	-0.021
Experimental tasks obscure	·	-0.007	-0.007
Financial wealth ∈ (10000 €, 30000 €]	0.059	0.010	0.068
Financial wealth ∈ (30000 €, ∞)	0.268	0.101	0.375
Net income > 2500 €	0.026	-0.038	-0.012
Net income missing	-0.043	0.028	-0.020
High education	-0.005	0.083	0.078
30 < Age ≤ 50	0.045	-0.042	-0.001
50 < Age ≤ 65	0.011	0.012	0.024
Age > 65	-0.019	0.032	0.009

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.3 in the main text. The model excludes respondents with missing information on financial wealth.

Figure A.21. Predicted probability to hold risky assets, sample restricted to individuals with available information on financial wealth



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index. Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

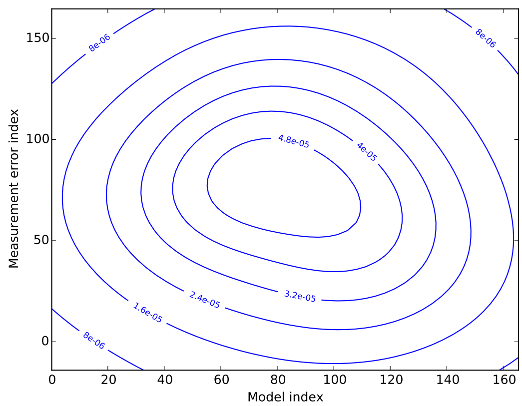
A.2.5 Alternative Belief Measure

Table A.11. Coefficient estimates for the economic model index and the measurement index, Philips instead of AEX

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs: $\mu_{t+1}^{\text{Philips}} - \mu_{t+1}^{\text{sav. acc.}}$	1.00	·	·	·
Subjective beliefs: $\sigma_{t+1}^{\text{Philips}}$	-0.57	0.66	·	·
Risk aversion	-14.05	5.21	·	·
Absolute difference between belief measures	·	·	1.00	·
Lack of confidence in Philips return estimate	·	·	3.65	18.13
Lack of confidence in sav. acc. return estimate	·	·	32.04	20.87
Experimental tasks difficult	·	·	54.23	19.41
Experimental tasks obscure	·	·	19.89	16.68
Financial wealth $\in$ (10000 €, 30000 €]	47.93	21.12	2.04	30.91
Financial wealth $\in$ (30000 €, ∞)	72.04	27.57	-60.64	38.99
Financial wealth missing	50.37	20.90	-46.67	30.85
Net income > 2500 €	29.14	13.83	52.94	17.97
Net income missing	-9.02	10.39	17.44	19.06
High education	27.47	11.23	-2.90	12.20
30 < Age ≤ 50	44.24	19.13	53.61	29.27
50 < Age ≤ 65	30.75	15.15	5.49	22.40
Age > 65	6.23	11.92	-11.83	21.96

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.2 in the main text, except for the belief measures pertaining to Philips N.V..

Figure A.22. Joint density of the two indices, Philips instead of AEX



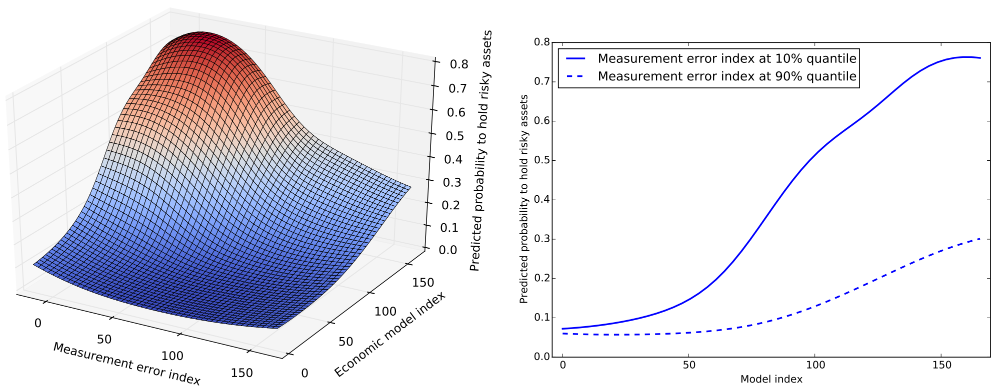
Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model.

Table A.12. Average partial effects, Philips instead of AEX

	Model	Meas. Err.	Combined
Subjective beliefs: $\mu_{t+1}^{\text{Philips}} - \mu_{t+1}^{\text{sav. acc.}}$	0.022	·	0.022
Subjective beliefs: $\sigma_{t+1}^{\text{Philips}}$	-0.005	·	-0.005
Risk aversion	-0.039	·	-0.039
Absolute difference between belief measures	·	-0.026	-0.026
Lack of confidence in Philips return estimate	·	-0.001	-0.001
Lack of confidence in sav. acc. return estimate	·	-0.011	-0.011
Experimental tasks difficult	·	-0.027	-0.027
Experimental tasks obscure	·	-0.007	-0.007
Financial wealth ∈ (10000 €, 30000 €]	0.144	-0.003	0.107
Financial wealth ∈ (30000 €, ∞)	0.240	0.131	0.373
Financial wealth missing	0.154	0.099	0.233
Net income > 2500 €	0.095	-0.066	0.020
Net income missing	-0.026	-0.026	-0.050
High education	0.088	0.004	0.093
30 < Age ≤ 50	0.129	-0.088	0.038
50 < Age ≤ 65	0.090	-0.010	0.089
Age > 65	0.017	0.020	0.036

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.3 in the main text, except for the belief measures pertaining to Philips N.V..

Figure A.23. Predicted probability to hold risky assets, Philips instead of AEX



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index. Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

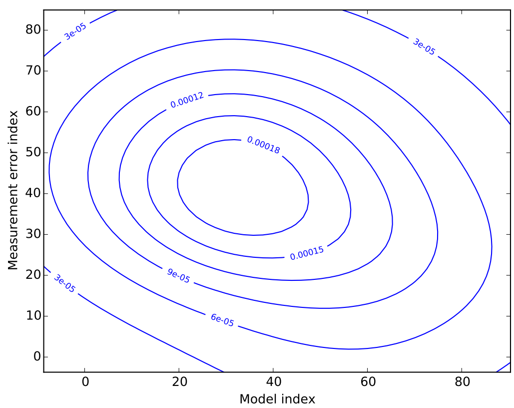
A.2.6 Disaggregated Risk Aversion Measures

Table A.13. Coefficient estimates for the economic model index and the measurement index, separate risk measures

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	1.00	·	·	·
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.81	0.42	·	·
Aversion to risks in general	4.66	2.18	·	·
Aversion to financial risks	-15.79	3.74	·	·
Risk aversion index based on staircase lottery task	-0.15	1.39	·	·
Absolute difference between belief measures	·	·	1.00	·
Lack of confidence in AEX return estimate	·	·	15.16	11.92
Lack of confidence in sav. acc. return estimate	·	·	14.65	15.61
Experimental tasks difficult	·	·	9.37	8.93
Experimental tasks obscure	·	·	17.42	9.19
Financial wealth ∈ (10000 €, 30000 €]	24.06	7.41	0.78	10.15
Financial wealth ∈ (30000 €, ∞)	39.92	10.50	-35.93	14.90
Financial wealth missing	38.77	9.96	-8.01	13.43
Net income > 2500 €	6.79	3.32	7.37	4.18
Net income missing	-7.33	5.33	5.26	6.24
High education	20.49	6.02	24.61	6.39
30 < Age ≤ 50	17.23	7.20	9.46	10.36
50 < Age ≤ 65	16.28	6.48	5.17	8.47
Age > 65	5.52	5.91	-2.47	8.81

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.2 in the main text, except for the disaggregated risk aversion measure.

Figure A.24. Joint density of the two indices, separate risk measures



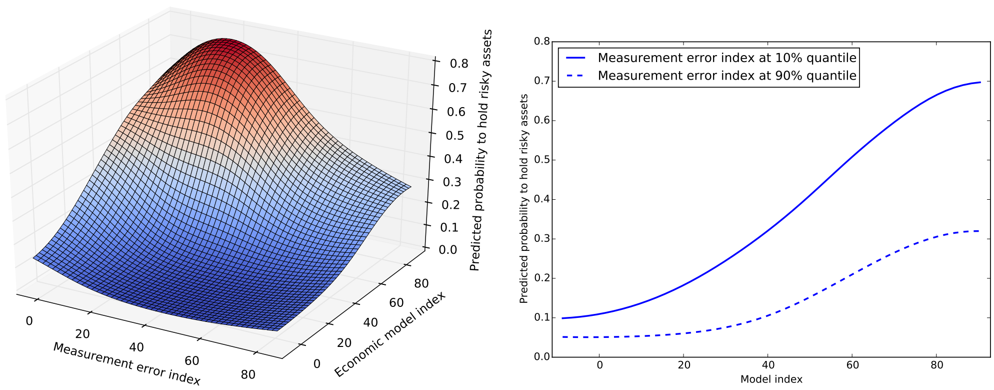
Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model.

Table A.14. Average partial effects, separate risk measures

	Model	Meas. Err.	Combined
Subjective beliefs: $\mu_{t+1}^{\text{AEX}} - \mu_{t+1}^{\text{sav. acc.}}$	0.038	·	0.038
Subjective beliefs: $\sigma_{t+1}^{\text{AEX}}$	-0.015	·	-0.015
Aversion to risks in general	0.022	·	0.022
Aversion to financial risks	-0.069	·	-0.069
Risk aversion index based on staircase lottery task	-0.000	·	-0.000
Absolute difference between belief measures	·	-0.027	-0.027
Lack of confidence in AEX return estimate	·	-0.007	-0.007
Lack of confidence in sav. acc. return estimate	·	-0.007	-0.007
Experimental tasks difficult	·	-0.006	-0.006
Experimental tasks obscure	·	-0.008	-0.008
Financial wealth ∈ (10000 €, 30000 €]	0.115	-0.002	0.093
Financial wealth ∈ (30000 €, ∞)	0.218	0.139	0.358
Financial wealth missing	0.210	0.026	0.224
Net income > 2500 €	0.033	-0.014	0.019
Net income missing	-0.034	-0.011	-0.044
High education	0.116	-0.031	0.077
30 < Age ≤ 50	0.080	-0.019	0.063
50 < Age ≤ 65	0.075	-0.010	0.067
Age > 65	0.024	0.005	0.030

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.3 in the main text, except for the disaggregated risk aversion measure.

Figure A.25. Predicted probability to hold risky assets, separate risk measures



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index. Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

A.2.7 Moments of the Belief Distribution Calculated Using Uniformly Distributed Expectations within Bins

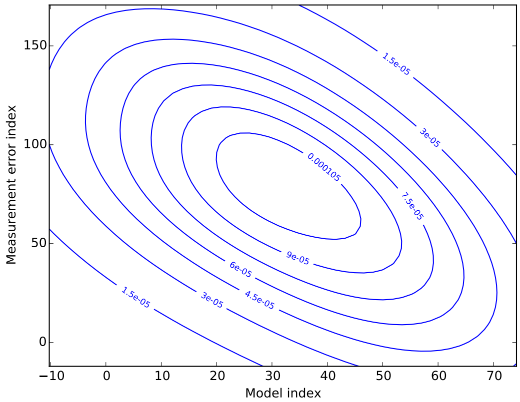
The simplest way to approximate the individual-specific distribution of beliefs is to assume that respondents' expectations are uniformly distributed within bins. To calculate moments under this assumption, we need to assign values to the outer bounds of the exterior bins. We fix these bounds at the values a 100 € investment would have had at the 2.5% and 97.5% percentile of the AEX's historical return distribution, 49.6 € and 151.3 €. We then compute the moments of the distribution assuming that the balls are uniformly distributed within each of the resulting 8 intervals.

Table A.15. Coefficient estimates for the economic model index and the measurement index, moments of beliefs calculated assuming uniform distributions within bins

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs (uniform): Expected excess return	1.00	·	·	·
Subjective beliefs (uniform): Expected standard deviation	-0.75	0.24	·	·
Risk aversion	-8.06	1.76	·	·
Absolute difference between belief measures	·	·	1.00	·
Lack of confidence in AEX return estimate	·	·	72.82	24.75
Lack of confidence in sav. acc. return estimate	·	·	57.22	21.68
Experimental tasks difficult	·	·	31.45	15.41
Experimental tasks obscure	·	·	13.43	14.45
Financial wealth $\in$ (10000 €, 30000 €]	22.85	6.47	-7.40	15.21
Financial wealth $\in$ (30000 €, ∞)	46.75	9.74	-54.92	23.20
Financial wealth missing	32.75	7.85	-40.87	19.52
Net income > 2500 €	6.35	2.69	34.35	11.47
Net income missing	-5.97	3.98	2.25	12.42
High education	5.48	2.81	-44.49	13.39
30 < Age ≤ 50	14.59	6.14	37.30	18.67
50 < Age ≤ 65	12.77	5.72	6.70	15.98
Age > 65	4.76	5.44	-6.96	15.01

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.2 in the main text, except for the estimated moments of the belief distribution.

Figure A.26. Joint density of the two indices, moments of beliefs calculated assuming uniform distributions within bins



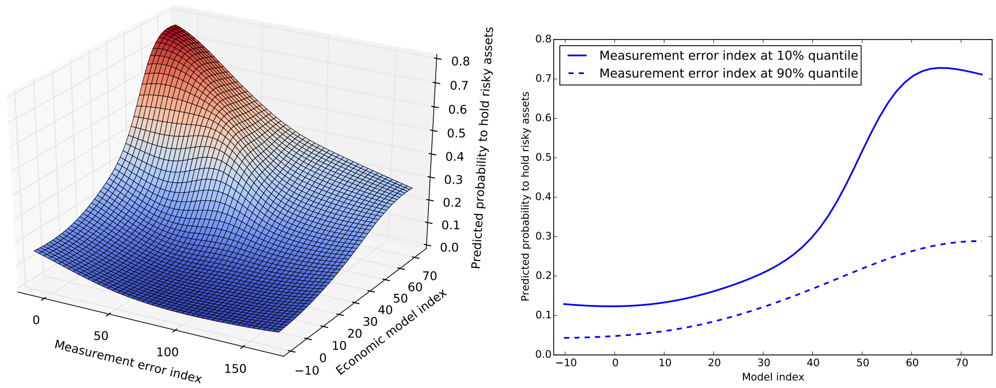
Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model.

Table A.16. Average partial effects, moments of beliefs calculated assuming uniform distributions within bins

	Model	Meas. Err.	Combined
Subjective beliefs (uniform): Expected excess return	0.041	·	0.041
Subjective beliefs (uniform): Expected standard deviation	-0.016	·	-0.016
Risk aversion	-0.042	·	-0.042
Absolute difference between belief measures	·	-0.013	-0.013
Lack of confidence in AEX return estimate	·	-0.016	-0.016
Lack of confidence in sav. acc. return estimate	·	-0.013	-0.013
Experimental tasks difficult	·	-0.010	-0.010
Experimental tasks obscure	·	-0.003	-0.003
Financial wealth ∈ (10000 €, 30000 €]	0.112	0.008	0.109
Financial wealth ∈ (30000 €, ∞)	0.290	0.069	0.364
Financial wealth missing	0.191	0.051	0.224
High education	0.029	0.049	0.079
Net income > 2500 €	0.035	-0.028	0.007
Net income missing	-0.032	-0.002	-0.033
30 < Age ≤ 50	0.080	-0.040	0.041
50 < Age ≤ 65	0.070	-0.007	0.070
Age > 65	0.027	0.006	0.033

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.3 in the main text, except for the estimated moments of the belief distribution.

Figure A.27. Predicted probability to hold risky assets, moments of beliefs calculated assuming uniform distributions within bins



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index. Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

A.2.8 Moments of the Belief Distribution Calculated Using Piecewise Cubic Hermite Interpolating Splines

We also approximate individual belief distributions using piecewise cubic Hermite interpolating splines, very similar to the method proposed in Bellemare, Bissonnette, and Kröger (2012). For each respondent, we first calculate a discrete cumulative distribution function by successively summing the probabilities assigned to each of the 8 bins. The method is less sensitive to the assumptions concerning the support of the exterior bins, so we fix these at more conservative values, the minimum and maximum of the AEX's historical return distribution over a calendar year, i.e., 47.0 € and 176.9 €. We then use a Hermite spline to connect the 9 points on the resulting CDF. The spline interpolates the CDF between each pair of neighbouring points by a monotonically increasing cubic polynomial, whose first derivative at each of the 7 interior points coincides with the respective first derivative of the polynomial in the next-higher interval. We employ the resulting estimate of an indivual's belief distribution to calculate the mean and standard deviation of the individual's return estimate.[1]

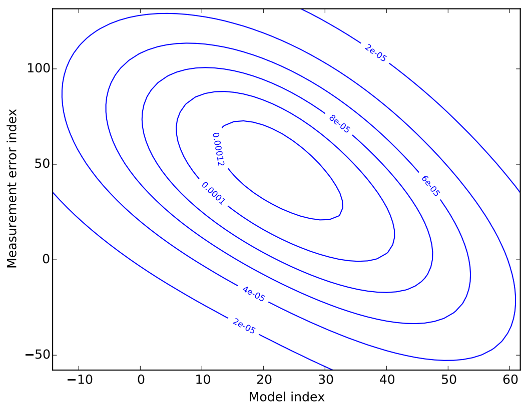
[1] We use the SciPy functions `scipy.interpolate.PchipInterpolator` to fit the splines and `scipy.integrate.quad` to calculate their moments.

Table A.17. Coefficient estimates for the economic model index and the measurement index, moments of beliefs calculated by approximating the distribution using splines

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs (Splines): Expected excess return	1.00	·	·	·
Subjective beliefs (Splines): Expected standard deviation	-0.73	0.17	·	·
Risk aversion	-7.15	1.43	·	·
Absolute difference between belief measures	·	·	1.00	·
Lack of confidence in AEX return estimate	·	·	54.82	24.96
Lack of confidence in sav. acc. return estimate	·	·	23.28	21.51
Experimental tasks difficult	·	·	49.76	17.83
Experimental tasks obscure	·	·	13.43	16.79
Financial wealth ∈ (10000 €, 30000 €]	19.98	5.11	-3.65	20.03
Financial wealth ∈ (30000 €, ∞)	40.82	7.18	-63.10	29.48
Financial wealth missing	28.21	5.92	-36.99	24.25
Net income > 2500 €	6.58	2.46	26.56	10.78
Net income missing	-6.15	4.07	-8.72	13.58
High education	3.94	2.90	-54.08	17.93
30 < Age ≤ 50	10.50	5.28	22.88	16.45
50 < Age ≤ 65	8.06	5.25	-12.07	14.03
Age > 65	0.25	5.04	-22.64	14.94

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.2 in the main text, except for the estimated moments of the belief distribution.

Figure A.28. Joint density of the two indices, moments of beliefs calculated by approximating the distribution using splines



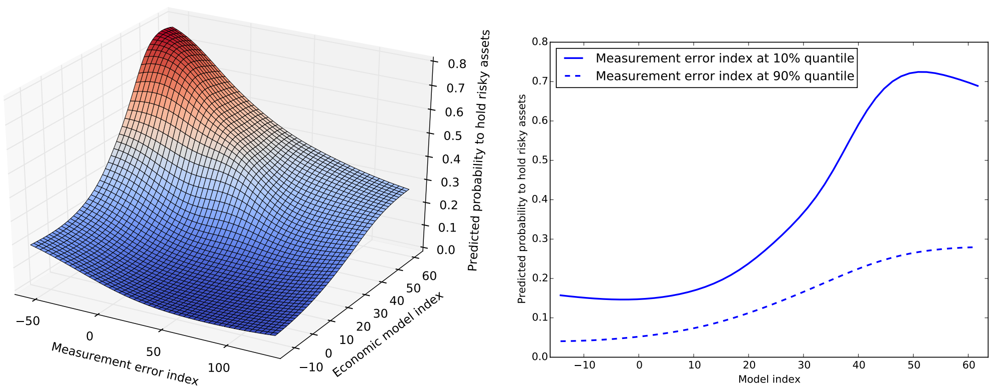
Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model.

Table A.18. Average partial effects, moments of beliefs calculated by approximating the distribution using splines

	Model	Meas. Err.	Combined
Subjective beliefs (Splines): Expected excess return	0.043	·	0.043
Subjective beliefs (Splines): Expected standard deviation	-0.020	·	-0.020
Risk aversion	-0.038	·	-0.038
Absolute difference between belief measures	·	-0.014	-0.014
Lack of confidence in AEX return estimate	·	-0.013	-0.013
Lack of confidence in sav. acc. return estimate	·	-0.006	-0.006
Experimental tasks difficult	·	-0.017	-0.017
Experimental tasks obscure	·	-0.003	-0.003
Financial wealth ∈ (10000 €, 30000 €]	0.110	0.004	0.102
Financial wealth ∈ (30000 €, ∞)	0.265	0.090	0.367
Financial wealth missing	0.177	0.049	0.211
High education	0.021	0.072	0.093
Net income > 2500 €	0.037	-0.026	0.011
Net income missing	-0.033	0.010	-0.025
30 < Age ≤ 50	0.057	-0.025	0.026
50 < Age ≤ 65	0.044	0.015	0.058
Age > 65	0.001	0.027	0.022

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.3 in the main text, except for the estimated moments of the belief distribution.

Figure A.29. Predicted probability to hold risky assets, moments of beliefs calculated by approximating the distribution using splines



Sources: LISS panel and own calculations. The left panel presents the predicted probability of stock market participation for varying levels of the economic model and measurement error indices. The right panel plots the relation between the predicted probability of participation and the economic model index for the 10 and 90% quantiles of the measurement error index. Ranges are limited to the interval between the 5% and 95% quantiles of the marginal distributions.

A.3 Specification with Less Customised Data

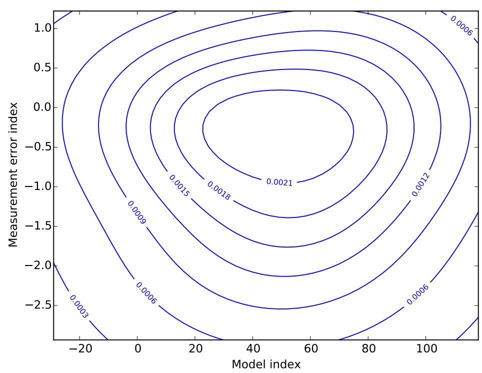
This section reports additional results for the specification with less customised data described in Section 3.4.3 of the main text. The specification is restricted to (i) the point estimate of AEX returns, (ii) one qualitative question to elicit risk attitudes, (iii) two simple qualitative measurement error proxies, and (iv) sociodemographics.

Table A.19. Coefficient estimates for the economic model index and the measurement index, specification with less customised data

	Model		Measurement Error	
	Estimate	Std. Err.	Estimate	Std. Err.
Subjective beliefs (direct question): Log expected excess return	1.00	·	·	·
Aversion to risks in general	-15.42	4.26	·	·
Experimental tasks difficult	·	·	1.00	·
Experimental tasks obscure	·	·	0.38	0.32
Financial wealth ∈ (10000 €, 30000 €]	37.52	16.10	-0.52	0.44
Financial wealth ∈ (30000 €, ∞)	24.81	30.76	-2.33	0.67
Financial wealth missing	46.98	22.68	-1.10	0.54
Net income > 2500 €	13.98	11.20	0.11	0.22
Net income missing	-49.30	19.05	-0.41	0.35
High education	-0.69	14.69	-0.94	0.29
30 < Age ≤ 50	54.41	19.00	0.79	0.43
50 < Age ≤ 65	28.49	15.03	-0.11	0.30
Age > 65	-17.08	15.09	-0.50	0.31

Sources: LISS panel and own calculations. The analyses in this table are analogous to those in Table 3.2 in the main text. The model includes the point estimate to measure beliefs, a qualitative question to elicit risk attitudes, two qualitative measurement error proxies, and sociodemographics.

Figure A.30. Joint density of the two indices, specification with less customised data

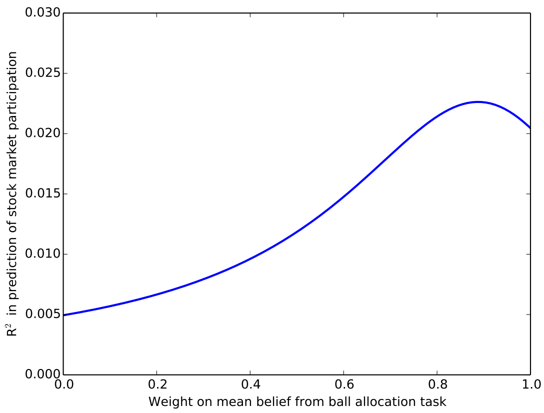


Sources: LISS panel and own calculations. The figure plots the joint density of the estimated indices of the Klein and Vella (2009) model.

A.4 Can We Correct for Measurement Error Using Multiple Measures?

This section argues that correcting for measurement error in subjective beliefs through multiple measures is infeasible. To this end, Figure A.31 presents the R^2 of an OLS regression of a stock market participation dummy on various linear combinations of the mean belief constructed from the ball allocation task and the point estimate. The figure shows that—contrary to what one would expect if repeated measurements reduce measurement error—the variance explained is maximised by putting almost maximal weight on the belief from the ball allocation task. This suggests that traditional methods of correcting for measurement error do not apply in the case of subjective beliefs because there is not necessarily a "true" quantity that an analyst can elicit.

Figure A.31. Variance in stockholdings explained by different linear combinations of two belief measures



Bibliography

Achleitner, Ann-Kristin, André Betzer, and Jasmin Gider (2010). "Do Corporate Governance Motives Drive Hedge Fund and Private Equity Fund Activities?" In: *European Financial Management* 16.5, pp. 805–828.

Alessie, Rob, Stefan Hochgürtel, and Arthur van Soest (2004). "Ownership of Stocks and Mutual Funds: A Panel Data Analysis". In: *The Review of Economics and Statistics* 86.3, pp. 783–796.

Ameriks, John and Stephen P. Zeldes (2004). "How Do Household Portfolio Shares Vary With Age?" Working Paper, Columbia Graduate School of Business.

Amihud, Yakov (2002). "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects". In: *Journal of Financial Markets* 5.1, pp. 31–56.

Andres, Christian (2008). "Large Shareholders and Firm Performance - An Empirical Examination of Founding-Family Ownership". In: *Journal of Corporate Finance* 14.4, pp. 431–445.

Aragon, George O. (2007). "Share Restrictions and Asset Pricing: Evidence from the Hedge Fund Industry". In: *Journal of Financial Economics* 83.1, pp. 33–58.

Artmann, Sabine, Philipp Finter, Alexander Kempf, Stefan Koch, and Erik Theissen (2012). "The Cross-Section of German Stock Returns: New Data and New Evidence". In: *Schmalenbach Business Review* 64.1, pp. 20–43.

Asquith, Paul, Michael B. Mikhail, and Andrea S. Au (2005). "Information Content of Equity Analyst Reports". In: *Journal of Financial Economics* 75.2, pp. 245–282.

Barber, Brad M., Reuven Lehavy, and Brett Trueman (2010). "Ratings Changes, Ratings Levels, and the Predictive Value of Analysts' Recommendations". In: *Financial Management* 39.2, pp. 533–553.

Barber, Brad M. and John D. Lyon (1997). "Detecting Long-Run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics". In: *Journal of Financial Economics* 43.3, pp. 341–372.

Barber, Brad M. and Terrance Odean (2008). “All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors”. In: *Review of Financial Studies* 21.2, pp. 785–818.

Battalio, Robert H. and Richard R. Mendenhall (2011). “Post-Earnings Announcement Drift: Bounds on Profitability for the Marginal Investor”. In: *Financial Review* 46.4, pp. 513–539.

Bebchuk, Lucian A. and Alma Cohen (2005). “The Costs of Entrenched Boards”. In: *Journal of Financial Economics* 78.2, pp. 409–433.

Bellemare, Charles, Luc Bissonnette, and Sabine Kröger (2012). “Flexible Approximation of Subjective Expectations Using Probability Questions”. In: *Journal of Business & Economic Statistics* 30.1, pp. 125–131.

Bessler, Wolfgang, Wolfgang Drobetz, and Julian Holler (2013). “The Returns to Hedge Fund Activism in Germany”. In: *European Financial Management, forthcoming.*

Binswanger, Johannes and Martin Salm (2013). “Does Everyone Use Probabilities? Intuitive and Rational Decisions about Stockholding”. IZA Discussion Paper No. 7265.

Black, Bernard S. (1990). “Shareholder Passivity Reexamined”. In: *Michigan Law Review* 89.3, pp. 520–608.

Boehmer, Ekkehart, Jim Musumeci, and Annette B. Poulsen (1991). “Event-Study Methodology under Conditions of Event-Induced Variance”. In: *Journal of Financial Economics* 30.2, pp. 253–272.

Boyson, Nicole M. and Robert M. Mooradian (2011). “Corporate Governance and Hedge Fund Activism”. In: *Review of Derivatives Research* 14.2, pp. 169–204.

Bradley, Daniel, Jonathan Clarke, Suzanne Lee, and Chayawat Ornthanalai (2014). “Are Analysts’ Recommendations Informative? Intraday Evidence on the Impact of Time Stamp Delays”. In: *The Journal of Finance* 69.2, pp. 645–673.

Bratton, William W. (2007). “Hedge Funds and Governance Targets”. In: *Georgetown Law Journal* 95.5, pp. 1375–1433.

Brav, Alon, Wei Jiang, Frank Partnoy, and Randall Thomas (2008). “Hedge Fund Activism, Corporate Governance, and Firm Performance”. In: *The Journal of Finance* 63.4, pp. 1729–1775.

Bruine de Bruin, Wändi and Katherine G. Carman (2012). “Measuring Risk Perceptions: What Does the Excessive Use of 50% Mean?” In: *Medical Decision Making* 32.2, pp. 232–6.

Bruine de Bruin, Wändi, Baruch Fischhoff, Susan G. Millstein, and Bonnie L. Halpern-Felsher (2000). “Verbal and Numerical Expressions of Probability: "It’s a Fifty-Fifty Chance"”. In: *Organizational Behavior and Human Decision Processes* 81.1, pp. 115–131.

Campbell, John Y. and Luis M. Viceira (2002). *Strategic Asset Allocation*. Clarendon Lectures in Economics. Oxford University Press.

Carhart, Mark M. (1997). “On Persistence in Mutual Fund Performance”. In: *The Journal of Finance* 52.1, pp. 57–82.

Chen, Xia, Qiang Cheng, and Kin Lo (2010). “On the Relationship between Analyst Reports and Corporate Disclosures: Exploring the Roles of Information Discovery and Interpretation”. In: *Journal of Accounting and Economics* 49.3, pp. 206–226.

Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman (2014). “Who Is (More) Rational?” In: *American Economic Review* 104.6, pp. 1518–1550.

Clement, Michael B. (1999). “Analyst Forecast Accuracy: Do Ability, Resources, and Portfolio Complexity Matter?” In: *Journal of Accounting and Economics* 27.3, pp. 285–303.

Clement, Michael B. and Senyo Y. Tse (2005). “Financial Analyst Characteristics and Herding Behavior in Forecasting”. In: *The Journal of Finance* 60.1, pp. 307–341.

Clifford, Christopher P. (2008). “Value Creation or Destruction? Hedge Funds as Shareholder Activists”. In: *Journal of Corporate Finance* 14.4, pp. 323–336.

Cronqvist, Henrik and Rüdiger Fahlenbrach (2009). “Large Shareholders and Corporate Policies”. In: *Review of Financial Studies* 22.10, pp. 3941–3976.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers (1997). “Measuring Mutual Fund Performance with Characteristic-Based Benchmarks”. In: *The Journal of Finance* 52.3, pp. 1035–1058.

Davis, Gerald F. and E. Han Kim (2007). “Business Ties and Proxy Voting by Mutual Funds”. In: *Journal of Financial Economics* 85.2, pp. 552–570.

Del Guercio, Diane and Jennifer Hawkins (1999). “The Motivation and Impact of Pension Fund Activism”. In: *Journal of Financial Economics* 52.3, pp. 293–340.

Delavande, Adeline and Susann Rohwedder (2008). "Eliciting Subjective Probabilities in Internet Surveys". In: *Public Opinion Quarterly* 72.5, pp. 866–891.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner (2011). "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences". In: *Journal of the European Economic Association* 9.3, pp. 522–550.

Drerup, Tilman H. (2014a). "Long-Term Effects of Hedge Fund Activism in Germany". Working paper, available at SSRN (Abstract ID 1718365).

– (2014b). "Recommendation Revisions and Differences in the Interpretation of Earnings". Mimeo, University of Bonn.

Drerup, Tilman H., Benjamin Enke, and Hans-Martin von Gaudecker (2014). "Measurement Error in Subjective Expectations and the Empirical Content of Economic Models". Working paper, available at SSRN (Abstract ID 2509323).

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde (2014). "An Experimentally Validated Preference Module". Mimeo, University of Bonn.

Fama, Eugene F. (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work". In: *The Journal of Finance* 25.2, pp. 383–417.

Fama, Eugene F. and Kenneth R. French (1993). "Common Risk Factors in the Returns on Stocks and Bonds". In: *Journal of Financial Economics* 33.1, pp. 3–56.

Foster, George (1977). "Quarterly Accounting Data: Time-Series Properties and Predictive-Ability Results". In: *The Accounting Review* 52.1, pp. 1–21.

Francis, Jennifer and Leonard Soffer (1997). "The Relative Informativeness of Analysts' Stock Recommendations and Earnings Forecast Revisions". In: *Journal of Accounting Research* 35.2, pp. 193–211.

Franks, Julian and Colin Mayer (2001). "Ownership and Control of German Corporations". In: *Review of Financial Studies* 14.4, pp. 943–977.

Freeman, Robert N. and Senyo Y. Tse (1989). "The Multiperiod Information Content of Accounting Earnings: Confirmations and Contradictions of Previous Earnings Reports". In: *Journal of Accounting Research* 27, pp. 49–79.

Gächter, Simon and Elke Renner (2010). "The Effects of (Incentivized) Belief Elicitation in Public Goods Experiments". In: *Experimental Economics* 13.3, pp. 364–377.

Gaudecker, Hans-Martin von. "How does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice?" In: *The Journal of Finance, forthcoming.*

Gaudecker, Hans-Martin von, Arthur van Soest, and Erik Wengström (2011). "Heterogeneity in Risky Choice Behavior in a Broad Population". In: *American Economic Review* 101.2, pp. 664–694.

Grossman, Sanford J. and Oliver D. Hart (1980). "Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation". In: *The Bell Journal of Economics* 11.1, pp. 42–64.

Haliassos, Michael and Carol C. Bertaut (1995). "Why Do so Few Hold Stocks?" In: *The Economic Journal* 105.432, pp. 1110–1129.

Hoderlein, Stefan and Joachim Winter (2010). "Structural Measurement Errors in Nonseparable Models". In: *Journal of Econometrics* 157.2, pp. 432–440.

Hossain, Tanjim and Ryo Okui (2013). "The Binarized Scoring Rule". In: *The Review of Economic Studies* 80.3, pp. 984–1001.

Hu, Henry T.C. and Bernard S. Black (2007). "Hedge Funds, Insiders, and the Decoupling of Economic and Voting Ownership: Empty Voting and Hidden (Morphable) Ownership". In: *Journal of Corporate Finance* 13.2-3, pp. 343–367.

Huck, Steffen, Tobias Schmidt, and Georg Weizsäcker (2014). "The Standard Portfolio Choice Problem in Germany". Working paper, available at SSRN (Abstract ID 2432045).

Hudomiet, Péter, Gábor Kézdi, and Robert J. Willis (2011). "Stock Market Crash and Expectations of American Households". In: *Journal of Applied Econometrics* 26.3, pp. 393–415.

Hurd, Michael D. (2009). "Subjective Probabilities in Household Surveys". In: *Annual Review of Economics* 1.1, pp. 543–562.

Hurd, Michael D. and Susann Rohwedder (2012). "Stock Price Expectations and Stock Trading". NBER working paper series: http://www.nber.org/papers/w17973.

Hurd, Michael D., Maarten van Rooij, and Joachim Winter (2011). "Stock Market Expectations of Dutch Households". In: *Journal of Applied Econometrics* 26.3, pp. 416–436.

Ivković, Zoran and Narasimhan Jegadeesh (2004). "The Timing and Value of Forecast and Recommendation Revisions". In: *Journal of Financial Economics* 73.3, pp. 433–463.

Jegadeesh, Narasimhan and Jason Karceski (2009). "Long-Run Performance Evaluation: Correlation and Heteroskedasticity-Consistent Tests". In: *Journal of Empirical Finance* 16.1, pp. 101–111.

Jegadeesh, Narasimhan, Joonghyuk Kim, Susan D. Krische, and Charles M. C. Lee (2004). "Analyzing the Analysts: When Do Recommendations Add Value?" In: *The Journal of Finance* 59.3, pp. 1083–1124.

Johnson, Norman J. (1978). "Modified t Tests and Confidence Intervals for Asymmetrical Populations". In: *Journal of the American Statistical Association* 73.363, pp. 536–544.

Kahan, Marcel and Edward B. Rock (2007). "Hedge Funds in Corporate Governance and Corporate Control". In: *University of Pennsylvania Law Review* 155.5, pp. 1021–1093.

Kahneman, Daniel (2011). *Thinking, Fast and Slow.* New York, NY: Farrar, Straus and Giroux.

Karpoff, Jonathan M. (2001). "The Impact of Shareholder Activism on Target Companies: A Survey of Empirical Findings". Working paper, available at SSRN (Abstract ID 885365).

Kézdi, Gábor and Robert J. Willis (2011). "Household Stock Market Beliefs and Learning". NBER working paper series: http://www.nber.org/papers/w17614.

Klein, April and Emanuel Zur (2009). "Entrepreneurial Shareholder Activism: Hedge Funds and Other Private Investors". In: *The Journal of Finance* 64.1, pp. 187–229.

Klein, Roger W. and Richard H. Spady (1993). "An Efficient Semiparametric Estimator for Binary Response Models". In: *Econometrica* 61.2, p. 387.

Klein, Roger W. and Francis Vella (2009). "A Semiparametric Model for Binary Response and Continuous Outcomes under Index Heteroscedasticity". In: *Journal of Applied Econometrics* 24.5, pp. 735–762.

Kleinjans, Kristin J. and Arthur van Soest (2014). "Rounding, Focal Point Answers and Nonresponse to Subjective Probability Questions". In: *Journal of Applied Econometrics* 29.4, pp. 567–585.

Koch, Adam S. and Amy X. Sun (2004). “Dividend Changes and the Persistence of Past Earnings Changes”. In: *The Journal of Finance* 59.5, pp. 2093–2116.

Kothari, S.P. (2001). “Capital Markets Research in Accounting”. In: *Journal of Accounting and Economics* 31.1-3, pp. 105–231.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny (2000). “Investor Protection and Corporate Governance”. In: *Journal of Financial Economics* 58.1-2, pp. 3–27.

Lipton, Martin and Steven A. Rosenblum (1991). “A New System of Corporate Governance: The Quinquennial Election of Directors”. In: *The University of Chicago Law Review* 58.1, pp. 187–253.

Livnat, Joshua and Yuan Zhang (2012). “Information Interpretation or Information Discovery: Which Role of Analysts Do Investors Value More?” In: *Review of Accounting Studies* 17.3, pp. 612–641.

Loh, Roger K. and Rene M. Stulz (2011). “When Are Analyst Recommendation Changes Influential?” In: *Review of Financial Studies* 24.2, pp. 593–627.

Lyon, John D., Brad M. Barber, and Chih-Ling Tsai (1999). “Improved Methods for Tests of Long-Run Abnormal Stock Returns”. In: *The Journal of Finance* 54.1, pp. 165–201.

Manski, Charles F. (2004). “Measuring Expectations”. In: *Econometrica* 72.5, pp. 1329–1376.

Maurer, Jürgen (2009). “Who Has a Clue to Preventing the Flu? Unravelling Supply and Demand Effects on the Take-Up of Influenza Vaccinations”. In: *Journal of Health Economics* 28.3, pp. 704–17.

McArdle, John J., James P. Smith, and Robert Willis (2011). “Cognition and Economic Outcomes in the Health and Retirement Survey”. In: *Explorations in the Economics of Aging*. University of Chicago Press, pp. 209–233.

Mendenhall, Richard R. (1991). “Evidence on the Possible Underweighting of Earnings-Related Information”. In: *Journal of Accounting Research* 29.1, pp. 170–179.

Mietzner, Mark and Denis Schweizer (2014). “Hedge Funds versus Private Equity Funds as Shareholder Activists in Germany - Differences in Value Creation”. In: *Journal of Economics and Finance* 38.2, pp. 181–208.

Mietzner, Mark, Denis Schweizer, and Marcel Tyrell (2011). "Intra-Industry Effects of Shareholder Activism in Germany - Is There a Difference between Hedge Fund and Private Equity Investments?" In: *Schmalenbach Business Review* 63.2, pp. 151–185.

Mikhail, Michael B., Beverly R. Walther, and Richard H. Willis (1997). "Do Security Analysts Improve Their Performance with Experience?" In: *Journal of Accounting Research* 35, pp. 131–157.

– (2003). "The Effect of Experience on Security Analyst Underreaction". In: *Journal of Accounting and Economics* 35.1, pp. 101–116.

Mitchell, Mark L. and Erik Stafford (2000). "Managerial Decisions and Long-Term Stock Price Performance". In: *The Journal of Business* 73.3, pp. 287–329.

Palfrey, Thomas R. and Stephanie W. Wang (2009). "On Eliciting Beliefs in Strategic Games". In: *Journal of Economic Behavior & Organization* 71.2, pp. 98–109.

Park, Chul W. and Morton Pincus (2000). "Market Reactions to Changes in Analyst Consensus Recommendations Following Quarterly Earnings Announcements". Working paper.

Parrino, Robert, Richard W. Sias, and Laura T. Starks (2003). "Voting with their Feet: Institutional Ownership Changes around Forced CEO Turnover". In: *Journal of Financial Economics* 68.1, pp. 3–46.

Patell, James M. (1976). "Corporate Forecasts of Earnings per Share and Stock Price Behavior: Empirical Tests". In: *Journal of Accounting Research* 14.2, pp. 246–276.

Rabin, Matthew (2000). "Risk Aversion and Expected-Utility Theory: A Calibration Theorem". In: *Econometrica* 68.5, pp. 1281–1292.

Rock, Edward B. (1991). "The Logic and (Uncertain) Significance of Institutional Shareholder Activism". In: *Georgetown Law Journal* 79.3, pp. 445–506.

Romano, Roberta (1993). "Public Pension Fund Activism in Corporate Governance Reconsidered". In: *Columbia Law Review* 93.4, pp. 795–853.

Rooij, Maarten van, Annamaria Lusardi, and Rob Alessie (2011). "Financial Literacy and Stock Market Participation". In: *Journal of Financial Economics* 101.2, pp. 449–472.

Rust, John (2014). "The Limits of Inference with Theory: A Review of Wolpin (2013)". In: *Journal of Economic Literature* 52.3, pp. 820–850.

Schennach, Susanne M. (2013). “Measurement Error in Nonlinear Models - A Review”. In: *Advances in Economics and Econometrics: Tenth World Congress: Economic Theory Econometric Society Monographs*. Cambridge University Press, pp. 296–337.

Seifert, Werner G. and Hans-Joachim Voth (2007). *Invasion der Heuschrecken: Intrigen, Machtkämpfe, Marktmanipulationen; Wie Hedge-Fonds die Deutschland-AG attackieren*. Ullstein.

Shiller, Robert J. (1981). “Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?” In: *The American Economic Review* 71.3, pp. 421–436.

Shleifer, Andrei and Robert W. Vishny (1986). “Large Shareholders and Corporate Control”. In: *Journal of Political Economy* 94.3, pp. 461–488.

Stickel, Scott E. (1985). “The Effects of Value Line Investment Survey Rank Changes on Common Stock Prices”. In: *Journal of Financial Economics* 14.1, pp. 121–143.

– (1991). “Common Stock Returns Surrounding Earnings Forecast Revisions: More Puzzling Evidence”. In: *The Accounting Review* 66.2, pp. 402–416.

– (1995). “The Anatomy of the Performance of Buy and Sell Recommendations”. In: *Financial Analysts Journal* 51.5, pp. 25–39.

Stiglitz, Joseph E. (2002). “Information and the Change in the Paradigm in Economics”. In: *American Economic Review* 92.3, pp. 460–501.

Sutton, Clifton D. (1993). “Computer-Intensive Methods for Tests About the Mean of an Asymmetrical Distribution”. In: *Journal of the American Statistical Association* 88.423, pp. 802–810.

Veenman, David (2012). “Disclosures of Insider Purchases and the Valuation Implications of Past Earnings Signals”. In: *The Accounting Review* 87.1, pp. 313–342.

Vissing-Jørgensen, Annette (2002). “Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures”. NBER working paper series: http://www.nber.org/papers/w8884.

Wahal, Sunil (1996). “Pension Fund Activism and Firm Performance”. In: *The Journal of Financial and Quantitative Analysis* 31.1, pp. 1–23.

Wang, Stephanie W. (2011). “Incentive Effects: The Case of Belief Elicitation from Individuals in Groups”. In: *Economics Letters* 111.1, pp. 30–33.

Wansbeek, Tom J. and Erik Meijer (2000). *Measurement Error and Latent Variables in Econometrics*. Amsterdam: Elsevier North Holland.

Weber, Peter and Heinz Zimmermann (2013). “Hedge Fund Activism and Information Disclosure: The Case of Germany”. In: *European Financial Management* 19.5, pp. 1017–1050.

White, Halbert (1980). “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”. In: *Econometrica* 48.4, pp. 817–838.

Wilcoxon, Frank (1945). “Individual Comparisons by Ranking Methods”. In: *Biometrics Bulletin* 1.6, pp. 80–83.

Womack, Kent L. (1996). “Do Brokerage Analysts’ Recommendations Have Investment Value?” In: *The Journal of Finance* 51.1, pp. 137–167.