This PDF is a selection from a published volume from
the National Bureau of Economic Research

Volume Title: NBER Macroeconomics Annual 2003,
Volume 18

Volume Author/Editor: Mark Gertler and Kenneth
Rogoff, editors

Volume Publisher: The MIT Press

Volume ISBN: 0-262-07253-X

Volume URL: http://www.nber.org/books/gert04-1

Conference Date: April 4-5, 2003

Publication Date: July 2004

Title: Perspectives on Behavioral Finance: Does "Irrationality"
Disappear with Wealth? Evidence from Expectations
and Actions

Author: Annette Vissing-Jorgensen

URL: http://www.nber.org/chapters/c11443

Annette Vissing-Jorgensen

KELLOGG SCHOOL OF MANAGEMENT, NORTHWESTERN
UNIVERSITY, NBER, AND CEPR

Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions

1. Introduction

The contributions of behavioral finance are many. The field:

1. Documents price patterns that seem inconsistent with traditional finance models of efficient markets and rational investors.
2. Documents behaviors by investors that seem inconsistent with the advice of traditional finance theory.
3. Provides new theories for these patterns and behaviors, often based on behaviors documented in the psychology literature or observed in experiments.
4. Argues that if prices deviate from fundamentals due to the behavior of irrational investors, arbitrage by rational investors may not be able to force prices back to fundamentals. This part of the behavioral finance literature is referred to as the limits to arbitrage literature.

The most influential work on price patterns within the behavioral finance literature has concerned initial underreaction and (in some cases) subsequent overreaction of prices to new information. This work is described

in Shleifer (2000) as well as in recent surveys of behavioral finance by (insiders) Barberis and Thaler (2003); Daniel, Hirshleifer, and Teoh (2002); and Hirshleifer (2001).

Defenders of the standard rational expectations, efficient markets asset pricing approach have argued that the evidence on underreaction and overreaction is unconvincing because (a) there are as many cases of initial overreaction as initial underreaction, and the evidence is not that solid statistically (Fama, 1998), and (b) if the documented price predictability was statistically solid and stable over time, mutual fund managers should be able to outperform the market substantially, on average, but are not (Rubinstein, 2001). On the modeling side, many have found references to the psychology literature or the experimental literature unconvincing. In some cases, it seems that too much is possible, in the sense that the literature provides evidence both in favor of a given behavioral bias as well as for the opposite bias. Furthermore, many have been skeptical of whether behavioral biases are present in real-world cases where individuals have had time to learn (by themselves or from prior generations), and in particular whether the wealthiest investors with large amounts at stake exhibit behavioral biases. The behavioral side has defended itself by arguing that prices may be far from the predictions of standard models even if (risk-adjusted) profit opportunities are not present. This is the case because arbitrage is limited. First, the mispricing may get worse in the short run (noise trader risk). This is especially a problem when investment is delegated to portfolio managers with short investment horizons (Shleifer and Vishny, 1997). Second, arbitrage is risky when it involves the whole stock market or when it involves individual stocks with no close substitutes. Third, arbitrage may involve substantial transactions costs or be hindered by costs of shorting stocks or restrictions on shorting stocks (e.g., by mutual funds). Barberis and Thaler (2003) provide a discussion of these limits to arbitrage. Abreu and Brunnermeier (2001, 2002) provide an additional argument for limited arbitrage. They show theoretically that it may be optimal for rational arbitrageurs to ride bubbles started by other investors.[1]

In this paper, I argue that direct evidence on investor beliefs and actions is valuable for determining whether assumptions made in behavioral asset pricing models are valid, and thus for determining which (if any) of the models are convincing explanations of the facts they set out to explain. Furthermore, to understand the causes of nonstandard beliefs and/or

1. In support of this theory, Brunnermeier and Nagel (2002) show that hedge fund portfolios were tilted heavily toward technology stocks during the stock-market boom of the late 1990s and that hedge funds started to reduce their exposure in the quarter prior to the price peaks of individual technology stocks.

actions, it is useful to distinguish between beliefs and actions that are present for wealthy investors and thus unlikely to be due to information or transactions costs, and beliefs and actions that are observed predominantly among less wealthy investors. This is also informative for determining whether a given bias is likely to have a substantial pricing impact. It is important to emphasize that biases affecting mostly low-wealth investors are nonetheless also important because these biases may have large effects on the utility of these investors.

I start in Section 2 by discussing the types of evidence about investors that would be valuable for understanding pricing anomalies. In Section 3 I provide new evidence on investor beliefs based on a dataset covering the period 1998–2002. Section 4 then turns to a set of investor behaviors that are inconsistent with the recommendations of standard finance theory and reviews evidence on whether these biases diminish substantially with investor wealth. Section 5 provides a rough calculation of how large information and transactions costs would be needed to explain one particular type of seemingly irrational investor behavior: limited participation in stock markets. Section 6 concludes.

2. The Value of Direct Evidence on Investor Beliefs and Actions

At the aggregate level, stock returns are predictable by the dividend-to-price ratio, the earnings-to-price ratio, the market-to-book ratio, the consumption-to-wealth ratio, and a host of other aggregate variables (see Campbell [2000] for a list of references). The direction of predictability indicates that future stock returns tend to be lower when the stock price is high relative to dividends and earnings. Within a rational agent framework, the interpretation of this is that investors' expected (and required) returns are low at such times. The alternative theory proposed by behavioral finance is overreaction of stock prices to news at the level of the aggregate stock market. According to overreaction theories, the returns expected by market participants are not unusually low when the price-to-dividend ratio is high.

The literature on the cross section of stock returns has identified many return patterns not predicted by standard models. Examples of overreaction include: (1) the market-to-book effect (low market-to-book or low price-to-dividend stocks have historically outperformed high market-to-book stocks [Fama and French, 1992, and earlier references cited therein]), (2) the small firm effect (small stocks have outperformed large stocks [Banz, 1981]), (3) long-run reversal (winners in the past three years perform worse than past three year losers over the following three years [DeBondt and Thaler, 1985]), and (4) the poor long-run performance of the

stock of firms issuing new stock (Loughran and Ritter, 1995). Examples of underreaction include: (1) momentum (winners from the past 3 to 12 months continue to outperform losers from the past 3 to 12 months during the following six months (Jegadeesh and Titman, 1993), and (2) the post-earnings announcement drift (Bernard and Thomas, 1989).

Several risk-based models have been proposed for the cross-sectional return patterns. Berk, Green, and Naik (1999) provide a rational model based on growth options and time-varying risk that generates the market-to-book effect, the size effect, and the momentum effect. Gomes, Kogan, and Zhang (2003) provide a related investment-based explanation of the market-to-book effect and the size effect.

The behavioral finance literature also provides several possible explanations. In Barberis, Shleifer, and Vishny (1998), earnings are generated by a random walk process. However, investors think that shocks to earnings either are negatively correlated (regime 1) or positively correlated (regime 2). Investors update their beliefs based on observed earnings. Regime 1 is motivated by experimental evidence that people overweight their prior (conservatism), while regime 2 is based on experimental evidence that people believe in a law of small numbers (that is, they expect even short samples to reflect the population probabilities). The model generates momentum, long-term reversal, and cross-sectional forecasting power for scaled price ratios (i.e., initial underreaction followed by subsequent overreaction). A related model based on the law of small numbers is given in Rabin (2002). Daniel, Hirshleifer, and Subrahmanyam (1998) provide an alternative model based on overconfidence in private signals plus biased self-attribution. Overconfidence implies initial overreaction of prices to private information, while biased self-attribution implies that new public information supporting the investor's private information leads to even more overconfidence. This, on average, leads to further overreaction. The model thus explains medium-term momentum as well as long-run reversals. A third model is that of Hong and Stein (1999), in which two groups of agents interact to produce the same facts. Private information diffuses slowly among news watchers who therefore generate initial underreaction. Momentum traders in turn generate overreaction.

The above behavioral models base their main assumptions on experimental evidence or simply assume certain trading strategies of investors. They all rely on expectational errors of investors. To provide evidence on whether the pricing anomalies reflect mispricing due to expectational errors, a few papers have studied whether a large part of the profits from value and momentum strategies occur at subsequent earnings announcement dates (the following references are from Daniel, Hirshleifer, and

Teoh [2002]). La Porta, Lakonishok, Shleifer, and Vishny (1997) find that differences in postformation earnings announcement returns account for about one-quarter of the value effect. Jegadeesh and Titman (1993) find that during the first 7 months following the portfolio formation date a similar fraction of the profits from their momentum strategies is due to expectational errors. Jegadeesh (2000) shows that firms that issue seasoned equity do especially poorly around subsequent earnings announcement dates.

While these findings are suggestive of some mispricing, they do not conclusively rule out the possibility that rational stories based on time-varying expected returns could provide most of the explanation. They also do not help sort among different behavioral explanations. Directly analyzing investor expectations would be valuable. Several papers analyze measures of expected returns of equity analysts (see Brav, Lehavy and Michaely [2002] and references to earlier work therein). Careful modeling of analyst incentives is needed to interpret such evidence because analyst forecasts and the forecasts of professional macroeconomic forecasters have been shown to depend on the incentives provided by existing payment schemes (Hong, Kubik, and Solomon, 2000; Lamont, 2002). Consistent with this, Brav et al. (2002) find substantial differences between independent analysts and analysts with investment banking ties. For the independent analysts, they find that expected returns are higher for small stocks, consistent with the small-firm effect being a rational phenomenon driven by risk, whereas book-to-market has little effect on expected returns and momentum affects expected returns with the opposite sign of what a risk-based explanation of the momentum effect would suggest.

Evidence on the beliefs of investors gets around any incentive problems involved in interpreting analyst forecasts, and also does not need to assume that investor beliefs are driven by or correlated with analyst forecasts. This turns out to be important because my evidence based on investor beliefs suggests that investors' expected returns were high during the last part of the stock-market boom in the late 1990s, which is the opposite of what Brav et al. (2002) find for the independent analysts. The data on investor beliefs also allows me to provide evidence regarding some of the behavioral stories told to explain momentum and reversals. Specifically I provide evidence in favor of a version of the law of small numbers (an ingredient in the model of Barberis, Shleifer, and Vishny [1998]) by analyzing the cross section of investor beliefs. I also find support for biased self-attribution (an ingredient in the model of Daniel, Hirshleifer, and Subrahmanyam [1998]). I show that these biases are present and fairly strong even for high-wealth investors and thus that some pricing impact is likely.

As an alternative to evidence based on beliefs, analysis of investment patterns is informative for determining whether return puzzles are due to mispricing or to time-varying risk and expected returns. Grinblatt and Keloharju (2000) confirm that the momentum effect is present in Finland and then analyze whether more sophisticated investors tend to be more momentum oriented or less contrarian in their trades. They find strong support for this, suggesting either that momentum represents mispricing and that this mispricing is better understood by more sophisticated investors, or that high-momentum stocks are riskier in some yet to be identified way and that more sophisticated investors are better able to bear this risk. Cohen, Gompers and Vuolteenaho (2002) find that institutions buy shares from (sell shares to) individuals in response to positive (negative) cash-flow news. Again, this has two possible interpretations. Either institutions attempt to exploit under-reaction of prices to earnings announcements, the post-earnings announcement drift, or stocks with large positive (negative) earnings surprises are by some measure riskier (less risky) than stocks with small earnings surprises, and those who invest through institutions are better able to bear high risk. In the context of both studies, the ideal evidence would be a combination of these facts on trades, with evidence on whether or not the expected returns of institutions and households differed. If they did, that would provide further support for the mispricing interpretation.

Before turning to the evidence on investor beliefs, it is important to emphasize that the dataset I use covers only a short time period, 1998–2002, and mainly focuses on the aggregate stock market. While the large price movements makes this period particularly interesting, I view my results as simply suggestive. My evidence indicates that (1) expected returns were high at the peak of the market; (2) many investors thought the market was overvalued but would not correct quickly; (3) investors' beliefs depend on their own investment experience (a version of the law of small numbers); (4) the dependence of beliefs on own past portfolio performance is asymmetric, consistent with theories of biased self-attribution; and (5) investor beliefs do affect investors' stockholdings. Mainly, the purpose of providing this evidence is to illustrate the value that direct evidence on investor beliefs and actions can have in distinguishing rational theories of pricing anomalies from irrational ones, as well as for testing the assumptions of behavioral models using data for actual investors. While experimental evidence and references to the psychology literature are suggestive, such evidence certainly is more convincing if supplemented with facts about the beliefs and actions of investors.

3. *Investor Beliefs from 1998 to 2002*

3.1 DATA

My study of investor beliefs is based on the household level data under-
lying the Index of Investor Optimism. Since 1996, UBS and Gallup have
conducted monthly telephone surveys of U.S. individual investors (an
international dimension was added starting in 2002). Until February 2003,
the UBS/Gallup data were proprietary. The data can now be purchased
via the Roper Center at the University of Connecticut. UBS granted me
access to the data in late 2002 so that I could undertake this study.

To be included in the survey, investors must have at least $10,000 in
household financial assets defined as "stocks, bonds, or mutual funds in
an investment account, or in a self-directed IRA or 401(k) retirement
account." In 1996, about 1 in 3 households qualified as potential partici-
pants in the survey based on this criteria, increasing to about 4 in 10
households by the start of 2003. Using data from the 1998 Survey of
Consumer Finances, households with $10,000 or more in financial assets
owned more than 99% of stocks owned directly or indirectly by U.S.
households, more than 99% of household financial wealth, and about 95%
of household net worth.

The UBS Index of Investor Optimism is based on qualitative responses
to a series of questions about optimism or pessimism regarding the
investor's own investment and income outlook as well as about the stock
market and other macroeconomic variables. In this study I focus on the
more quantitative questions also included in the survey.

Each month about 1,000 investors are interviewed. The survey is not a
panel, but given the relatively large number of investors interviewed each
month, cohort analysis is possible. Information is collected about a host of
expectational and demographic variables. Four questions about returns
are of particular interest:

1. One-year own past return: "What was the overall percentage rate of
   return you got on your portfolio in the past twelve months?"
2. Expected one-year own return: "What overall rate of return do you
   expect to get on your portfolio in the next twelve months?"
3. Expected one-year market return: "Thinking about the stock market
   more generally, what overall rate of return do you think the stock mar-
   ket will provide investors during the coming twelve months?"
4. Expected ten-year market return: "And, what annual rate of return do you
   think the stock market will provide investors over the next ten years?"

Information on these variables is available for June, September, and
December 1998, and then monthly from February 1999 to December 2002,

with the exception that the expected ten-year market return is not asked about in June 1998 and various months of 2002. For 1998 and 1999, responses of less than 1% (including negative responses) are coded as one category. I set these values to zero.[2] I drop observations of expected market or own portfolio returns and of own past portfolio returns that are below −95% or above 95%.[3] I supplement the answers to these questions with background information on age, years of investing experience ("How long have you been investing in the financial markets?"), financial wealth (categorical), and household income (categorical).

To determine if expectations affect investment decisions, I consider special topical modules with information about portfolio shares (available for September 1998, February 2001, and May 2001), and about Internet stockholdings and expectations (available for March, June, and September 1999, and February, April, June, and July 2000).

Finally, to analyze investors' perceptions about misvaluation of the stock market and whether this is expected be corrected soon, I consider three additional questions:

1. Overvaluation perception: "Do you think the stock market is overvalued/valued about right/undervalued, or are you unsure?"[4]
2. Expected three-month market change: "Over the next three months, do you think the stock market will go up, go down, or remain about the same?"
3. Expected one-year market change: "A year from now, do you think the stock market will be higher than it is now, lower, or about the same?"

The overvaluation perception is available for most months of the survey since June 1998. The expected three-month market change is available from December 1998 to August 2000, and the expected one-year market change is available for September 1998 and from March 2000 onward.

3.2 WERE EXPECTED RETURNS HIGH IN THE LATE 1990s?

The UBS/Gallup data provide an opportunity to address several questions central to behavioral finance as well as traditional finance theory.

2. In the 1998 and 1999 data, fewer than 3% of responses for each of the four variables listed are in the less-than-1% category, suggesting that the lack of negative values for these years is not a substantial problem. To confirm this, I considered the data for January 2000, the first month where zero and negative values are available in noncategorical form. The average expected market return calculated by setting responses of less than 1% to zero differed by less than one-quarter percentage point from the value using the actual responses.
3. This approach was followed in some months in the data I received by Gallup. Also, it is not clear how responses of 100% or above were coded before year 2000.
4. This question is one of the few where respondents explicitly are allowed an "unsure" category.

I start by considering what the data from the recent stock-market experience can teach us about the reasons for predictability of aggregate stock returns. If investors have rational expectations and understand the historical relation between price-dividend ratios and future stock returns, then their expected stock returns should be low during the last years of the market boom when both price-dividend and price-earnings ratios reached historical highs (and appropriate measures of risk should be low at that time). On the contrary, if expected stock returns were high toward the end of the market boom, this would lend support to behavioral stories of overreaction. Prior work on this issue includes Shiller, Kon-Ya, and Tsutsui (1996), who used expectational data for institutional investors in Japan to help analyze expectations at and after the peak of the Nikkei index. Their results are hard to interpret. Japanese institutional investors expected one-year capital gains on the Nikkei index of about 10% at the peak of the market, which seems neither unusually high or not unusually low, but expectations then increased to levels of around 20% after the first year and a half of the Nikkei's decline.

Using the UBS/Gallup data, Figure 1a shows average expected one-year stock-market returns from June 1998 to December 2002. The graph uses survey weights to make results representative of the population. For reference, Figure 1b and c shows the time series for the NASDAQ and NYSE market indices.

The average expected one-year stock-market return increased from an average of 11.8% in 1998 to 15.8% in January 2000, and then declined dramatically to around 6% at the end of 2002. Thus, expected returns were high when the market was at its highest, counter to what the historical statistical relation would have predicted. The correlation at the monthly frequency between the average expected one-year stock-market return and the level of the NYSE is 64.6, and the corresponding correlation with the NASDAQ index is 78.0.[5] An ordinary least squares (OLS) regression (not shown) of the average expected one-year stock return on the NYSE index results in a coefficient of 0.035 with t statistic of 5.9. Using the NASDAQ index, the regression coefficient is 0.0024 with a t statistic of 8.6. Splitting the sample into investors with less than $100,000 in financial assets and investors with $100,000 or more (not shown), the average expected one-year stock returns are about 1% lower throughout the period for those with $100,000 or more in financial assets, but the time pattern is similar for the two groups.

This evidence suggests that, at least for this particular historical experience, prices and expected returns move together positively and thus that

5. These correlations are calculated using the NYSE and NASDAQ indices at the start of the month. Survey interviews are conducted during the first two weeks of the month.
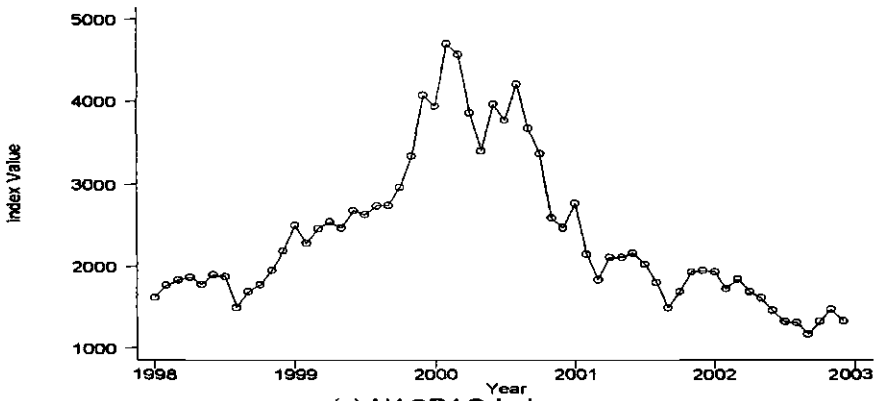
Figure 1 AVERAGE EXPECTED ONE-YEAR AND TEN-YEAR STOCK MARKET
RETURNS, UBS/GALLUP DATA, AND THE LEVEL OF THE NYSE AND
NASDAQ INDICES 1998–2002

(a) Average Expected Stock Market Returns
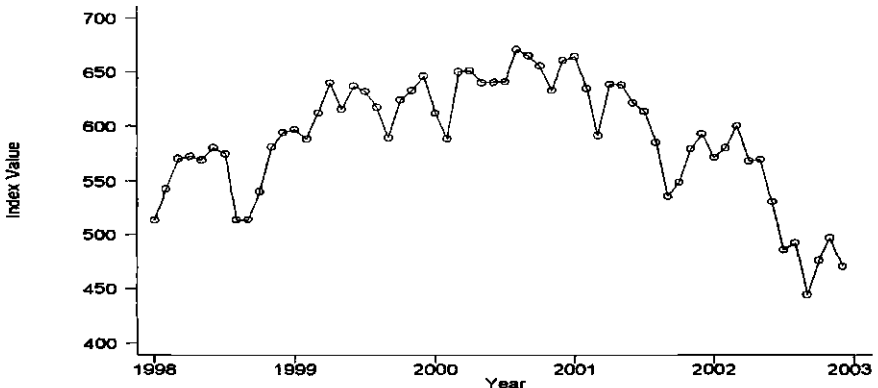
(b) NYSE Index

(c) NASDAQ Index

some amount of overreaction of prices may have been present.[6] The average expected ten-year stock-market returns, also shown in Figure 1, are much more stable over time. Given the small number of ten-year periods for which we have data, and the uncertainty about return predictability at this frequency, stable beliefs at the ten-year horizon seem rational.[7] Graham and Harvey (2001) study the stock-market return expectations of a smaller sample of chief financial officers (CFOs) for six quarters, starting in the second quarter of 2000. They also find that ten-year return expectations are more stable than one-year return expectations and that one-year return expectations move together positively with the realized market return.

3.3 DISAGREEMENT AND NOISE TRADER RISK

Standard finance theory suggests that expected stock-market returns should be similar across investors. While some investors may have private information about the returns on individual stocks, private information about the return on the whole market is less likely. Furthermore, to the extent that trading by better informed investors lead prices to reflect their information, others can learn from prices (Grossman and Stiglitz, 1980), reducing any belief heterogeneity further. In essence, because everyone, by assumption, believes in the same model of how expected stock returns are generated and are equally able to process information, any heterogeneity in beliefs requires both that some investors have private information about market returns and that noise traders or other impediments to learning prevent prices from revealing this information to all investors.

Behavioral finance theory, on the other hand, suggests that differences in expected returns across investors are likely. There is no presumption that all investors use the same model to form expected stock-market returns. Since Miller (1977), several models have considered the possible equilibrium effects of disagreement, emphasizing that in the presence of short-sales constraints, high disagreement leads to high prices and subsequent low returns. See Diether, Malloy, and Scherbina (2002) for references to this literature and for empirical evidence in favor of this theory based on analysts' earnings forecasts and the cross section of stock returns. Less is known about disagreement concerning aggregate stock-market returns and how investor beliefs begin to differ.

6. After completion of the final version of the paper I became aware that Fisher and Statman (2002) use the aggregate UBS/Gallup averages to emphasize this feature of investor beliefs.
   The following facts about investor beliefs exploit the household level UBS/Gallup data and are thus more novel. The household level data also allows one to confirm that the pattern shown in Figure 1(a) is present even for high wealth investors.
7. The average expected ten-year market returns are surprisingly high, however, relative to the average expected one-year market returns. Median ten-year return expectations also exceed median one-year return expectations, although not quite as dramatically.

Figure 2 CROSS-SECTIONAL STANDARD DEVIATION OF ONE-YEAR
EXPECTED STOCK RETURN (DISAGREEMENT), UBS/GALLUP DATA
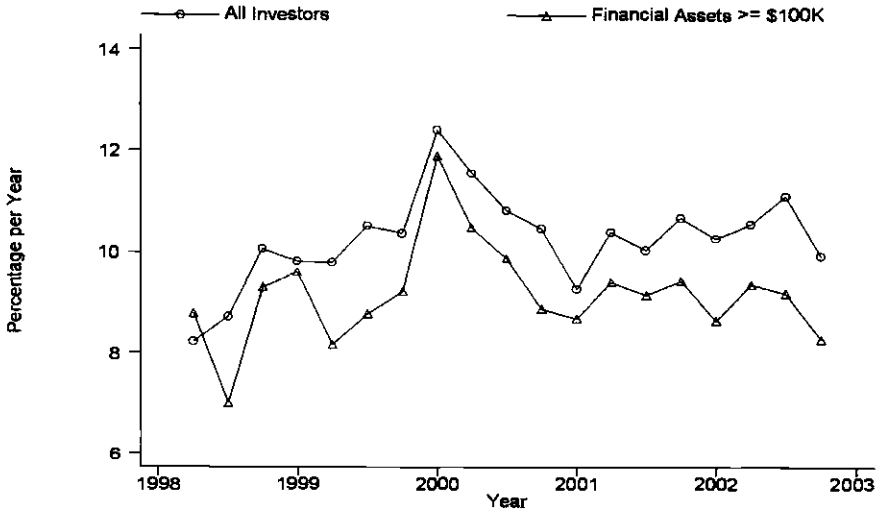
The time series of cross-sectional standard deviations is shown in Figure 2.[8] In a cross section of investors, it is likely that some of the observed differences in responses for expected stock-market returns simply reflect lack of knowledge about stock-market returns, rather than firmly held beliefs on which the investor would place substantial trades. This is confirmed by the fact that the cross-sectional standard deviation of expected one-year stock-market returns is 10.3% for all investors in the sample compared to 9.2% for those with $100,000 or more in financial assets, who have a greater incentive to be informed about returns. These numbers are averages over time of the quarterly cross-sectional standard deviations. Figure 2 therefore shows both the disagreement across all investors as well as the disagreement among those with $100,000 or more in financial assets. Consistent with the findings of Diether, Malloy, and Scherbina (2002) for the cross section of stocks, disagreement was highest just prior to the market decline. It is important to emphasize, however, that finding a positive relation between disagreement and subsequent returns does not necessarily reflect the importance of short sales constraints. Did a significant number of investors in fact want to short the market in the late 1990s? De Long, Shleifer, Summers, and Waldmann (1990) and

8. The figure uses a quarterly data frequency. Results are similar when monthly average returns are subtracted before calculating the quarterly cross-sectional standard deviations. Monthly cross-sectional standard deviations show the same patterns, but are a bit more erratic, likely because a large number of observations is needed to estimate cross-sectional standard deviations accurately.

Shleifer and Vishny (1997) emphasize how noise trader risk can limit arbitrageurs' willingness to take market-stabilizing positions. The risk that misvaluation may worsen will lead rational arbitrageurs to bet less heavily against the mispricing, more so the shorter the horizon of the arbitrageur. The argument of Abreu and Brunnermeier (2002) that it may even be optimal for arbitrageurs to attempt to ride bubbles rather than bet against them only serves to limit arbitrage further.

Shiller has collected expectations data that provide useful information on this issue. His data cover U.S. institutional investors and U.S. individual investors with net worth generally $250,000 or more.[9] While between 50 and 70% thought that U.S. stock prices were overvalued in 1998 and 1999 (calculated excluding those with "do not know" responses), about 70% expected the Dow Jones Industrial index to increase over the next year.

Figure 3 provides related information for the UBS/Gallup data, which covers a broader sample of individual investors and does not include institutions. As shown in Figure 3a, about 50% of investors thought the stock market was overvalued during the last two years of the boom, and typically less than 10% thought it was undervalued. Despite this, Figure 3b shows that only about 20% thought that the market would decline over the next three months/one year. (I use the three-month horizon from December 1998 up to February 2000, and the one-year horizon when it becomes available from March 2000 onward and for September 1998.) As shown in Figure 3c, even among those thinking the market was overvalued in 1999–2000, only about 25% thought it would decline. A similar pattern (not shown) is present for investors with $100,000 or more in financial assets. Along with the evidence on hedge fund holdings from Brunnermeier and Nagel (2002) mentioned above, the expectations data support the idea that noise trader risk matters.

Of course, there is an identification problem here. Short sales constraints could be the reason that few thought the market would go down in the near future. Mankiw, Reis, and Wolfers (2003, this volume) provide an interesting study of disagreement about inflation expectations based on data from the Survey of Consumer Attitudes and Behavior (SCAB). They find the same positive relation between the level of inflation and disagreement about next year's inflation rate that is present for stock-market returns. Since a high level of inflation is unrelated to short sales constraints for stocks, it is possible that a positive relation between the level of a series and the disagreement about the series in the future is a more general feature of expectations formation, for example, because households have less history on which to base their expectations when the series is at an unusually high value.

9. Shiller's analysis is available at http://icf.som.yale.edu/confidence.index (accessed June 3, 2003).
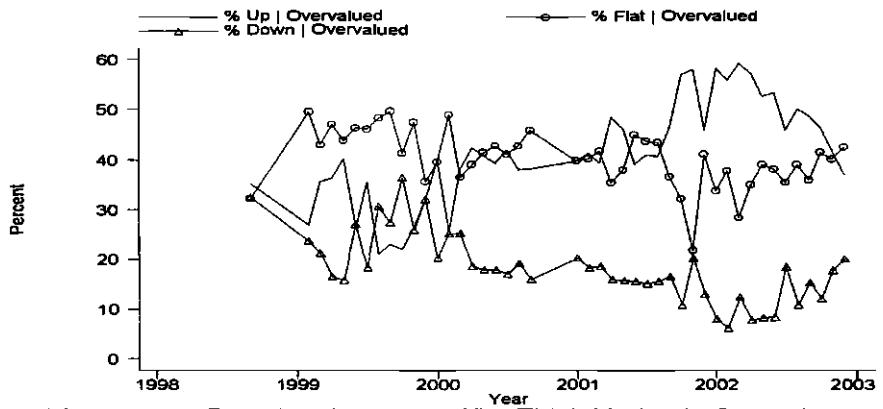
Figure 3 PERCEPTION OF MARKET VALUATION AND EXPECTED
DIRECTION, UBS/GALLUP DATA
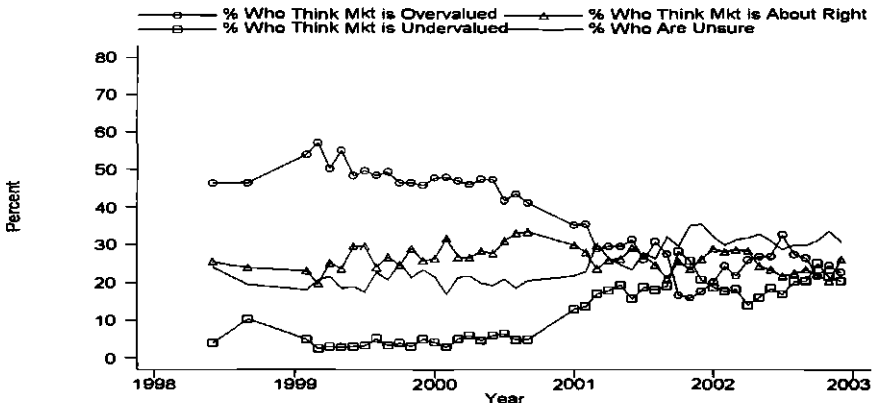
(a) Perception of Market Valuation

(b) Expected Direction, All Investors

(c) Expected Direction, Investors Who Think Market Is Overvalued

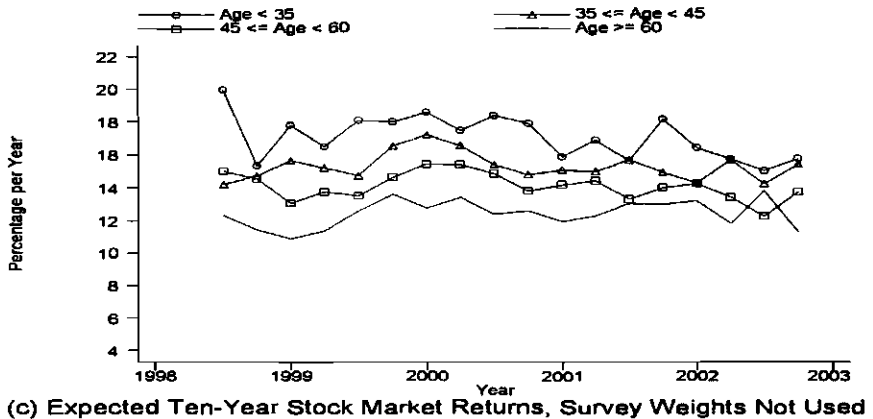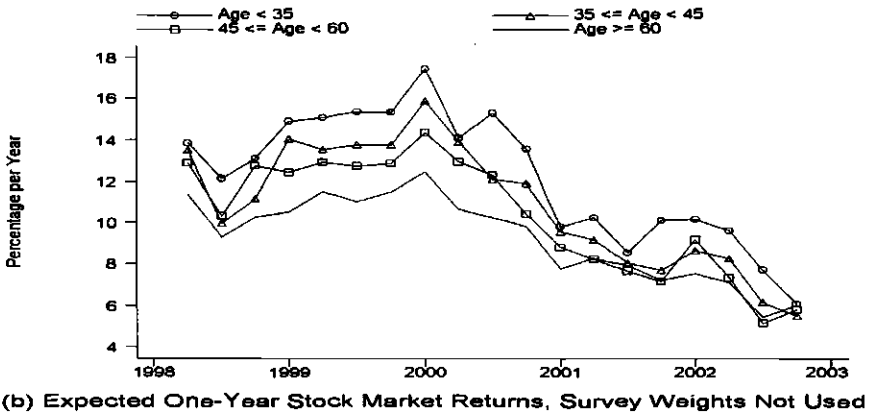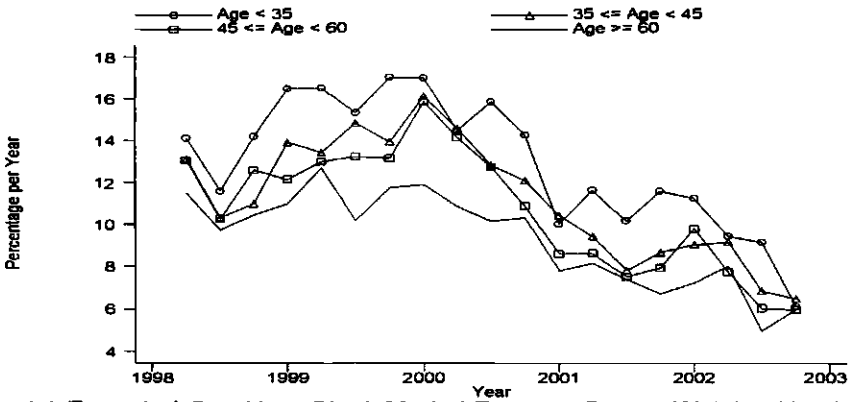3.4 THE DEPENDENCE OF INVESTOR BELIEFS ON THEIR OWN
INVESTMENT EXPERIENCE

A unique feature of the UBS/Gallup data is that they provide a host of information about each individual investor in terms of demographics and past portfolio performance. These data allow further analysis of differences in beliefs. In this section I document that an investor's belief about future stock-market returns depends on the investor's own experience measured by age, years of investment experience, and own past (self-reported) portfolio returns. A behavioral interpretation of these facts is that they provide support for the law of small numbers emphasized by Barberis, Shleifer, and Vishny (1988) and Rabin (2002). (I discuss a possible rational story below.) Investors subject to this bias will expect even short samples to reflect the properties of the parent population and will thus have high expected returns after a period of high realized returns. However, the dependence of expected returns on investor age and experience makes this bias more precise by pointing to what defines the beginning of the (more) relevant small sample—the date the investor started investing in the market. The data also allow a more detailed analysis of how investors' expected market return and expected own portfolio return depend on the past return on their own portfolio to determine whether investors exhibit biased self-attribution, a key ingredient in the model of momentum and reversal of Daniel, Hirshleifer, and Subrahmanyam (1998).

Figure 4a plots the expected one-year stock-market returns of different investor age groups against time.[10] A strong relation between beliefs and age is apparent with young investors expecting substantially higher returns than middle-aged investors, who in turn are more optimistic than older investors. At the peak of the market, young investors, defined as those younger than 35 years, on average expected the market to do about 5 percentage points better over the next year than did older investors, those age 60 years or older. The difference narrows as the market declines. One would expect such narrowing because new data points should be weighted more by young investors who effectively have a shorter data sample. How much the gap narrows during the market downturn depends on whether one uses sample weights or not. Figure 4a uses sample weights, Figure 4b does not. Differences between age groups narrow more consistently when the data are not weighted within age groups.[11] Analyzing medians rather than means leads to similar patterns,

10. To have a reasonably large number of observations per age group per period, the figure shows quarterly average expectations rather than monthly average expectations.
11. It is not clear whether weighting is preferred. For calculating overall average expectations for each time period, weighting is appropriate. When considering the effect of a given investor characteristic on beliefs in a regression context, however, we know that OLS is efficient (other problems aside), and weighting observations by sampling probabilities leads to a less efficient estimator. Therefore, I do not use sample weights in the rest of the analysis.

Figure 4 AVERAGE EXPECTED ONE- AND TEN-YEAR STOCK-MARKET
RETURNS BY INVESTOR AGE, UBS/GALLUP DATA

(a) Expected-One Year Stock Market Returns, Survey Weights Used

(b) Expected One-Year Stock Market Returns, Survey Weights Not Used

(c) Expected Ten-Year Stock Market Returns, Survey Weights Not Used

although the difference in median expectations of young and older households at the peak of the market was around 2% compared to about 5% when focusing on means (except in the first quarter of 2000, where even the median difference increases to 5%). Figure 4c shows large age differences in ten-year expected stock returns as well. Table 1, regression 1, shows that the age effect is statistically significant at the 5% level in almost all

Table 1 DETERMINANTS OF ONE-YEAR STOCK-MARKET RETURN EXPECTATIONS, UBS/GALLUP DATA, OLS REGRESSIONS[1]

Dependent variable: one-year expected stock-market return

Regressor	Regression 1		Regression 2		Regression 3	
	β	t-stat	β	t-stat	β	t-stat
Age*d983	−0.057	−2.03	−0.049	−1.25	−0.012	−0.40
Age*d984	−0.069	−2.27	−0.096	−2.24		
Age*d991	−0.115	−5.79	−0.158	−5.55	−0.020	−0.66
Age*d992	−0.082	−5.06	−0.088	−3.74	0.029	1.22
Age*d993	−0.104	−6.30	−0.098	−4.11	−0.037	−1.91
Age*d994	−0.088	−5.35	−0.084	−3.63	0.024	1.24
Age*d001	−0.125	−7.60	−0.113	−4.74	0.011	0.54
Age*d002	−0.091	−5.73	−0.081	−3.60	−0.001	−0.07
Age*d003	−0.102	−6.03	−0.114	−4.77	−0.006	−0.31
Age*d004	−0.100	−6.00	−0.055	−2.32	−0.029	−1.48
Age*d011	−0.055	−3.40	−0.056	−2.40	0.004	0.21
Age*d012	−0.053	−3.21	−0.053	−2.18	−0.031	−1.62
Age*d013	−0.026	−1.61	−0.020	−0.87	−0.034	−1.85
Age*d014	−0.061	−3.82	−0.066	−2.84	−0.019	−1.03
Age*d021	−0.049	−3.01	0.004	0.19	0.007	0.38
Age*d022	−0.056	−3.43	−0.050	−2.14	−0.033	−1.72
Age*d023	−0.052	−3.22	−0.019	−0.79	−0.038	−2.05
Age*d024	−0.001	−0.09	0.028	1.16	0.018	0.93
Experience*d983					−0.042	−1.06
Experience*d984						
Experience*d991					−0.072	−1.80
Experience*d992					−0.122	−4.05
Experience*d993					−0.042	−1.63
Experience*d994					−0.116	−4.52
Experience*d001					−0.114	−4.63
Experience*d002					−0.089	−3.84
Experience*d003					−0.093	−3.79
Experience*d004					−0.088	−3.57
Experience*d011					−0.103	−4.09
Experience*d012					−0.027	−1.13
Experience*d013					−0.019	−0.78
Experience*d014					−0.064	−2.62
Experience*d021					−0.060	−2.40

Table 1 CONTINUED

Regressor	Regression 1		Regression 2		Regression 3	
	β	t-stat	β	t-stat	β	t-stat
Experience*d022					−0.056	−2.32
Experience*d023					−0.016	−0.64
Experience*d024					−0.012	−0.49
Own past*d983					0.278	8.97
Own past*d984						
Own past*d991					0.301	10.86
Own past*d992					0.335	15.30
Own past*d993					0.374	22.42
Own past*d994					0.419	24.82
Own past*d001					0.369	28.77
Own past*d002					0.267	20.37
Own past*d003					0.311	20.16
Own past*d004					0.264	17.03
Own past*d011					0.117	9.74
Own past*d012					0.110	10.27
Own past*d013					0.114	9.75
Own past*d014					0.101	9.62
Own past*d021					0.173	14.92
Own past*d022					0.179	14.27
Own past*d023					0.167	15.73
Own past*d024					0.106	9.81
d983	13.038	9.42	12.367	5.95	6.618	4.55
d984	15.138	10.29	16.202	7.19		
d991	18.404	18.81	20.762	13.82	9.392	6.45
d992	17.118	21.20	16.834	13.34	7.609	7.15
d993	18.099	22.33	17.509	13.89	8.819	10.29
d994	17.461	21.42	17.213	14.05	6.457	7.42
d001	20.916	25.50	20.533	16.27	8.377	9.63
d002	17.311	21.83	16.398	13.75	9.106	11.05
d003	17.203	20.32	17.321	13.56	8.703	9.96
d004	16.052	19.31	13.357	10.58	10.392	12.42
d011	11.631	14.20	11.160	8.80	9.197	11.65
d012	11.404	13.64	11.081	8.60	10.200	12.61
d013	9.210	11.33	8.689	6.95	9.116	11.84
d014	10.745	13.08	10.601	8.38	9.388	12.10
d021	11.278	13.54	7.901	6.09	9.084	11.38
d022	10.605	12.87	9.496	7.48	9.875	12.17
d023	8.403	10.10	5.639	4.25	8.157	10.23
d024	5.890	7.00	3.764	2.89	5.379	6.53
	N = 39391		N = 17138		N = 31106	
	Adj. R2 = 0.503		Adj. R2 = 0.527		Adj. R2 = 0.590	

1. "Experience" refers to years of investment experience, "Own past" refers to the self-reported return on the investor's portfolio over the past year, 'dYYQ' is a dummy equal to 1 for year YY, quarter Q. Regression 1 and regression 3 is based on all investors, while regression 2 is for those investors with $100,000 or more in financial assets.

quarters. The coefficient on age is allowed to vary over time (by year and quarter), and time dummies are included separately. The negative effect of age on the expected stock-market return is strongest around the peak of the market. Table 1, regression 2, shows that the age effect is as strong for those with financial assets of $100,000 or more as for the full sample.

Additional evidence regarding the effect on beliefs of the stock-market returns observed by the investor him- or herself can be gained by considering the effect of years of investment experience within age groups. If the dependence of beliefs on age in fact is due to investors weighting stock-market returns they have observed more, then after a series of good stock returns, expected returns should be higher for those with low investment experience for a given age than for those with more years of experience. In Figure 5,

Figure 5 AVERAGE EXPECTED ONE-YEAR STOCK-MARKET RETURNS BY INVESTMENT EXPERIENCE WITHIN AGE GROUPS, UBS/GALLUP DATA
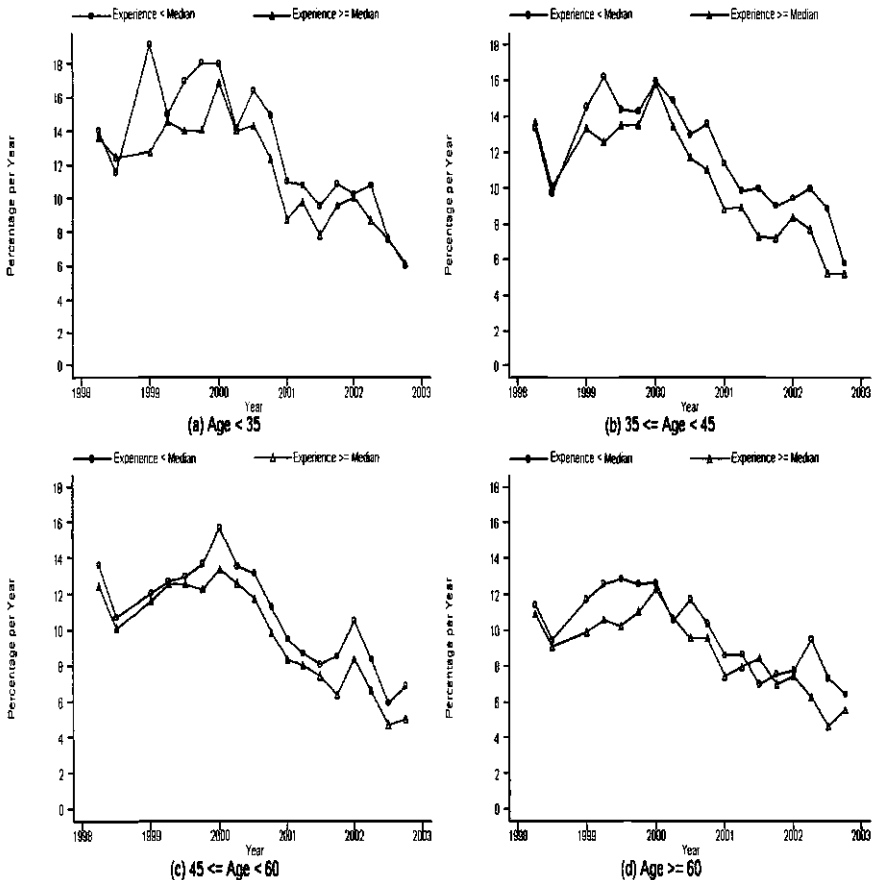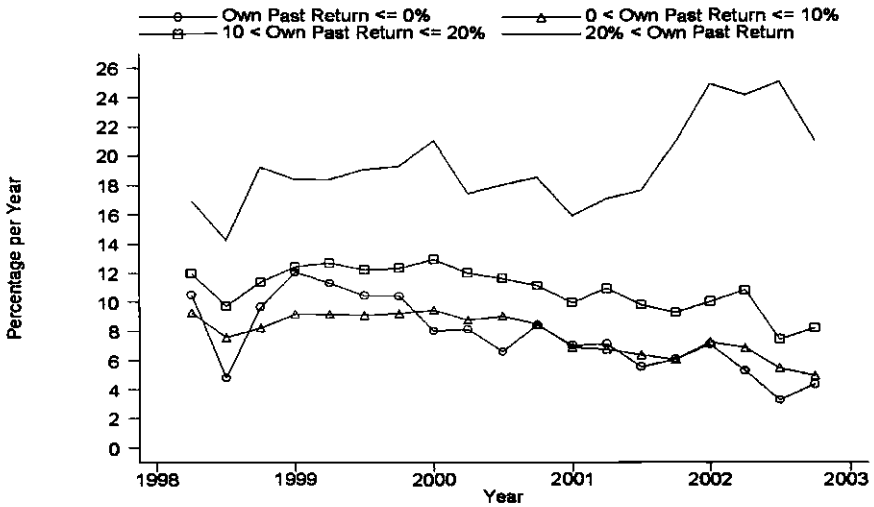
Figure 6 AVERAGE EXPECTED ONE-YEAR STOCK-MARKET RETURNS BY
INVESTOR OWN PAST PORTFOLIO RETURN, UBS/GALLUP DATA

households are split into those with less than and more than median years of
investment experience, within each age group. Over the time period covered
by the UBS sample, the less experienced investors expect about 1 to 2%
higher market returns, with no clear time pattern in this difference.

A final approach to analyzing how observed returns affect beliefs is to
consider whether there is an effect of own past portfolio returns on expecta-
tions about the market return. This is strongly the case. I sort the respon-
dents into four groups based on their reported own portfolio return over the
past year. Figure 6 shows that, compared to those with reported own past
returns between 0 and 10%, those with own past returns between 10 and
20% expected the market return over the next year to be about 3 to 4 per-
centage points higher, with the difference increasing to about 10 percentage
points for those with own past returns above 20%.[12] To determine whether
age, years of investment experience, and own past returns have independ-
ent effects on market expectations, Table 1, regression 3, provides regression
results with all three variables included. Once experience and own past
returns are included, the effect of age largely disappears. Thus, the higher
expected returns of young investors seem to be driven mainly by their
shorter average investment experience and the higher (actual or perceived)
returns on their own portfolios during the stock-market boom. This leads to

12. Some of this effect could be due to measurement error in reported own past portfolio
returns if those who exaggerate their past returns expect high market returns. The asym-
metry of the effect of own past returns on expected market returns (and expected own
returns) documented below is less subject to such problems.

two possible interpretations of the age effects on stock-market return expectations. The first is that investors are rational, but information about past market returns is costly. Then investors may rationally form expectations about future market returns based on their own actual past returns or, if such information is also costly, own perceived past portfolio returns. Because the young report higher own past returns, this would provide a rational explanation of the age effect. The second interpretation is that investors of different ages are equally informed about past market returns but, due to a behavioral bias, nonetheless use their own portfolio returns in forming beliefs about future market returns. The fact that the age effect is equally strong for the wealthiest half of the sample suggests that information costs are unlikely to be driving it and thus that a behavioral story is needed.

The data on inflation expectations from the SCAB can be used to determine whether the age dependence of expectations about stock-market returns generalizes to other aggregate variables.[13] The data also include each respondent's perception of what inflation was for things he or she buys during the past year. This is useful for distinguishing the above two interpretations of the age dependence of stock-market return expectations. If there is an age effect in inflation expectations but no difference in past perceived inflation (or the difference is the opposite of what is needed to explain the age effect in inflation expectations), that would be evidence against the rational costly information explanation of age effects in expectations about aggregate variables.

The SCAB asks respondents whether they think prices will go up, down, or stay the same over the next 12 months. From 1966–1979, respondents who expect price increases are asked for their expected inflation rate as a percentage. From 1980 onward, all respondents are asked for their expected inflation rate. Before the third quarter of 1977, all or some of the percentage responses are categorical. To construct a comparable time series of expected inflation rates, I assume that inflation is normally distributed in the cross section of respondents in each quarter or month (the survey is quarterly up to 1977 and monthly after that). I then estimate the cross-sectional mean and standard deviation based on the percentage of respondents who expect inflation to be below 5%, including those expecting no or negative inflation, and the percentage of respondents who expect inflation to be below 10%.[14] Figure 7 shows the expected inflation

13. See Souleles (2001) and Mankiw, Reis, and Wolfers (2003) for additional description of the data and analysis of heterogeneity in inflation expectations.
14. In principle, it would be more efficient to use all the inflation categories provided rather than only two pieces of information. In practice, a lot of the responses are at inflation rates of 0%, 3%, 5%, 10%, etc. A more sophisticated statistical approach would therefore need either to use a different distribution than the normal distribution or to model the rounding of the responses to popular values.

rate for the next 12 months, by age of respondent, for the period 1966–2001. The expectation plotted for a given year and age group is the cross-sectional average based on responses from all months of that year. Figure 7b shows similar series for 1975–2001 based on expected (annual) inflation over the next five to ten years. The actual inflation rate (based on the consumer price index for all urban consumers) is plotted in Figure 7c. Because survey interviews are spread out over the year, the actual infla-tion rate plotted is the annual inflation rate from July of the current year to July of the following year.

Figure 7 AVERAGE EXPECTED ONE-YEAR AND FIVE-TO-TEN-YEAR INFLATION RATES BY INVESTOR AGE, SURVEY OF CONSUMER ATTITUDES AND BEHAVIOR, AND THE ACTUAL INFLATION RATE FOR THE YEAR, 1966–2001
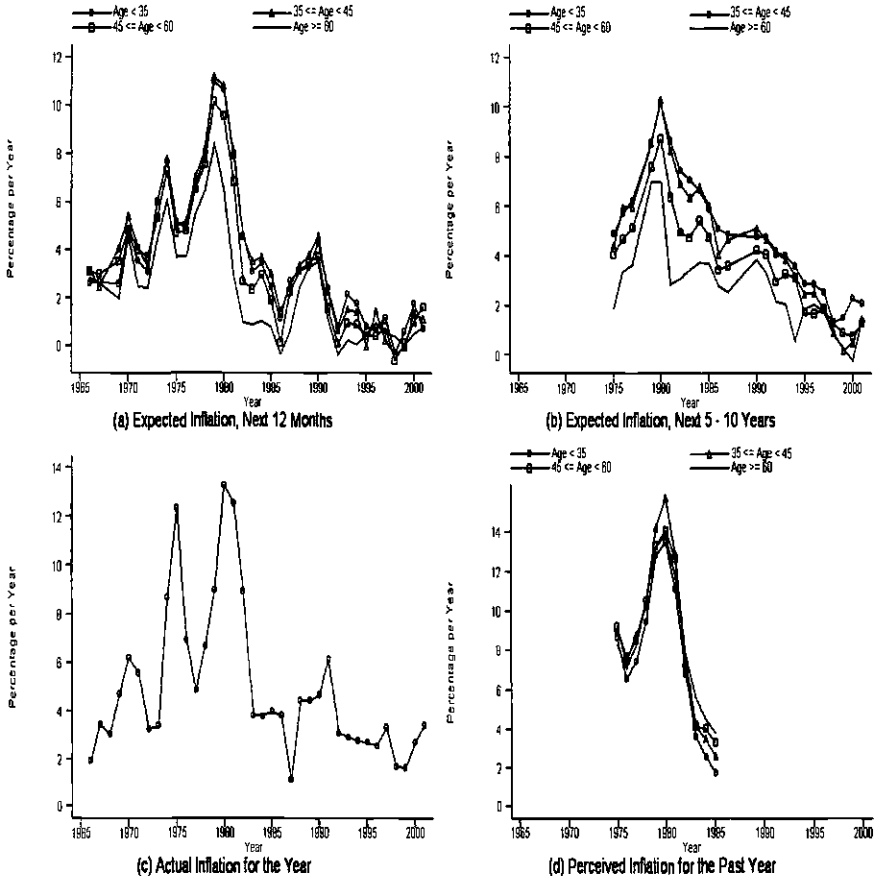
Inflation expectations for the coming year peak in 1979 after a period of high actual inflation rates. In that year, the average expected inflation rate for the next 12 months of those under age 35 exceeded that of those age 60 or older by 2.5 percentage points. The difference widens to 4.9 percentage points in 1981, due to a more dramatic drop in expected inflation for older respondents in 1980 and 1981, and then gradually diminishes during the 1980s. For the years starting with 1980, where the expected percentage inflation rates are available for all households, a simple approach to test whether the age differences are significant is to run a pooled OLS regression of expected inflation rates on year dummies and on age interacted with year dummies, thus allowing the coefficient on age to differ by year (this approach is similar to that used for stock returns in Table 1). The regression, not included in a table for brevity, shows that age is significant at the 5% level in all years from 1980 to 1987. Overall, the age differences in expectations around the period of high inflation are quite similar to the evidence for stock-market return expectations. Figure 7b shows that a strong age pattern is also present in expectations about the level of inflation over the next five to ten years.

In some periods of the survey, households are asked for the inflation rate (for items they buy) over the past 12 months. Quantitative data, consistently defined across years, are available for 1975–1985. Time series for average perceived inflation rates are constructed using the same method as was used for the two forward-looking variables and are illustrated in Figure 7d. Notably, the youngest group generally have the *lowest* perceived inflation, while the ordering of the other three age groups depends on the year in question. A regression (not included in a table) of perceived inflation over the past 12 months on year dummies and on age interacted with year dummies can be run for 1980–1985 (again, percentage responses are available only in a noncategorical form for all respondents from 1980 onward). Perceived inflation is significantly *positively* related to age in each of these six years. Consistent with this, the negative effect of age on expected 12-month inflation is a bit stronger when controlling for perceived past inflation, which itself has a strong positive effect on expected inflation. Thus, the finding that the old expected much lower inflation than the young around 1980 is not driven by different perceptions about inflation over the past year. This again suggests that costs of acquiring information about the inflation level is not likely to explain the age difference in beliefs, consistent with the finding for stock return expectations that the age effect was equally strong for wealthier investors.

Further study of whether the young or the old have more accurate inflation and/or stock-market expectations would be interesting. Whether weighting recent data more is advantageous depends on the persistence

of the series being predicted and thus could be expected to lead to improved accuracy for inflation but possibly decreased accuracy for stock returns. Given the quite short series of expectations on stock returns available in the UBS/Gallup data, I do not pursue the issue of forecast accuracy further.

3.5 BIASED SELF-ATTRIBUTION

Figure 8 illustrates that the dependence of expected one-year stock-market returns on the investor's own past portfolio return is asymmetric. Figure 8a is based on a regression (not included in a table) of market return expectations on age, experience, own past portfolio return, and time dummies.[15] The age and experience effects are allowed to vary by year and quarter, as in Table 1. The effect of own past portfolio return is now allowed to differ depending on whether the return was positive or negative and is allowed to vary by amount of financial wealth. The regression is estimated using data only from 2000–2002 where responses (for market and own return expectations and for own past portfolio returns) of less than 1% are not combined into one category. Figure 8a plots the predicted effect of own past portfolio returns on expected one-year market return. For those with financial wealth less than $100,000, an own past portfolio returns of 25% increases the expected market return by 10.4%, while an own past portfolio returns of −25% leads to an increase of 1.6%.[16] Thus, while positive own past portfolio returns leads to higher expected market returns, negative own past portfolio returns have a quite small and positive effect on expected market returns. The effect of positive own past returns is weaker for wealthier investors, but a 25% own past portfolio return still leads to an increase in the expected market return of as much as 6.7%, even for those with financial wealth of $500,000 or more. The difference to the lowest wealth group is significant at the 1% level. A diminished effect of positive own past returns was also found for higher income or higher education groups.

Several robustness checks are needed to determine if these findings reflect biased self-attribution. If they do, then the asymmetry results should be stronger for expected own portfolio returns than for expected market returns because the investor presumably is more likely to think that high own past portfolio returns are indicative of high future returns on his or her own portfolio than on the stock market as a whole. Figure 8b shows that this is indeed the case. The effect of positive own past returns

---

15. Note that by including time dummies, the effect of own past return on expected market returns is identified based on cross-sectional differences in own past returns, not based on time variation in own past returns.
16. Both these effects are significant at the 1% level. About 24% of own past portfolio returns for 2000–2002 are negative.

Figure 8 ESTIMATED EFFECT OF OWN PAST PORTFOLIO RETURN ON EXPECTED ONE-YEAR STOCK-MARKET RETURN AND EXPECTED ONE-YEAR OWN PORTFOLIO RETURN, UBS/GALLUP DATA
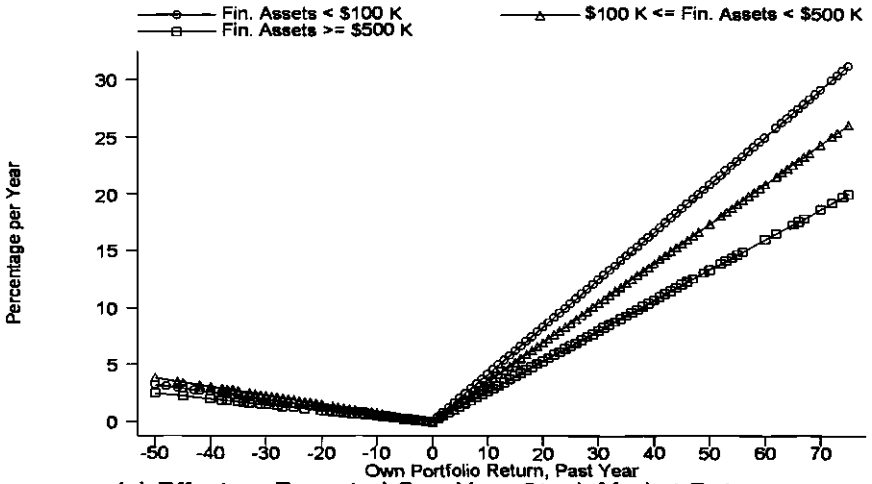
(a) Effect on Expected One-Year Stock Market Return
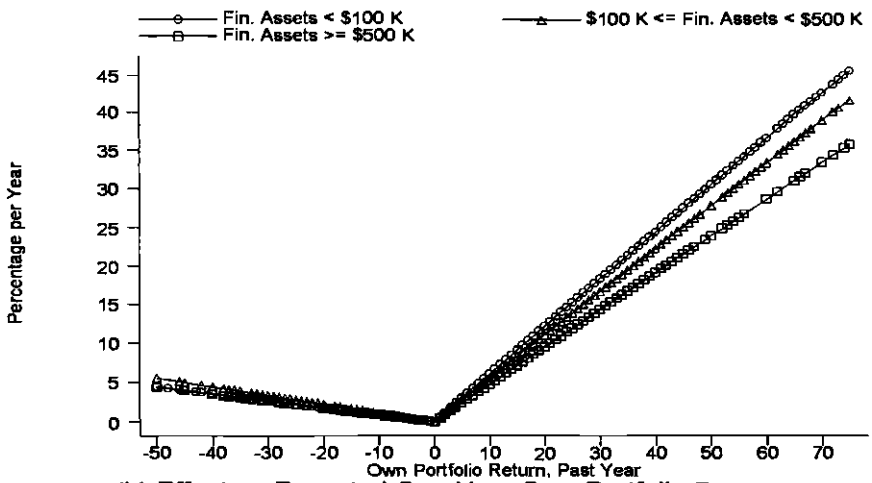
(b) Effect on Expected One-Year Own Portfolio Return

on expected own portfolio returns over the next year is about 50% larger than the effect on the expected one-year market return. When focusing on the expected own portfolio returns, a potentially important concern is whether the positive slope in the region of negative past own returns could be due to lack of controls for portfolio choice. The question in the survey refers to the investor's entire portfolio of financial assets, not just the return on stockholdings. To get a substantial negative return, an investor likely had invested a lot in stocks that are likely to have a higher expected return than other assets. Investors' portfolio shares for each of the categories "stocks, stock mutual funds," "bonds, bond mutual funds," "cash, CDs, money market funds," and "real estate investments" are available for September 1998 and February and May 2001. Using data for these three months, controlling for portfolio shares has only a negligible impact on the effect of own past portfolio returns on expected own portfolio returns (or expected market returns). Another concern may be that an own past return of zero may not be the most reasonable comparison point against which to evaluate whether own past portfolio performance is high or low. Allowing for a kink at an own past portfolio returns of 10% leads to strong positive effects for own past returns above 10% and a flat relation for own past returns below 10%.

Overall the results support the assumption of biased self-attribution made by Daniel, Hirshleifer and Subrahmanyam (1998). The finding that the effect diminishes in investor wealth or other measures of investor sophistication does suggest, however, that more work is needed to understand why some investors are more subject to this bias than others.

3.6 DO BELIEFS AFFECT ACTIONS?

The above results regarding investor beliefs would be of little interest if expectations reported to the survey are not correlated with investor choices. For the three months for which portfolio shares for broad invest-ment categories are available, it is possible to determine whether investors with higher expected stock returns did in fact have higher equity portfolio shares.[17] Table 2 shows that this is strongly the case in the region of expected market returns up to 20%. This range covers over 95% of the investors used in the regression. As another piece of information about the link between expectations and portfolio holdings, Table 3 turns

17. Since the observed equity portfolio shares are in the range from 0 to 100%, I estimate the relationship using a two-sided Tobit model. The estimation also controls for age, invest-ment experience, financial assets, education, and income because these factors may affect portfolio choice directly and, as discussed earlier, are correlated with expectations.

Table 2   EFFECT OF EXPECTATIONS ON STOCK HOLDINGS, 1998
(SEPTEMBER), 2001 (FEBRUARY, MAY), UBS/GALLUP DATA, TOBIT
REGRESSIONS

Dependent Variable: Percentage of Portfolio Held in Stocks

Regressor	β	t-statistic
Expected market return dummies (omitted = d $(E(r_M) \leq 0)$)		
d$(0 < E(r_M) \leq 5)$	2.729	0.84
d$(5 < E(r_M) \leq 10)$	4.648	1.51
d$(10 < E(r_M) \leq 15)$	10.164	2.88
d$(15 < E(r_M) \leq 20)$	10.191	2.30
d$(E(r_M) \geq 20)$	5.016	1.14
Time dummies (omitted = d9809)		
d0102	5.699	2.93
d0105	−7.883	−4.20
Age dummies (omitted = d(age < 30))		
d$(30 \leq$ age $< 40)$	0.517	0.15
d$(40 \leq$ age $< 50)$	−6.009	−1.68
d$(50 \leq$ age $< 60)$	−7.411	−1.98
d$(60 \leq$ age $< 70)$	−14.847	−3.48
d(age ≥ 70)	−22.754	−4.82
Experience dummies (omitted = d(experience ≤ 5 Years))		
d$(5 <$ experience $\leq 10)$	−2.195	−0.91
d$(10 <$ experience $\leq 15)$	1.354	0.49
d$(15 <$ experience $\leq 25)$	−5.686	−2.03
d(experience > 25)	−0.374	−0.11
Financial asset dummy (omitted = d(financial assets < 100 K))		
d(financial assets $\geq$ 100 K)	3.793	2.06
Education dummies (omitted = d($\leq$ high school graduate))		
d(some college/technical college)	4.444	1.63
d(college graduate)	10.931	4.05
d(> college graduate)	10.146	3.64
Income dummies (omitted = d(income < 40 K))		
d(40 K $\leq$ income < 50 K)	−1.627	−0.45
d(50 K $\leq$ income < 60 K)	2.839	0.83
d(60 K $\leq$ income < 75 K)	−3.057	−0.94
d(75 K $\leq$ income < 100 K)	1.171	0.38
d(income $\geq$ 100 K)	−0.875	−0.28
Constant	51.601	9.64
N	2026	
N censored at 0	123	
N censored at 100	221	

to the relation between Internet stockholdings and expectations about
Internet stock returns. Information about Internet stockholdings are
included in six months of the survey spread out over 1999 and 2000.
Investors who expected Internet stocks to have much higher returns
than the stock market on average held as much as 25% more of their

Table 3   INTERNET STOCK HOLDINGS: TOBIT REGRESSION FOR
PERCENTAGE OF PORTFOLIO HELD IN INTERNET STOCKS, 1999 (MARCH,
JUNE, SEPTEMBER) AND 2000 (FEBRUARY, APRIL, JULY), UBS/GALLUP DATA

Dependent Variable: Percentage of Portfolio Held in Internet Stocks[1]

Regressor	β	t-stat	β	t-stat
Expected Internet stock return dummies (omitted = d(much higher))				
d(somewhat higher)	−7.787	−4.65	−8.049	−3.39
d(about same)	−18.039	−8.40	−18.462	−6.14
d(somewhat/much lower)	−25.151	−8.81	−22.560	−5.49
Perceived Internet stock risk dummies (omitted = d(much more risky))				
d(somewhat more risky)	4.674	2.89	4.290	1.83
d(about same risk)	4.955	2.31	4.723	1.58
d(somewhat less/much less risky)	7.276	2.27	9.325	2.19
Expected one-year stock market return (omitted = d($0 \leq E(r_M) < 5$))				
d($0 < E(r_M) \leq 5$)	−3.751	−0.80	4.961	0.69
d($5 < E(r_M) \leq 10$)	−0.209	−0.05	5.135	0.79
d($10 < E(r_M) \leq 15$)	3.835	0.91	8.672	1.32
d($15 < E(r_M) \leq 20$)	7.147	1.60	11.896	1.71
d($E(r_M) \geq 20$)	8.337	1.84	12.427	1.73
Age dummies (omitted = d(age < 30))				
d($30 \leq$ age < 40)	−8.730	−3.23	−9.565	−2.51
d($40 \leq$ age < 50)	−17.359	−6.17	−15.247	−3.85
d($50 \leq$ age < 60)	−15.801	−5.21	−9.877	−2.30
d($60 \leq$ age < 70)	−21.640	−5.85	−17.218	−3.18
d(age $\geq$ 70)	−23.780	−5.33	−13.569	−2.10
Experience dummies (omitted = d(experience $\leq$ 5 years))				
d(5 < experience $\leq$ 10)	−6.816	−3.31	−7.961	−2.61
d(10 < experience $\leq$ 15)	−4.259	−1.80	−1.684	−0.49
d(15 < experience $\leq$ 25)	−2.669	−1.09	−2.770	−0.77
d(experience > 25)	−0.824	−0.27	3.995	0.91
Education dummies (omitted = d($\leq$ high school graduate))				
d(some college/technical college)	5.051	1.79	9.146	2.18
d(college graduate)	10.743	4.04	10.532	2.61
d(> college graduate)	11.755	4.33	8.810	2.13
Financial wealth dummy (omitted = d(financial assets < 100 K))				
d(financial assets $\geq$ 100 K)	9.617	5.94	6.039	2.61
Income dummies (omitted = d(income < 40 K))				
d(40 K $\leq$ income < 50 K)	6.829	1.83	5.155	0.93
d(50 K $\leq$ income < 60 K)	6.296	1.78	1.381	0.26
d(60 K $\leq$ income < 75 K)	10.491	3.14	6.825	1.35
d(75 K $\leq$ income < 100 K)	9.429	2.85	4.905	0.97
d(income $\geq$ 100 K)	19.768	6.08	13.134	2.63
Internet use dummies (omitted = d(never gets on Internet))				
d(gets on Internet, never purchased online)			8.209	2.57
d(gets on Internet, purchased online)			20.500	6.40
Constant	−32.806	−5.40	−44.982	−4.85
N/N cens. at 0/N cens. at 100	4164/1076/9		2084/445/2	

1. Regressions include time dummies. The table omits these for brevity.

portfolio in Internet stocks than those expecting Internet stock returns to be somewhat lower or much lower than the return on other stocks.[18] Lower perceived risk of Internet stocks relative to the risk of the market similarly has the expected positive effect on Internet stockholdings. Overall, the portfolio data thus show that investor actions are linked to their beliefs.

4. The Value of Correlating Irrational Actions with Wealth

In this section, I turn to the other main strand of the behavioral finance literature, which has focused on types of investor behavior that are inconsistent with the recommendations of standard finance models. Part of this literature is separte from the literature on pricing anomalies, while the pricing impact of other of these behaviors has been studied and linked to the pricing puzzles. Of course, even the behaviors listed below that may not have significant price impact are still important because such behaviors could have large effects on the utility of investors who act in supposedly irrational ways.

Investor behaviors that contradict the predictions of traditional finance models have been surveyed elsewhere (among others, see Barberis and Thaler [2003] and Daniel, Hirshleifer, and Teoh, [2002]). What I would like to focus on here is whether a given type of seemingly irrational behavior diminishes with investor wealth or with other measures of investor sophistication.

If the frequency or intensity of such behaviors diminishes substantially with wealth or sophistication, then two possibilities arise. The first possibility is that these behaviors are driven by information costs that likely have a large fixed component (once you understand diversification, you can apply your insights without cost to a larger portfolio). If they are driven in this way, then investors may be acting rationally given the costs they face. To confirm this, we would need to establish that the required information costs are not implausible, and to argue that the behavior exhibited is a reasonable response to lack of information. The latter is more likely to be satisfied in cases where the behavior involves too little action or too simple an action (e.g., lack of investment in some securities, lack of reallocation) than in cases where the behavior involves too much action (excessive trading). If information costs are to blame for seemingly irrational investor behavior, the policy recommendation would be increased investor education, especially for low-wealth and low-sophistication

18. It is not clear from the question asked whether the Internet portfolio share is the share of Internet stocks in the investor's equity portfolio or in his or her total financial asset portfolio.

investors who may not choose to become informed at their own cost. Of course, from an efficiency perspective, this would be the policy recommendation only if such education has positive externalities (i.e., that one educated investor can help another improve his or her choices) or can be provided more cheaply than the cost at which investors could have acquired the information on their own.

If behaviors that look irrational based on traditional finance theory diminish with wealth, a second interpretation is that psychological biases differ across individuals. Additional analysis of such cases would improve our understanding of the more fundamental determinants of the biases in beliefs and behavior and such correlations would need to be accounted for in models and calibrations of the likely pricing impact of such biases. To draw a parallel to the traditional finance literature, absolute risk aversion is typically thought and estimated to be decreasing in wealth. This does not mean that risk aversion is not a fundamental element of preferences or that risk-averse behavior is due to information costs, but it does mean that it is crucial for modeling and calibration whether or not this wealth dependency is accounted for. Correlating investor choices with other investor characteristics would also be helpful in this context.

Conversely, if a given irrational action remains equally frequent for high-wealth investors, then it is unlikely to be driven by information costs and is likely to have substantial impact on equilibrium prices. The behavioral finance literature is still not at the point where calibration of theoretical general equilibrium asset pricing models is done to determine the magnitude of the effects of nonstandard types of behavior on asset prices. I hope that work will progress to this stage as more information about investor expectations and actions becomes available and we get increasingly accurate estimates of the strength of the various biases.

Of course, it is important when considering the relation between biases and wealth to determine whether reverse causality could be driving the results. Some of the biases listed below are known to generate poor returns and thus low wealth. This means that one has to consider investors with vastly different wealth for comparisons to be robust to endogeneity issues; look at more exogenous measures of investor wealth and sophistication, such as labor income or education; or compare the behavior of different investor types, as in the earlier mentioned studies of trading behavior of households versus institutions, households versus foreign investors, and households versus hedge funds. An even better approach would be to consider the effects of exogenously provided information on investor behavior (examples of such studies are given in the next section).

A partial list of investor behavior not in accordance with standard finance theory includes those discussed in the following subsections. Some of these facts were documented by researchers in the rational camp. I include them to provide a more complete picture.

4.1 THE DISPOSITION EFFECT

This refers to a tendency of investors to delay selling investments on which they have incurred losses in the hope that they will recover their losses. This has been documented in the stock trades of individuals in the United States (Shefrin and Statman, 1985, and Odean, 1998), in the stock trades of Israeli individuals (Shapira and Venezia, 2001), in the stock trades of Finnish individuals and institutions (Grinblatt and Keloharju, 2001b), in the option exercise patterns of employees in the United States (Heath, Huddart, and Lang, 1999), and in sales patterns for homes (Genesove and Mayer, 2001).

The leading argument against the disposition effect being a rational phenomenon is that winners sold by individual investors subsequently outperform losers not sold (Odean, 1998). Behavioral researchers typically attribute the disposition effect to prospect theory (Kahneman and Tversky, 1979), with a reference price equal to the investor's purchase price. Based on experimental evidence, Kahneman and Tversky argued that utility should be defined not over wealth or consumption but over gains and losses, and that people are risk averse in the region of gains, but risk loving in the region of losses. Such preferences can induce the disposition effect because investors become risk loving in a security's payoff after a loss but not after a gain. An alternative behavioral story is a mistaken belief in mean-reversion. Odean (1999) argues against this by showing that the stocks purchased by individuals tend to be past winners. Grinblatt and Han (2002) consider the general equilibrium implications of the disposition effect. They construct a model where the momentum effect is driven by some investors exhibiting the disposition effect in their trading behavior. Goetzmann and Massa (2003) provide evidence that a disposition effect factor is priced in the cross section of daily stock returns.

Dhar and Zhu (2002) provide evidence about how the strength of the disposition effect depends on investor sophistication. Using U.S. data from a discount brokerage firm, they find that the disposition effect is only about half as strong for high-income, retired investors as for low-income investors working in nonprofessional jobs. Controlling for income and occupation, they also find a significant weakening of the disposition effect in investor age and in investor trading experience. Twenty percent of investors in their sample exhibit no disposition effect or exhibit a reverse disposition effect. Brown, Chapel, da Silva Rosa, and Walter

(2002) analyze the disposition effect using Australian data and find that the effect is weaker but still significant for investors taking large trading positions compared to others. Shapira and Venezia (2001) compare the disposition effect for accounts of independent investors and accounts of investors who have delegated portfolio management to a professional portfolio manager. The trades decided on by the investment professionals exhibit a weaker but still substantial disposition effect. In their study of the disposition effect in real estate transactions, Genesove and Mayer (2001) find that the disposition effect is twice as strong for owner-occupants as for (likely wealthier/more sophisticated) real estate investors. The evidence overall suggests that the disposition effect weakens substantially with investor wealth.

4.2 LIMITED DIVERSIFICATION OF STOCK PORTFOLIOS

French and Poterba (1991) emphasize that investors concentrate the vast majority of their equity portfolios in domestic stocks (the home bias puzzle). Coval and Moskowitz (1999) document a local equity preference in domestic portfolios of U.S. investment managers (home bias as home). Huberman (2001) reports a similar local stock preference by showing that the amount invested in local regional Bell phone companies far exceeds the amount invested in out-of-state regional Bell phone companies in most states. Grinblatt and Keloharju (2001a) report that home bias at home is also present among Finnish stockholders, while Massa and Simonov (2003) document it for Swedish investors. Benartzi (2001) analyzes stockholdings in employer stock and shows that employees invest 23% of their discretionary retirement plan contributions in company stock. Blume, Crockett and Friend (1974) and many subsequent papers have emphasized the low number of stocks held by many investors.

Coval and Moskowitz (2001) argue that informational advantages may motivate local holdings in their sample because fund managers earn an extra 2.67% per year from their local investments relative to their nonlocal investments. Benartzi (2001) shows that this is not the case for own company stockholdings. While employees tend to allocate more to company stock in firms that have done well in the past, retirement plans with higher discretionary contributions to own company stock do not outperform other plans. Benartzi also provides survey evidence that employees on average think high past returns will continue in the future and that only 16.4% of the respondents believe company stock is riskier than the overall stock market, measured by the likelihood of either investment losing half its value over the next five years.

The UBS/Gallup data provide a new opportunity to analyze the relationship between familiarity, expectations, and investments. For three

months in 1999, the survey contains information about both Internet stockholdings, Internet stock return expectations, and Internet use. Table 4 shows that of those reporting that they use the Internet and have purchased something online, about 69% expected higher returns on Internet stocks than on other stocks, compared to 40% for those who did not use the Internet. Internet users also perceived Internet stocks to be riskier. This finding could be consistent with an information story where Internet use leads to cheaper or free information about Internet stocks because Internet stocks probably were riskier and therefore may have had higher expected returns than other stocks in 1999. The second regression in Table 3 shows, however, that even controlling for expected returns and risk (and a host of other variables), Internet use has a strong effect on Internet stockholdings, with those getting on the Internet and having purchased something online investing about 20% more in Internet stocks than those who do not use the Internet. This may be suggestive of an attention effect, where investors simply do not know about all stocks and invest in those stocks they—partly by accident—become aware of. Barber and Odean (2002) and Frieder and Subrahmanyam (2002) find evidence of an attention effect in the stock purchases of individual investors. If information is costly, the attention effect could be rational, although one could argue that any deviations of an investor's equity portfolio from the market portfolio is irrational.

I turn now to the relation between diversification and wealth/sophistication. Table 5 documents a relationship between home bias and investor income. The numbers are from the New York Stock Exchange (2000) and are based on a survey of 4842 investors in early 1999 (see Investment Company of America and the Securities Industry Association [1999]). The home bias is seen to diminish quite strongly with investor income, especially when it comes to directly held, non-U.S. stock or holdings of foreign stock through equity mutual funds in retirement accounts. Addressing home bias as home, Grinblatt and Keloharju (2001) show that the preference of Finnish investors for local stocks or for stocks with a chief executive officer (CEO) of their own cultural origin diminishes in investor sophistication as measured by the number of stocks held by the investors. Massa and Simonov (2003) find that the local stock preference of Swedish investors is driven purely by low-wealth investors.

Table 6 uses data from the 1998 and 2001 Survey of Consumer Finances to document that the number of stocks held in directly held equity portfolios is strongly increasing in the wealth of the household. While households with net worth below $100,000 hold on average just a couple of stocks in directly held stock portfolios (conditional on having any directly held stock), the average number of stocks increases to about 14 for

Table 4 PERCEPTIONS ABOUT RISK AND RETURN OF INTERNET STOCKS, BY INTERNET USE, 1999 (MARCH, JUNE, SEPTEMBER), UBS/GALLUP DATA

Overall, compared to investing in other common stocks, do you think that investing in selected Internet companies is . . .

	Much more risky	Somewhat more risky	About the same risk	Somewhat/ much less risky	Do not know	Number of observations
Never gets on Internet	21.22	37.08	24.65	9.65	7.40	933
Gets on Internet, never purchased online	26.52	37.55	25.81	8.10	2.02	988
Gets on Internet, purchased online	36.37	38.45	19.04	5.32	0.81	1108
Overall	28.49	37.74	22.98	7.56	3.24	3029

Again, compared to the return one can get from investing in other common stocks, do you think that the percentage return from investing in selected Internet companies is . . .

	Much higher	Somewhat higher	About the same	Somewhat/ much lower	Do not know	Number of observations
Never gets on Internet	12.33	28.08	33.90	11.58	15.11	933
Gets on Internet, never purchased online	17.61	36.34	29.05	10.73	6.28	988
Gets on Internet, purchased online	28.79	39.80	19.22	8.66	3.52	1108
Overall	20.07	35.06	26.64	10.23	7.99	3029

Table 5 EFFECT OF HOUSEHOLD INCOME ON HOME BIAS

Income	<$15 K	$15–25 K	$25–50 K	$50–75 K	$75–100 K	$100–250 K	> $250 K
Percentage of shareowners who have directly held, non-U.S. stock or an equity mutual fund holding non-U.S. stock inside a retirement account	11.9	11.5	27.7	35.9	38.1	42.3	38.5
Percentage of shareowners who have equity mutual funds holding non-U.S. equities, outside retirement accounts	26.2	23.3	21.8	29.3	30.0	34.8	39.2

Table 6 EFFECT OF HOUSEHOLD NET WORTH ON STOCK-MARKET PARTICIPATION AND DIVERSIFICATION OF DIRECTLY HELD EQUITY

Net worth	< $10 K	$10–50 K	$50–100 K	$100–250 K	$250–1 M	> $1 M	All
Percentage who hold stocks							
1998	16.2	42.3	46.9	59.7	81.0	91.9	48.9
2001	18.0	38.8	48.4	62.0	79.3	92.9	51.9
Mean (median) number of directly held stock conditional on owning stock directly							
1998	1.5 (1)	2.4 (1)	2.5 (2)	3.1 (2)	5.3 (3)	14.9 (8)	5.7 (2)
2001	1.7 (1)	1.9 (1)	2.6 (1)	3.0 (2)	6.0 (0)	13.3 (8)	6.3 (3)
Mean percentage of stocks held directly							
1998	12.9	12.1	15.9	20.3	25.7	35.9	20.6
2001	20.3	14.6	12.5	15.7	21.8	33.1	19.7

Source: Calculated using data from the 1998 and 2001 Survey of Consumer Finances (using survey weights).

households with a net worth of $1 million or more. In an earlier study (Vissing-Jorgensen, 1999), I argue that the percentage of equity owned by very poorly diversified investors in terms of the number of stocks is quite small. Goetzmann and Kumar (2001) analyze equity portfolio diversification using investor accounts at a particular brokerage firm and conclude that the majority of such investors are very poorly diversified. While analysis of brokerage accounts can be useful (for example, for analyzing the disposition effect) it is less compelling for analyzing diversification. Investors may use multiple brokers or hold most of their equity portfolios in mutual funds. Overall, investors with larger amounts of wealth or income, and thus greater incentives to become informed, hold better diversified portfolios than others.

4.3 LIMITED ASSET MARKET PARTICIPATION

A more extreme example of poor diversification is limited asset market participation. Many households have zero holdings of certain asset classes. The most well known is limited participation in stock markets. Other examples include holding no bonds or no investment real estate. The papers in the volume edited by Guiso, Haliassos, and Jappelli (2002) provide evidence that limited participation in markets for risky assets is prevalent in many countries. In Vissing-Jorgensen (2002) and in Section 5 below, I consider the role that costs of stock-market participation may play in providing a rational explanation for this. Heaton and Lucas (1999), Polkovnichenko (2001), I (Vissing-Jorgensen, 1998), and others have considered the equilibrium impact of limited participation on the equity premium. The consensus is that, in standard models where the equity premium is small with full participation, limited participation on its own will have some but not a dramatic effect on the equilibrium equity premium.

Table 6 illustrates that stock-market participation is strongly increasing in investor wealth and income. I return to this fact in Section 5.

4.4 NAÏVE DIVERSIFICATION OF RETIREMENT ACCOUNT CONTRIBUTIONS

Benartzi and Thaler (2001) document that the relative number of equity-type investment options offered in 401(k) plans affects the mean allocation to equities of plan participants. Investors in plans that are in the highest third in terms of percentage of equity-type investment options invest 64% on average in stocks, compared to 49% for investors in plans in the bottom third in terms of equity-type options. Experimental evidence roughly confirms these magnitudes and also suggests that this is driven by some investors choosing portfolio shares of $1/n$ for each plan

option. A $1/n$ rule seems like a reasonable response to diversifying for an investor who understands the basic idea of diversification but not the exact differences between asset classes. Correlating this type of behavior with income or wealth would be informative for determining if a simple information explanation is likely.

4.5 STATUS QUO BIAS IN RETIREMENT ACCOUNT ALLOCATIONS

Ameriks and Zeldes (2001) analyze a ten-year panel of TIAA-CREF participants. Consistent with earlier findings of Samuelson and Zeckhauser (1988), they find that both changes in flow allocations and reallocation of accumulated assets are rare: 47% of individuals made no changes in flow allocations over a ten-year period; 73% made no changes in the allocation of accumulated assets. Ameriks and Zeldes suggest a rational explanation, namely, that individuals may face a nonmonetary fixed cost per transaction. If so, we would expect the status quo bias to diminish with the dollar amount invested, and thus with employee salary. The bias would also be expected to diminish with age or years of employment because a certain amount of free information about the value of reallocating arrives over time from interaction with colleagues and friends. Table 7 shows that these predictions are borne out in the data. The table is from Agnew, Balduzzi, and Sunden (2003), who analyze data from a large 401(k) plan. They find that employees with higher income and older employees place substantially more trades (changes in flow contributions and allocation of existing assets) and have a higher retirement asset turnover than younger employees and employees with lower income.

4.6 EXCESSIVE TRADING

In sharp contrast to the trading behavior in retirement plans, some investors trading though brokers or online trade frequently and on average

Table 7   EFFECT OF HOUSEHOLD SALARY AND AGE ON STATUS QUO BIAS IN RETIREMENT PLANS

Salary	Annual number of trades	Annual turnover, percentage	Age	Annual number of trades	Annual turnover, percentage
< $25 K	0.11	7.78	< 35	0.17	10.40
$25–50 K	0.16	10.80	35–44	0.27	17.14
$50 K–75 K	0.22	14.18	45–54	0.36	22.28
$75 K–100 K	0.39	23.11	55–64	0.60	36.93
≥ $100 K	0.66	39.43	65+	0.03	2.78

Source: Agnew, Balduzzi, and Sunden (2003), Tables 4 and 5.

lose money by trading as a result of the transaction costs involved. Odean (1999) finds an average monthly turnover rate of 6.5% in a sample of discount brokerage customers. He argues that trading by these investors is excessive because the stocks purchased perform worse on average than the stocks sold, implying that the trades are disadvantageous even before payment of commissions. Using a sample of accounts at a discount brokerage firm, Barber and Odean (2000) find that the average investor in their sample performs about the same as the S&P500 index before costs but underperforms the index by 1.5% per year after costs. Within the sample, those in the top quintile in terms of turnover underperform the index by 5.5% per year after costs. The authors argue that overconfidence motivates frequent trading.

Table 8 provides evidence on the dependence of trading in directly held stocks on wealth and income. Wealthier households trade much more than less wealthy households, with about one-quarter of the wealthiest group (those with 1 million or more in net worth) trading more than ten times per year. This could be interpreted as evidence that wealthier investors are more overconfident than others. This interpretation would

Table 8   EFFECT OF HOUSEHOLD NET WORTH ON TRADING
FREQUENCY FOR DIRECTLY HELD STOCKS

Net worth	< $10 K	$10–50 K	$50–100 K	$100–250 K	$250–1 M	> $1 M	All
Percentage who hold stocks directly							
1998	3.2	8.3	13.8	21.5	43.1	68.6	19.2
2001	4.8	8.2	11.2	21.0	41.3	67.4	21.3
Percentage who bought or sold stocks in the last year, conditional on owning stock directly							
1998	6.4	17.8	12.9	16.7	35.7	61.3	22.8
2001	2.2	17.1	11.2	15.4	33.4	64.1	24.9
Percentage who traded 1–2 times							
1998	4.2	8.9	8.7	7.7	14.6	11.1	9.1
2001	0.7	7.5	4.7	9.2	11.6	12.7	8.5
Percentage who traded 3–10 times							
1998	0.9	5.7	3.6	6.3	12.1	22.6	7.8
2001	0.6	5.7	3.8	4.0	13.7	26.9	9.7
Percentage who traded >10 times							
1998	1.2	3.1	0.6	2.7	9.0	27.6	5.9
2001	0.9	3.8	2.6	2.2	7.8	24.5	6.7

Source: Calculated using data from the 1998 and 2001 Survey of Consumer Finances (using survey weights).

be consistent with the earlier evidence provided on biased self-attribution. Alternatively, high-wealth investors were seen to hold more shares on average, and that fact could be driving the results (portfolio turnover cannot be calculated in the survey of Consumer Finances). It would be interesting to correlate the dependence of underperformance due to frequent trading in Barber and Odean's (2000) study with wealth (or, better, labor income or education) to determine whether the wealthy are in fact trading more excessively than others with direct stockholdings or whether their frequent trading is rational. Coval, Hirshleifer, and Shumway (2002) document significant persistence in the performance of invididual investors buying and selling directly held stocks through a particular brokerage firm, suggesting that some do seem to have investment skill.

In sum, the evidence suggests that most of the seeming irrational investor behaviors are weaker for investors with higher wealth or income (frequent trading of directly held stocks being the main exception). This points to information or transactions costs as a potentially important contributing factor for these behaviors. I now turn to a simple calculation of the costs needed to explain one such behavior, namely, limited stock-market participation.

5. Costs of Stock-Market Participation

Information and/or transaction costs are a possible explanation for investor behavior that consists of inaction/too infrequent action/too simple action relative to the predictions of traditional finance theory. For each such behavior, however, it must be shown that the necessary costs are not implausibly large. In this section I give an example of how one might approach such a calculation in the case of stock-market participation. I start by considering which types of costs may be involved and then turn to an estimation of how large a per-period cost of stock-market participation would be needed to explain the choices of a substantial fraction of those who do not participate in the stock market.

5.1 COSTS FACED BY STOCK-MARKET INVESTORS

Consider the optimization problem of a household that maximizes expected lifetime utility given an exogenous stream of nonfinancial income and that faces the opportunity to invest in two assets: a risky asset and a riskless asset. The risky asset represents the stock market. The riskless asset is a catchall for less risky financial assets such as bonds, T-bills, bank accounts, etc. I assume it is free to invest in the riskless asset, whereas investing in stocks may involve several types of costs. First-time buyers likely incur an initial cost F^I representing the time/money spent

understanding basic investment principles as well as acquiring enough information about risks and returns to determine the household's optimal mix between stocks and riskless assets. Add to that the cost of time spent setting up accounts. Subsequently, a per-period stock-market participation cost F^P may be incurred. This cost would include the value of time spent throughout the year determining if trading is optimal. With time-varying conditional asset return distributions, theory suggests that households should actively follow the stock market to form more precise expectations of future returns and change their portfolios accordingly. For households who attempt to gather information and thus benefit from buying individual stocks or subcomponents of the stock-market index, the cost of this would also be included in F^P. A more subtle part of F^P is that stocks complicate tax returns. According to Internal Revenue Service (IRS) numbers for 2002, households who have to fill out schedules D and D1 (the schedules for capital gains and losses) spend 8 hours and 34 minutes on average doing so. In addition to F^I and F^P, stock-market investors face a fixed cost of trading stocks, including the fixed part of brokerage commissions as well as the value of time spent implementing the trade. Investors also face variable (proportional) costs of trading stocks. For directly held stocks, this cost represents the bid-ask spread and the variable part of brokerage commissions.[19] Indirect holding of stocks also involve transaction costs. For load mutual funds, the front load paid on entry into the fund would enter the proportional trading costs. In addition, or as an alternative, some funds have contingent deferred sales loads requiring investors to pay a certain percentage of their initial investment if they sell their mutual fund shares before a given number of years. These again work as a variable cost. Annual expenses on mutual funds also reduce investor returns.[20]

The above discussion emphasizes the costs of acquiring and processing information as an important element of F^I and F^P. Several recent papers find evidence that households who report to be better informed about financial issues make portfolio decisions more in line with theoretical predictions by having a higher probability of owning risky financial assets and holding a larger number of financial asset classes (see Guiso and Jappelli [2002] for evidence based on Italian data; Alessie, Hochguertel, and van Soest [2002] for results based on Dutch data; and Eymann and

---

19. Jones (2001) documents a quite strong decline in NYSE average one-way transaction costs (commissions plus half of the bid-ask spread) since the mid 1970s, from around 1.10 percentage points in 1970 to around 0.20 percentage point in the late 1990s. Consistent with the importance of trading costs, turnover has increased dramatically over the same period (of course, reverse causality cannot be ruled out based on these aggregate data).

20. Investment Company Institute (2002) estimates average annual total shareholder costs (operating expenses plus distribution costs) for equity mutual funds of 2.26% in 1980, gradually declining to 1.28% in 2001.

Borsch-Supan [2002] for findings from German data). While these relations may not be causal, other papers suggest a causal effect of information on savings and portfolio choice. Chiteji and Stafford (2000) find that parental stockholding has a strong effect on the probability that children become stockholders, controlling for economic and demographic characteristics of the children as well as for bequests. This suggests an effect of education about financial matters on stock-market participation. Duflo and Saez (2002) study retirement plan choices among the employees in various departments of a particular university. They find that the decision to enroll in a tax-deferred account plan (and the choice of mutual fund vendor for people who enroll) is affected by the decisions of other employees in the same department. Information flow from colleagues is a plausible explanation for such effects. Hong, Kubik, and Stein (2003) provide related evidence of peer effects. Bernheim, Garrett, and Maki (2001) find that the savings rate of households who grew up in a state with a high school financial curriculum mandate is about 1.5 percentage points higher than for others, controlling for income and demographics. Bernheim and Garrett (2003) find similar effects for employer-based retirement education plans.

5.2 HOW LARGE ARE THE COSTS NEEDED TO EXPLAIN NONPARTICIPATION?

I now turn to a simple estimation of how large costs are needed to explain nonparticipation in the stock market by many households. I focus on the case with a fixed per-period participation cost F^P only, but discuss how an entry cost or transaction costs may affect the results. I first estimate how large a value of F^P is needed for participation costs to explain the majority of nonparticipants' choices not to participate in the stock market. This assumes that all nonparticipating households face the same value of F^P. Then I allow F^P to differ across households and estimate its cross-sectional distribution. The advantage of allowing heterogeneity in F^P is that it enables the framework to explain different participation choices of households with similar wealth and other observable characteristics.

Both estimations are based on estimating the benefits of stock-market participation for each household, taking as given its current level of financial wealth. The advantage of this simple approach over a more structural one is that it allows me to use the actual distribution of financial wealth in the data without providing a detailed model able to generate the observed distribution. The most closely related paper on costs of stock-market participation is Mulligan and Sala-i-Martin (2000). They focus on all interest-bearing assets and the per-period cost of investing in such assets. At an interest rate of 5%, they estimate the median cost of holding interest-bearing assets to be $111 per year. I focus on stockholdings only

and present a theoretical argument to clarify the assumptions needed for the analysis of per-period participation costs. In addition, I consider a case where the cost is restricted to be the same for all nonparticipants so that one can identify the smallest cost needed to explain the choice of a given percentage of nonparticipants. Other related papers on investment costs and asset pricing are Luttmer (1999) and Paiella (1999), who focus on the costs needed to prevent households from adjusting their consumption from its current value (as opposed to reallocating existing financial wealth, as emphasized here).

5.2.1 *Theoretical Framework* My approach to estimating the benefits of stock-market participation relies on the definition of the certainty equivalent return to a portfolio. Start by considering a one-period setting with utility defined over end-of-period wealth and with no nonfinancial income. Consider a portfolio with stochastic net return r. If household i invests an amount W_i in the portfolio at the beginning of the period, end-of-period wealth is $W_i(1 + r)$. The certainty equivalent end-of-period wealth W_i^{ce} is given by[21]:

$$EU[W_i(1 + r)] = U(W_i^{ce}) \tag{1}$$

Correspondingly, the certainty equivalent return to the portfolio r_i^{ce} can be defined as:

$$EU[W_i(1 + r)] = U[W_i(1 + r_i^{ce})] \tag{2}$$

with the interpretation that the investor is indifferent between investing W_i in the risky portfolio with stochastic return r and investing it in a riskless portfolio with return r_i^{ce}. In a setting with participation costs of investing in the risky portfolio, replace initial wealth by $W_i^{Post} = W_i - F^P$. This wealth level then enters on the right side of the equation as well:

$$EU[W_i^{Post}(1 + r)] = U[W_i^{Post}(1 + r_i^{ce})] \tag{3}$$

If the risky portfolio consists of stocks and riskless assets in the fractions α_i and $1 - \alpha_i$, the above equation says that:

$$EU\left\{W_i^{Post}[1 + r_f + \alpha_i(r_s - r_f)]\right\} = U[W_i^{Post}(1 + r_i^{ce})] \tag{4}$$

where r_s is the stock return and r_f the riskless rate. Since the only risk in the portfolio of stocks and riskless assets stems from stocks, the certainty equivalent return to stocks $r_{s,i}^{ce}$ can be defined by the equation:

21. In the terminology of Pratt (1964), W_i^{ce} is given by $E[W_i(1 + r)] - \pi_i$, where π_i is the risk premium that makes the investor indifferent between receiving the stochastic amount $W_i(1 + r)$ and receiving the certain amount $E[W_i(1 + r)] - \pi_i$.

$$EU\{W_i^{\text{Post}}[1 + r_f + \alpha_i(r_s - r_f)]\} = U\{W_i^{\text{Post}}[1 + r_f + \alpha_i(r_{s,i}^{ce} - r_f)]\} \tag{5}$$

which states that the investor is indifferent between investing fractions α_i and $1 - \alpha_i$ in a portfolio of stocks and riskless bonds and investing all of W_i^{Post} in a riskless asset with net return $r_f + \alpha_i(r_{s,i}^{ce} - r_f)$. If the household is risk averse, then r_i^{ce} is a number smaller than the expected net return on the risky portfolio $E[r_f + \alpha_i(r_s - r_f)]$. Therefore, if $\alpha_i > 0$ $r_{s,i}^{ce} < E(r_s)$. A household choosing $\alpha_i > 0$ furthermore reveals that $r_{s,i}^{ce} > r_f$.

Consider now the more realistic case where households live for multiple periods and have nonfinancial income. In this case, we can define the certainty equivalent stock return $r_{s,i,t+1}^{ce}$ by the following equation:

$$\max_{C_{it}} U(C_{it}) + \beta E_t V_{t+1}\{(W_{it}^{\text{Post}} - C_{it})[1 + r_{f,t+1} + \alpha_{it}(r_{s,t+1} - r_{f,t+1})] + Y_{i,t+1}\} =$$

$$\max_{C_{it}} U(C_{it}) + \beta E_t V_{t+1}\{(W_{it}^{\text{Post}} - C_{it})[1 + r_{f,t+1} + \alpha_{it}(r_{s,i,t+1}^{ce} - r_{f,t+1})] + Y_{i,t+1}\} \tag{6}$$

where $V_{t+1}(W_{i,t+1})$ denotes the value function defined over date $t + 1$ wealth and β is the discount factor. On the left side of this equation, the expectation is taken over $r_{s,t+1}$ and $Y_{i,t+1}$. On the right side, it is taken over $Y_{i,t+1}$ only because $r_{s,i,t+1}^{ce}$ is nonstochastic. In the above definition, consumption in period t is allowed to differ depending on whether the risky portfolio or the riskless portfolio is held. Below, however, I will need to assume that the chosen consumption for period t (but not for future periods) is approximately unaffected by the portfolio choice.

The certainty equivalent stock return can now be used to determine the value of participating in the stock market. Given the definition of $r_{s,i,t+1}^{ce}$, the household will choose to participate in the stock market in the current period if:

$$\max_{C_{it}} U(C_{it}) + \beta E_t V_{t+1}\left\{(W_{it}^{\text{Post}} - C_{it})\left[1 + r_{f,t+1} + \alpha_{it}(r_{s,i,t+1}^{ce} - r_{f,t+1})\right] + Y_{i,t+1}\right\}$$

$$> \max_{C_{it}} U(C_{it}) + \beta E_t V_{t+1}\left[(W_{it} - C_{it})(1 + r_{f,t+1}) + Y_{i,t+1}\right] \tag{7}$$

where, as earlier, $W_{it}^{\text{Post}} = W_{it} - F^p$.

Below I consider two estimations. The first, estimation A, estimates the per-period cost sufficient to explain the choices of a given percentage of nonparticipants. The second more ambitious approach, estimation B, estimates the distribution of participation costs in the population.

5.2.2. *Estimation A: Homogeneous F^p* From equation (7), it follows that the gross benefit, as of time $t + 1$, of participating in the stock market in period t is:

$$\text{Benefit}_{it} = \left(W_{it}^{\text{Post}} - C_{it}\right) \alpha_{it} \left(r_{s,i,t+1}^{ce} - r_{f,t+1}\right) \tag{8}$$

under the simplifying assumption that period t consumption (but not future consumption) is unaffected by whether or not the household decides to enter the stock market.

On the cost side, the per-period cost of stock-market participation that could be avoided in period t by not entering, or entering in a subsequent period, reduces $W_{i,t+1}$ by:

$$\text{Avoidable cost}_{it} = F^P \left(1 + r_{f,t+1}\right) \tag{9}$$

A value of $F^P (1 + r_{f,t+1})$ greater or equal to Benefit$_{it}$ is sufficient to deter the household from participating in this period.[22] A lower value will also be sufficient if there are transactions costs (because the household would need to be able to recover these additional costs either in this period or in future periods of stock-market participation). In other words, if $x\%$ of nonparticipants have benefits less than y dollars in period t, then it is conservative to say that a per-period cost of $F^P = y$ is sufficient to explain the nonparticipation of $x\%$ of nonparticipants.

Under an additional assumption one can be more precise.

Assumption A: The per-period benefits of stock-market participation for observed nonparticipants are approximately the same across time periods for a given household i.

Most important, this assumes approximately constant holdings of financial wealth across periods for this group.[23] Then the entry condition states that the household should participate if:

Benefit$_{it}$ $> F^P$ + annuity value of all stock-market
transaction costs for household i (10)

I will refer to the right side as the total participation cost, F_i^{Total}. The advantage of this is that it no longer ignores the potential importance of any initial entry cost F^I or transaction costs (but at the cost of the extra assumption needed). The annuity value is calculated over years of stock-market participation.[24] Under assumption A, one can then estimate the annualized value of total stock-market participation costs which is sufficient to explain the nonparticipation of $x\%$ of nonparticipants. One problem with assumption A is that there may be a life-cycle component to

22. Because $1 + r_{f,t+1}$ is close to 1, for simplicity I replace $F^P (1 + r_{f,t+1})$ by F^P in what follows.
23. This assumption clearly would make less sense for participants because they decided to enter the stock market at some point, which suggests that their financial wealth likely increased to make this optimal.
24. Note that, unlike F^I, F^P and the transaction costs, which are exogenous parameters in the household's problem, the total participation cost has an endogenous element because the number of periods of stock-market participation is chosen by the household.

financial wealth even for relatively low-wealth households. One could consider repeating the estimations below with middle-aged households to provide a more conservative estimate of the costs needed to explain nonparticipation.

The data for the estimation come from the Panel Study of Income Dynamics (PSID). I use the Survey Research Center sample of the PSID, which was representative of the civilian noninstitutional population of the United States when the study was started in 1968. The PSID tracks all original family units and their adult offspring over time. With low attrition rates, the sample therefore remains representative as long as offspring are included. To keep the sample representative of the U.S. population, I exclude the poverty sample and the Latino sample. Wealth information from the 1984, 1989, and 1994 supplements is used to calculate financial wealth, defined as the sum of cash (checking or savings accounts, money market bonds, or Treasury bills, including such assets held in individual retirement accounts [IRAs]), bonds (bond funds, cash value in life insurance policies, valuable collections, rights in trusts or estates), and stocks (shares of stock in publicly held corporations, mutual funds, or investment trusts, including stocks in IRAs). To identify entries for which imputations were used, I use the wealth information as given in the family files instead of the wealth supplement files. Imputed values for cash, bonds, or stocks can then be coded as missing. Topcoding of wealth or income variables is very rare in the PSID, and topcoded variables were left at their topcodes. Although in reality households can have a portfolio share for a given asset above one, the PSID wealth data does not allow one to observe this due to the way the wealth questions are formulated. For example, the questions asked concerning stockholdings are "Do you (or anyone in your family living there) have any shares of stock in publicly held corporations, mutual funds, or investment trusts, including stocks in IRAs?" and "If you sold all that and paid off anything you owed on it, how much would you have?" Thus, a household who had borrowed to invest more than its total financial wealth in stocks would be recorded as having a portfolio share for stocks of one. Similarly, it is not possible to identify short sales from these questions. To allow comparison of amounts for different years, wealth variables are deflated by the consumer price index (CPI) for all urban consumers, with 1982–1984 as the basis year. My earlier paper (Vissing-Jorgensen, 2002) contains summary statistics for the sample. Among households with positive financial wealth, the percentage who owns stocks is 28.4% in 1987, 37.0% in 1989, and 44.2% in 1994. Among all households in the sample, 23.7% own stocks in 1987, compared to 29.1% in 1989, and 36.4% in 1994.

To implement estimation A, I make three additional assumptions. First, I calculate the benefit of stock-market participation as $W_{it}\alpha_t\,(r^{ce}_{s,i,t+1} - r_{f,t+1})$

Figure 9 STOCK-MARKET PARTICIPATION BENEFITS FOR NONPARTICIPANTS, PSID
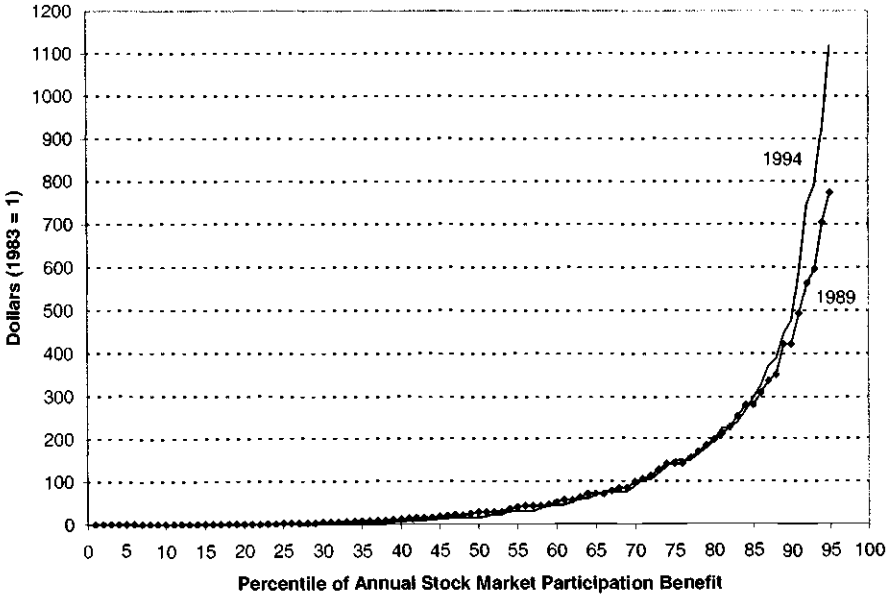
rather than $(W_{it}^{Post} - C_{it})\, \alpha_t\, (r_{s,i,t+1}^{ce} - r_{f,t+1})$. This overstates the benefits both by assuming that no wealth must be set aside for current period consumption and by replacing W_{it}^{Post} (financial wealth after participation costs) with observed financial wealth. Second, I assume a value of $r_{s,i,t+1}^{ce} - r_{f,t+1}$ of 0.04. With a historical equity premium around 7% and a tax rate of, for example, 20 percentage points, the after-tax equity return will be 5.6%.[25] Since the certainty equivalent excess return on stocks is risk adjusted, 4% seems, if anything, to be a high value. Thus, both these assumptions are conservative because they most likely overstate the benefits of stock-market participation and thus the costs needed to explain nonparticipation. Third, for the values of α_{it}, I assume that each nonparticipant would have had a value of α_{it} equal to the average value for participants in the PSID in that year (43.6 for 1989, 55.2 for 1994). Having calculated the period t benefit of stock-market participation for each of the nonparticipants as $W_{it}\, \alpha_t\, 0.04$, I calculate the percentiles of the cross-sectional distribution of this benefit in the set of nonparticipants. Figure 9 illustrates these

25. The exact tax rate is difficult to calculate because some stockholdings are in pensions plans where returns accumulate tax-free and are taxed only on withdrawal.

percentiles and thus gives the minimum dollar amount necessary to explain the choices of various percentages of nonparticipants.

The curve labeled 1989 in Figure 9 shows the percentiles of the benefit distribution for those who were nonstockholders in 1989 (and in 1984 to be reasonably confident that the household did not participate in earlier periods). The benefits are calculated based on the households' 1989 financial wealth. Similarly the curve labeled 1994 is based on those who were nonparticipants in 1994 and 1989. For readability, the figure leaves out percentiles above the 95th percentile.

In both 1989 and 1994, half of nonparticipants had estimated real annual stock-market participation benefits of less than $30. The price index used to calculate the real values has a basis value of one on average over the years 1982–1984. Multiply dollar values in the figure by 1.817 to adjust them to January 2003 dollars. Thus, a per-period stock-market participation cost (or a total participation cost under assumption A) of around $55 in 2003 prices is enough to explain the nonparticipation of half the nonparticipants. This reflects the fact that these households had little or no financial wealth to invest. Of the nonparticipants in 1989 (and 1984), around 21% had no financial wealth. Of the nonparticipants in 1994 (and 1989), about 29% had no financial wealth.

Interpreting the per-period participation cost as the cost of additional time spent following the market and doing more complicated taxes, a cost of $55 translates into less than 4 hours at an hourly wage of $15 per hour. For both 1989 and 1994, a cost of $150 per year (about $275 in 2003 prices) is enough to explain the choices of 75% of nonparticipants.

5.2.3. *Estimation B: Heterogeneous* F^P Suppose now that F^P is allowed to differ across households and time. This improves the models' ability to explain different choices by households with similar observable characteristics. For now, assume again that F^I and transactions costs are zero. I return to the possible effects of these costs below.

Given the definition of the benefit of stock-market participation in equation (8), a simple approach to estimating the cross-sectional distribution of F^P at date t is as follows. Suppose that $\alpha_{it} = \alpha_t$ for all i, that $r^{ce}_{s,i,t+1} - r_{f,t+1}$ 0.04 $\forall i$, and that F^P_{it} is uncorrelated with financial wealth in the cross section of households. Given these assumptions, the stock-market participation condition states that household i should participate in period t if:

$$(W^{Post}_{it} - C_{it})\, \alpha_t\, 0.04 > F^P_{it}\, (1 + r_{f,t+1}) \tag{11}$$

This condition is similar to the condition used by Mulligan and Sala-i-Martin (2000) in the context of the demand for interest-bearing assets more generally.

Since the incentive to participate is linear in financial wealth, one can estimate the cross-sectional distribution of F_{it}^P directly from the wealth distribution at date t. A simple nonparametric approach consists of calculating the percentage of households in different financial wealth groups who participate in the stock market. For example, if 27% of households with financial wealth of \$10,000 participate, then 27% of these households must have participation costs below \$10,000α_t0.04 = \$400α_t [as in estimation A, replace $(W_{it}^{Post} - C_{it})$ with W_{it} and $F_{it}^P (1 + r_{f,t+1})$ with F_{it}^P]. Given the assumption that F_{it}^P is cross-sectionally uncorrelated with W_{it}, this implies that 27% of all households must have had participation costs below \$400α_t. By splitting the sample into 10 wealth deciles and using this approach for each decile, one obtains 10 estimates of points on the cumulative distribution function (CDF) for the cross-sectional distribution of F_{it}^P.

How will the presence of an initial entry cost F^I or of transaction costs affect this estimation? Such costs imply that stock-market participation status becomes a state variable in the household's value function. This is the case because participating today affects the choices available tomorrow given the entry or transaction costs. In the example above where 27% of those with approximately \$10,000 in financial wealth were stock-market participants, one can no longer be sure that this implies that 27% of the draws of F_{it}^P are below \$400α_t. Let S_{it} be an indicator variable for whether household i participates in the stock market in period t. At date t, households can be split into four groups according to their participation choices at $t - 1$ and t: $(S_{i,t-1} = 0, S_{it} = 0)$, $(S_{i,t-1} = 0, S_{it} = 1)$, $(S_{i,t-1} = 1, S_{it} = 0)$, $(S_{i,t-1} = 1, S_{it} = 1)$. We would like to determine the percentage of the draws of F_{it}^P that are less than \$400α_t. The group $(S_{i,t-1} = 0, S_{it} = 1)$ poses no difficulties. We can be sure that their F_{it}^P draw is less than \$400α_t (because their choice reveals that the period t benefit exceeds F^P plus any part of entry or transaction costs that must be covered by the period t benefits for entry to have been worthwhile). With the group $(S_{i,t-1} = 1, S_{it} = 0)$, we can be sure that their F_{it}^P draw is above \$400α_t because they have revealed that F_{it}^P exceeds their current period benefit of \$400α_t plus any future transaction costs they may save by staying in the market during this period.[26] The possible misclassifications arise for the groups choosing $(S_{i,t-1} = 0, S_{it} = 0)$ or $(S_{i,t-1} = 1, S_{it} = 1)$. Those choosing $(S_{i,t-1} = 0, S_{it} = 0)$ reveal only that \$400α_t is not sufficient to cover F_{it}^P plus any part of the entry and transaction costs that must be covered by a period t gain for entry to have been optimal. Thus, they reveal $F_{it}^P \geq \$400\alpha_t - z_{it}^{00}$ for some positive value z_{it}^{00}. Using the approach outlined above and classifying them all as having $F_{it}^P \geq \$400\alpha_t$

---

26. For example, a household with a temporary increase in consumption needs may decide to run down only nonstock wealth in this period and thus save the transaction costs involved in trading stocks.

leads one to overestimate the deciles of the cost distribution. However, those choosing $(S_{i,t-1} = 1,\ S_{it} = 1)$ lead to a counterbalancing bias: they reveal only that $\$400\alpha_t$ plus any future transaction costs they save by staying in the market during this period exceeds F_{it}^P; i.e., that $F_{it}^P \leq \$400\alpha_t + z_{it}^{11}$ for some positive value z_{it}^{11}. Thus, classifying them all as having $F_{it}^P \leq \$400\alpha_t$ leads one to understimate the deciles of the cost distribution. Overall, if an equal number of each of the $(S_{i,t-1} = 0,\ S_{it} = 0)$ households and the $(S_{i,t-1} = 1,\ S_{it} = 1)$ households are misclassified, then the approach outlined, assuming the absence of transaction costs, will lead to an unbiased estimate of the cross-sectional distribution of F_{it}^P. Because it is difficult to evaluate whether the two biases are likely to cancel each other, the cost distributions estimated below should be interpreted with some caution.

In both estimation A and B, one can allow for heterogeneity in α_{it} across households (rather than only across time). Nonparticipants may have chosen to stay out of the market due to a low optimal stock share conditional on participation. Accounting for heterogeneity based on a sample selection model has only small effects on the results, however, and for simplicity is therefore omitted from the results shown. Essentially, this is due to the fact that while the fit in models of the stock-market participation decision is quite high, models of the share invested in stocks conditional on participation typically has low explanatory power (possibly due to transaction costs of portfolio adjustment leading to substantial differences between optimal and observed portfolio shares for equity).

The results of estimation B are shown in Figure 10 for the sample of all households with positive financial wealth. Again assume that α_{it} (the actual or potential share of financial wealth invested in stocks) for each household equals the average value for participants in the PSID in that year (43.4 for 1984, 43.6 for 1989, and 55.2 for 1994). Households with no financial wealth provide no information about the participation cost in this approach because their benefit of stock-market entry is zero, assuming they cannot borrow or change their current consumption to invest in the stock market. The median per-period participation cost is around \$350 (real 1982–1984 dollars) for 1994, around \$500 for 1989, and around \$800 for 1984. Even among very rich households, not all hold stocks, so the estimated CDF does not reach 1 at any wealth level (the point corresponding to the last wealth decile is not included in the graph but is also far below 1). This emphasizes the advantage of using a nonparametric approach because a parametric approach would impose the requirement that the CDF reaches 1. The economic implication is that participation costs are unlikely to be the explanation for nonparticipation among high-wealth households. More generally, if some of the nonparticipants at each wealth level have chosen not to hold stocks for reasons other than participation

Figure 10 ESTIMATED CDF OF PER-PERIOD STOCK-MARKET
PARTICIPATION COST, PSID
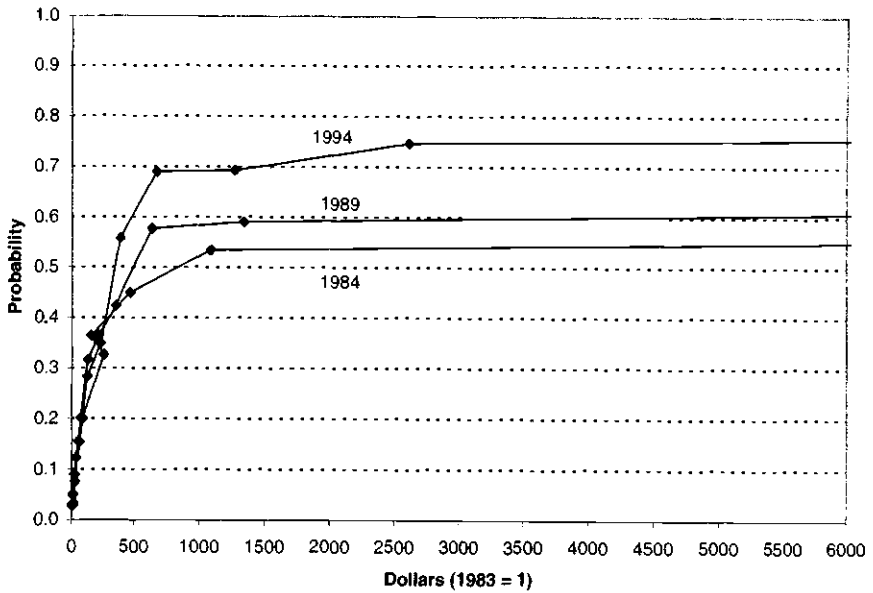
costs, my estimated CDF of the cost distribution will be shifted down
compared to the true CDF.

Overall the results of the estimations of stock-market participation costs
show that, while it is not reasonable to claim that participation costs can
reconcile the choices of all nonparticipants, modest costs are sufficient to
understand the choices of a large part of these households due to their
fairly low amounts of financial wealth.

6. Conclusion

Behavioral finance, and behavioral economics more generally, is a very
active area of research. The state of the literature is still one of exploration,
with little agreement among researchers on what the most important
investor biases are from an asset-pricing perspective.

In this paper I have argued that more direct evidence about the beliefs
and actions of investors would make behavioral theories more convincing
to outsiders, many of whom remain unconvinced that any of the multi-
tude of biases documented in the psychology literature and the experi-
mental literature have much impact on asset prices (see Hirshleifer [2001]
for a thorough discussion of the evidence from psychology and experi-
ments). To exemplify the potential value of such direct evidence, I have

analyzed new data from UBS/Gallup on investor expectations and stock-holdings for 1998–2002. The evidence suggests that, even for wealthy investors, (1) expected returns were high at the peak of the market; (2) many investors thought the market was overvalued but would not correct quickly; (3) investors' beliefs depend on their own investment experience (a version of the law of small numbers); (4) the dependence of beliefs on own past portfolio performance is asymmetric, consistent with theories of biased self-attribution; and (5) investor beliefs do affect their stockholdings, suggesting that understanding beliefs is in fact useful for understanding prices.

I then turned to existing evidence about investor behaviors that are inconsistent with traditional finance theory recommendations. Information and/or transaction costs represent a possible rational explanation of behaviors that involve too little action or too simple actions relative to the theoretical benchmark. I argued that many such behaviors tend to diminish with investor wealth and sophistication and thus that information and transaction costs should be seriously considered as an explanation. As an example, a simple calculation showed that, given the observed distribution of financial wealth, an annual cost of about $55 is enough to explain the choices of half of those who do not invest in the stock market.

REFERENCES

Abreu, D., and M. Brunnermeier. (2001). Bubbles and crashes. *Econometrica* 71(1):173–204.

Abreu, D., and M. Brunnermeier. (2002). Synchronization risk and delayed Arbitrage. *Journal of Financial Economics* 66:341–360.

Agnew, J., Balduzzi, P. L. and A. Sunden. (2003). Portfolio choice and trading in a large 401(k) plan. *American Economic Review* 93(1):193–215.

Alessie, R., S. Hochguertel, and A. van Soest. (2002). Household portfolios in the Netherlands. In *Household Portfolios*, L. Guiso, M. Haliassos, and T. Jappelli (eds.). Cambridge, MA:MIT Press, pp. 341–388.

Ameriks, J., and S. P. Zeldes. (2001). How do household portfolio shares vary with age? Working Paper, Columbia University.

Banz, R. W. (1981). The relationship between return and the market value of common stocks. *Journal of Financial and Quantitative Analysis* 14:421–441.

Barber, B. M., and T. Odean. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* 55(2):773–806.

Barber, B. M., and T. Odean. (2002). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. University of California, Davis. Working Paper.

Barberis, N., A. Shleifer, and R. Vishny. (1998). A Model of investor sentiment. *Journal of Financial Economics* 49:307–343.

Barberis, N., and R. Thaler. (2002). A survey of behavioral finance. In *Handbook of the Economics of Finance*, G. Constantinides, M. Harris, and R. Stulz (eds.). Forthcoming.

Benartzi, S. (2001). Excessive extrapolation and the allocation of 401(k) accounts to company stock. *Journal of Finance* 56(5):1747–1764.

Benartzi, S., and R. H. Thaler. (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review* March: 79–98.

Berk, J. B., R. C. Green, and V. Naik. (1999). Optimal investment, growth options and Security Returns. *Journal of Finance* 54:1553–1608.

Bernard, V. L., and J. K. Thomas. (1989). Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27:1–36.

Bernheim, B. D., and D. M. Garrett. (2003). The effects of financial education in the workplace: Evidence from a survey of households. *Journal of Public Economics* 87(7–8):1487–1519.

Bernheim, B. D., D. M. Garrett, and D. M. Maki. (2001). Education and saving: The long-term effects of high school financial curriculum mandates. *Journal of Public Economics* 80(3):435–465.

Blume, M. E., J. Crockett, and I. Friend. (1974). Stock ownership in the United States: Characteristics and trends. *Survey of Current Business* 54(11):16–40.

Brav, A., R. Lehavy, and R. Michaely. (2002). Expected return and asset pricing. Duke University. Working Paper.

Brown, P., N. Chappel, R. da Silva Rosa, and T. Walter. (2002). The reach of the disposition effect: Large sample evidence across investor classes. University of Sydney. Working Paper.

Brunnermeier, M. K., and S. Nagel. (2002). Arbitrage at its limits: Hedge funds and the technology bubble. Princeton University. Working Paper.

Campbell, J. Y. (2000). Asset pricing at the millennium. *Journal of Finance* 55(4):1515–1567.

Chiteji, N. S., and F. P. Stafford. (2000). Asset ownership across generations. Population Studies Center at the Institute for Social Research, University of Michigan. Research Report No. 00–454.

Cohen, R. B., P. A. Gompers, and T. Vuolteenaho. (2002). Who underreacts to cash-flow news? Evidence from trading between individuals and institutions. *Journal of Financial Economics* 66(2–3):409–462.

Coval, J. D., D. Hirshleifer, and T. Shumway. (2002). Can individual investors beat the market? Harvard Business School. Working Paper.

Coval, J., and T. J. Moskowitz. (1999). Home bias at home: Local equity preference in domestic portfolios. *Journal of Finance* 54:2045–2074.

Coval, J., and T. J. Moskowitz. (2001). The geography of investment: Informed trading and asset prices. *Journal of Political Economy* 4(109):811–841.

Daniel, K., D. Hirshleifer, and A. Subrahmanyam. (1998). Investor psychology and security market under- and overreactions. *Journal of Finance* 53(5):1839–1886.

Daniel, K., D. Hirshleifer, and S. H. Teoh. (2002). Investor psychology and capital markets: Evidence and policy implications. *Journal of Monetary Economics* 49:139–209.

DeBondt, W. F. M., and R. H. Thaler. (1985). Does the stock market overreach? *Journal of Finance* 40:793–808.

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

Dhar, R., and N. Zhu. (2002). Up close and personal: An individual level analysis of the disposition effect. Yale School of Management. Working Paper.

Diether, K., C. J. Malloy, and A. Scherbina. (2002). Differences of opinion and the cross section of stock returns. *Journal of Finance* 57:2113–2141.

Duflo, E., and E. Saez. (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. *Journal of Public Economics* 85:121–148.

Eymann, A., and A. Borsch-Supan. (2002). Household portfolios in Germany. In *Household Portfolios*, L. Guiso, M. Haliassos, and T. Jappelli (eds.) Cambridge, MA: MIT Press, pp. 291–340.

Fama, E. F. (1998). Market efficiency, long-term returns and behavioral finance. *Journal of Financial Economics* 49:283–306.

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

Fisher, K. L., and M. Statman. (2002). Blowing bubbles. *Journal of Psychology and Financial Markets* 3(1):53–65.

French, K. R., and J. M. Poterba. (1991). Investor diversification and international equity Markets. *American Economic Review* 81(2):222–226.

Frieder, L., and A. Subrahmanyam. (2002). Brand perceptions and the market for common stock. UCLA. Working Paper.

Genesove, D., and C. Mayer. (2001). Loss aversion and seller behavior: Evidence from the housing market. 116(4):1233–1260.

Goetzmann, W. N., and A. Kumar. (2001). Equity portfolio diversification. Cambridge, MA: National Bureau of Economic Research. NBER Working Paper No. 8686.

Goetzmann, W. N., and M. Massa. (2003). Disposition matters: Volume, volatility and price impact of a behavioral bias. Cambridge, MA: National Bureau of Economic Research. NBER Working Paper No 9499.

Gomes, J. F., L. Kogan, and L. Zhang. (2003). Equilibrium cross-section of returns. *Journal of Political Economy* 111:693–732.

Graham, J. R., and C. R. Harvey. (2001). Expectations of equity risk premia, volatility and asymmetry from a corporate finance perspective. Fuqua School of Business. Working Paper.

Grinblatt, M., and B. Han. (2002). The disposition effect and momentum. UCLA. Working Paper.

Grinblatt, M., and M. Keloharju. (2000). The investment behavior and performance of various investor types: A study of Finland's unique data set. *Journal of Financial Economics* 55:43–67.

Grinblatt, M., and M. Keloharju. (2001a). How distance, language, and culture influence stockholdings and trades. *Journal of Finance* 56(3):1053–1073.

Grinblatt, M., and M. Keloharju. (2001b). What makes investors trade? *Journal of Finance* 56(2):589–616.

Grossman, S. J., and J. E. Stiglitz. (1980). On the impossibility of informationally efficient markets. *American Economic Review* 70(3):393–408.

Guiso, L., M. Haliassos, and T. Jappelli. (2002). Household portfolios. Cambridge, MA: MIT Press.

Guiso, L., and T. Jappelli. (2002). Households' portfolios in Italy. In *Household Portfolios*, L. Guiso, M. Haliassos and T. Jappelli (eds.). Cambridge, MA:MIT Press, pp. 251–290.

Heath, C., S. Huddart, and M. Lang. (1999). Psychological factors and stock option exercise. *Quarterly Journal of Economics* 114(2):601–627.

Heaton, J., and D. Lucas. (1999). Stock prices and fundamentals. In *NBER Macroeconomics Annual 1999*, B. Bernanke and J. J. Rotemberg (eds.). Cambridge, MA: MIT Press, pp. 213–242.

Hirshleifer, D. (2001). Investor psychology and asset pricing. *Journal of Finance* 64:1533–1597.

Hong, H. G., J. D. Kubik, and A. Solomon. (2000). Security analysts' career concerns and herding of earnings forecasts. *Rand Journal of Economics* 31:121–144.

Hong, H., J. D. Kubik, and J. Stein. (2003). Social interaction and stock market participation. *Journal of Finance* Forthcoming.

Hong, H., and J. C. Stein. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance* 54(6):2143–2184.

Huberman, G. (2001). Familiarity breeds investment. *Review of Financial Studies* 14(3):659–680.

Investment Company of America and the Securities Industry Association. (1999). Equity ownership in America. Washington, D. C.: Investment Company Institute.

Investment Company Institute. (2002). Total shareholder cost of mutual funds: An update. *Fundamentals*, 11(4):1–8.

Jegadeesh, N. (2000). Long-run performance of seasoned equity offerings: Benchmark errors and biases in expectations. *Financial Management* 29(3):5–30.

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

Jones, C. M. (2001). A century of stock market liquidity and trading costs. Columbia University. Working Paper.

Kahneman, D., and A. Tversky. (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 46:171–185.

Lamont, O. A. (2002). Macroeconomic forecasts and microeconomic forecasters. *Journal of Economic Behavior and Organization* 48:265–280.

La Porta, R., J. Lakonishok, A. Shleifer, and R. W. Vishny. (1997). Good news for value stocks: Further evidence on market efficiency. *Journal of Finance* 52:859–874.

Loughran, T., and J. Ritter. (1995). The new issues puzzle. *Journal of Finance* 50:23–52.

Luttmer, E. G. J. (1999). What level of fixed costs can reconcile asset returns and consumption choices. *Journal of Political Economy* 107(5):969–997.

Mankiw, N. G., R. Reis, and J. Wolfers. (2003). Disagreement about inflation expectations. In *Macroeconomics Annual 2003*, M. Gertler and K. Rogoff (eds.). Cambridge, MA: MIT Press, pp. xxx–xxx.

Massa, M., and A. Simonov. (2003). Behavioral biases and portfolio choice. INSEAD. Working Paper.

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

Mulligan, C., and X. Sala-i-Martin. (2000). Extensive margins and the demand for money at low interest rates. *Journal of Political Economy* 108(5):961–991.

New York Stock Exchange. (2000). Shareownership 2000. Available at http://www.nyse.com/pdfs/shareho.pdf (accessed fund 3, 2003).

Odean, T. (1999). Do investors trade too much? *American Economic Review* 89(5):1279–1298.

Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance* 53(5):1775–1798.

Paiella, M. (1999). Transaction costs and limited stock market participation to reconcile asset prices and consumption choices. University College London. Working Paper.

Polkovnichenko, V. (2001). Limited stock market participation and the equity premium. University of Minnesota. Working Paper.

Pratt, J. W. (1964). Risk aversion in the small and the large. *Econometrica* 32:122–136.

Rabin, M. (2002). Inference by believers in the law of small numbers. *Quarterly Journal of Economics* 117(3):775–816.

Rubinstein, M. (2001). Rational markets: Yes or no? The affirmative case. *Financial Analysts Journal*, May/June, 57(3):15–29.

Samuelson, W., and R. J. Zeckhauser. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1 (March, 1):7–59.

Shapira, Z., and I. Venezia. (2001). Patterns of behavior of professionally managed and independent investors. *Journal of Banking and Finance* 25:1573–1587.

Shefrin, H., and M. Statman. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance* 40(3): 777–790.

Shiller, R. J., F. Kon-Ya, and Y. Tsutsui. (1996). Why did the Nikkei crash? Expanding the scope of expectations data collection. *Review of Economics and Statistics* 78(1):156–164.

Shleifer, A. (2000). *Inefficient Markets: An Introduction to Behavioral Finance*, Oxford, England: Oxford University Press.

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

Souleles, N. S. (2001). Consumer sentiment: Its rationality and usefulness in forecasting expenditure—Evidence from the Michigan micro data. *Journal of Money, Credit, and Banking*, forthcoming.

Vissing-Jorgensen, A. (1998). Limited stock market participation. Department of Economics, MIT. Ph.D. Thesis.

Vissing-Jorgensen, A. (1999). Comment on stock prices and fundamentals. *NBER Macroeconomics Annual 1999*. Cambridge, MA: The MIT Press, pp. 242–253.

Vissing-Jorgensen, A. (2002). Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures. Northwestern University. Working Paper.

Comment

JOHN Y. CAMPBELL
Harvard University and NBER

Almost 20 years ago, Robert Shiller, Lawrence Summers, and Richard Thaler challenged the finance profession to take seriously the possibility that investor behavior and asset prices deviate from the predictions of simple rational models. Since that time, behavioral finance has become one of the most active areas in financial economics, maturing to the point where it can be summarized in both popular and professional books (Shiller, 2000; Shleifer, 2000). Behavioral economics has had great success more generally, as illustrated by the award of the 2001 Clark Medal to Matthew Rabin and the 2002 Nobel Prize to Daniel Kahneman and

Vernon Smith, and behavioral finance is probably the most successful application of this approach.

In asset pricing, it is often hard to draw clear distinctions between behavioral and other research. Empirical researchers document systematic tendencies for some types of assets to outperform others, or for assets to perform better at some times than others. Very weak restrictions on asset markets ensure that these patterns can be explained by the properties of a stochastic discount factor that summarizes the rewards for taking on different kinds of risks. Behavioral finance models may derive the stochastic discount factor from nonstandard models of investor preferences, such as the prospect theory of Kahneman and Tversky (1979), but this can be hard to distinguish from more conventional models with features such as habit formation.

Behavioral finance is more distinctive in its insistence that we should try to measure the beliefs and actions of particular investors. We should not assume that investors' beliefs are homogeneous or rational, or that they deviate only idiosyncratically from a common set of rational beliefs. Rather, we should identify meaningful groups of investors and explore the possibility that these groups have different beliefs that induce them to trade with one another. Equilibrium asset prices emerge from the interactions of these heterogeneous investors.

Initially, the behavioral literature distinguished two groups of investors: rational investors and irrational noise traders. This raises the question of which investors play the role of noise traders. Much recent work has emphasized the distinction between individual investors, who may be particularly susceptible to cognitive limitations and psychological biases, and institutions, which seem likely to be more rational but may be limited in their risk-taking capacity. Vissing-Jorgensen's paper follows this tradition and examines a fascinating new dataset on the expectations of individual investors.

1. The UBS/Gallup Survey

The UBS/Gallup telephone survey has some inherent limitations. The most serious issue is whether respondents answer the survey questions accurately. Any survey that involves telephoning people at home in the evening is likely to elicit hasty or flippant responses. This is a particular problem here because the survey is relatively ambitious, going far beyond simple questions with binary answers such as yes/no or Republican/ Democrat. Accurate answers require both effort and comprehension. Questions about recent portfolio performance, for example, may require respondents to aggregate information from multiple brokerage

and retirement accounts, while questions about expected future long-term returns require respondents to understand the difference between annual and cumulative returns. These problems do not mean that the data are worthless, but they do limit the weight that can be placed on the results.

A second limitation of the UBS/Gallup survey is that it is a series of cross sections and not a panel; thus, it cannot be used to track the expectations of particular individuals through time. Previous research has shown that the beliefs of different market participants may evolve in very different ways. Brav, Lehavy, and Michaely (2002), for example, look at the price targets issued by stock analysts and use these to construct analysts' return expectations. They argue that these numbers represent the expectations not only of the analysts themselves but also of the investors who follow them. They compare the price targets of analysts employed by sell-side brokerage firms (First Call data over the period 1997–2001, covering 7000 firms) with the price targets of independent analysts (Value Line data over a longer period—1987–2001—covering just under 3000 firms). They find that sell-side analysts' return forecasts increased with the level of the stock market in the late 1990s, while independent analysts' forecasts decreased throughout the 1990s. These discrepancies may be caused by honest differences of opinion, by differences in the horizon of the return forecast (one year for sell-side analysts, four years for independent analysts), or by the investment banking ties of sell-side analysts that induced them to tout the stocks of client companies. Regardless of the source, differences in analysts' opinions might well have led some individual investors to increase their return expectations in the late 1990s even while other investors were reducing their expectations.

2. The Distribution of Return Expectations

Although the UBS/Gallup survey is not a panel, there is much that can be learned from the cross-sectional distribution of return expectations within each month. Figure 2 in the paper shows that the cross-sectional standard deviation of return forecasts averages around 10%, comparable to the cross-sectional mean in Figure 1. Individual investors clearly do not have homogeneous expectations. In addition, the cross-sectional standard deviation appears to increase in the late 1990s, peaking in 2000, and then declines modestly. This pattern would be expected if some investors reacted to high returns in the late 1990s by increasing their return expectations, in the manner of sell-side analysts, while other investors decreased their expectations, in the manner of independent analysts.

Figure 2 also shows a great deal of variability in the cross-sectional standard deviation from month to month, and this tends to obscure the lower-frequency variation in disagreement. It would be good to know more about the possible sources of this high-frequency variation in the cross-sectional standard deviation of return expectations. For example, how much time-series variation would be expected just from sampling error if the true cross-sectional standard deviation is constant and 1000 households are interviewed each month?

The heterogeneity of investors' expectations raises difficult issues when one tries to summarize the survey results in a single average return expectation. The paper emphasizes an equal-weighted average, sometimes with an adjustment for the sampling methods used in the survey. This average, shown in Figure 1, increases in the late 1990s and declines after 2000. For the determination of asset prices, however, a wealth-weighted average is more relevant because wealthy investors have a much greater effect on asset demands than poor investors do. Vissing-Jorgensen reports that wealthy investors, with more than $100,000 in assets, have lower return expectations throughout the sample period but that their average expectations have the same time pattern shown in Figure 1.

If investors are constrained from selling shares short, or if they are reluctant to do so, then the most optimistic investors have a disproportionate influence on prices. The high level of disagreement about future stock returns throughout the sample, and particularly in 2000, indicates that this problem is relevant and that a wealth-weighted average return expectation understates the average demand for stocks by individual investors. Overall, the UBS/Gallup data suggest that optimism among individual investors was an important source of demand for stocks in the late 1990s. This raises the question of why individual investors were so optimistic in this period.

3. Irrational Extrapolation?

One plausible story is that individuals overreact to their recent past experience, irrationally extrapolating it into the future. According to this story, a series of favorable shocks during the 1990s set the stage for a speculative bubble at the end of the decade.

Although the UBS/Gallup survey does not follow investors through time, it does ask them to report their age, the number of years for which they have been investing, and their recent past portfolio returns. This feature of the data allows Vissing-Jorgensen to ask whether investors irrationally extrapolate their own past experience. She argues that if this is the case, then young and inexperienced investors should be more optimistic

than older and experienced investors at the market peak because they place more weight on recent high returns in forming their expectations. The evidence, reported in Figures 4 and 5, turns out to be mixed. Young and inexperienced investors do have higher return expectations than do older and experienced investors at the market peak in 2000, but as the market falls in 2001 and 2002, the gap in expected returns narrows only for young investors and not for inexperienced investors.

Vissing-Jorgensen argues that reported past portfolio returns provide additional evidence of irrational extrapolation. Investors who report high past portfolio returns also expect higher future returns on the market (Figure 6), and the effect of past portfolio returns on expectations is stronger when those past returns are positive (Figure 8). Vissing-Jorgensen interprets the latter result as evidence for biased self-attribution; the human tendency to treat past success as meaningful evidence about one's skill and to treat past failure as random bad luck.

I believe that the results using past portfolio returns should be treated with caution. A first problem is that past returns are self-reported and may well reflect an investor's general optimism. A respondent who receives the UBS/Gallup telephone call when she is in a good mood may say that her portfolio has been doing well and that the market will do well in the future, whereas another respondent who is in a bad mood may give more pessimistic responses. This sort of correlated measurement error could account for the patterns shown in Figure 6.

A second problem is more subtle. The results shown in Figure 8 are based on a regression of the following form:

$$R_{it}^e = \alpha_t + \beta_t \text{age}_{it} + \gamma_t \, \text{experience}_{it} + \theta_1 R_{i,t-1}^{P+} + \theta_2 R_{i,t-1}^{P-} + u_{it}$$

Here i denotes an individual respondent and t denotes a time period. The intercept and the coefficients on age and experience are all time-varying, but the coefficients on lagged positive portfolio return $R_{i,t-1}^{P+}$ and lagged negative portfolio return $R_{i,t-1}^{P-}$ are fixed. The regression is estimated over the period 2000–2002. Figure 8 reflects the fact that θ_1 is estimated to be large and positive, while θ_2 is estimated to be small and negative.

This regression is hard to interpret because past portfolio returns are not exogenous and are likely to be correlated with other determinants of return expectations that are omitted from the regression. Also, more investors had positive past portfolio returns in 2000 than in 2002. Thus, the separate coefficients on $R_{i,t-1}^{P+}$ and $R_{i,t-1}^{P-}$ may capture a change over time in the correlation between past portfolio performance and return expectations, rather than a true structural difference between the

effects of positive and negative past returns. The signs of the coefficients could be explained, for example, if investors are exogenously optimistic or pessimistic and invest accordingly (optimists invest in stocks and pessimists invest in Treasury bills). In 2000, optimists had high past returns and reported high return expectations, while pessimists had mediocre past returns and reported low return expectations, generating a positive coefficient θ_1. In 2002, optimists had low past returns and reported high return expectations, while pessimists had mediocre past returns and reported low return expectations, generating a negative coefficient θ_2.

4. The Limitations of Behavioral Finance

Stepping back from the details of the empirical work in the paper, the difficulty in interpreting individual investors' optimism at the end of the 1990s illustrates the challenges that face behavioral finance. Compared with traditional models in financial economics, behavioral models often have a degree of flexibility that permits reinterpretation to fit new facts. Such flexibility makes it hard either to disprove or to validate behavioral models. For example, the theory of biased self-attribution does not make a clear prediction about individual investors' expectations of returns on the aggregate stock market. The theory says that investors interpret their past success as evidence of their skill, but it is not clear why people who have earned high returns in stocks and believe themselves to be skillful investors should necessarily expect the stock market to keep rising. They might just as well switch from one asset class to another in the belief that they have identified the next new trend.

The lack of theoretical discipline would not be a problem if empirical research on investor behavior indicated that individual investors are consistently biased in a particular direction. Unfortunately, this is not the case. While some behavior patterns are consistent with irrational extrapolation, others contradict it. Individual investors are keen to put their money in mutual funds that have performed well recently (Chevalier and Ellison, 1997; Sirri and Tufano, 1998), but they also tend to sell stocks that have performed well and hold on to those stocks that have performed badly (the disposition effect of Shefrin and Statman [1985] and Odean [1998]). While the mutual fund evidence is consistent with irrational extrapolation, it may also reflect the mechanism by which skillful fund managers are compensated (Berk and Green, 2002). The disposition effect is hard to reconcile with either the principles of rational investing or irrational extrapolation. It is sometimes attributed to prospect theory, combined with stock-level mental accounting, but this leaves an open

question about the balance between these forces and the opposing force of irrational extrapolation.

The finance profession has learned a great deal from the detailed and careful empirical research on investor behavior that has been promoted by behavioral finance. Vissing-Jorgensen's paper is an excellent example of this type of research. I do not believe, however, that behavioral finance has yet been able to offer a coherent theoretical framework comparable to traditional finance theory. It is better thought of as a set of observed behaviors, particularly prevalent among individual investors with less experience and wealth, that can affect asset prices and, just as important, the financial well-being of these investors. Financial economists should take such behaviors seriously and should try to use financial education to reduce their incidence.

REFERENCES

Berk, J., and R. Green. (2002). Mutual fund flows and performance in rational markets. University of California Berkeley. Unpublished Paper.
Brav, A., R. Lehavy, and R. Michaely. (2002). Expected return and asset pricing. Duke University. Unpublished Paper.
Chevalier, J., and G. Ellison. (1997). Risk taking by mutual funds as a response to incentives. *Journal of Political Economy* 105:1167–1200.
Kahneman, D., and A. Tversky. (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 46:171–185.
Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance* 53:1775–1798.
Shefrin, H., and M. Statman. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance* 40:777–790.
Shiller, R. (2000). *Irrational Exuberance*. Princeton, NJ: Princeton University Press.
Shleifer, A. (2000). *Inefficient Markets: An Introduction to Behavioral Finance*, Oxford, England: Oxford University Press.
Sirri, E., and P. Tufano. (1998). Costly search and mutual fund flows. *Journal of Finance* 53:1589–1622.

Comment

OWEN A. LAMONT
School of Management, Yale University, and NBER

The U.S. stock market has experienced amazing upheaval in the past five years. Valuations seemed absurdly high in the period 1998–2000, especially for technology-related stocks. Many previously identified anomalies, particularly those relating to new issues, grew larger during the tech-stock mania episode. Other previously identified patterns relating to

scaled prices (that is, price expressed as a ratio such as price/dividend or price/book value) seemed to go away entirely in the late 1990s, only to return with a vengeance in 2000–2002. This period will be studied by financial economists for years to come because it is an extraordinarily revealing episode full of important clues.

Vissing-Jorgensen discusses some fascinating evidence about this period, most of it derived from a continuing survey of investors from 1998–2002. Like all data, this data has limitations. First, all survey data should be regarded with skepticism and this type of survey more than most (I discuss this point further below). Second, it is a real shame this dataset starts in 1998 and does not include more of the pre-mania period. Despite these limitations, the data reveal many interesting facts that are useful for sorting out different hypotheses about the tech-stock mania.

Vissing-Jorgensen has done a great job in revealing the features of the data, taking what was undoubtedly a complicated and demanding task and making it look easy. There are many different ways this data could have been used, and I am convinced that the graphs, regressions, and statistical tests are accurately telling us what we need to know. Vissing-Jorgensen does many different things in the paper, but I am going to focus only on the main survey results and interpret them from my own perspective of what the tech-stock mania period was about.

1. Limitations of Survey Data

To me, survey data about expectations and beliefs is one of the weakest forms of data, just one rung above anecdotes in the quality ladder. I think we should always be suspicious of survey data on beliefs, especially involving abstract and intangible concepts (such as expected stock returns) that are unfamiliar to the respondents. This data is most useful when it is possible to cross-verify with data on actual (not self-reported) behavior observed by objective external measurement. I see survey evidence of this type as suggestive, but not definitive.

Fortunately, there is ample cross-verification for many of the patterns documented in the paper. For example, Figure 8 shows striking evidence for biased self-attribution (individuals believe good performance is due to their skill, but bad performance is due to luck). The fact that the performance is self-reported makes interpreting the results problematic. Fortunately, other evidence documents biased self-attribution. Barber and Odean (2002) document, using actual portfolio performance and actual trading, that investors who have done well in the past tend to increase their trading, consistent with the hypothesis that they believe themselves to be more skillful.

Having stated these reservations, let me say that I find nothing at all implausible in the results shown. They seem revealing about what was going through people's minds in real time. This survey is also potentially more reliable than surveys involving analyst expectations because these analysts have their own problems (most notably a pronounced optimistic bias).

2. Aggregate Expected Returns

There are three possible explanations for why the aggregate market was so extraordinarily high in the tech-stock mania period. The first explanation is the honest mistake hypothesis: investors believed that future profits would be extraordinarily high and set prices accordingly. As it turns out, this high-profit scenario did not occur, but perhaps at that time it was reasonable to forecast high profits. According to the honest mistake hypothesis, there is no special reason to think that investors believed that expected returns were either particularly high or particularly low in, say, March 2000. The honest mistake hypothesis appears to be the preferred explanation for true believers in the efficient market hypothesis.

The second explanation is the low expected return hypothesis. Under this hypothesis, everyone knew expected equity returns had fallen, but they were happy to hold stocks despite their lower returns. At the time, some asserted that the equity premium had fallen for various reasons: the increasingly broad ownership of stocks, lower economic risk, higher risk tolerance, more institutions to share risk, and demographic changes. The low expected return hypothesis is a bit shakier on explaining the dramatic fall in stock prices from 2000 to 2002, but perhaps for some reason the equity premium rose again. Like the honest mistake hypothesis, the low expected return hypothesis is consistent with frictionless efficient markets.

The third explanation is overpricing: investors set prices too high, either knowingly or unknowingly, and this overpricing was obvious to some set of rational and informed investors at the time. One particular version of the overpricing hypothesis is that some optimistic investors extrapolated returns into the future, not realizing that the market was overvalued. It could also be that many investors knew the market was overpriced but chose to buy stocks anyway. In either case, because mispricing is eventually corrected, the overpricing hypothesis predicts that when stocks are overpriced, subsequent long-term returns will be low as the correction takes place, which is exactly what happened.

Figure 1 refutes the low expected return hypothesis and also casts substantial doubt on the honest mistake hypothesis. Reported expected one-year returns were wildly optimistic in early 2000 and fell sharply in 2001

and 2002. This is exactly the opposite pattern required for the low expected return hypothesis. In reality, investors reported expectations were undoubtedly simply chasing past returns. This pattern is confirmed by a strong pattern in mutual fund flows: inflows also chase past returns. This return-chasing pattern occurs both in the cross section (top funds have big inflows) and in the time series (all stock funds have inflows when the stock market has done well). Indeed, one piece of confirming evidence for Figure 1 is that net flows to stock funds during this period roughly match the pattern of reported expected returns. In summary, naïve adaptive expectations appear to be an accurate model for many investors.

The honest mistake hypothesis takes another hit in Figure 3, which shows in early 2000 that about 50% of the respondents thought the market was overvalued, while less than 10% thought it was undervalued. If pessimists outnumber optimists by 5 to 1, that does not sound like an honest mistake. It sounds like many knew the market was too high, but for some reason they went along with the ride.

3. Heterogeneous Expectations and Short Sale Constraints

One of the most important contributions of the paper is its examination of differences of opinion among investors. Heterogeneity is an increasingly important topic in asset pricing as well as in macroeconomics (see the paper by Mankiw, Reis, and Wolfers in this volume, for example).

Combined with short sale constraints, differences of opinion can create overpricing. Short sale constraints are anything that inhibits investors from short selling securities; in the case of tech-stock mania, the main constraint was probably that pessimists thought that shorting tech stocks was too risky. In Figure 1, it is clear in hindsight that NASDAQ was too high at 3000 in 1999. But anyone shorting NASDAQ then would have suffered severe losses as NASDAQ went to 5000 in March 2000. As hedge fund manager Cliff Asness has commented about short sale constraints, "Our problem wasn't that we couldn't short NASDAQ in 1999, our problem was that we could and did."

As Miller (1977) pointed out, with short sale constraints, stock prices reflect only the views of the optimists. Thus, differences of opinion plus short sale constraints can lead to overpricing. Now, one reason opinions may differ is that some investors are irrationally optimistic—this is what I would call the behavioral finance explanation, and there is substantial evidence to support this view of tech-stock mania. However, Harrison and Kreps (1978) showed that even when all investors are rational but have different beliefs, overpricing can occur. Let me give an example of

the Harrison and Kreps (1978) story. A remarkable property of this example, and one that fits well with the evidence given in the paper, is that everybody agrees that stocks are overpriced but they are still willing to hold stocks.

Suppose investor A and investor B have different beliefs about the prospects for the level of NASDAQ. Each investor knows what the other one believes, but they agree to disagree, so there is no asymmetric information. Now, it is a controversial issue in economic theory whether rational agents can agree to disagree, but let's leave that aside. Assume a simple setup with three dates, date 0, date 1, and date 2. For simplicity assume risk-neutral agents behaving competitively and a discount rate of zero. Assume also that there are sufficient numbers of type A and type B investors for each type to hold all of NASDAQ by themselves. Suppose it is currently date 0 and both investor A and investor B believe that NASDAQ is worth 2000 today. Specifically, they both believe that at date 2 it will be at 3000 with 50% probability and at 1000 with 50% probability. However, investor A thinks that at date 1, some news will arrive that will resolve all uncertainty, while investor B thinks there will be no relevant news released until date 2. This belief about the timing of news is the only disagreement between investor A and investor B (it is not necessary to state who, if either, is right in their beliefs). The Harrison and Kreps (1978) model has the remarkable property that, in the presence of short sale constraints, both investor A and investor B would be willing to hold NASDAQ at 2500 at date 0, despite the fact that they both think it is worth only 2000.

To get to this result, work backward from date 1, using the principle that with short sale constraints the optimist always sets the price. At date 1, if good news has arrived, then investor A will value NASDAQ at 3000, while investor B still thinks it is worth 2000; thus, the price will be 3000, investor A will hold all the asset, and investor B will hold none of it. If bad news arrives at date 1, the price will be 2000 and investor B will hold all of it. Because these two states happen with 50–50 probability, the date 0 expected price for date 1 is 2500. Thus at date 0, both investor A and investor B are willing to hold NASDAQ at a price of 2500. Although everyone thinks it is overvalued at date 0, they are willing to buy at date 0 because they believe they are following a dynamic trading strategy that will take advantage of the other guy. This example formalizes the notion of the greater fool theory of asset pricing. Note that, in this example, everyone agrees that long-term expected returns between date 0 and date 2 are low, and that a buy-and-hold strategy is a bad idea. If surveyed at date 0, both investor A and investor B would say that NASDAQ was overvalued relative to date 2 but fairly valued relative to date 1.

Key predictions of this story are that overpricing is highest when differences of opinion are highest, everyone agrees that prices are too high, and trading volume is high because everyone is following dynamic trading strategies. Vissing-Jorgensen's evidence supports the first two predictions. Figure 2 shows disagreement peaked in early 2000, around the time when stock prices peaked. Figure 1 shows that the majority of those who had an opinion about the market thought it was overvalued. The third prediction, about volume, is also supported by the events during this period. Not only did tech stocks have high prices, they also had very high volume. Volume on NASDAQ more than doubled between January 1999 and its peak in January 2001. Volume certainly seems like a key part of the tech-stock mania story, and one that the honest mistake and low expected return hypotheses cannot explain.

Another fact explained by the overpricing hypothesis is the high level of stock issuance that occurred in 1998–2000. One interpretation is that issuers and underwriters knew that stocks were overpriced and so rushed to issue. Although Vissing-Jorgensen's survey does not include issuers, evidence arising out of subsequent legal action against underwriters (such as emails sent by investment-bank employees) is certainly consistent with the hypothesis that the underwriters thought the market was putting too high a value on new issues. One way to think about issuance is as a mechanism for overcoming short sale constraints. Both short selling and issuance have the effect of increasing the amount of stock that the optimists can buy; both are examples of supply increasing in response to high prices. Suppose you think lamont.com is overpriced in 1999. One way to take advantage of this fact is to short the stock. In doing so, you are selling overpriced shares to optimists. This action is risky, however, because lamont.com might well double in price. A safer alternative action is for you to start a new company that competes with lamont.com, call it lamont2.com, and issue stock. This issue is another way to sell overpriced shares to optimists.

4. Sources of Heterogeneity

Why might differences of opinion be more pronounced in 1998–2000 than at other times? Miller (1977) presciently lists many of the characteristics that lead to differences of opinion. The first is that the firm has a short track record or has intangible prospects: "The divergence of opinion about a new issue are greatest when the stock is issued. Frequently the company has not started operations, or there is uncertainty about the success of new products or the profitability of a major business expansion" (p. 1156).

The second is that the company has high visibility, so that there are many optimists: "Some companies are naturally well known because their products are widely advertised and widely consumed. . . . Of course, the awareness of a security may be increased if the issuing company receives much publicity. For instance, new products and technological breakthroughs are news so that companies producing such products receive more publicity" (p. 1165).

Tech stocks certainly fit both these criteria. Stocks like Amazon or AOL were familiar to the investing classes who used them, but unlike other familiar products (such as Coca-Cola), they had a short operating history, so that optimists could construct castles in the sky without fear of contradiction by fact. Vissing-Jorgensen reports survey data on Internet use that seems to fit this story. Those who used the Internet thought Internet stocks had higher expected returns than other stocks, and they were more likely to include Internet stocks in their portfolio.

Vissing-Jorgensen also documents another interesting fact: young investors expected higher returns than older (and wiser) investors. This fact illustrates another key principle of behavioral finance: there's a sucker born every minute. Folk wisdom on Wall Street often claims that overvaluation occurs when young and inexperienced investors (who did not live through the last bear market) come to dominate. Apparently there is some truth to this claim.

5. Conclusion

We need to understand the events of 1998–2002 if we are to have any hope of understanding how stock markets work. Economists have expended a lot of effort studying episodes such as the stock market crash of 1987, the crash of 1929, and various alleged bubbles such as tulip mania. We are fortunate that this particular episode is well documented. Any satisfactory explanation will need to explain the high level of prices, the high level of volume, the high level of stock issuance, and the forces that prevented pessimists from correcting prices. Vissing-Jorgensen's paper is a good first step in arriving at an explanation. Her results suggest to me that tech stocks were indeed identifiably overpriced in the tech-stock mania period.

REFERENCES

Barber, Brad M., and Terrance Odean. (2002). Online investors: Do the slow die first? *Review of Financial Studies* 15:455–487.

Harrison, J. Michael, and David M. Kreps. (1978). Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics* 92:323–336.
Miller, Edward M. (1977). Risk, uncertainty, and divergence of opinion. *Journal of Finance* 32:1151–1168.

Discussion

Annette Vissing-Jorgensen agreed with the discussants that data reliability is an issue. However, she mentioned several pieces of evidence to support the fact that responses in the survey are not entirely noise. For example, respondents with higher market expectations put more in stocks than those with lower market expectations. Vissing-Jorgensen also suggested that the reliability of the survey data could be examined by looking at the length of time respondents took to complete the survey.

Rick Mishkin expanded on the comments of Owen Lamont about the structure of investment banking and the stock market bubble. He noted that bubbles occur when there is nonfundamental pricing of assets, but that this nonfundamental pricing can in fact be driven by institutions. He held that the U.S. stock market bubble of the 1990s was driven in part by a conflict of interest in investment banking between advice and sales. He mentioned that in Scandinavia and Japan, a combination of financial liberalization, government safety nets for the financial system, and poor prudential supervision had led to real estate bubbles. He recommended that policymakers should focus on institutions and regulation to understand and learn how to prevent bubbles. On this issue, Vissing-Jorgensen noted that the datasets described by John Campbell could be used to examine how price movements subsequent to earnings announcements depend on the independence of the analysts concerned.

Annamaria Lusardi was concerned by the implications of the paper's findings on the speed of learning for wealth accumulation. She speculated that the consequences for retirement savings could be substantial if investors do indeed learn very slowly. Vissing-Jorgensen responded that the fact that payoffs increase with wealth could be explained by costs of obtaining information. She noted that if this were indeed the case, it would suggest a role for investor education.

Mark Gertler was curious about whether survey respondents could be distinguished by frequency of trading. He pointed out that investors who adopt buy-and-hold strategies have less of an incentive to pay attention to the market and hence might have beliefs determined by outdated

information. Vissing-Jorgensen said that it might be possible to perform this exercise eventually.

Kjetil Storesletten questioned the suggestion that short sale constraints could explain investor underperformance. He noted that if two-thirds of people expect the market to fall, there should be a substantial demand for short sales, and that this should lead to financial innovation. He suggested that it might not have been clear to investors at the time that there was a bubble. Vissing-Jorgensen responded that investors might not have been comfortable shorting the market because they thought that it would continue to go up before going down. She said that the bigger question was, Why are investor horizons so short? Owen Lamont commented that, for whatever reason, there are very few institutions allowing investors to go short. His view was that there was little demand for such institutions because psychological constraints made people reluctant to take negative positions, although they were quite happy to take zero positions.

Eva Nagypal was curious about how savings rates correlated with investor performance and behavior. Vissing-Jorgensen responded that those expecting high portfolio returns, i.e., of at least 20%, did save more but that otherwise there was no correlation between performance and savings.