

Expecting a Stock Market Miracle*

Markku Kaustia

Aalto University School of Business, Department of Finance

Joni Kokkonen

Católica-Lisbon School of Business and Economics

Vesa Puttonen

Aalto University School of Business, Department of Finance

March 23, 2016

Abstract

We develop two new methods for calibrating subjective expectations regarding the return generating process (RGP) of financial assets without resorting to noisy realized returns. Using finance professionals' expectations of average and extreme returns, volatilities, and probabilities of stocks beating bonds, we investigate what these expectations imply of other key aspects of the RGP, namely stock-bond correlation, stock mean-reversion, and tails of the return distribution. We find a high degree of confidence in stocks beating bonds, with moderate returns and volatility, and no extreme returns. For most subjects these expectations imply implausible RGP's given established empirical facts, or else, a miracle.

Keywords: Return expectations, Optimism, Miscalibration

* E-mails: markku.kaustia@aalto.fi, joni.kokkonen@ucp.pt, and vesa.puttonen@aalto.fi. We thank John Paul Broussard, Jan Ericsson, Markku Lanne, Peter Sørensen, and seminar participants at Aalto University, CFA Institute, Nordea Bank and University of Gothenburg for their comments.

1. Introduction

According to a widely held belief stocks are very likely to outperform bonds over long horizons. In addition to higher expected returns, a possible mean reversion in stock returns would further help stocks to outperform (Poterba and Summers, 1988). The historical record of stocks over bonds is indeed impressive. For example, Siegel (2002) reports that the percentage of 20-year periods over which stocks outperformed bonds is 92% in the US during 1802-2001. To produce comparable outperformance going forward, the equity premium would need to be high, or mean-reversion relatively strong, or both. This may be challenging, as forward-looking models tend to suggest equity premia much lower than the realized returns in the US market.¹ Also the tendency of stock returns to mean revert has received mixed support.²

How optimistic are market participants regarding the future outperformance of stocks over bonds? Under what assumptions are expectations of future outperformance of stocks internally consistent with return expectations? To answer these questions, we develop two new methods for assessing key implicit assumptions embedded in subjective expectations of asset returns, namely stock-bond correlation, stock mean-reversion, and fatness of tails of the return distribution. Our approach differs from existing studies in that it allows us to infer these results directly from the subjects' estimates and so we do not need to rely on comparing return expectations to noisy realized returns.

¹ Claus and Thomas (2001), Fama and French (2002), Ibbotson and Chen (2003) and Donaldson, Kamstra, and Kramer (2010) employ different forward-looking models, and all estimate a US equity premium in the 3-3.5% range.

² Poterba and Summers (1988) find mean reversion in US and international stock returns. Jorion (2003) expands the sample significantly and finds no robust evidence for mean reversion, similar to Dimson, Marsh, and Staunton (2004). Barberis (2000) and Pastor and Stambaugh (2012) argue that stocks are actually riskier from an investor's perspective over long horizons because of the uncertainty concerning expected returns.

The first method models the joint process of stock and bond (or bill) returns as a vector autoregressive (VAR) process. It relates subjective estimates on returns and volatilities to estimates of the probability that stocks beat bonds (bills) via unobserved implicit stock-bond correlation and stock mean-reversion. As an example, consider an expectation of an 8% mean return and 16% volatility for stocks, and 3.5% and 7% for bonds, respectively, as well as a belief that stocks outperform bonds 80% of the time for a 10-year horizon. These numbers would be consistent with certain combinations of the stock-bond correlation and stock mean-reversion, but for other combinations the expectations will be mis-calibrated, i.e., the 80% outperformance likelihood will be too optimistic or too pessimistic given other parameters.

The second method relates volatility expectations and expectations of sample minimum and maximum returns to unobserved implicit tail index of the return distribution using extreme value theory. For example, continuing with the previous 8% mean return and 16% volatility expectation for stocks, say someone expects stocks to return 35% in the best year, and -20% in the worst year over the next 10 years. This would be consistent with a particular kurtosis, or a tail index, for the return distribution. For other values of the tail index, the range between the minimum and maximum estimates will be either too narrow or too wide.

To apply these methods to subjective estimates of a relevant group of people, we survey finance professionals such as analysts and investment advisors in practitioner seminars obtaining very high response rates.³ The results of calibrating the models to subjects' expectations shows extreme

³ A growing number of studies turn to survey data for analyzing aspects of the return generating process that are difficult to glean from market data. See Frankel and Froot (1987) and Shiller (1990) for early examples, and Hommes et al. (2005), Ben-David, Graham, and Harvey (2013), Greenwood and Shleifer (2014), and Amromin and Sharpe (2014) for more recent work.

optimism regarding stocks outperforming bonds with quite moderate stock return expectations *per se*. Rather than uncovering realistic estimates of mean-reversion implied by these expectations, we found it impossible to reconcile the moderate return expectations with a high degree of confidence in stocks' outperformance for most subjects. It would require extreme degrees of stock return mean reversion, or, alternatively extreme stock-bond correlation, and sometimes both. This result is more pronounced for long-horizon (20-year) expectations but is found also for five and 10-year horizons. Comparing samples collected before and after the financial crisis of 2008 shows that the subjects have become even more optimistic about the relative performance of stocks versus bonds, again, given their return and volatility expectations of the respective asset classes.

Why do the professionals we survey hold such seemingly inconsistent beliefs? The securities industry has traditionally been optimistic about stocks, perhaps because equity-based investment products tend to produce more fee income (see, e.g., Gennaioli, Shleifer, and Vishny, 2015). It would be easy to express optimistic beliefs about stocks by expecting a high average annual return, say, 15%. However, this could be a hard sell, even to the person themselves expressing this forecast. A more subtle and seemingly more realistic way of optimism towards stocks might involve a lower return expectation, yet an exaggerated belief in stocks outperforming other assets. One way to describe this thinking is as a Halo effect for stocks: stocks are universally great no matter what the relative expected returns are. An example of this view taken to an extreme is the book "Dow 36,000" (Glassman and Hasset, 1999), where the authors argue that the risk premium of stocks over bonds should in fact be zero because there is no real risk of stocks ever underperforming.

If the subjects overestimate the chances of stocks outperforming conditional on their return estimates, do they also underestimate the risk of stocks conditional on their volatility estimates?

To investigate this issue, we compare the subjects' estimates of the minimum and maximum (annual) returns on stocks over a 20-year horizon to their volatility estimates over the same horizon. Using Extreme Value Theory to model these expectations within each subject, we show that even for the most thin-tailed distribution that we consider (the normal distribution), about two thirds of the subjects underestimate the minimum and maximum stock returns conditional on their own volatility forecast. Assuming more realistic, and thus more fat tailed distributions, increases these proportions significantly. Subjective expectations are thus either miscalibrated or imply return distributions with substantially thin tails. If one is willing to rule out the possibility that the true return generating processes actually would generate thin-tailed distributions, then one can conclude that the subjects are miscalibrated. Specifically, they underestimate the chance that stocks would provide a large negative surprise, consistent with a Halo effect. They also underestimate the chance of large positive surprises, perhaps driven by need for some symmetry in the expectations on the subjects' part.

We contribute mainly to three strands of literature. First, a growing literature that explores finance professionals' return expectations and possible behavioral biases. For example, Ben-David, Graham, and Harvey (2013) document that Chief Financial Officers are miscalibrated in their return expectations. Bodnaruk and Simonov (2015) analyze the private portfolios of mutual fund managers and conclude that they do not outperform non-expert individual investors, nor are they more diversified. In an audit study, Mullainathan, Noeth, and Schoar (2012) argue that financial advisors may even reinforce their clients' behavioral biases that are in their interests. Other papers that present evidence that also finance professional may suffer from behavioral biases include Kaustia, Alho and Puttonen (2008), Haigh and List (2005), and Glaser, Langer, and Weber (2007).

Second, our results relate closely to the specific biases of cognitive dissonance and groupthink.⁴ Bénabou (2013) proposes a model of groupthink, where agents face a tradeoff between realism (accepting negative public signals on a project's value) and denial (ignoring these signals). As an implication of the model, the author shows that contagious wishful thinking can lead to financial market frenzies and crashes. Goetzmann and Peles (1997) argue that mutual fund investors' memories have a positive bias that is conditional on previous investor choices. That is, investors tend to be overly optimistic of past fund performance, and this phenomenon can partially explain the convex nature of the return-flow relation. Chang, Solomon, and Westerfield (2016) conjecture that cognitive dissonance among individual investors can generate the disposition effect in non-delegated assets such as stocks. Antoniou, Doukas, and Subrahmanyam (2013) propose that cognitive dissonance can contribute to the slow diffusion of signals opposing current investor sentiment and thus be a driving force behind the profitability of momentum strategies. Cheng, Raina, and Xiong (2014) analyze the personal home transaction data of mid-level managers in securitized finance between 2004 and 2006 to study whether these professionals were aware of the problems in the housing markets before the burst of the bubble. The authors do not find support for any type of market timing (some groups of agents were actually aggressively increasing their exposure) and conclude that understanding the interaction between incentives and beliefs is crucial, especially in environments that may foster groupthink and cognitive dissonance.

Finally, we complement the fairly recent literature that uses survey data to explore the return expectations of individual investors. Greenwood and Shleifer (2014) show that investors'

⁴ Cognitive dissonance refers to the tendency of altering one's beliefs to be consistent with past actions (Festinger (1957)). Groupthink refers to the harmful conformity to group values that can lead to symptoms such as, for example, collective denial and willful blindness (Bénabou (2012)).

expectations tend to be extrapolative, and negatively correlated with model-based expected returns. To account for these observations, Barberis, Greenwood, Jin and Shleifer (2015) propose a consumption-based asset pricing model where a fraction of agents have extrapolative expectations. Other related papers include Amromin and Sharpe (2014) who find that survey-based expectations are negatively correlated with proxies for a ‘recession’ state, and Bacchetta, Mertens and van Wincoop (2009) who argue that the predictability of excess returns is related to the predictability of expectational errors.

In the remainder of this paper, Section 2 introduces the methodology and Section 3 discusses the data on finance professionals’ expectations. Section 4 presents the results on the relative performance of stocks and bonds, and Section 5 discusses results on tail expectations. Section 6 returns to the issue of relative performance expectations in a more recent data set after the financial crisis. Section 7 concludes.

2. Methodology

This section describes the two new methods of uncovering implicit parameters in subjective expectations of financial asset returns.

A. Relative Performance Expectations

We model the joint process of stock and bond returns as a bivariate vector autoregressive (VAR) process assuming that stock and bond returns do not depend on each other’s lagged values, as follows

$$\begin{bmatrix} r_{S,t} \\ r_{B,t} \end{bmatrix} = \begin{bmatrix} \mu_S \\ \mu_B \end{bmatrix} + \begin{bmatrix} \phi & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_{S,t-1} - \mu_S \\ r_{B,t-1} - \mu_B \end{bmatrix} + \begin{bmatrix} \epsilon_{S,t} \\ \epsilon_{B,t} \end{bmatrix} \quad (1)$$

where ϕ is between -1 and 1, and the error vector follows a bivariate normal distribution with mean vector of zero and covariance matrix Σ . Similar specifications have been used by Hodrick (1992), Barberis (2000), and Campbell, Chan, and Viceira (2003), and others. We set the autoregressive (AR) coefficient for bonds to zero. In unreported analysis we relax this assumption and obtain results similar to the ones we report.

We then use subjective estimates of the means and volatilities as the parameters in the model, and use simulations to investigate what parameters of mean reversion and correlation would be needed to match the probability of stocks outperforming bonds to those given by the same subjects. One can think of this as producing “implied” correlation and mean reversion taking the return and outperformance expectations as given.

B. Tail Expectations

A key insight in Extreme Value Theory (EVT) is that the extreme outcomes from any distribution (called parent distribution) conform to only three basic types of extreme value distributions, and the assignment to those types depends on the fatness of the parent distributions’ tails (see Leadbetter, Lindgren, and Rootzén, 1983, and Embrechts, Klüppelberg, and Mikosch, 1997, for reviews on EVT). The novelty of this methodology in our context is that it allows us to compare the respondents’ direct estimates of the minimum and the maximum return to their theoretical expected values given subjects’ volatility expectations. We thus avoid using noisy realized return outcomes.

Let X_1, \dots, X_n denote a sample of i.i.d. random variables with a common distribution function $F(x)$, and let $M_n = \max\{X_1, \dots, X_n\}$ denote the maximum of this sample (the case for the minimum

is analogous). If we can find sequences of normalizing constants a_n and $b_n > 0$ so that the sequence of normalized maxima converges in distribution, that is,

$$\Pr \left[\frac{M_n - a_n}{b_n} \leq x \right] = F^n(b_n x + a_n) \xrightarrow{d} G(x), \quad x \in \mathbb{R} \quad (2)$$

then F is said to belong to the maximum domain of attraction of G , denoted by $F \in D(G)$. The normalizing constants, a_n and b_n are called the location parameter and the scale parameter, respectively. Then, G must belong to one of three possible limiting distributions of maxima: Gumbel (Type I), Fréchet (Type II), or Weibull (Type III).

The limiting distribution of the (appropriately standardized) maxima, regardless of the parent distribution, can thus take only one of the three specific types. Fréchet type distributions have a polynomially decaying tail and are thus suited to model heavy-tailed distributions. Common distributions belonging to the Fréchet domain of attraction include the Pareto and stable distributions as well as the Student's t . Distributions belonging to the Gumbel type limit laws, such as the normal distribution, have an exponentially decaying tail. The Weibull limit corresponds with parent distributions with a finite end point. The three types of limiting distributions can be represented by a single equation as the Generalized Extreme Value (GEV) distribution with distribution function

$$\Pr(X \leq x) = GEV(\xi, \mu, \sigma) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \quad (3)$$

where

$$1 + \xi \frac{x - \mu}{\sigma} > 0, \quad \xi \neq 0$$

and μ is a location parameter, σ is a scale parameter, and ξ is a shape parameter.

To model subjective tail expectations, we use the expected value of the GEV distribution, given by

$$E(X) = \begin{cases} \mu + \frac{\sigma}{\xi} (\Gamma(1 - \xi) - 1), & \xi \neq 0 \\ \mu + \sigma\gamma, & \xi = 0 \end{cases} \quad (4)$$

where $\Gamma(\cdot)$ is the gamma function and γ is the Euler-Mascheroni constant ≈ 0.577 .

As discussed above, EVT assumes i.i.d observations. The independence condition most likely holds for expectations data collected from individual respondents. However, subjective probability distributions are unlikely to be identical. To correct for non-identical distributions, we standardize the maxima and the minima as

$$Z_{M,i} = \frac{M_i - \mu_i}{\sigma_i} \quad (5)$$

where $Z_{M,i}$ is the standardized minimum (maximum), M_i is the original minimum (maximum), and σ_i is the volatility estimate of respondent i . We then consider various alternative distribution types for the parent distribution of returns, such as fat tailed t-distributions and a thin tailed normal distribution.

3. Data on subjective expectations

To apply the methods we collect data on subjective expectations using field surveys of financial market professionals. The data collection is done in connection to internal investment seminars organized by three different asset management firms, two in Finland and one in Sweden. The participants come to the seminars without knowing that data on expectations will be collected. At the beginning of the event, after all the participants have arrived and are seated in an auditorium they are asked to participate in a voluntary study on stock market return expectations. We were

able to reach the entire audience attending these events, and get a 75% response rate. This is very high compared to the response rates in conventional surveys which are typically in the range of 5 to 10%. The most frequent job title among the respondents is financial adviser (19%). Other typical backgrounds include analyst, private banker, investments expert, broker, wealth manager, and stock specialist. Director and manager levels are also represented. In the initial sample collected in 2004 we have 100 responses.

We also collect data in 2009, 2010, 2011, and 2012 which we pool into a single post-2008 sample. The total number of respondents in these later surveys is 251. The subjects in the post-2008 sample are, on average less experienced than those in the initial 2004 sample (mean experience of 5 versus 9 years). To address this systematic difference we also analyze a subsample of experienced professionals with at least 5 years of financial markets experience.

Prior to filling out the questionnaires, the participants were given detailed instructions on what is being asked, and the relevant terms were defined. Specifically, we explained that geometric real total return of stocks means that the invested capital is compounded and that this corresponds to how investment returns are usually measured, real return is what is left of nominal returns after inflation has been deducted, and total return means capital gains plus reinvested dividends, and that taxes and other costs are ignored. One of the authors was present in all sessions to answer any questions. The survey took about 15 minutes to complete, of which 5 minutes were spent on instructions.

We rely on the intrinsic motivation of the subjects to perform well and offer no compensation for participating in the study. We believe this has little impact on the quality of the answers. People with low intrinsic motivation can elect not to participate, as some did. Furthermore, conditional on

participating, we expect the subjects to be motivated to perform well in a task that is related to central concepts in their work. There was also no incentive to lie.

The surveys contained the following questions, asked separately for the returns on stocks, bonds, and bills:

- Your estimate of the expected annual real return over the next 20 years?
- Your estimate of the annual volatility of the real returns over the next 20 years?
- Your estimate of the worst year's real return over the next 20 years?
- Your estimate of the best year's real return over the next 20 years?
- Your estimate of the probability of stocks outperforming bonds (and similarly for bills), asked for three different horizons (5, 10, and 20 years).

Table 1 gives the descriptive statistics for the base sample of 100 respondents. The mean expected real return on stocks is 7.6% and the mean estimate for volatility is 16.3%. Dimson, Marsh, and Staunton (2002) report that the geometric average real return for the US stock market over the period from 1900 to 2000 was 6.7% with a volatility of 20.2%. The corresponding figures for Sweden were 7.6% and 22.8%, respectively. Overall the subjective expectations thus seem reasonable. However, as we will show later, reasonable estimates of expected returns can still lead to extreme optimism regarding the expected performance of the stock market.

In column six, $\hat{\sigma}$ is an approximation of the volatility calculated for respondent i as⁵

⁵ See, for example, Keefe and Bodily (1983)

$$\hat{\sigma}_i = \sqrt{\left(\frac{MIN - MAX}{6}\right)^2} \quad (6)$$

Since this volatility approximation is based on the respondents' estimates of minimum and maximum, it allows us to compare the consistency of the respondents' direct estimates of volatility with their estimates of the minimum and the maximum. In general, the respondents' direct estimates of volatility are higher than the min-max approximations from (6). This already indicates that the respondents' estimates of the range between the maximum and minimum returns might be too narrow to be consistent with their volatility estimates.

4. Results on the Relative Performance of Assets

The estimates of mean returns for the different asset classes imply that the professionals in our sample are fairly neutral on stocks. We next analyze whether their estimates of the means and volatilities of different asset classes are consistent with their estimates of the probability that stocks outperform bonds.

First, as an illustration, consider the probability that stocks outperform bonds given a fixed equity premium. Figure 1 shows the probabilities implied by the VAR model given in equation (6) for a horizon of 20 years. The probabilities are calculated for different combinations of stock-bond correlation as well as autoregressive coefficient in the case of stocks. The upper panel shows the probabilities of stocks beating bonds given a risk premium of 6.6% (based on Dimson, Marsh, and Staunton, 2002). The lower panel assumes a risk premium of 4.0%. In general, the probability that stocks outperform bonds increases for the higher risk premium and for lower values of the AR coefficient. The average probability implied by the model is 84.0% (70.2%) for the higher (lower) risk premium. The historical frequency reported in Siegel (2002) for the 20-year horizon of stocks

beating bonds is 92% for the period of 1802 to 2001. Reaching this 92% consistent with Siegel (2002) requires a moderate negative autocorrelation for stocks with an equity premium of 6.6%. Given an equity premium of 4%, however, would make the 92% probability unattainable even with extreme levels of autocorrelation and stock/bond correlation.

Next, we turn to analyzing the subjects' estimates of the probability of stocks outperforming bonds. The aim is to compare these directly elicited estimates of probability, to ones implied by the subjects' return and volatility expectations. In addition, as we saw earlier, the probability of stocks outperforming bonds depends on the stock-bond correlation, and the autocorrelation of stock returns. We did not ask the subjects to estimate these parameters. Instead, the idea is to see what must be assumed of these parameters to make the subjects' outperformance estimates compatible with their return and volatility estimates.

Figure 2 gives the histogram of the respondents' estimates of the probability that stocks outperform bonds. Three forecast horizons are considered: 5, 10, and 20 years, respectively. As expected, estimates on the probability of stocks outperforming bonds (bills) become higher for longer horizons. At five years, the mean (median) probability is 56% (50%). When the horizon is increased to 20 years, the mean (median) increases to 82% (90%). Overall, the respondents are highly optimistic about the relative performance of stocks. The 75th percentile at 20 years is 100%.

Figure 3 shows the histograms of the best-fit parameter pairs (correlation and mean reversion) corresponding to the probabilities given by the respondents. The probability estimates lead to a corner solution in a large proportion of the cases. Roughly for one third of the respondents, the implied correlation (as well as mean reversion) needed to match their subjective probabilities is between 0.8 and 1 for the 20-year horizon (Panel (a)). For the 10-year horizon (Panel (b)), this proportion is slightly lower. Extreme negative cases are also well represented. The estimates of 10

respondents require almost perfectly negative correlation. Panel (c) gives the corresponding histogram for the 5-year horizon, and as before, both negative and positive extremes have the highest frequency. In this case almost half of the respondents' estimates would imply a correlation exceeding 0.8.

Are these implied stock-bond correlations and stock return mean reversion parameters realistic? First, consider the stock-bond correlation. Estimates based on realized returns typically show moderate positive long-term correlations.⁶ Campbell and Viceira (2005) calculate the stock-bond correlation implied by a VAR(1) model for different horizons. They show that correlation increases for short horizons, but starts declining when the horizon exceeds 10 years. In their model, the correlation is between 0.5 and 0.6 for horizons of five and 10 years, respectively. At 20 years, the implied correlation given by their model is slightly below 0.4. Implied correlations in excess of 0.8 in absolute terms thus seem highly unrealistic.

Second, consider autocorrelation of stock returns. Fama and French (1988) and Poterba and Summers (1988) find negative autocorrelation (mean reversion) in stock returns over long horizons. In the model of Campbell and Viceira (2005), predictability from the dividend yield induces mean reversion in stock returns. However, Kim, Nelson, and Startz (1991) argue that mean reversion in stock returns was a pre-war phenomenon. Using long data series from 16 countries, Dimson, Marsh, and Staunton (2004) find the 3-year auto-correlation of stock returns to be -0.07 on average. It is negative for 11 countries, but not statistically significant for any. Though the case

⁶ Shocks to discount rates affect stock and bond returns similarly, while cash flow shocks do not affect (government) bond returns. This implies a positive correlation between stock and bond returns. However, empirical estimates over some samples have shown slightly negative values. For example, using data from 1990 to 2012, Carhart et al. (2014) document a correlation coefficient of -0.05 between US stock and bond returns, where bonds include both treasury and corporate bonds.

for mean reversion may appear weak, we nevertheless entertain the possibility that it exists. However, the extreme values of the autoregressive coefficients implied by our subjects' estimates are not consistent with realistic mean reversion.

Using the estimates discussed above as benchmarks, we calculate the proportion of respondents for whom the implied correlation, autocorrelation, or both, given by the simulation model lies outside a realistic, yet rather wide range. For correlation we use a range of $[0, 0.7]$ and for autocorrelation a range of $[-0.1, 0.1]$. Panel A of Table 2 shows the results for the full sample. For the 5-year horizon, virtually all (99%) respondents have implied coefficients that are outside of at least one of the two ranges. For the 10 and 20-year horizons the corresponding fraction is 96%. The proportions for which both coefficients lie outside the range are also substantial: 67% for the 5-year horizon, 70% for the 10-year horizon, and 72% for the 20-year horizon.

In the analysis reported on Panel B of Table 2 we consider two subsamples with the hope of identifying subjects who are more consistent in their estimates. First we consider the effect of financial expertise. The idea is that the more sophisticated finance professionals may have a better intuition on the interplay of expected returns and probability of outperformance, and may thus be more consistent in their responses. To investigate this issue we single out professionals with more than five years of work experience, and label them *experienced pros*. Second, we single out respondents with a relatively higher (greater than 20%) volatility expectation for stocks. As we discussed earlier, it seemed like the average respondent underestimated stock volatility. We limit to the 10-year horizon in these analyses. Counter to our expectation, both of these groups exhibit more inconsistency compared to the full sample. For example, 77% of the experienced pros, versus 70% in the full sample, need both of the implied parameters to be outside of realistic ranges to

match their belief in stocks beating bonds. Extreme stock market optimism does not diminish with experience.

Figure 4 repeats the simulation exercise with no mean reversion. Thus, now we simulate from a random walk with drift and correlated error terms. The results are similar to those with mean reversion. Again for the majority of respondents, either extreme negative or positive correlation is needed to match their subjective probability estimates. For half of the respondents, matching their estimates requires positive correlation between 0.8 and 1. Using the same range as for the VAR(1) model above (correlation between 0.2 and 0.7), less than 5% of the implied correlation parameter estimates would fit in the range.

The results in this section thus show that for a majority of the respondents, their estimates of the probability that stocks outperform bonds are much higher than implied by their estimates of the expected returns and volatilities. Based on the return and volatility estimates, the subjects seem to be fairly neutral on bonds but still they expect stocks to outperform with an extremely high probability. This discrepancy leads to extreme values of correlation and autocorrelation that are needed to match the probabilities.

5. Results on Tail Expectations

We begin this section by investigating subjects' estimates on minimum and maximum returns as a whole, as if they were independent draws from the same distribution. Under this assumption, we can characterize the tails of the asset class distributions based on expectations only. We then move to the level of an individual subject, and compare and contrast each subject's direct estimates of extreme returns, to those derived from Extreme Value Theory given the subject's expectations on volatility. This shows whether subjects' expectations of extremes are in line with their volatility

expectation. More specifically, it tells us what kind of an underlying return distribution we must assume to bring these two types of expectations in line.

A. The Return Distribution Implied by Pooled Estimates of Extreme Returns

The idea is to treat each respondent as providing a single tail event observation for each of the tails of each asset class. We assume these data points are independent, and identically distributed after standardization (see section 3.A.). After pooling the data over all respondents, we have two data series for each asset class, one for the minimum, and one for the maximum. We then fit the Generalized Extreme Value (GEV) distribution to each data series. The outcome of interest is the shape parameter of GEV characterizing the tail thickness of the underlying distribution of returns implied by these expected minima and maxima. The normal distribution has a shape parameter of zero. Negative values mean that tails are thinner than with the normal distribution, and positive values signify fatter tails. A Student's t distribution with n degrees of freedom has a shape parameter equal to $1/n$. For example, the shape parameter of a $t(4)$ distribution is then 0.25.

Table 3 gives the maximum likelihood estimates of the parameters and the 95% confidence intervals. The shape parameters for stocks are quite reasonable (0.22 for the minimum and 0.31 for the maximum), roughly corresponding to a t -distribution with four degrees of freedom, i.e., relatively fat tails. The shape parameters for both the minimum and the maximum of bills (0.60 and 0.69) are considerably higher than for stocks, and close to zero for bonds (0.13 and 0.04). There are no statistically significant differences in the shape parameter estimates of the minima and the maxima for any of the assets. Overall, these results show that when each respondent is considered as providing a single tail event observation, the implied return distribution is fat tailed

for stocks and bills. However, this does not say anything about the internal consistency of the expectations on a level of each subject, a question we turn next.

B. Respondent Level Calibration Analysis

We next analyze the relation between the respondents' estimates of max/min return (direct estimates) and compare them to the ones implied by their own volatility estimates using Extreme Value Theory (EVT estimates). We are interested in whether these estimates are internally consistent, or whether, for example, the direct estimates are too small given the volatility estimate. In this case, a subject would be surprised by extreme return realizations more often than they should. Direct estimates that are closer to zero, that is, smaller in absolute value, than the EVT estimates imply subjective probability distributions that are too narrow. To calculate the proportions of direct estimates that are too small in this sense, we use several different parent distributions to avoid making any assumptions on the "correct" underlying distribution of asset returns. The distributions we consider are the normal distribution, and three different Student's t distributions with the degree of freedom ranging from three to five. These distributions differ in the thickness of their tails with the normal distribution having the thinnest tail and the Student's t with three degrees of freedom the thickest tails.

Table 4 gives the proportion of respondents whose estimate of the minimum, maximum, or both, are too small (in absolute terms) with respect to the EVT estimates for a given parent distribution. As the tails of the underlying return distribution get fatter, the EVT estimates get further from zero. The most fat-tailed distribution we consider is the t -distribution with three degrees of freedom, $t(3)$. Assuming this distribution, the results show that 82% of the respondents gave smaller direct min/max estimates for stocks than are implied by EVT given the $t(3)$

distribution. Even more respondents give such estimates for bonds (88%), particularly for the positive side of bond returns. While the $t(3)$ might be quite a realistic model of asset returns, it is of course difficult to conclude whether the respondents rationally expect the return generating processes to be less fat tailed, or whether the respondents' expectations are miscalibrated. Stronger conclusions can be made when we move toward thinner tailed distributions, such as the normal distribution. The EVT estimates are then closer to zero, and so the proportion of respondents whose direct estimates are smaller than the EVT estimates naturally gets smaller as well.

Assuming a normal distribution for stocks, the top panel of Table 4 shows that 62% (68%) of the direct minimum (maximum) estimates are still too low compared to the EVT estimates. For 52% of the respondents, this is true for both minimum and maximum. The true data generating process for stock returns almost certainly implies fatter tails and more frequent extreme events than the normal distribution (see, e.g., Fama, 1965 and Longin, 1996).⁷ Since it is very unlikely that stock returns actually follow a distribution with such thin tails, this result shows that a considerable fraction of the professionals in the sample are miscalibrated. Our methodology provides a way of inferring this result directly from the respondents' estimates, without having to compare expectations to noisy realized returns.

The results are similar for bonds (Panel B) and bills (Panel C). Assuming normally distributed returns the proportion of respondents underestimating the minimum (maximum) is 66% (77%) for

⁷ Fama (1965) shows that the log-price changes of stocks in the Dow Jones Industrial Average follow stable Paretian distributions with fatter tails than the normal distribution. Longin (1996) shows that the extreme returns of the most traded stocks in the New York Stock Exchange follow a fat-tailed Fréchet distribution. Bali (2003) demonstrates that this result holds also for the extreme changes in US Treasury yields. Jondeau and Rockinger (2003) study a sample of 20 countries (developed and emerging), and conclude that the thin-tailed normal distribution is rejected for all countries.

bonds, and 67% (67%) for bills, respectively. Column 4 shows the z -statistic for the difference in proportions test that can be used to test for the symmetry of extreme expectations. For bonds the proportion of respondents underestimating the maximum is significantly higher when assuming any of the t -distributions. The differences for stocks and bills are not significant.

Figure 5 illustrates these results by plotting the ordered minimum and maximum estimates (from the highest to the lowest) together with the theoretical expected values of normal and $t(4)$ parent distributions. It is worth noting that when looking for miscalibration we only focus on the estimates that lie below the theoretical expected value. Of course, some estimates clearly above the theoretical expected value could also be classified miscalibrated. However, such estimates could be rationalized by subjective expectations implying very fat tails, reflecting, for example, the chance of an extremely low return occurring due to a market crash. Our approach is conservative in the sense that we use parent distributions that very likely have thinner tails (the normal distribution) or realistic tails (Student's t distributions) compared to the empirical distribution. Even with this conservative approach, we are still able to establish that most professionals' expectations imply even thinner tails.

The results presented so far imply that a large fraction of the professionals in the sample are miscalibrated in the sense that their estimates of the minimum and maximum are too close to zero given their volatility estimates. This statement can be made based on expectations alone, without resorting to realized returns. However, a look at the realized returns of the MSCI World index in 2008 and 2009, shows that the minimum (maximum) estimates of yearly return have already been exceeded for 97% (73%) of the respondents, and for both minimum and maximum for 72% of the subjects.

To examine whether the subjects' miscalibration results from too small estimates of the minima and maxima, or possibly from too large estimates of volatility, we repeat the exercise using estimates of historical realized volatility instead of the respondents' own volatility estimates. The estimates of historical volatility are based on real returns on the US market and are from Dimson, Marsh, and Staunton (2002, p. 60). The estimates are $\hat{\sigma}_{Stocks} = 20.2\%$, $\hat{\sigma}_{Bonds} = 10.0\%$, and $\hat{\sigma}_{Bills} = 4.7\%$. Table 5 shows that when historical realized volatility is used in place of the respondents' volatility estimates, the proportions of direct max/min estimates that fall below the EVT estimates actually increase. This makes sense, as the average stock market volatility expectation among the subjects is lower than the historical volatility (mean of 16.3%, see Table 1). This shows it is not the subjects' volatility estimates, but rather the narrow direct max/min estimates that drive the results.

Finally, Table 6 repeats the same exercise using both historical volatility and historical mean returns.⁸ The respondents' estimates of the expected return on stocks (mean of 7.55%) are fairly close to the historical geometric average of 6.7%. However, even so, the results are consistent with those presented above and the proportion of respondents whose minimum or maximum estimates (or both) fall below the expected value exceeds 80% in all cases considered, and is mostly over 90%.

In general, the results in this section imply that the range between the respondents' estimates of minima and maxima is too narrow to be consistent with the respondents' own volatility estimates. An alternative explanation could be that subjects are well calibrated, but expect the underlying

⁸ Again, the estimates of the mean returns are taken from Dimson, Marsh, and Staunton (2002) and are $E(R_{Stocks}) = 6.7\%$, $E(R_{Bonds}) = 1.6\%$, and $E(R_{Bills}) = 0.9\%$.

distribution of returns to have thin tails. Such expectations would nevertheless be in contrast to what is empirically known about returns.

6. Did the Financial Crisis Curb Excessive Optimism?

Prior literature suggest that experiencing negative stock returns reduces optimism (Vissing-Jørgensen, 2003 and Ben-David, Graham, and Harvey, 2013). To investigate this, we use responses collected in 2009 to 2012, that is, during and after the financial crisis and the European debt crisis. The respondents were asked to consider a horizon of 10 years and give their estimates of the probability that stocks outperform bonds and the probability that stocks outperform bills.

Table 7 gives the descriptive statistics for these data. The estimates of expected real returns are clearly lower for all assets compared to the pre-crisis sample. However the mean expected equity premium has increased from a pre-crisis level of 3.4% to 4.5%. In line with a higher expected equity premium, the probability estimates of stocks outperforming bonds or bills have also grown. The median (mean) probability estimate of stocks outperforming bonds is now 90% (86%).

Are the outperformance expectations now quantitatively better in line with the return and volatility expectations? Figure 6 shows the histogram of the coefficients of correlation and mean reversion implied by these post-crisis estimates of mean returns and probabilities. As before, the probability estimates of the professionals lead to a corner solution in a clear majority of cases. However, the evidence for mis-calibrated beliefs is even stronger than in the pre-crisis sample, and the distribution of implied correlation and mean reversion even more concentrated on the most extreme cases than before. In the case of stocks outperforming bonds, the probability estimates of 58% of the respondents imply correlation coefficients in excess of 0.9. All in all, Panel (a) of the Figure shows that the respondents' probability estimates imply implausible levels of correlation

between stocks and bonds, and very strong mean-reversion for stocks. The results for stocks outperforming bills given in Panel (b) of the Figure show a similar pattern. Stock outperformance expectations have thus increased even more than the expected equity premium has.

One possible explanation for the even more extreme results in the post-crisis sample is that the sample of respondents is less experienced: the mean of financial markets work experience is 5 years, while the 2004 sample mean is 9 years. To investigate this hypothesis we again form a subsample of experienced pros from those with more than 5 years of financial markets experience. Table 8 shows that the median estimate of the probability that stocks outperform bonds stays at 90%, and the mean declines only slightly to 85%. This implies that the result is not merely due to the post crisis sample being less experienced.

In sum, the financial crisis had, in line with expectations, a decreasing effect on stock return forecasts but it did not decrease the professionals' optimism on the relative long-term performance of stocks versus bonds. Conditional on equity premium and other expectations, the professionals are even more certain that stocks beat bonds.

7. Conclusion

This paper develops two new methods for assessing the optimism as well as the plausibility of return expectations without having to rely on noisy realized returns as benchmarks. We define a bivariate process for the returns on stocks and bonds (or bills) to explore the plausibility of the subjects' return expectations across different asset classes. This is achieved by contrasting the subjective probability of stocks beating bonds on one hand, to expected returns and volatilities on the other hand. We also model the tails of the respondents' subjective probability distributions utilizing their estimates of the 20-year minimum and maximum annual return. We use techniques

of Extreme Value Theory for several different parent distributions to model the tail expectations and calculate the proportions of respondents whose range of the minimum and maximum are too narrow compared to the theoretical expected value. We apply these methods on field surveys of financial market professionals' expectations regarding asset returns, volatilities, and outperformance probabilities.

Our empirical results are summarized as follows. First, subjects' estimates of the probability of stocks outperforming bonds are way too optimistic given their estimates of expected returns and volatilities of these asset classes. Alternatively, implausible mean-reversion or correlation parameters would be required to make the expectations internally consistent. This extreme form of stock market optimism exists before the financial crisis, and is even stronger in a sample collected after the financial crisis in 2009-2013, even when accounting for a somewhat higher equity premium expectation. Second, the volatility estimates of the finance professionals are too high to be consistent with their estimates of the minimum and the maximum returns, or else imply return distributions with thin tails. Specifically, they would imply tails even thinner than in the normal distribution.

On the whole there are two broad ways of interpreting our results. One is to take the expectations as unbiased, and conclude that asset class return generating processes must be very different from what was previously thought. The other is to say that expectations of finance professionals are outside of reasonable bounds. We lean toward this latter explanation.

References

- Amromin, G., Sharpe, S.A., 2014. From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science* 60, 845-866.
- Antoniou, C., Doukas, J.A., Subrahmanyam, A., 2013, Cognitive dissonance, sentiment, and momentum, *Journal of Financial and Quantitative Analysis*, 48, 245-275.
- Bacchetta, P., Mertens, E., van Wincoop, E., 2009, Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance*, 28, 406 – 426.
- Bali, T.G., 2003. An extreme value approach to estimating volatility and Value at Risk. *Journal of Business* 76, 83-108.
- Barberis, N., 2000. Investing for the long run when returns are predictable. *Journal of Finance* 55, 225-264.
- Barberis, N., Greenwood, R., Jin, L., Shleifer, A., 2015, X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics*, 115, 1-24.
- Bénabou, R., 2013, Groupthink: Collective delusions in organizations and markets, *Review of Economic Studies*, 80, 429-462.
- Ben-David, I., Graham, J.R., Harvey, C., 2013. Managerial miscalibration. *Quarterly Journal of Economics* 128, 1547-1584.
- Bodnaruk, A., Simonov, A., 2015, Do financial experts make better investment decisions? *Journal of Financial Intermediation*, 24, 514 – 536.
- Campbell, J.Y., Chan, Y.L., Viceira, L.M., 2003. A multivariate model of strategic asset allocation. *Journal of Financial Economics* 67, 41-80.
- Campbell, J.Y., Viceira, L.M., 2005. The term structure of the risk-return tradeoff. *Financial Analysts Journal* 61, 34-44.
- Carhart, M., Cheah, U., De Santis, G., Farrell, H., Litterman, R., 2014. Exotic beta revisited. *Financial Analysts Journal* 70, 24-52.
- Chang, T.Y., Solomon, D.H., Westerfield, M.M, 2016, Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect, *Journal of Finance*, 71, 267-302.
- Cheng, I.-H., Raina, S., Xiong, W., 2014, Wall Street and the housing bubble, *American Economic Review*, 104, 2797-2829.
- Claus, J., Thomas, J., 2001. Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. *Journal of Finance* 56, 1629-1666.
- Dimson, E., Marsh, P., Staunton, M., 2002. *Triumph of the Optimists: 101 Years of Global Investment Returns*. Princeton University Press, Princeton, NJ.
- Dimson, E., Marsh, P., Staunton, M., 2004. *Global Investment Returns Yearbook 2004*. ABN Amro and London Business School.

- Donaldson, G.R., Kamstra, M.J., Kramer, L.A., 2010. Estimating the equity premium. *Journal of Financial and Quantitative Analysis* 45, 813-846.
- Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag.
- Fama, E.F., 1965. The behavior of stock market prices. *Journal of Business* 38, 34-105.
- Fama, E.F., French, K.R., 1988. Permanent and temporary components of stock prices. *Journal of Political Economy* 96, 246-273.
- Fama, E.F., French, K.R., 2002. The equity premium. *Journal of Finance* 57, 637-659.
- Festinger, L., 1957, *A Theory of Cognitive Dissonance*, Stanford University Press, Stanford, CA.
- Frankel, J.A., Froot, K.A., 1987. Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review* 77, 133-153.
- Gennaioli, N., Shleifer, A., Vishny, R.W., 2015. Money doctors. *Journal of Finance* 70, 91-114.
- Glaser, M., Langer, T., Weber, M., 2007. On the trend recognition and forecasting ability of professional traders. *Decision Analysis* 4, 176-193.
- Glassman, J.K., Hasset, K.A., 1999. *Dow 36,000: The new strategy for profiting from the coming rise in the stock market*. Times Books.
- Goetzmann, W.N., Peles, N., 1997, Cognitive dissonance and mutual fund investors, *Journal of Financial Research*, 20, 145-158.
- Greenwood, R., Shleifer, A., 2014. Expectations of returns and expected returns. *Review of Financial Studies* 27, 714-746.
- Haigh, M.S., List, J.A., 2005. Do professional traders exhibit myopic loss aversion? An experimental analysis. *Journal of Finance* 60, 523-534.
- Hodrick, R.J., 1992. Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies* 5, 357-386.
- Hommes, C., Sonnemans, J., Tuinstra, J., van de Velden, H., 2005. Coordination of expectations in asset pricing experiments. *Review of Financial Studies* 18, 955-980.
- Ibbotson, R., Chen, P., 2003. Long-run stock returns: Participating in the real economy. *Financial Analysts Journal* 59, 88-98.
- Jondeau, E., Rockinger, M., 2003. Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance* 10, 559-581.
- Jorion, P., 2003. The long-term risks of global stock markets. *Financial Management* 32, 5-26.
- Kaustia, M., Alho, E., Puttonen, V., 2008. How much does expertise reduce behavioral biases? The case of anchoring effects in stock return estimates. *Financial Management* 37, 391-411.
- Keefer, D.L., Bodily, S.E., 1983. Three-point approximations for continuous random variables. *Management Science* 29, 595-609.

- Kim, M.J., Nelson, C.R., Startz, R., 1991. Mean reversion in stock prices? A reappraisal of the empirical evidence. *Review of Economic Studies* 58, Special Issue: The Econometrics of Financial Markets, 515-528.
- Leadbetter, M.R., Lindgren, G., Rootzén, H., 1983. *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag.
- Longin, F.M., 1996. The asymptotic distribution of extreme stock market returns. *Journal of Business* 69, 383-408.
- Mullainathan, S., Noeth, M., Schoar, A., 2012, The market for financial advice: an audit study, NBER Working paper 17929, NBER, Cambridge, MA.
- Pastor, L., Stambaugh, R.F., 2012. Are stocks really less volatile in the long run? *Journal of Finance* 67, 431-478.
- Poterba, J.M., Summers, L., 1988. Mean reversion in stock prices: Evidence and implications. *Journal of Financial Economics* 22, 27-59.
- Shiller, R.J., 1990. Speculative prices and popular models. *Journal of Economic Perspectives* 4, 55-65.
- Siegel, J.J., 2002. *Stocks for the long run*. McGraw-Hill, NY.
- Vissing-Jørgensen, A., 2003. Perspectives on behavioral finance: Does “irrationality” disappear with wealth? Evidence from expectations and actions. *NBER Macroeconomics Annual*.

Table 1
Summary Statistics

This table gives descriptive statistics for expectations for stocks (Panel A), bonds (Panel B), and bills (Panel C). The columns MIN and MAX give the means over all subjects for their estimated worst and best single year real return over 20 years. $E(R)$ denotes the estimate for the expected return (average annual real return over 20 years), and σ_i denotes the volatility estimates. $\hat{\sigma}$ is the volatility approximation calculated for each respondent by equation (1). $\sigma_i - \hat{\sigma}$ and $\sigma_i / \hat{\sigma}$ give the difference and the ratio between the respondents' volatility estimates and the volatility approximation.

Panel A: Stocks							
	MIN	MAX	$E(R)$	σ_i	$\hat{\sigma}$	$\sigma_i - \hat{\sigma}$	$\sigma_i / \hat{\sigma}$
Mean	-17.68	31.37	7.55	16.26	8.17	8.08	2.72
Median	-15.00	30.00	7.00	18.50	7.33	7.58	1.83
St dev	12.68	18.87	2.57	7.51	4.84	8.26	2.33
Min	-70.00	6.00	3.50	2.00	1.17	-11.33	0.50
Max	8.00	100.00	16.00	35.00	28.33	30.83	12.50
Panel B: Bonds							
	MIN	MAX	$E(R)$	σ_i	$\hat{\sigma}$	$\sigma_i - \hat{\sigma}$	$\sigma_i / \hat{\sigma}$
Mean	-1.81	8.67	4.13	4.96	1.75	3.22	4.06
Median	-0.25	7.50	4.00	5.00	1.33	2.25	2.40
St dev	4.85	4.56	1.73	3.36	1.39	3.21	4.41
Min	-25.00	1.50	1.00	0.00	0.17	-2.00	0.00
Max	4.00	25.00	10.00	15.00	8.33	13.00	20.00
Panel C: Bills							
	MIN	MAX	$E(R)$	σ_i	$\hat{\sigma}$	$\sigma_i - \hat{\sigma}$	$\sigma_i / \hat{\sigma}$
Mean	-0.10	4.82	2.51	3.18	0.82	2.31	4.25
Median	0.00	4.35	3.00	2.00	0.67	0.92	2.75
St dev	1.72	2.53	1.37	3.79	0.50	3.63	4.51
Min	-7.00	1.00	0.00	0.00	0.17	-1.00	0.00
Max	3.00	15.00	6.00	20.00	2.67	18.33	22.50

Table 2**Proportions of respondents with unrealistic implied stock-bond correlation or/and stock autocorrelation**

This table gives the proportions of respondents whose implied coefficients of correlation, autocorrelation are outside pre-specified ranges. The implied stock-bond correlation and autocorrelation parameters are obtained by calibrating a bivariate stock-bond model for each subject so that it produces a probability of stocks beating bonds equal to the subject's estimate of that quantity while also using the subjects' estimates of expected returns and volatilities for stocks and bonds. Here we consider as unrealistic a stock-bond correlation outside the range [0, 0.7] and a stock autocorrelation outside the range [-0.1, 0.1]. Panel A gives the results for the base sample for three different horizons; 5 (N = 99), 10 (N = 100) and 20 (N = 100) years. Panel B gives the results for the two subsamples; experienced pros (N = 48) and the high volatility group (N = 51). The subsample results are based on the horizon of 10 years.

PANEL A: Base sample			
Horizon	5-year	10-year	20-year
Correlation	0.71	0.73	0.74
Autocorrelation	0.95	0.93	0.94
Either correlation or autocorrelation	0.99	0.96	0.96
Both	0.67	0.70	0.72

PANEL B: Subsamples (10-year horizon)		
	Experienced pros	High volatility
Correlation	0.81	0.77
Autocorrelation	0.94	0.96
Either correlation or autocorrelation	0.98	0.98
Both	0.77	0.75

Table 3
Maximum Likelihood Estimates of GEV Parameters

This table gives the results of fitting the GEV distribution to the series of minimum and maximum estimates pooled over all respondents. The fitting is done by maximum likelihood. The entries in the table are the MLE estimates for the shape parameter (ξ), the scale parameter (σ), and the location parameter (μ), respectively. The shape parameter measures the tail thickness of the distribution. $\xi = 0$ corresponds to the Gumbel (type I), and $\xi > 0$ to the Fréchet (type II), and $\xi < 0$ to the Weibull (type III) distribution families, respectively. The higher is the value of ξ , the fatter is the tail. The values given in brackets below the point estimates are the 95% confidence bounds.

	Stocks		Bonds		Bills	
	MIN	MAX	MIN	MAX	MIN	MAX
Shape parameter (ξ)	0.22	0.31	0.13	0.04	0.60	0.69
95% CI	[0.01 0.44]	[0.06 0.56]	[-0.14 0.40]	[-0.23 0.32]	[0.30 0.90]	[0.36 1.02]
Scale parameter (σ)	0.84	0.81	0.62	0.62	0.68	0.64
95% CI	[0.67 1.05]	[0.63 1.03]	[0.64 1.02]	[0.49 0.78]	[0.51 0.92]	[0.47 0.88]
Location parameter (μ)	1.27	1.02	1.03	0.79	0.73	0.64
95% CI	[1.03 1.50]	[0.80 1.25]	[0.79 1.27]	[0.61 0.98]	[0.54 0.93]	[0.45 0.84]

Table 4**Proportion of Estimates that are Below Expected Extremes: Individual Volatility Estimates**

This table shows the proportions of respondents' minimum (MIN) and maximum (MAX) return estimates that are smaller in absolute value than those implied by the subjects' volatility estimates. The column labeled Z-value gives the Z-statistic for the difference in proportions test between MIN and MAX. The column labeled BOTH gives the proportion of respondents whose estimates for both the minimum and the maximum are below the expected values. Panel A gives the results for stocks, Panel B for bonds, and Panel C for bills, respectively.

Panel A: Stocks				
	MIN	MAX	Z-value	BOTH
Normal	0.62	0.68	-0.73	0.52
<i>t</i> (5)	0.71	0.77	-0.80	0.67
<i>t</i> (4)	0.76	0.79	-0.42	0.73
<i>t</i> (3)	0.83	0.85	-0.24	0.82
Panel B: Bonds				
	MIN	MAX	Z-value	BOTH
Normal	0.66	0.77	-1.37	0.62
<i>t</i> (5)	0.74	0.92	-2.90	0.71
<i>t</i> (4)	0.82	0.97	-2.92	0.80
<i>t</i> (3)	0.89	0.98	-2.23	0.88
Panel C: Bills				
	MIN	MAX	Z-value	BOTH
Normal	0.67	0.67	0.00	0.59
<i>t</i> (5)	0.78	0.81	-0.44	0.71
<i>t</i> (4)	0.79	0.86	-0.94	0.76
<i>t</i> (3)	0.86	0.89	-0.54	0.83

Table 5**Proportion of Estimates that are Below Expected Extremes: Historical Volatility Estimates**

This table gives the proportion of respondents' estimates that are below the theoretical expected value for the minimum (column labeled MIN) or the maximum (MAX), and are based on historical volatility taken from Dimson, Marsh, and Staunton (2002). The values used are 20.2% for stocks, 10.0% for bonds, and 4.7% for bills, respectively. The column labeled Z-value gives the Z-statistic for the difference in proportions test. The column labeled BOTH gives the proportion of respondents whose estimates for both the minimum and the maximum are below the expected values. Panel A gives the results for stocks, Panel B for bonds, and Panel C for bills, respectively.

Panel A: Stocks				
	MIN	MAX	Z-value	BOTH
Normal	0.89	0.82	1.25	0.77
<i>t</i> (5)	0.97	0.92	1.17	0.91
<i>t</i> (4)	0.97	0.92	1.17	0.91
<i>t</i> (3)	0.98	0.97	0.58	0.97
Panel B: Bonds				
	MIN	MAX	Z-value	BOTH
Normal	0.97	0.98	-0.58	0.97
<i>t</i> (5)	0.97	1.00	-1.44	0.97
<i>t</i> (4)	0.98	1.00	-1.01	0.98
<i>t</i> (3)	1.00	1.00	--	1.00
Panel C: Bills				
	MIN	MAX	Z-value	BOTH
Normal	0.98	0.98	0.00	0.97
<i>t</i> (5)	0.98	1.00	-1.01	0.98
<i>t</i> (4)	1.00	1.00	--	1.00
<i>t</i> (3)	1.00	1.00	--	1.00

Table 6**Proportion of Estimates that are Below Expected Extremes: Historical Mean and Volatility**

This table gives the proportion of respondents' estimates that are below the theoretical expected value for the minimum (column labeled MIN) or the maximum (MAX), and are based on historical volatilities and mean returns taken from Dimson, Marsh, and Staunton (2002). The volatilities are 20.2% for stocks, 10.0% for bonds, and 4.7% for bills, whereas the mean returns are 6.7% for stocks, 1.6% for bonds, and 0.9% for bills. The column labeled Z-value gives the Z-statistic for the difference in proportions test. The column labeled BOTH gives the proportion of respondents whose estimates for both the minimum and the maximum are below the expected values. Panel A gives the results for stocks, Panel B for bonds, and Panel C for bills, respectively.

Panel A: Stocks				
	MIN	MAX	Z-value	BOTH
Normal	0.91	0.82	1.54	0.77
<i>t</i> (5)	0.97	0.91	1.47	0.89
<i>t</i> (4)	0.97	0.92	1.17	0.91
<i>t</i> (3)	0.98	0.97	0.58	0.97
Panel B: Bonds				
	MIN	MAX	Z-value	BOTH
Normal	0.97	0.97	0.00	0.95
<i>t</i> (5)	0.98	1.00	-1.01	0.98
<i>t</i> (4)	0.98	1.00	-1.01	0.98
<i>t</i> (3)	1.00	1.00	--	1.00
Panel C: Bills				
	MIN	MAX	Z-value	BOTH
Normal	1.00	0.94	2.07	0.94
<i>t</i> (5)	1.00	0.98	1.01	0.98
<i>t</i> (4)	1.00	0.98	1.01	0.98
<i>t</i> (3)	1.00	1.00	--	1.00

Table 7

Summary Statistics for the Post-2008 Sample

This table gives summary statistics for the sample collected in 2009, 2010, 2011, and 2012. The first three columns show the statistics for the expected annual real return on stocks, bonds, and bills for a horizon of 10 years. The real returns are calculated using the respondents' estimates of expected nominal returns and inflation. The last two columns give statistics for the probability of stocks outperforming bonds and stocks outperforming bills. The number of observations is 251.

	Stocks	Bonds	Bills	Stocks vs. bonds	Stocks vs. bills
Mean	5.76	1.22	0.45	86.3	89.7
Median	5.37	0.98	0.00	90.0	95.0
St dev	2.71	1.45	1.68	15.5	14.9
Min	-6.09	-8.70	-11.30	20.0	10.0
Max	17.07	11.65	8.74	100.0	100.0

Table 8

Summary Statistics for the Post-2008 Experienced Pros -Sample

This table gives summary statistics for the sample collected in 2009, 2010, 2011, and 2012. The first three columns show the statistics for the expected annual real return on stocks, bonds, and bills for a horizon of 10 years. The real returns are calculated using the respondents' estimates of expected nominal returns and inflation. The last two columns give statistics for the probability of stocks outperforming bonds and stocks outperforming bills. Experienced pros are defined as those that have more than 5 years of experience. The number of observations is 81.

	Stocks	Bonds	Bills	Stocks vs. bonds	Stocks vs. bills
Mean	5.880	1.125	0.277	85.363	91.613
Median	5.882	0.980	0.000	90.000	98.000
St dev	2.233	1.179	1.088	16.133	11.805
Min	0.478	-1.887	-2.913	25.000	50.000
Max	14.286	3.922	3.922	100.000	100.000

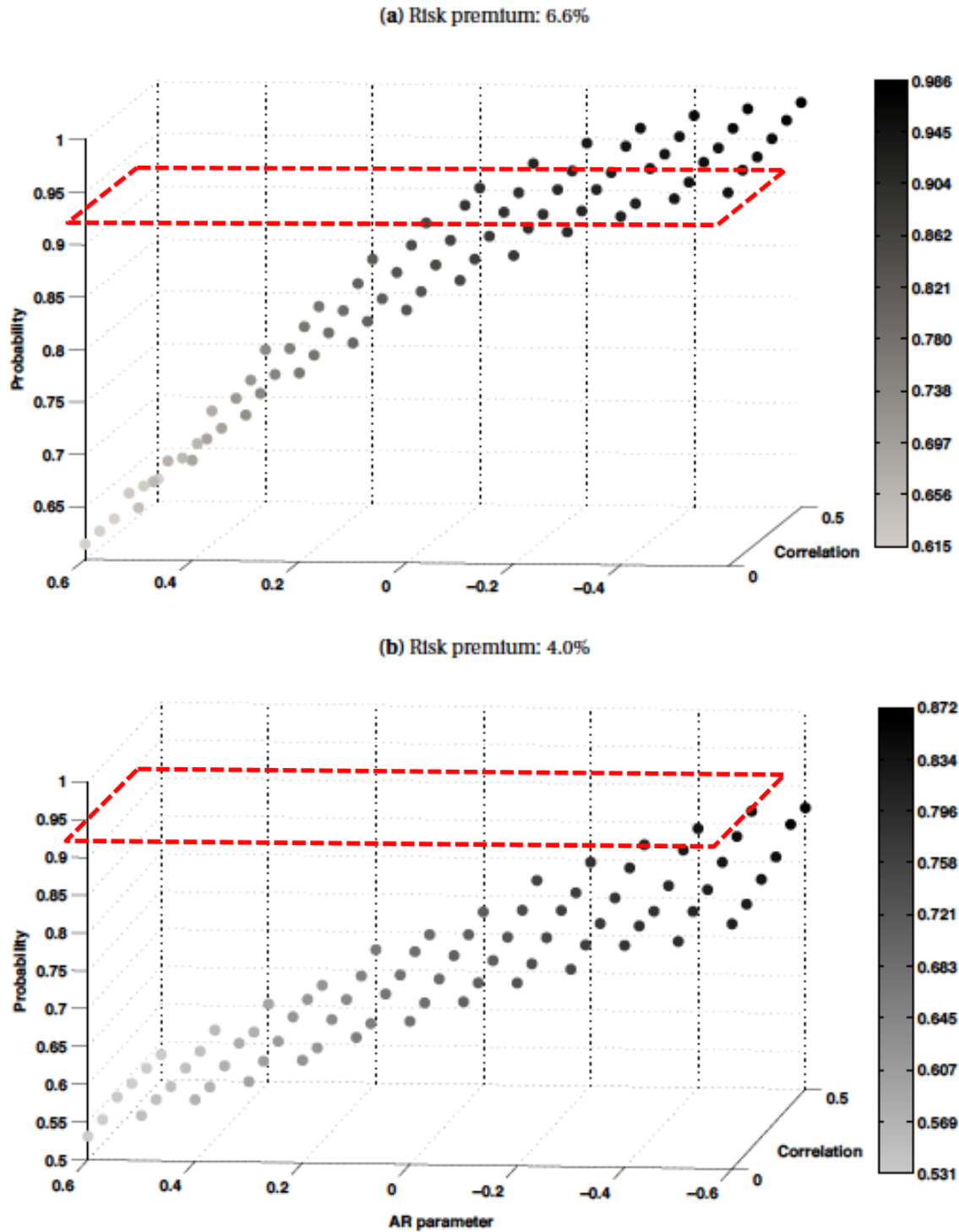


Figure 1. Probabilities of Stocks Outperforming Bonds Implied by the VAR-Model. This figure plots on the Y-axis probabilities that stocks outperform bonds over an investment horizon of 20 years, as given by the 10,000 rounds of simulation from the VAR-model of Equation (6). The X-axis (left-right) measures the autocorrelation of stock returns, and the Z-axis measures the stock-bond correlation. Panel a (b) uses an equity risk premium of 6.6% (4%). The parallelograms composed of red dashed lines depict 3-D planes indicating a probability of 0.92, corresponding to the historical frequency reported in Siegel (2002).

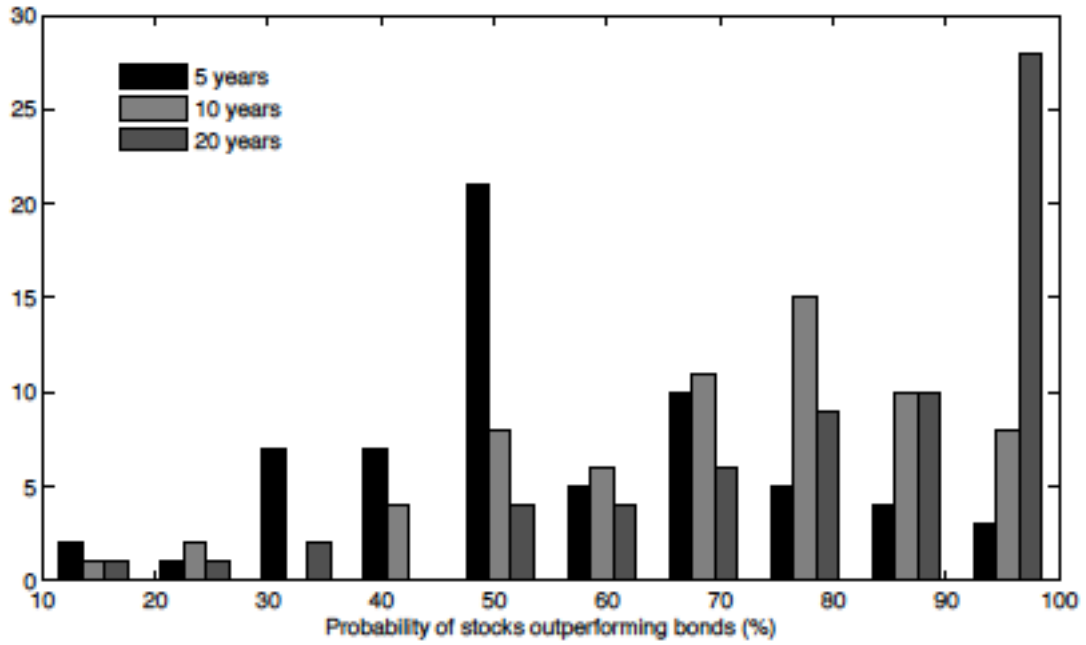
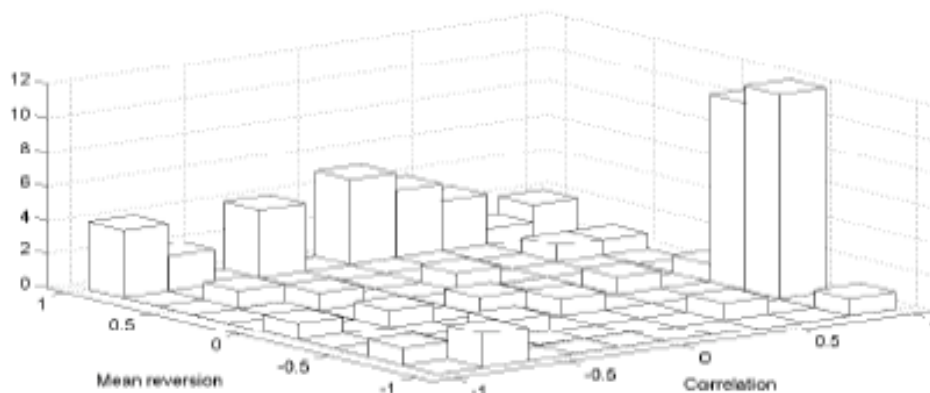
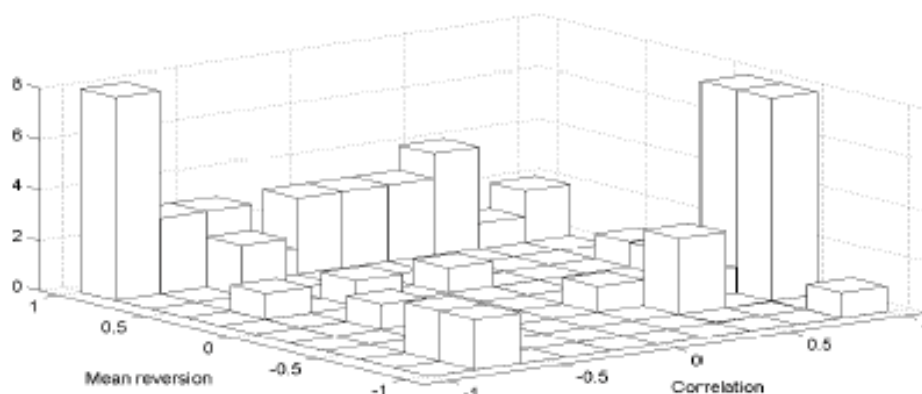


Figure 2. Histogram of Estimates of Stocks Outperforming Bonds. This figure plots the histogram of the respondents' estimates of the probability that stocks outperform bonds. Three different forecast horizons are considered: 5, 10, and 20 years.

(a) Horizon: 20 years



(b) Horizon: 10 years



(c) Horizon: 5 years

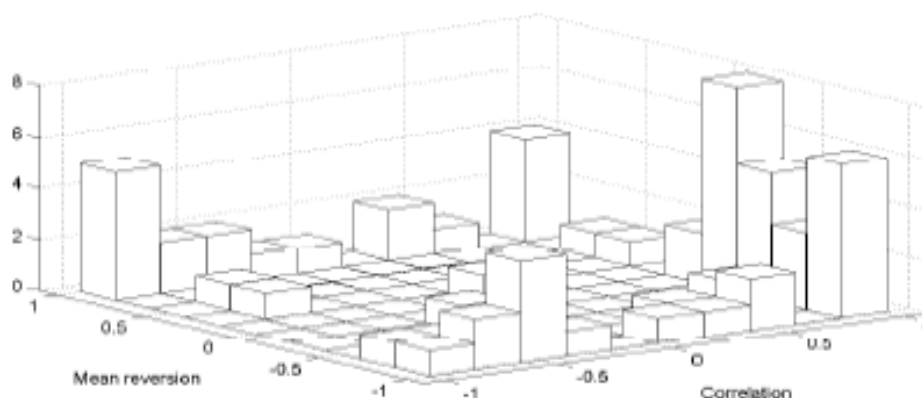


Figure 3. Implied Correlation and Mean Reversion: VAR(1). This figure gives the histogram of the correlation and mean reversion parameters implied by the respondents' estimates of the expected returns and volatilities of stocks and bonds. The implied parameters are calculated by the simulation algorithm described in Appendix A. The simulation is done with 10,000 replications.

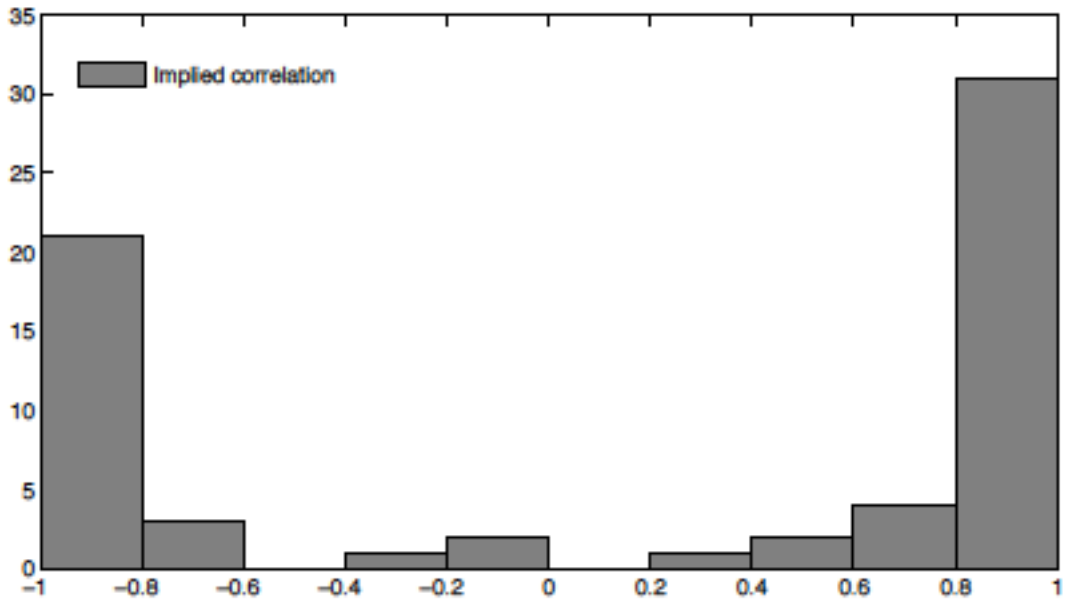


Figure 4. Implied Correlation: Random Walk with Drift. The figure plots the histogram of the correlation parameters implied by the respondents' estimates of the expected returns on stocks and bonds, and their volatilities. The implied parameters are calculated using the simulation algorithm described in Appendix A. The figure is based on 50,000 replications

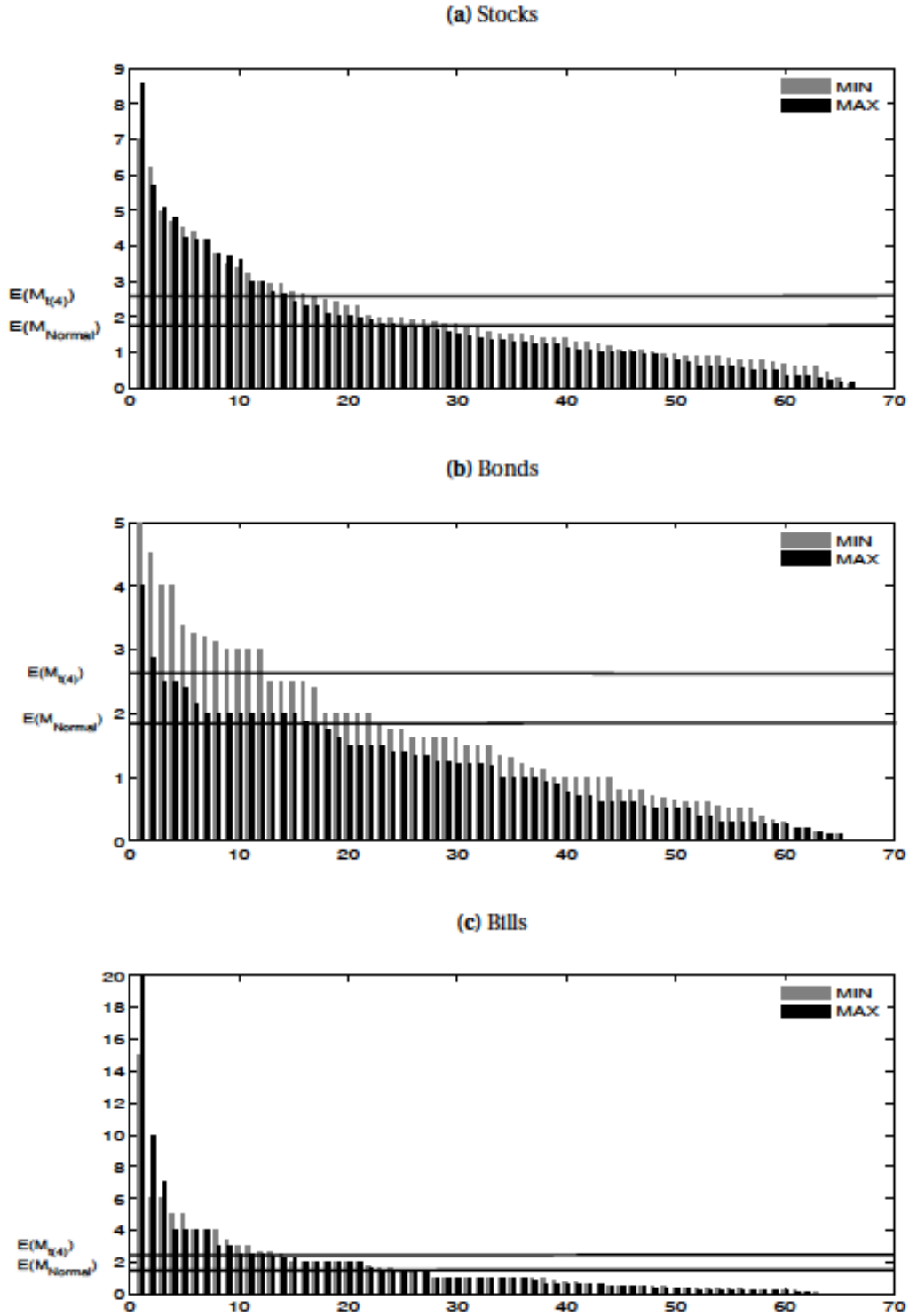
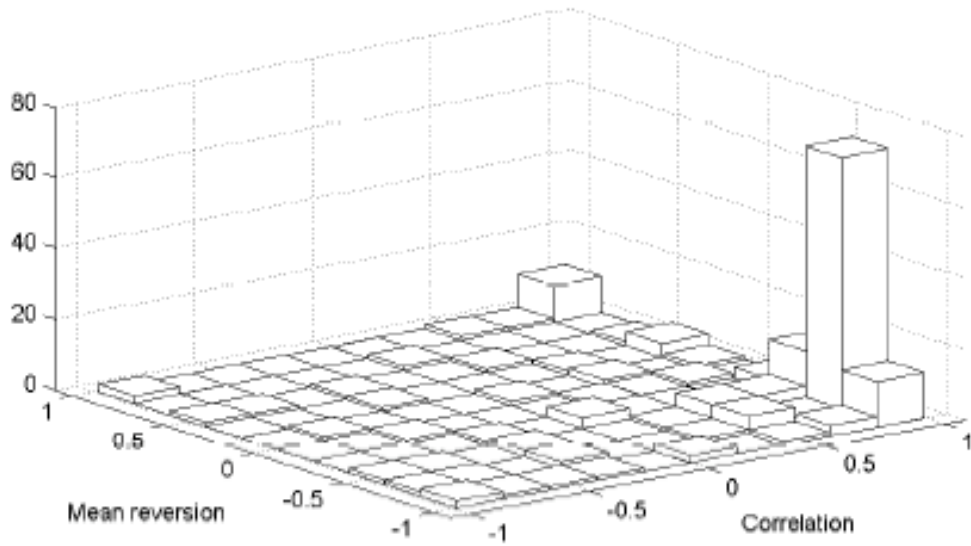


Figure 5. Ordered Minimum and Maximum Estimates. This figure plots the ordered series of the minimum (grey bars) and maximum (black bars) estimates of the respondents together with the expected values for a normal and $t(4)$ parent (the two straight horizontal lines). Panel (a) gives the results for stocks, Panel (b) for bonds, and Panel (c) for bills, respectively.

(a) Stocks vs. bonds



(b) Stocks vs. bills

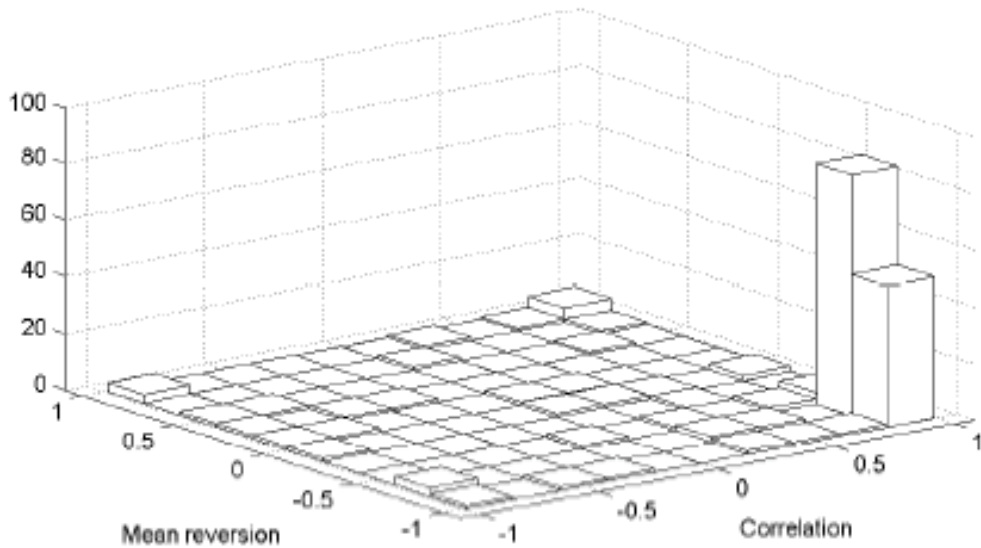


Figure 6. Implied Correlation and Mean Reversion: Post-2008 Sample. This figure plots the histogram of the correlation and mean reversion parameters implied by the respondents' estimates of the expected returns on stocks, bonds, and bills in the sample collected after the financial crisis of 2008. Panel (a) shows the results for stocks outperforming bonds, and Panel (b) for stocks outperforming bills. The implied parameters are calculated by the simulation algorithm described in the Appendix. The figure is based on 10,000 replications.

Appendix A: Simulation Algorithm

In the simulations, we assume that the returns on stocks and bonds jointly follow the bivariate VAR(1) model given in Equation (1) in the text, reproduced below:

$$\begin{bmatrix} r_{S,t} \\ r_{B,t} \end{bmatrix} = \begin{bmatrix} \mu_S \\ \mu_B \end{bmatrix} + \begin{bmatrix} \phi & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} r_{S,t-1} - \mu_S \\ r_{B,t-1} - \mu_B \end{bmatrix} + \begin{bmatrix} \epsilon_{S,t} \\ \epsilon_{B,t} \end{bmatrix}$$

We then proceed as follows.

For each respondent we use their individual estimates of the expected returns and volatilities of the assets in the simulations. The simulation algorithm for respondent j can be summarized in the following five steps.

1. First we discretize the continuous range for the mean reversion parameter and the correlation coefficient, and form a two-way grid of all their possible combinations⁹.
2. Then, for parameter pair i and respondent j , we generate 10,000 time-series of 120 (annual) return observations, discard the first 100 to mitigate the effect of the starting values (we use zero as the first lagged return), and calculate the cumulative return over the horizon of 20 years for both stocks and bonds.
3. At the end of each replication for parameter pair i , we store an indicator variable that is equal to one if stocks outperformed bonds in that replication.
4. Then we divide the sum of the indicator variable by the total number of replications to get an estimate of the probability that stocks beat bonds for parameter pair i .

⁹ Both variables lie between -1 and 1. We discretize the range such that we take values in intervals of 0.1, i.e., for each variable we take the values -1, -0.9, ..., 0.9, 1. This leads to $21^2 = 441$ different combinations.

5. Next we calculate the squared deviation of the estimated probability from the respondents estimate for each grid value, that is

$$D = (\hat{p}_i - \hat{p}_j)^2$$

where \hat{p}_i denotes the probability of stocks beating bonds for parameter pair i in the grid and \hat{p}_j denotes the probability estimates given by respondent j . We then choose the parameter pair in the grid that minimizes this distance.

This way we are able to get, for each individual, the implied mean reversion parameter and correlation coefficient that would result in their estimated probability of stocks beating bonds.