Revealed Expectations and Learning Biases: Evidence from the Mutual Fund Industry

Francesco Nicolai
The London School of Economics

Simona Risteska *
The London School of Economics

January 15, 2020 – This Version
January 31, 2018 – First Version

Abstract

By inverting the optimal portfolios of mutual fund managers in a fairly general setting, which allows us to partial out the effect of risk aversion and hedging demands, we provide an estimate of perceived expected excess returns and show that they are significantly affected by experienced returns. The effect of past returns is non-monotone: we provide reduced-form and structural evidence of managers displaying *recency* and *primacy bias*. Finally, we estimate an average coefficient of relative risk aversion close to unity.

*We benefited from helpful comments from Nicholas Barberis, Pasquale Della Corte, Boyan Jovanovic, Christian Julliard, Samuli Knupfer, Ralph Koijen, Avner Langut, Dong Lou, Ian Martin, Cameron Peng, Asaf Razin, Andrew Redleaf, Andrea Tamoni, Michela Verardo and the participants at the LSE seminar, the 2019 Yale Whitebox Conference, the 2019 Belgrade Young Economists Conference. Any errors or omissions are the responsibility of the authors. Francesco Nicolai thanks the Paul Woolley Centre for the Study of Capital Market Disfunctionality at the LSE Financial Market Group for financial support. Francesco Nicolai, Department of Finance, The London School of Economics and Political Science, Houghton St, WC2A 2AE, London, UK, email: f.nicolai@lse.ac.uk. Simona Risteska, Department of Finance, The London School of Economics and Political Science, Houghton St, London, WC2A 2AE, UK, email: s.risteska@lse.ac.uk. Send correspondence to Francesco Nicolai.

How do investors form their return expectations? Do they take all available information into account? Does personal experience play a crucial role in the formation of expectations? We attempt to answer these questions by looking at mutual fund managers' stock return expectations as revealed by their portfolio holdings. We exploit the fact that, under a large class of models, the optimal portfolio rule has a similar functional form; using a three dimensional panel consisting of the portfolio holdings of mutual fund managers over a period of thirty-five years, we are able to extract a measure of subjective expected returns for every manager in our panel by exploiting the variation across stocks over time between and within managers. To see this, consider a mean-variance investor for whom the vector of physical expected returns is given by the following formula:

$$\mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f \mathbf{1}] = \gamma_{i,t} \Sigma_t \boldsymbol{w^*_{i,t}} \tag{1}$$

where $\mathbb{E}_{i,t}[\cdot]$ is the conditional expectation operator taken under investor i's information set at time t, $\boldsymbol{r_{t+1}} - r_f \mathbf{1}$ is a vector of excess returns, $\gamma_{i,t}$ is the coefficient of relative risk aversion of manager i at time t, Σ_t is the conditional covariance matrix of stock returns and $\boldsymbol{w^*_{i,t}}$ is the time t vector of optimal portfolio weights of investor i. The above expression for expected excess returns is obtained by inverting the first-order condition of a mean-variance investor provided that we have a good measure of conditional covariances Σ_t[1]. In Section 2.1 we show that - as long as we correctly interpret the manager-specific time-varying parameter $\gamma_{i,t}$ - many optimal portfolio models give rise to a subjective expected return similar to (1); whenever that is not the case, we can saturate the model with fixed effects in order to split the total demand into a mean-variance component and a hedging component; to isolate the effect of risk aversion from the effect of subjective expected returns we resort to the very general principle that, given the cross-section of assets the manager invests in, risk aversion is a manager-specific quantity,

[1]In these regards we follow Merton (1980) and argue that investors' disagreement should mainly regard expected returns and not variances and covariances. We will show that empirically this is a good approximation.

while expected returns are at the same time asset-specific. The information contained in the cross section of holdings, therefore, greatly reduces the issue of separating the variation due to the manager's preferences from the one due to beliefs.

We start by providing evidence in Section 3 that more than 50% of the variation in expected returns is explained by a common time-varying factor and we are able to explain almost 90% of the variation with manager-time and stock-time components. This suggests that a saturated regression will likely allow us to isolate the idiosyncratic part of expected returns affected by manager-stock-time specific effects; we will focus on this component to explore the extent to which managers' beliefs are affected by experience. In particular, we investigate whether fund managers put more emphasis on past stock returns that they have personally experienced over their investment career. To begin, we consider the effect of the simple average of past observed returns on portfolio holdings decisions. Having experienced a one standard deviation higher average return on a given stock causes the manager to inflate his expected excess return by between 10.3 and 15.1 basis points, partialling out the effect of common stock and manager characteristics. The result is both statistically and economically significant and it is almost an order of magnitude larger than other commonly used predictors. Nonetheless, the effect of average experienced returns masks important heterogeneity in the effect of past returns observed a different points in time: when we move on to examining the particular shape of the learning curve we find evidence of a differential effect. We start by providing non-parametric results that let us avoid taking a stance on the precise functional form that investors use to weight past experienced returns. Mutual fund managers in our sample are subject to the so-called serial-position effect: the tendency to predominantly remember the initial and the last observations in a series. More precisely, managers' investment decisions and beliefs are particularly affected by the returns they have experienced early on during their stock-specific experience and those they have experienced most recently. In other words, professional investors seem to exhibit the *primacy* and *recency* bias.

As one would expect, the effect is stronger for single-managed funds and decays fast as the number of managers increases: the effect of recently experienced returns on managers in a single managed fund is twice as large compared to managers working with at least one other professional; the effect of early returns is an order of magnitude larger.

We also show that the differential effect of taxes on capital gains and losses cannot explain these findings, as proven by the fact that the effect of early career experience is still present even when the manager switches to a different fund. At most, tax considerations can explain 20% of the estimated influence of past returns on portfolio choices and expected returns.

Armed with the reduced-form evidence, we provide a tentative structural estimation of the managers' learning function. In particular, the results in the reduced-form estimation seem to suggest a non-monotonic learning function. For this reason we adopt a variation of the parametrisation of the learning function in Malmendier and Nagel (2016) that allows for a variety of decreasing and increasing, convex and concave, monotone and non-monotone learning weights. We find that fund managers on average do indeed place a disproportionate weight on personal past experience and that this biases the expected returns recovered from their stock holdings, after having adjusted for risk and risk aversion. When we allow for time-varying weights on past stock returns, we show that mutual fund managers tend to place excessive weight on returns experienced at the beginning of their careers and in the most recent quarters compared to those in the middle period, suggesting that both early-career and recent experience seem to be important determinants of the investment behaviour of a large class of professional investors. As an example, a manager with the median stock-specific experience of 9 quarters assigns around 1.84 times larger weight on the return experienced in the most recent quarter compared to the benchmark of 1/9, while the weight on the first experienced return is 3.13 times larger compared to the benchmark. We thus reconcile two conflicting strands of the literature: similarly to Malmendier and Nagel (2011) and Malmendier and Nagel (2016), we confirm that investors do overweight their personal experience and manifest a *recency bias*, but - at the same time - we show that professional investors also place a disproportionately large weight on returns that have been experienced in the early part of their investing career, similarly to the findings of Kaustia and Knüpfer (2008) and Hirshleifer et al. (2019). When looking at co-managed funds, we show that a large fraction of the effect of early experience washes out while the effect of recently experienced returns persists; this might be due to the fact that, while there is large heterogeneity in early experience, recently experienced returns are mostly shared among managers within a team.

Finally, in the last part of the paper we briefly focus our attention on risk preferences. Notice from equation (1) that, while risk aversion varies at the manager-time level, beliefs vary at the manager-time-stock level. This lets us separate *variation* in adjusted portfolio holdings that is due to the managers' risk appetite from differences in believes, but does not guide us regarding their *level*. Once we make some minimal assumptions to pin down their level, we show that individual expected returns tend to be quite biased and that preferences display significant heterogeneity across individuals and time. Moreover, on average, mutual fund managers display an Arrow (1965)-Pratt (1964) coefficient of relative risk aversion between 0.915 and 1.283.

The rest of the paper is organized as follows: Section 1 provides an overview of recent literature. We proceed by showing that most of the literature relies on evidence from surveys obtained from non-professional investors or, when not affected by these concerns, on a relatively limited amount of data. We argue that the present paper tries to solve the aforementioned issues. Section 2 describes how we can separate the variation in expected returns from the variation in risk aversion or other factors in a wide class of models. Section 3 gives details of the data used in our empirical work and provides some summary statistics. Section 4 provides the non-parametric results of our analysis, while Section 5 describes and show the results of our parametric approach. In Section 6, we tackle the question of the level of risk aversion of investment professionals. Finally, Section 7 provides concluding remarks.

1 Previous Literature

The issue of whether economic agents learn with experience has been explored to some extent by the existing literature. Evidence from the literature in psychology and economics shows that personal experience exerts a larger influence on behaviour compared to other shared sources of information[2], especially very recent and very early experience. These two phenomena are usually referred to as the *recency* and the *primacy* effect and they generate what is known to researchers in psychology as the U-shaped serial-position curve[3].

Diving deeper into the field of finance there is growing evidence that personal experience affects financial behaviour. Kaustia and Knüpfer (2008) and Chiang et al. (2011) show that the likelihood of participating in subsequent IPOs is affected by returns experienced in previous offerings. Choi et al. (2009) provide evidence that investors with high return or low volatility on their 401(k) savings tend to invest a larger fraction of their wealth. Using data from the Survey of Consumer Finances from 1960 to 2007, Malmendier and Nagel (2011) find that individuals born before the 1920s who have experienced the lackluster stock market returns during the Great Depression report higher risk aversion, lower expected returns and are less likely to invest in the stock market. Those that happened to experience lower bond market returns tend to reduce their bond holdings. They also find that returns experienced in the previous year contribute four to six times more to future investment decisions than those experienced thirty years ago. In a similar vein, Malmendier and Nagel (2016) analyse the effect of life-time experience on

[2]For early evidence on the concept of reinforcement learning, see the seminal study by Thorndike (1898). A large body of theoretical and empirical literature studies the role of personal experience in learning, see, for instance, Tversky and Kahneman (1973) for a discussion of the availability bias, Fazio et al. (1978) for experimental evidence on the differential processing of information that results from direct versus indirect experience, Roth and Erev (1995) and Erev and Roth (1998) for experimental data and theory regarding learning in sequential games, Camerer and Ho (1999) for a combined model of reinforcement and belief-based learning, Simonsohn et al. (2008) for experimental analysis of the effect of personal experience in a game theory context.

[3]The psychology literature on these topics goes beyond the scope of this paper. Among others, see Nipher (1878), Ebbinghaus (1913) and Murdock (1962) for evidence on the serial-position effect; for evidence on the primacy effect, see Asch (1946); the recency effect is explored by Deese and Kaufman (1957). See Murdock (1974) for a survey.

inflation expectations using the Reuters/Michigan Survey of Consumers; they show that the effect is stronger for younger respondents, and has a direct effect on their borrowing and savings decisions. Malmendier et al. (2017) analyse the effect of experienced inflation on members of the FOMC board and find similar results. Greenwood and Nagel (2009) investigate the effect of experience on mutual fund managers during the dot.com bubble of the late 1990s. The authors use age as a proxy for experience and show that younger managers were investing more in technology stocks compared to similar older managers and displaying a more pronounced trend-chasing behavior. Chernenko et al. (2016) study the effect of experience on a panel of mutual funds holdings of MBS during the 2003-2007 mortgage boom and show that less experienced managers had larger positions in these securities, especially those backed by subprime mortgages; moreover they show that personal experience outside of the fund had an effect on portfolio choice behaviour. Andonov and Rauh (2018) analyse the effect of experienced returns on a cross-section of U.S. Pension Fund managers, showing a significant effect of past experience on the expected returns that these investors report in annual target asset allocations; in particular, earlier experiences have a stronger effect on investment behaviour. Giglio et al. (2019) look at retail investors' portfolio allocations and match them to beliefs elicited from surveys. They find that stated beliefs have a low explanatory power for the timing of trades, however, they are able to predict the direction and size of those trades that do occur. Finally, there is evidence that experienced risk affects financial behaviour: Knüpfer et al. (2017) show that experienced labour market distress affects portfolio choices, while Lochstoer and Muir (2019) find that individuals have extrapolative beliefs about market volatility.

While the contribution of the above papers is substantial, we argue that most of them are affected by one or more of the following issues: reliance on evidence obtained from surveys where agents report their subjective expected returns, focus on non-professional investors who spend limited time investing and, usually, invest relatively small amounts, and reliance on limited time-series or cross-sections implying that it is harder to perform statistical inference.

Regarding the first issue, the task of recovering investors' expectations is a particularly tricky one. It is well known at least since Harrison and Kreps (1979) that asset prices reveal only risk-neutral expectations of market participants; a way to circumvent this problem is, therefore, to focus attention on expectations elicited from surveys. Most of these measures seem to display high correlations as Greenwood and Shleifer (2014) point out. However Cochrane (2017) argues that there is no guarantee that people report their "true-measure unconditional mean" in surveys. In these regards, Klaus et al. (2018) provide evidence that surveyed expected returns are inconsistent with risk-neutral expected returns, ambiguity averse/robust expected returns or any other risk-adjusted returns[4]. However, nothing guarantees that the reported expected returns are exactly representative of the mathematical physical expectation of investors. Consider for instance a survey respondent that interprets the question as asking "what is the most likely return" instead of "what is the expected return". In that case, the respondent will provide a measure of the modal return rather than its average taken across states of the world. Although the previous example may seem far-fetched, Martin (2017) shows that - for a log investor who holds the market - the physical distribution of returns is asymmetric and, for instance, at the height of the crisis, while the expected return on the S&P 500 was above 20% per year, the author recovers a probability of almost 20% of a 20% decline in the index. Large probability masses far from the mean imply large discrepancies between modal, median and average returns. Beliefs reflected in portfolio choices are more informative and represent the primary object of interest, given that it is ultimately changes in demand and supply that determine the variation in prices. Malmendier and Nagel (2011), Andonov and Rauh (2018) and Giglio et al. (2019) show that portfolio choices are consistent with stated beliefs, but the explanatory power is only partial, while - by construction - our beliefs are fully consistent with trading behaviour.

Regarding the second issue, we argue that there are reasons to believe that sophisticated

[4]Appendix A shows that our framework can also deal with this type of preferences.

professional investors might behave differently compared to households and, for this reason, we have decided to focus our attention on mutual fund managers; they also routinely follow the stock market and therefore there might be reasons to expect them to be less prone to biases or memory issues. While this seems to be true in the case of IPO subscriptions (Chiang et al., 2011), we show that our investors display large biases even though we cannot provide a direct comparison to households. It should also be noted that, to the extent that these financial intermediaries represent a large fraction of total stock market activity, their beliefs will be an important driver of stock price movements. Finally, concerning the third issue, many of the papers dealing with institutional investors focus on specific events (e.g., Greenwood and Nagel (2009) or Chernenko et al. (2016)) or rely on limited time series data (e.g., Andonov and Rauh (2018)). The aim of the present paper is to be more general and explore whether the effect of experienced returns is common across periods and stocks and represents a permanent trait of professional investors' behaviour.

2 Methodology

In this section we provide a detailed description of our empirical strategy. We first explain how we obtain a measure of expected returns given portfolio holdings. We argue that in a wide set of models - including a mean-variance benchmark - we are able to separate the effect of risk and risk aversion from the effect of return beliefs by using the cross-section of manager holdings. We then describe the way we deal with the issue of estimating covariance matrices and, finally, our plan for identifying risk aversion.

2.1 Recovering Subjective Expected Returns

Portfolio choices reveal information about future stock return expectations: this is the main insight of Sharpe (1974) *indirect approach* to mean-variance optimisation whereby beliefs about expected returns are inferred from portfolio holdings, rather than the other way around[5]. Consider the problem of an investor trying to maximise his value function by choosing his portfolio allocations into N risky and one risk-free assets:

$$\max_{\{\boldsymbol{w_t},\dots\}} J(W_t) \tag{2}$$

where $J(.)$ is the value function of the investor evaluated at the current wealth W_t. When returns follow a geometric Brownian motion, the law of motion for wealth is:

$$\frac{dW_t}{W_t} = r_f dt + \boldsymbol{w_t'}(\boldsymbol{\mu_t} - r_f \mathbf{1})dt - \Delta C_t dt + \boldsymbol{w_t'} \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{3}$$

where r_f is the instantaneous risk-free rate (or the instantaneous rate of return of any other *reference* asset with respect to which excess returns are computed), $\boldsymbol{\mu_t}$ is an $N \times 1$ vector of stock return drifts, $\boldsymbol{w_t}$ is an $N \times 1$ vector of stock portfolio weights, $\Sigma^{\frac{1}{2}}$ is an $N \times N$ matrix of instantaneous loadings on the Brownian motion processes $\boldsymbol{Z_t}$, ΔC_t is the (net) outflow of resources[6], and $\mathbf{1}$ is an $N \times 1$ vector of ones.

The investor chooses his optimal portfolio by selecting $\boldsymbol{w_t}$. Notice that we have deliberately remained vague about other potential choice variables, i.e., in what follows, our analysis derives solely from the optimality conditions for the portfolio holdings and the fact that current wealth is the only state variable. Standard dynamic optimisation arguments (Back, 2017) give the following optimality condition:

$$\boldsymbol{w_t^*} = -\frac{J_{W_t}}{W_t J_{W_t W_t}} \Sigma_t^{-1} (\boldsymbol{\mu_t} - r_f \mathbf{1}) \tag{4}$$

where J_{W_t} and $J_{W_t W_t}$ are the first and second derivatives of the value function with

[5]Black and Litterman (1992) start from the same insight to obtain portfolio holdings that combine the manager's views with average realised returns in a consistent way; Cohen et al. (2008) and Shumway et al. (2011) use a similar approach to extract a measure of beliefs from portfolios holdings. The former paper measures the *best ideas* of mutual funds as the investment positions for which the authors can extract the largest expected returns, while the latter analyses the rationality implications of extracted beliefs.

[6]For a standard consumption maximisation problem we can interpret $\Delta C_t = \frac{C_t - Y_t}{W_t}$, i.e., the instantaneous flow of consumption C_t net of the income flow Y_t, expressed as a fraction of wealth W_t. In this setting ΔC_t can be loosely interpreted as the net outflow of money the mutual fund manager is subject to in each period because of redemptions/creation of new fund shares. Because of Markovianity we have that $\Delta C = \Delta C(W_t)$.

respect to current wealth and therefore $-\frac{J_{W_t}}{W_t J_{W_t W_t}}$ is the Arrow (1965)-Pratt (1964) coefficient of instantaneous relative risk aversion measuring the curvature of the value function with respect to wealth, which we will denote $\gamma_{i,t} \equiv -\frac{J_{W_t}}{W_t J_{W_t W_t}}$. Notice that equation (4) is a generalisation of the optimal demand employed by Koijen and Yogo (2019)[7]. We can invert the optimality condition (4) in order to get an expression for expected excess returns as a function of optimal holdings and Σ_t. In particular, we have that

$$\boldsymbol{\mu_{i,t}} - r_f \mathbf{1} = \frac{1}{\gamma_{i,t}} \Sigma_t \boldsymbol{w^*_{i,t}} \tag{5}$$

If we had information about the level of the investor's risk aversion $\gamma_{i,t}$ and the covariance matrix Σ_t, we could obtain an exact measure of his subjective expectations of future one-period ahead excess returns $\boldsymbol{\mu_{i,t}} - r_f \mathbf{1}$. We follow Merton (1980) in arguing that investors should share beliefs regarding Σ_t; we will provide later evidence in support of this assumption. To isolate the effect of $\gamma_{i,t}$, let us consider each element of the vector of excess returns $\boldsymbol{\mu_{i,t}} - r_f \mathbf{1}$. At each point in time t, for each stock j, each manager i forms a measure of expected excess return which we can denote by $(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})_j$[8]. By simply keeping track of the subscripts one can realise that there is variation in expected returns across managers, stocks and time, i.e., along the three dimensions i, j, t. On the other hand, the coefficient of relative risk aversion $\gamma_{i,t}$ varies only at the i-t level, implying that the cross-section of holdings for manager i at time t gives us enough information to isolate the variation in beliefs from the variation in the risk aversion which acts as a level shifter on the demand for risky assets[9]. When instantaneous returns are normally distributed and wealth is the only state variable, any utility function (and therefore any value function $J(W_t)$) gives rise to a demand as the one in (4). We can extend this approach to a wide

[7]The optimal demand in equation (7) of Koijen and Yogo (2019) is equivalent to our specification whenever $-\frac{J_{W_t}}{W_t J_{W_t W_t}} = 1$, i.e., investors have logarithmic utility. It is easy to incorporate short sale constraints in our setting as we show in Appendix A.

[8]$(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})_j$ is the $j-$th element of the vector of expected excess returns for manager i, time t, i.e., $\boldsymbol{\mu_{i,t}} - r_f \mathbf{1} = [(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})_1, ..., (\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})_j, ..., (\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})_N]'$.

[9]For the reader who is familiar with the textbook mean-variance optimisation, this is analogous to the fact that the selection of the tangency portfolio does not depend on the investor's risk aversion which merely influences the relative proportion of wealth invested in the risk-free and risky assets.

class of models where there is an $L \times 1$ vector of Markovian state variables $\boldsymbol{X_t}$ with the following law of motion:

$$\boldsymbol{dX_t} = \phi(\boldsymbol{X_t})dt + \Gamma(\boldsymbol{X_t})\boldsymbol{dZ_t} \tag{6}$$

Standard dynamic optimisation arguments imply that, in that case, the optimal demand will be:

$$\boldsymbol{w_t^*} = -\frac{J_{W_t}}{W_t J_{W_t W_t}} \Sigma_t^{-1} \left((\boldsymbol{\mu_t} - r_f \mathbf{1}) - \sum_{l=1}^{L} \frac{J_{W_t X_{l,t}}}{J_{W_t}} \boldsymbol{K_{l,t}} \right) \tag{7}$$

where $\frac{J_{W_t X_{l,t}}}{J_{W_t}} = \frac{\partial \log J_{W_t}}{\partial X_{l,t}}$ measures the semi-elasticity of the marginal utility of wealth J_{W_t} with respect to the Markovian state variable $X_{l,t}$, and $\boldsymbol{K_{l,t}} = \Sigma_t^{\frac{1}{2}} \boldsymbol{\Gamma_{l,t}}$ represents the vector of instantaneous covariances between returns and the state variable $X_{l,t}$. Let us denote the hedging demand $\boldsymbol{H_t} \equiv \sum_{l=1}^{L} \frac{J_{W_t X_{l,t}}}{J_{W_t}} \boldsymbol{K_{l,t}}$. There are many settings in which we can still disentangle variation in beliefs from variation in hedging demands[10]. First, we might consider the possibility that the mutual fund is facing borrowing constraints. In this case the expected return can be recovered from:

$$(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})_j = \frac{1}{\gamma_{i,t}} \left(\Sigma_t \boldsymbol{w_{i,t}^*} \right)_j + H_{i,t} \tag{8}$$

Similarly, suppose mutual funds managers are ranked according to a common summary statistic (e.g. alpha over a benchmark). The expected excess return can then be approximated by:

$$(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})_j = \frac{1}{\gamma_{i,t}} \left(\Sigma_t \boldsymbol{w_{i,t}^*} \right)_j + H_{j,t} \tag{9}$$

The previous examples show that, by saturating the regressions with the proper fixed effects, we are able to use the cross-section of assets of a particular investor to separate the effect of changes in beliefs (which vary at the i, j, t level) from the effect of changes in risk aversion (varying at the i, t level) and hedging demand (as long as the latter varies at a coarser level). As a caveat, notice that the only situation where we would be un-

[10] For more details, see Appendix A where we analyse the case of borrowing and short selling constraints, concerns about model misspecification and the issue of benchmarking.

able to separate changes in the hedging demand from changes in beliefs is if the hedging demand varied at the stock-manager-time level (i.e., we had $H_{i,j,t}$). This would undermine any attempt to recover variation in beliefs from variation in portfolio holdings; however, the results in the paper would not lose their relevance. First of all, as shown by Moreira and Muir (2019) in the case of time-varying expected returns and volatilities, optimal portfolios can be closely approximated by an affine transformation of the standard mean-variance portfolio. Second, even if expected excess returns cannot be separated from hedging demands, it is not easy to conceive of a story where past experience has a large impact on hedging demands. Third, even if this were the case, we could still interpret all the results in terms of scaled demands ($\Sigma_t \boldsymbol{w}^*_{\boldsymbol{i},\boldsymbol{t}}$) as opposed to beliefs. Asset prices are ultimately determined by investors' holdings and the variation thereof; it would be nice to know whether the effect on investors' demands goes through expected returns ($\boldsymbol{\mu}_{\boldsymbol{i},\boldsymbol{t}} - r_f \mathbf{1}$), risk aversion ($\gamma_{i,t}$) or hedging demands ($\boldsymbol{H}$), but ultimately what matters is the fact that part of the variation in the cross-section and the time-series of assets holdings is due to the returns that the agent has experienced. Having said that, in what follows, we are going to impose the previously discussed restrictions in order to disentangle between the different mechanisms. We are, therefore, going to assume that the issue of hedging demands can be solved by saturating the regression with the appropriate level of fixed effects. In the following two sections, we are going to tackle the two remaining problems, namely, the estimation of the conditional covariance matrix and level of risk aversion.

2.2 Estimating the covariance matrix

As can be seen in the previous section, in order to construct a measure of one-period ahead expected excess returns, we need to have a measure of the conditional covariance matrices. In this paper we will rely on an argument set forth by Merton (1980), which states that, in principle, all investors should agree on Σ_t since it can be very precisely estimated by using increasingly more granular data. In practice it is unavoidable to take a stance on how to estimate the conditional covariance matrix. To make sure that our results do not depend on the chosen estimator for Σ_t, we decide to take three different approaches for this exercise:

1. As a first measure, we compute the sample covariance matrix of stock returns:

$$\hat{\Sigma}_t^{d,1} = \frac{1}{t-1}(R_t - \bar{\boldsymbol{r}}_{\boldsymbol{t}}\mathbf{1}')(R_t - \bar{\boldsymbol{r}}_{\boldsymbol{t}}\mathbf{1}')'$$

   where $R_t = [\boldsymbol{r_{1,t}}, ..., \boldsymbol{r_{j,t}}, ..., \boldsymbol{r_{N,t}}]'$ is an $N \times t$ matrix that contains past realised returns as rows, $\bar{\boldsymbol{r}}_{\boldsymbol{t}}$ is an $N \times 1$ vector that collects sample average returns computed at time t, and $\mathbf{1}$ is a $t \times 1$ vector of ones. We estimate $\hat{\Sigma}_t^{d,1}$ from a one-year rolling window of daily returns[11] and we scale it by $K = \frac{\text{nb. obs.}}{\text{nb. quarters}} = 63.07$ days to obtain our first estimator as $\hat{\Sigma}_t^1 = K \times \hat{\Sigma}_t^{d,1}$. It is well known that it is extremely hard to estimate correlations between stocks and correlations close to unity in absolute value tend to give extreme long-short portfolios. For this reason we resort to the next two measures of the sample covariance matrix;

2. Our second estimate makes use of a Bayesian Stein Shrinkage estimator. We follow Touloumis (2015) and compute the daily covariance matrix $\hat{\Sigma}_t^{d,2}$ as a weighted-average of the sample covariance matrix $\hat{\Sigma}_t^{d,1}$ and a target matrix Σ_{target} which imposes zero correlations across stocks:

$$\hat{\Sigma}_t^{d,2} = \lambda\hat{\Sigma}_t^{d,1} + (1-\lambda)\Sigma_{target}$$

   where Σ_{target} is a diagonal matrix where the elements on the diagonal are the sample estimated variances, namely $\Sigma_{target} = \hat{\Sigma}_t^{d,1} * I_N$ where $*$ denotes the Hadamard product and I_N is a $N \times N$ identity matrix where N is the number of stocks. The

[11] The reader might be worried about the fact that we estimate expected returns employing covariance matrices that rely on past return data, to subsequently regress on past realised returns. However, notice that the same covariance estimates are shared in the cross-section of managers, which is not true for past experienced returns. Furthermore, our estimates of covariance matrices employ only one year of data while the average manager has more than three years of experience with a given stock.

estimator of quarterly covariances is then: $\hat{\Sigma}_t^2 = K \times \hat{\Sigma}_t^{d,2}$;

3. In our third and final approach, we again apply a similar Bayesian Stein Shrinkage Estimator:

$$\hat{\Sigma}_t^{d,3} = \lambda \hat{\Sigma}_t^{d,1} + (1-\lambda)\tilde{\Sigma}_{target}$$

   Following Ledoit and Wolf (2004), $\tilde{\Sigma}_{target}$ is a diagonal matrix with the constant average daily sample variance on the diagonal, namely $\tilde{\Sigma}_{target} = \frac{tr(\hat{\Sigma}_t^{d,1})}{N} I_N$, where $tr(\hat{\Sigma}_t^{d,1})$ is the trace of the covariance matrix, and I_N is a $N \times N$ identity matrix where N is the number of stocks. The estimator is then: $\hat{\Sigma}_t^3 = K \times \hat{\Sigma}_t^{d,3}$.

More details on the construction of $\hat{\Sigma}_t^2$ and $\hat{\Sigma}_t^3$ and the optimal choice of λ are provided in Appendix B. We will show in the rest of the paper that the way we compute the covariance matrices is not very relevant for our results. This should be expected, given that, as long as managers' estimates of covariances are very similar in the cross-section, up to the first order, the covariance matrix behaves like a stock-time fixed effect and therefore will be absorbed by those in the saturated regressions.

2.3 Recovering Risk Aversion

Having discussed the identification of hedging demands and the way we estimate covariance matrices, we now turn to the issue of risk aversion. Let us first disregard any hedging demand for simplicity. The portfolio choice in that case takes the form of (4). It is important to notice that, while we can separate changes in, we are unable to determine the level of $\gamma_{i,t}$, the investor's risk aversion. As a simple example, notice that $\tilde{\gamma}_{i,t} = 2 \times \gamma_{i,t}$ and $\tilde{\boldsymbol{\mu}}_{\boldsymbol{i,t}} - r_f\mathbf{1} = 2 \times (\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f\mathbf{1})$ would yield the exact same portfolio choice as that implied by $\gamma_{i,t}$ and $\boldsymbol{\mu}_{\boldsymbol{i,t}} - r_f\mathbf{1}$. In Section 6, we will impose a plausible restriction on the *level* of subjective expected returns and risk aversion, namely, that fund managers expectations are formed in such a way to minimise the difference with ex-post realised returns[12]. Start from the following identities:

$$\begin{aligned} \boldsymbol{r_{t+1}} - r_f \mathbf{1} &= \mathbb{E}_t[\boldsymbol{r_{t+1}} - r_f \mathbf{1}] + \boldsymbol{\epsilon_{t+1}} && (10)\\ &= (\boldsymbol{\mu_{i,t}} - r_f \mathbf{1}) + \boldsymbol{\epsilon_{i,t+1}} && (11)\\ &= \frac{\gamma_{i,t}}{\gamma_{i,t}}(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1}) + \boldsymbol{\epsilon_{i,t+1}} && (12)\\ &= \gamma_{i,t}(\Sigma_t \boldsymbol{w^*_{i,t}}) + \boldsymbol{\epsilon_{i,t+1}} && (13) \end{aligned}$$

The first line of the above expression is a definition for $\boldsymbol{\epsilon_{t+1}}$: realized returns have to be equal to expected returns plus an orthogonal prediction error. In the second line, we assume that the subjective expectation $(\boldsymbol{\mu_{i,t}} - r_f \mathbf{1})$ and the error $\boldsymbol{\epsilon_{i,t+1}}$ made by the investor are orthogonal. This can be interpreted as a requirement that the expected return is consistent with the law of iterated expectations [13]. The third line multiplies and divides this expectation by the investor's risk aversion $\gamma_{i,t}$. In the empirical counterpart, this will require that the instantaneous relative risk aversion is known to the manager at time t. Finally, we use equation (5) to rewrite (12) as (13). We can, therefore, pin down the level of risk aversion $\gamma_{i,t}$ by running multiple regressions across managers and/or time of stock realised returns on scaled portfolio weights. For instance, if we think that risk aversion is a manager-specific quantity we can run the following regression:

$$r_{j,t+1} - r_f = \alpha_i + \beta_i(\Sigma_t \boldsymbol{w^*_{i,t}})_j + \varepsilon_{i,j,t+1} \quad (14)$$

[12]Conditional expectations are the best predictor in a mean square sense, i.e. given the information set $\mathcal{F}_t$ and the random variable y_{t+1}, the conditional expectation $\mathbb{E}_t[y_{t+1}|\mathcal{F}_t]$ minimises $\mathbb{E}[(y_{t+1} - f_t)^2]$ over all the $\mathcal{F}_t$-measurable functions f_t.

[13]To see this remember that, according to our notation, the expected excess return of manager i using his information set at time t is $\mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f \mathbf{1}] = \boldsymbol{\mu_{i,t}} - r_f \mathbf{1}$. We can therefore rewrite (11) as $\boldsymbol{r_{t+1}} - r_f \mathbf{1} = \mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f \mathbf{1}] + (\boldsymbol{r_{t+1}} - r_f \mathbf{1} - \mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f \mathbf{1}])$. If the law of iterated expectations applies under manager i's expectation, i.e., if $\mathbb{E}_i[\mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f \mathbf{1}]] = \mathbb{E}_i[\boldsymbol{r_{t+1}} - r_f \mathbf{1}]$, it is easy to show that:

- $\mathbb{E}_i[(\boldsymbol{r_{t+1}} - r_f \mathbf{1} - \mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f \mathbf{1}])] = \mathbf{0}$, i.e., no unconditional bias,
- $\mathbb{E}_i[\mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f \mathbf{1}](\boldsymbol{r_{t+1}} - r_f \mathbf{1} - \mathbb{E}_{i,t}[\boldsymbol{r_{t+1}} - r_f \mathbf{1}])'] = 0_{N\times N}$, i.e., the perceived expected return and the error are uncorrelated.

where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{\boldsymbol{i,t}})_j$ is the expected excess return of stock j for manager i, at time t obtained by using the optimal demand $\boldsymbol{w}^*_{i,t}$ of the same manager scaled by the conditional covariance matrix Σ_t. The estimate for α_i will then be a measure of the bias or residual hedging demand. If $\alpha_i = 0$, i.e., the bias or hedging demand is not statistically different from zero, we would then be able to interpret the estimate for β_i as the average coefficient of relative risk aversion of manager i, that is $\beta_i = \gamma_i$. It is important to notice that, while it might be interesting to pin down the *level* of risk aversion and beliefs of each manager, the identification of the learning parameters comes from differential *changes* in beliefs in the cross-section of stocks held, hence it is not affected by our choice of the risk aversion parameter.

3 Data and Summary Statistics

In this section we will describe the data that we will use in the empirical analysis. Data on mutual funds and mutual fund managers' information are obtained from the Center for Research on Security Prices (CRSP) Mutual Fund database. Given that we aim to conduct our analysis at the fund manager level, as opposed to the fund level, we need to construct a dataset of managers' careers. To do this, we first obtain a list of the managers that at any point in time are managing at least one equity fund. We then split each occurrence of multiple managers managing a fund at the same time into separate observations. We also disregard all the cases in which no manager name is available and all the observations where we have words such as "team", "group", "partners" or others that do not allow us to infer who was managing the fund. The most challenging part, however, is to account for the cases in which a typo in the fund manager's name causes CRSP to treat the same manager as two different individuals. As an illustration, an individual named John Smith could, for example, appear as "John Smith", "J. Smith", "J Smith" or just "Smith". In order to tackle this issue, we first match names into pairs using a string matching algorithm. We match similar names using three different string distances: the cosine, Jaccard and Jaro-Wrinkler metrics, and we apply rather large distance-specific thresholds that allow us to keep the names which are sufficiently close. We subsequently proceed by manually checking the matched results which amount to more than 15,000 pairs of matched names. Out of these pairs, our manual exercise left us with roughly 20% of real matches which suggests that we have been quite flexible with the distance thresholds. We would also like to stress the fact that, although our manual check might have contained some errors, i.e., false positive matches and/or false match rejections, so long as these mistakes were random they will only introduce noise in our estimates and not cause any bias. More details on the process are provided in Appendix B. After having matched the names, we assign a unique index to each manager in order to build their careers. This exercise leaves us with 3,214 unique managers in our sample. We next match the above managerial data with CRSP mutual fund data based on the first and last date when a manager has been managing a given fund. We remove index funds, fixed-income funds and funds which mainly own foreign equities following Evans (2010), Benos et al. (2010) and Kacperczyk et al. (2006)[14]. We then match the fund information with mutual fund holdings data from the Thomson-Reuters Institutional Holdings database, using Russ Wermer's MFLinks tables. We finally merge the above data with CRSP data on stock returns and risk-free rates and Compustat-Capital IQ data on firm fundamentals. Since we have monthly mutual fund and return data while holdings data are only available on a quarterly basis, we compute quarterly stock returns from the CRSP monthly data and proceed by merging with Compustat quarterly data. The final dataset comprises of over 13 millions observations for 3,214 distinct managers in the period 1980-2015[15]. Table 1 provides descriptive statistics. The first panel reports summary statistics regarding average and median past returns experienced by

[14]Details on the funds that have been removed can be found in Appendix B.

[15]The number of observations includes a sizeble fraction of holdings that have zero weights but are included because they are part of the manager investment universe. The investment universe is constructed similarly to Koijen and Yogo (2019).

Table 1 : Summary Statistics

The table reports summary statistics for the data used. Column $\bar{x}$ reports the sample average of each variable, column σ its standard deviation, Min the smallest observation, Q1 the first quartile, Median the $50th$ percentile, Q3 the third quartile, Max the largest observation and N the number of observations. The first panel reports summary statistics regarding average and median past returns experienced by managers. The second panel reports six measures of expected excess returns computed as $\hat{\Sigma}_t \boldsymbol{w_t}$. Rows (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; rows (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Rows (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, rows (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and rows (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$. The third panel reports summary statistics on managers' careers; experience refers to the number of quarters since the first time a certain stock appeared in the manager's portfolio; max.experience refers to the maximum experience achieved for each manager-stock pair; tenure refers to the number of quarters since the first time the manager appeared in sample.

	$\bar{x}$	σ	Min	Q1	Median	Q3	Max	N
				Experienced Returns				
average	0.024	0.100	-0.557	-0.010	0.026	0.063	0.607	13,912,677
median	0.014	0.111	-0.871	-0.026	0.021	0.062	1.198	13,912,677
				Expected Excess Returns				
(1)	0.012	0.015	-0.282	0.004	0.007	0.014	1.336	5,416,032
(2)	0.011	0.014	-0.208	0.003	0.006	0.012	0.806	5,416,032
(3)	0.011	0.015	-0.161	0.003	0.006	0.013	0.764	5,416,032
(4)	0.012	0.015	-0.278	0.004	0.007	0.014	0.766	12,707,119
(5)	0.011	0.015	-0.292	0.003	0.006	0.012	1.086	12,707,119
(6)	0.011	0.015	-0.319	0.003	0.006	0.013	1.034	12,707,119
				Managers Careers				
experience	13.158	12.853	1	4	9	17	139	13,912,677
max. experience	13.884	11.981	1	6	11	17	139	1,223,610
tenure	26.896	21.943	1	10	21	39	139	75,179

managers. As one should expect, past experienced returns tend to be right skewed with mean average returns that are larger than mean median returns (2.4% and 1.4%, respectively). While the standard deviation of average experienced returns is similar to the one of median experienced returns (10% and 11% respectively), counterintuitively, the latter seem to be more dispersed, implying that negative experienced returns tend to be right skewed (so that the median is smaller than the average) and positive experienced returns tend to be left skewed (so that the median is larger than the average).

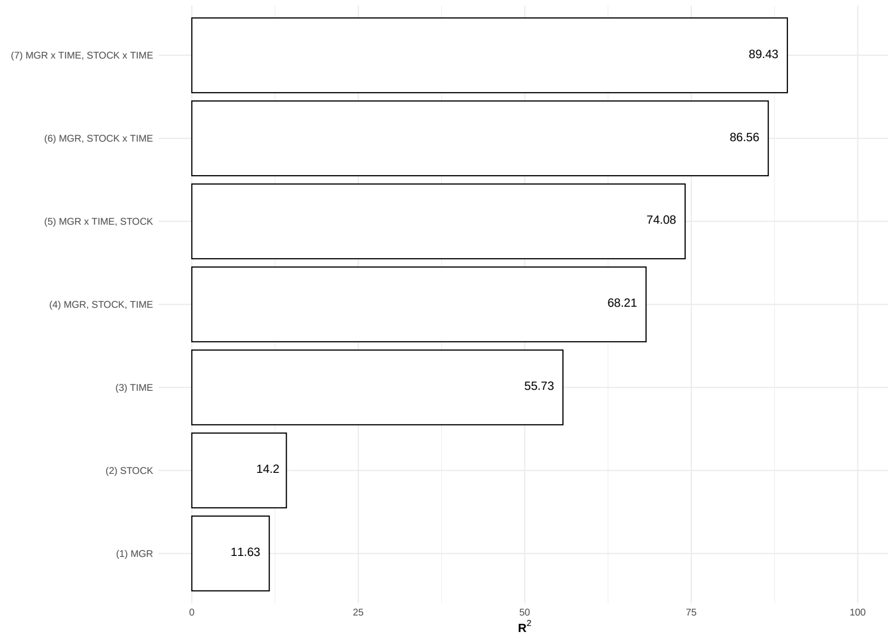
The second panel of Table 1 regards expected returns, which have been computed as explained in Section 2.1. In the rest of the paper we will provide six measures of expected excess returns which we will denote (1)-(6). The first issue regards the inclusion of zero weights[16]. Measures (1)-(3) include only positive weights, while measures (4)-(6) do include the zero weights[17]. Measures (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, measures (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and measures (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$. It is clear from the table that the measures are quite similar in terms of summary statistics. All the measures have an average expected excess return of about 1% per quarter and a median expected excess return of about 0.6%. It should also be noted that, while we have about 12.7 million data points if we consider the zero weights, the number of observations drops to about 5.4 million once we remove the zeros. Figure 1 sheds light on the sources of variation in beliefs. We provide a decomposition of the variation in expected excess returns according to measure (1) by regressing it against various fixed effects. Manager and stock fixed effects explain a small fraction of excess returns (11.63% and 14.20%, respectively), while time fixed effects explain more than half (55.73%) of the variation. This suggests that manager and stock immutable characteristics are relatively less important than aggregate time-varying

[16]Similarly to the present paper, Koijen and Yogo (2019) discuss how the analysis might be affected by including or excluding zero weights.

[17]It might be important to know whether zero weights arise by choice or because the manager cannot short sell stocks that would otherwise appear with negative weights. Appendix A shows how the optimal choice of a manager is affected by short selling constraints and how to deal with them when trying to recover beliefs.

Figure 1 : Explained R^2

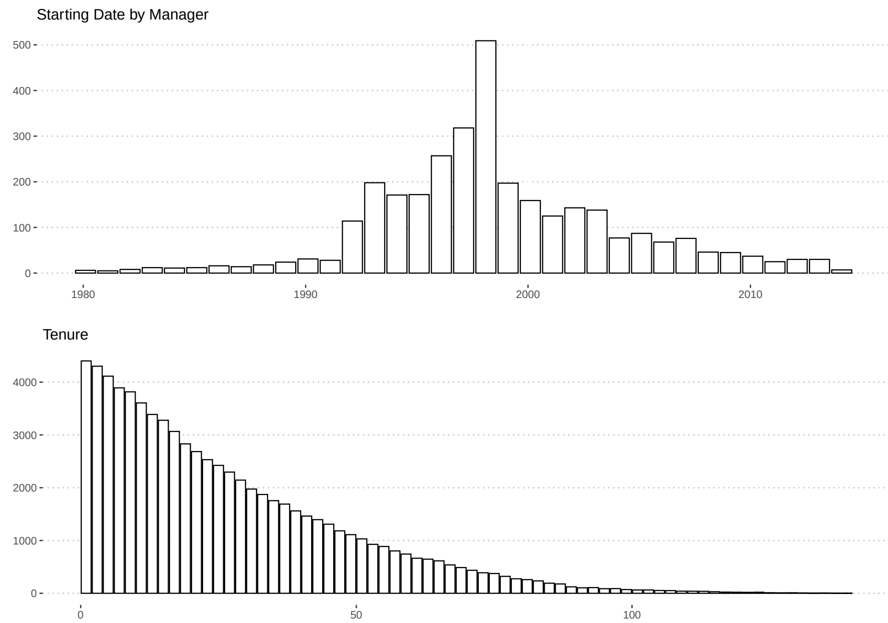
The figure reports the fraction of variation in expected excess returns explained by various fixed effects. For (1), (2) and (3) we report the R^2 of the following regression $\mu_{i,j,t} - r_f = H_k$. (1) reports results for manager fixed effects, i.e., $H_k = H_i$; (2) for stock fixed effects $H_k = H_j$; (3) for time fixed effects $H_k = H_t$. (4) reports the R^2 for separate manager, stock and time fixed effects, i.e., $\mu_{i,j,t} - r_f = H_i + H_j + H_t$. (5) reports the results for manager-time and stock fixed effects, i.e., $\mu_{i,j,t} - r_f = H_{i,t} + H_j$. (6) reports the results for manager and stock-time fixed effects, i.e., $\mu_{i,j,t} - r_f = H_i + H_{j,t}$. (7) reports the results for manager-time and stock-time fixed effects, i.e., $\mu_{i,j,t} - r_f = H_{i,t} + H_{j,t}$.



factors in the formation of expectations. When we separately include manager, stock and time fixed effects the explanatory power rises to almost seventy percent (68.21%). If we allow for interactions between fixed effects, we can explain up to almost ninety percent (89.43%) of the variation in expected excess returns when we include manager-time and stock-time fixed effects. From this decomposition we learn that the largest part of the changes in expected returns is due to time-varying factors, then stock specific characteristics and, finally, factors related to the manager. Adding manager-time and stock-time fixed effects will remove the greatest majority of the variation in expected excess returns and will, thus, ensure that the results will be driven by idiosyncratic variation in expected returns unexplained by systematic factors. This gives more credibility to our identification strategy.

Figure 2 : Managers' Careers

The upper panel shows the distribution of starting date for the managers' careers, as the first date we can track the manager in sample. The bottom panel shows the distribution of tenure across managers and dates as the difference between the current date and the starting date in quarters.



Finally, we consider the data related to the managers' careers which can be analysed with the help of the last panel of Table 1 and Figures 2 and 3. The upper panel of Figure 2 provides information regarding the experience of the managers in the sample. We plot the number of managers by the first time they appear in the sample, which we call the starting date of the fund manager and denote it by $t_{i,0}$. The sample extends from 1980 to 2015 and covers a period of 35 years. Notice, however, that there are fewer managers who start their career in the first ten years compared to the rest of the sample. This can be attributed to low data coverage during the 1980s. Most of the managers in our sample begin their career in the late 1990s. We can observe, however, a wide range of manager starting dates up until the last sample year. We then proceed to construct a tenure variable which measures how many quarters have passed since the start of the manager's career, i.e., for a given manager i and date t, tenure$_{i,t} = t - t_{i,0}$[18]. The lower panel of fig-

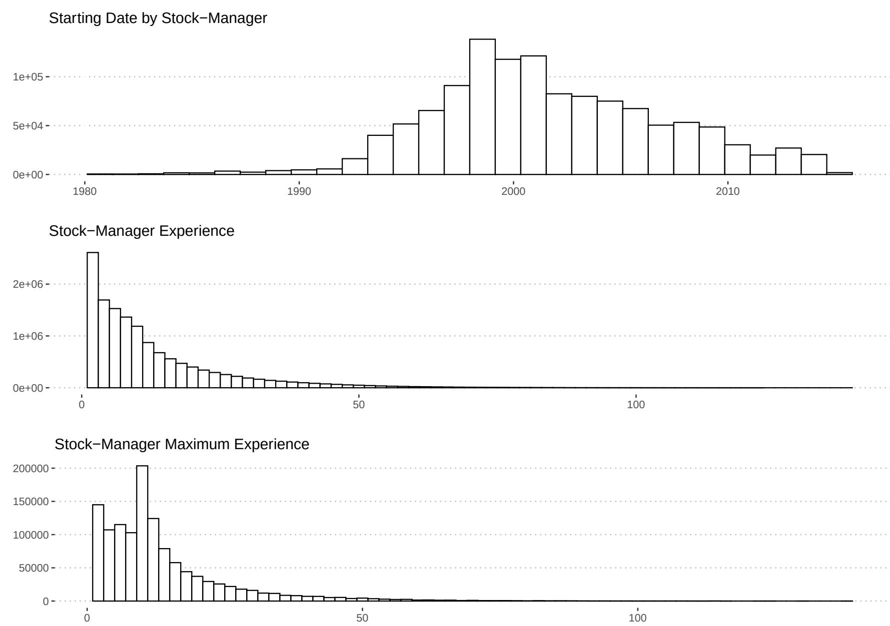
[18]Notice that for each manager we disregard the first quarter of experience, i.e., $t_{i,0}$, when computing the

ure 2 displays the number of managers with a given level of accumulated tenure over the sample period, i.e., the empirical distribution of $(t - t_{i,0})$ for all i, t. Most of the managers in our sample are relatively young and inexperienced, but again, there is quite a large variation in tenure as well, ranging from less than a year up to some managers that are present in the whole sample (i.e., for a period of 35 years). Note that, by construction, the number of observations with a given level of accumulated tenure should be decreasing as, for example, a manager who has 5 quarters of accumulated tenure must also have accumulated 4 quarters of experience previously. In practice, this could be violated for two reasons: the first reason is that mutual funds were required to report holdings at a semi-annual level up until 2003 and only later regulators enforced quarterly reporting, as a result, some funds used to report holdings on a quarterly basis while others did so only on a semi-annual basis prior to 2003; second, there might be some missing data in our sample which means that we might be able to observe a given manager's career and holdings in a particular quarter but not in the previous one. The bottom panel of Table 1 shows that the average tenure is of 26.9 quarters (almost 7 years), but because of the positive skewness manifested in Figure 2, a median of only 22 quarters (5.5 years). We then proceed to the main object of interest of the paper, which is the relationship between each manager and stock. Figure 3 describes the relationship between fund managers and individual stock holdings. The first panel displays the date when a given stock-manager pair has first appeared in our sample which we call the starting date. For each manager i and stock j we can denote the starting date as $t_{i,j,0}$. Unsurprisingly, the largest number of such initiations have occurred in the late nineties and early 2000s, i.e., when the number of managers in our sample significantly increases. There is, however, large variation in the stock-manager starting dates which we will exploit as part of our identification strategy. To see this, the second histogram depicts the length of the personal experience of a given manager with a given stock, i.e., for each manager i, stock j and date t, experience$_{i,j,t} = t - t_{i,j,0}$. It is clear from the histogram that there is a large varia-

statistic.

Figure 3 : Stock-Manager Experience

The upper panel depicts the starting date of each manager-stock pair $t_{i,j,0}$, as the first date in which we observe a certain manager i holding a certain stock j. The middle panel shows the distribution of stock-manager experience, i.e., for any date t, manager i and stock j experience$_{i,j,t} = t - t_{i,j,0}$. The bottom panel reports the distribution of the maximal experience achieved for each manager-stock pair, i.e., max. experience$_{i,j} = \max_t\{\text{experience}_{i,j,t}\}$.



tion in experience. The third panel of Table 1 shows that it ranges from 1 to 139 quarters, with a standard deviation of about 12.9 quarters. The standard deviation is of similar magnitude compared to the average (about 13.2 quarters) and the median experience (9 quarters). The main hypothesis of the paper is that this variation in stock-specific experience will be associated with a variation in expected returns across managers. Finally, we can look at the maximal experience achieved for each stock-manager pair, in the bottom panel of Figure 3 and Table 1[19]. While the average maximal experience and the standard deviation are similar to the above (13.9 and 12 quarters respectively), the median maximal experience is larger (11 quarters compared to 9 quarters of experience).

In the next section, we are going to present the reduced-form results of our empirical

[19]For each manager i and stock j, the maximal experience is defined as max. experience$_{i,j} = \max_t\{\text{experience}_{i,j,t}\}$.

analysis.

4 Reduced-Form Results

The main hypothesis of the paper is that past experienced returns affect expected future returns. Moreover, if that is the case, we would like to further explore whether certain periods carry more relevance than others. In what follows, we will show that differential stock-specific experience across managers indeed matters in the formation of expectations and, in particular, differences in the first and the most recent few quarters of experience play the most crucial role.

The empirical specification in this section will rely on the following argument: we conjecture that the manager will try to estimate future returns by looking at the returns he has experienced over his career. A manager i with $T_{i,j,t}$ quarters of experience with a given stock j at time t might use the average experienced return as a sufficient statistic when forming expectations, i.e., his expected return for that stock can be represented as:

$$\mathbb{E}_{i,t}[r_{j,t+1}] = \beta \bar{r}_{i,j,t} = \beta \left(\frac{1}{T_{i,j,t}} \sum_{k=1}^{T_{i,j,t}} r_{j,t+1-k} \right) \tag{15}$$

where $\bar{r}_{i,j,t}$ denotes the equal-weighted average of stock j returns observed over the investor's experience. Notice that the variation in the length of past experience $T_{i,j,t}$ allows us to exploit the cross-section of managers holding a given stock j as our source of differential treatment[20]. The coefficient β captures the average effect that past observed returns have on expectations formation, while the implicit constant weight $\omega_k = \omega = \frac{1}{T_{i,j,t}}$ means that all past observations are equally-weighted. This choice implies that investors attach equal importance to all observations, however, as the length of experience grows every observation receives a progressively lower weight. Note that this approach does not restrict managers from incorporating other sources of information in their estimation. This

[20]On the other hand, the variation in the length of past experience $T_{i,j,t}$ for a given manager i at time t across different stocks is what helps us in disentangling preferences from expected returns.

Table 2 : The Effect of Average Experienced Returns

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \beta \bar{r}_{i,j,t} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t}$ is the standardised equal-weighted average experienced return, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$..

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.103^{***}	0.103^{***}	0.105^{***}	0.149^{***}	0.148^{***}	0.151^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
N	$1,270,823$	$1,270,823$	$1,270,823$	$2,856,830$	$2,856,830$	$2,856,830$
R^2	0.781	0.765	0.773	0.709	0.692	0.695
Within-R^2	0.006	0.006	0.006	0.009	0.009	0.009
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

can be easily taken into account by saturating the regression with the proper fixed effects. To reiterate, in this case the fixed effects would account for the information that all managers, or all stocks in the portfolio of a given manager, have in common; the coefficient on the average experienced return would thus provide a measure of the incremental effect of experience[21]. We, therefore, show in Table 2 the results of the following regression:

$$\mu_{i,j,t} - r_f = \beta \bar{r}_{i,j,t} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{16}$$

where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t}$ is the previously defined equal-weighted average experienced return[22], $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To better disentangle the effect of experience we focus on the subsample of single-managed funds[23]. The results in the table confirm our main hypothesis: having experienced an increase of one standard deviation in average quarterly return leads to an increase in the expected excess return of between 0.103% and 0.151%; the results are both economically and statistically large and display very minor variation across specifications. This validates our intuition that the estimation method for the covariance matrix is not very consequential. Similarly, the inclusion of the zero weights has no effect on our main findings, even though the drop in R-squared shows that the zeros are indeed informative and cannot be fully explained by the fixed effects alone. The within R-squared shows that the average experienced returns explain between 0.6% and 0.9% of the variation in expected returns. While this might seem low, it is in fact in line with the findings of Koijen and Yogo (2019) that observable characteristics explain a small part of the variation in investors' demands which is mostly explained by latent factors. Table 1 in Appendix D reports the results of a similar regression with manager-time and stock fixed effects, and a number of

[21]Notice that this implies that managers could very well use all past realised returns when they form expectations and this would be absorbed by the stock-time fixed effects. In particular, β would then measure the relative over-weighting of experienced returns.

[22]All the regressions in the paper use standardised explanatory variables for ease of interpretation.

[23]Section 4.1 will analyse the case of co-managed funds, showing indeed that most of the effect washes out when we aggregate across managers.

time-varying stock characteristics, namely, profitability, investment, book-to-market ratio, market equity, and dividend-price ratio. The findings are similar in magnitude and statistically significant, and show that the effect of experienced returns is almost an order of magnitude larger than other known characteristics, confirming again the findings of Koijen and Yogo (2019) that standard predictors have a hard time explaining portfolio choices[24].

So far, we have assumed that the effect of experience is homogeneous. Alternatively, we could allow for more flexible weights in order to investigate whether certain periods matter more than others. Consider the following modified weight: $\omega_k = \frac{\delta_k}{T_{i,j,t}}$, such that $\frac{1}{T_{i,j,t}} \sum_{k=1}^{T_{i,j,t}} \delta_k = 1$. Namely, the manager estimates future returns from the *weighted* average of past experienced returns:

$$\mathbb{E}_{i,t}[r_{j,t+1}] = \beta \sum_{k=1}^{T_{i,j,t}} \frac{\delta_k}{T_{i,j,t}} r_{j,t+1-k} = \sum_{k=1}^{T_{i,j,t}} \beta \delta_k \frac{r_{j,t+1-k}}{T_{i,j,t}} = \sum_{k=1}^{T_{i,j,t}} \tilde{\beta}_k \tilde{r}_{j,t+1-k} \tag{17}$$

The weighting term δ_k is a number centered around one measuring the relative over- or under-weighting of a given past observation. If $\delta_k < 1$, then returns observed k-periods ago are under-weighted, while if $\delta_k > 1$ they are over-weighted relatively to the previous benchmark. The last equality in equation (17) shows that if we rewrite $\tilde{\beta}_k = \beta \delta_k$ and $\tilde{r}_{t+1-k} = \frac{r_{t+1-k}}{T_{i,j,t}}$, then we can run a regression on experience-adjusted returns and obtain:

$$\beta = \frac{1}{T_{i,j,t}} \sum_{k=1}^{T_{i,j,t}} \tilde{\beta}_k, \quad \delta_k = \frac{\tilde{\beta}_k}{\beta} \tag{18}$$

that is, the average effect of past experience can be obtained as the average of the k coefficients $\tilde{\beta}_k$, while the relative weight assigned to the k-periods ago return is given as the ratio of the coefficient on the k-th term and the equal-weighted average of all coefficients.

In practice, this approach breaks down if we have to deal with varying experience lengths

[24]We do not report results for median experienced returns which are virtually identical.

$T_{i,j,t}$, as the number of regressors would change together with $T_{i,j,t}$. For this reason, we group past returns into buckets as a means of fixing the number of regressors. In our first such specification we divide the stock-specific experience of the manager into five non-overlapping buckets of equal length, $\Delta T^q_{i,j,t}$, with $q = \{1, 2, 3, 4, 5\}$[25]. Table 3 reports the results of the following regression:

$$\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{19}$$

for $Q = 5$ and where $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1, 2, 3, 4, 5\}$, is the average return in the q-th bucket. Table 4 reports the results for ten non-overlapping buckets of equal length, i.e., the specification in equation (19) for $Q = 10$. In both cases we focus on the subsample of single-managed funds. To better visualise the results, the estimated coefficients of a regression with five buckets are reported in the upper panel of Figure 4, while the bottom panel reports the results for ten buckets. The picture immediately reveals that the effect of past experienced returns is clearly neither constant nor monotone. Consider, for instance, our first model of expected returns with $Q = 5$ for which we show results in column (1) of Table 3: an increase of a standard deviation in experienced average quarterly return in the most recent or in the earliest period of holding the stock increases the expected return by roughly 0.25% ($\beta_1 = 0.276$ and $\beta_5 = 0.238$); on the other hand, the effect of an increase of one standard deviation midway through the manager's experience has an effect lower by almost an order of magnitude ($\beta_3 = 0.041$). Figure 4 confirms that the effect of experienced returns is "U-shaped" regardless of whether we include the zero weights and independently from the estimator for the covariance matrix used. The lower panel of the figure reports the results for $Q = 10$, painting almost an identical picture. The coef-

[25]To cast this specification in terms of the previously discussed model, let us denote each bucket by $\Delta T^q_{i,j,t}$ and its length as $|\Delta T^q_{i,j,t}|$. We then have that $\delta_k = \beta_q \frac{T_{i,j,t}}{|\Delta T^q_{i,j,t}|}$, where for each time index k in bucket $\Delta T^q_{i,j,t}$ we assign a common effect β_q and take the average return $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} = \sum_{k \in \Delta T^q_{i,j,t}} \frac{r_{j,t+1-k}}{|\Delta T^q_{i,j,t}|}$. Notice that $\frac{T_{i,j,t}}{|\Delta T^q_{i,j,t}|} \approx 5$, where the approximation derives from the fact that we have to split ties when the experience length is not a multiple of five.

Table 3 : The Effect of Experienced Returns - Five Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{5} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1,2,3,4,5\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager-stock pair must have at least 5 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.276***	0.287***	0.272***	0.275***	0.273***	0.281***
	(0.008)	(0.013)	(0.008)	(0.006)	(0.007)	(0.006)
β_2	0.134***	0.132***	0.136***	0.134***	0.132***	0.136***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.004)
β_3	0.041***	0.043***	0.040***	0.042***	0.042***	0.046***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
β_4	0.073***	0.073***	0.077***	0.075***	0.072***	0.078***
	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.002)
β_5	0.238***	0.237***	0.241***	0.238***	0.237***	0.237***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
N	796,021	796,021	796,021	1,958,072	1,958,072	1,958,072
R^2	0.798	0.786	0.792	0.720	0.705	0.708
Within-R^2	0.042	0.043	0.043	0.043	0.042	0.042
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4 : The Effect of Experienced Returns - Ten Buckets

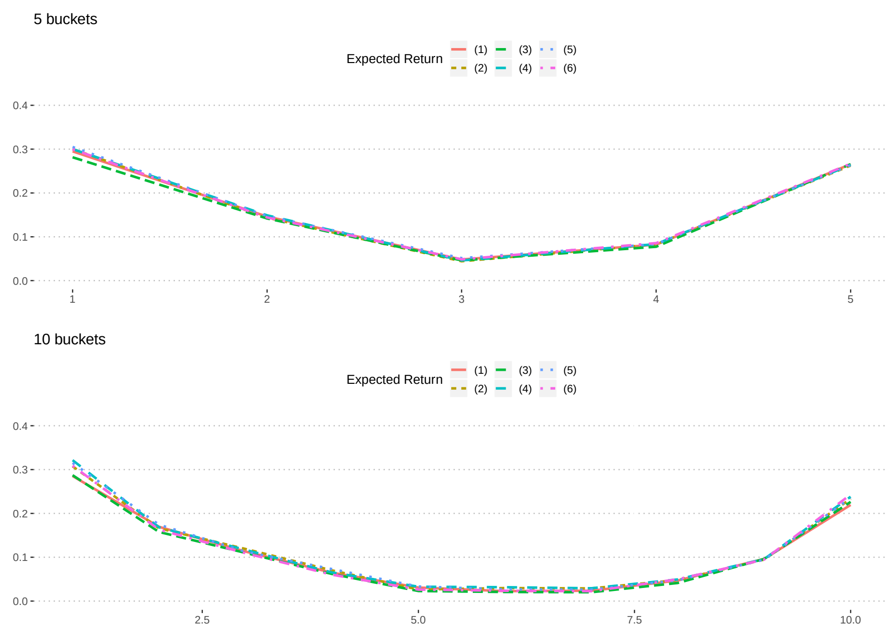
The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{10} \beta_q \bar{r}_{i,j,t\in\Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t\in\Delta T^q_{i,j,t}}$, $q \in \{1, 2, ..., 10\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager-stock pair must have at least 10 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.276***	0.293***	0.258***	0.271***	0.290***	0.268***
	(0.012)	(0.039)	(0.010)	(0.010)	(0.017)	(0.007)
β_2	0.147***	0.163***	0.141***	0.149***	0.157***	0.148***
	(0.008)	(0.023)	(0.009)	(0.006)	(0.009)	(0.005)
β_3	0.100***	0.102***	0.098***	0.100***	0.102***	0.096***
	(0.006)	(0.011)	(0.006)	(0.004)	(0.005)	(0.004)
β_4	0.060***	0.058***	0.067***	0.059***	0.066***	0.061***
	(0.006)	(0.008)	(0.006)	(0.004)	(0.004)	(0.003)
β_5	0.028***	0.023***	0.021***	0.029***	0.030***	0.025***
	(0.005)	(0.008)	(0.005)	(0.003)	(0.003)	(0.003)
β_6	0.022***	0.024***	0.019***	0.021***	0.027***	0.024***
	(0.004)	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)
β_7	0.020***	0.027***	0.026***	0.024***	0.020***	0.023***
	(0.004)	(0.005)	(0.004)	(0.002)	(0.002)	(0.003)
β_8	0.043***	0.045***	0.040***	0.045***	0.046***	0.046***
	(0.004)	(0.006)	(0.005)	(0.002)	(0.003)	(0.003)
β_9	0.080***	0.088***	0.087***	0.086***	0.088***	0.087***
	(0.006)	(0.007)	(0.004)	(0.004)	(0.003)	(0.003)
β_{10}	0.206***	0.204***	0.206***	0.208***	0.216***	0.215***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
N	442,353	442,353	442,353	1,073,779	1,073,779	1,073,779
R^2	0.824	0.812	0.820	0.750	0.736	0.738
Within-R^2	0.039	0.041	0.039	0.039	0.042	0.039
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Figure 4 : Weights on Past Experience

The figure reports the parameter estimates for β_q obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$ is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. The upper panel reports the results for $Q = 5$, while the bottom panel for $Q = 10$. To be included in the upper panel, a manager-stock pair must have at least 5 quarters of experience, while 10 quarters are needed for the bottom panel. Measures (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; measures (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Measures (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, measures (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and measures (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$.



ficients for the ten buckets are similar in magnitude to those for the regression with five buckets and follow the same "U-shaped" pattern. We report in Appendix D the results for various other specifications: Tables 2 and 3 report the results of the previous models with stock fixed effects and the previously mentioned controls, while Tables 4 and 5 describe the results for a model with three non-overlapping equal-sized buckets; finally Tables 6, 7 and 8, 9 report the results for three non-overlapping buckets of unequal length (with stock-time fixed effects or stock fixed effects and varying controls), where the first and last buckets consist of four and eight quarters, respectively. All these specifications confirm the previously discussed results: experienced returns are important in determining expected returns and most of the impact comes from most recent and earliest stock-specific observations. This is evidence in favour of the so-called serial-position effect, concept well studied among researchers in psychology (Murdock, 1974). Moreover, our findings reconcile two apparently distinct phenomena observed in previous research: on the one hand, Malmendier and Nagel (2011) show that economic agents are principally affected by recent experience, while on the other hand Kaustia and Knüpfer (2008) and Hirshleifer et al. (2019) report evidence in favour of the *primacy effect* or *first impression bias*. We show that both effects are present in mutual fund managers and that they need to be separately considered.

So far we have focused our attention on single-managed funds, but one might be interested to know whether the above findings are, in fact, weaker in cases when managers work in teams. We might expect this to be true to the extent that stock-manager specific experience should somewhat cancel out when we group together managers with different experiences. In the next section, we investigate the effect of the number of managers within a team on portfolio holdings and recovered beliefs.

4.1 Co-managed Funds

In this section we check the impact of the number of managers within a team on the effect of experience. Our hypothesis is that personal stock-specific experiences should partly offset each other within a team, so long as the managers that form part of the team have followed different career paths. To explore this hypothesis, we run the following regression:

$$\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_{q,n} \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{20}$$

We split managers into subsamples based on the number of co-managers they work with, i.e., $n_{i,t} \in$ {1, 2, 3, 4 or more} signifies that the manager works in a team of one, two, three or four or more people. We thus obtain a different set of coefficients $\beta_{q,n}$ for each combination of buckets and size of the management team. To save on space, we report in

Table 5 : The Effect of Experienced Returns by Number of Managers

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_{q,n} \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1, 2, 3, 4, 5\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Each column reports the results for the sub-sample of managers working in a team of $n_{i,t} \in \{1, 2, 3, 4$ or more$\}$, members at time t. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. The first four columns report results for measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; the last four columns report results for measure (4), using sample covariance matrices $\hat{\Sigma}^1_t$ and including zero weights on stocks that belong to the manager's investment universe.

	Expected Returns							
	(1)				(4)			
Nr. Managers	1	2	3	≥ 4	1	2	3	≥ 4
β_1	0.276***	0.114***	0.006**	0.015***	0.275***	0.114***	0.006***	0.005***
	(0.008)	(0.014)	(0.003)	(0.003)	(0.006)	(0.008)	(0.002)	(0.002)
β_2	0.133***	0.053***	0.004**	0.008***	0.134***	0.047***	0.004***	0.001
	(0.005)	(0.005)	(0.002)	(0.002)	(0.003)	(0.004)	(0.001)	(0.002)
β_3	0.040***	0.011***	0.004**	0.006***	0.041***	0.010***	0.006***	0.003**
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
β_4	0.072***	0.014***	0.000	0.001	0.074***	0.015***	0.003***	0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
β_5	0.237***	0.017***	0.002**	0.001	0.237***	0.019***	0.004***	0.001
	(0.004)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
N	796,021	580,367	1,000,968	790,078	1,958,072	1,455,284	2,773,180	2,181,406
R^2	0.798	0.912	0.991	0.989	0.720	0.866	0.984	0.978
Within-R^2	0.042	0.002	0.000	0.001	0.043	0.003	0.001	0.000
$w_{i,j,t} = 0$	No	No	No	No	Yes	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^1_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Figure 5 : Weights on Past Experience by Number of Managers

The figure reports the parameter estimates for $\beta_{q,n}$ obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_{q,n} \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. The horizontal axis refers to q, while each line to $n_{i,t} \in \{1, 2, 3, 4 \text{ or more}\}$. The top row reports the results for $Q = 5$, the bottom for $Q = 10$. The left column plots coefficients for measure (1), namely expected excess returns are computed without $w_{i,j,t} = 0$ and using the sample covariance matrix $\hat{\Sigma}^1_t$; the right column for measure (4), namely expected excess returns are computed with $w_{i,j,t} = 0$ and using the sample covariance matrix $\hat{\Sigma}^1_t$.

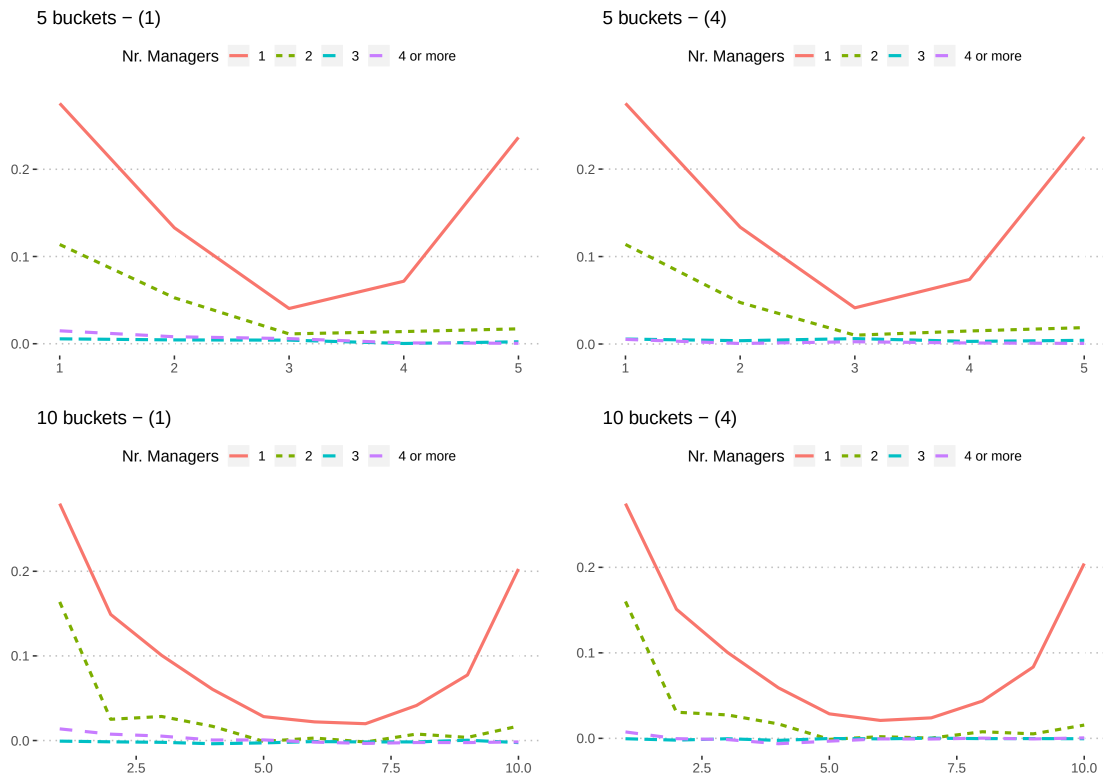


Table 5 of the main text the results of this exercise for $Q = 5$ buckets; the results for the specification with 10 buckets are reported in Table 10 in Appendix D. To better visualise the results, Figure 5 displays the coefficients $\beta_{q,n}$. The two plots on the left-hand-side of the figure show the results for measure (1) while the right-hand-side plots display the coefficients for measure (4). The first row reports the results for $Q = 5$ and the bottom row for $Q = 10$ buckets. As one can see in Table 5, the coefficient on the most recently experienced returns for single-managed funds is more than twice as large as the same for funds managed by two managers; the difference is even larger for the coefficient on the earliest bucket of returns, more specifically, the effect of returns observed at the beginning of a stock-specific experience is more than ten times greater for single-managed funds compared to funds managed by at least two people. The effect on managers working in funds with three or more managers is orders of magnitude lower, while still statistically significant for recent experienced returns. On the other hand, the effect of early returns loses significance. The above is visually confirmed by the plots in Figure 5 showing a rather steep decrease in the coefficients on the earliest bucket of returns across teams of different sizes, especially when going from a single-managed fund to a fund managed by two professionals. The findings are equally pronounced for the specification with ten buckets.

This seems to suggest that a considerable part of personal experience washes out in the cross-section of managers working in the same team, and more so the further we go in the past since managers are more likely to change teams over a longer period of time. On the other hand, recent returns presumably affect all co-managers in a similar way as they have gone through the same recent experience, having been working for the same fund. This could justify the difference in spreads observed between buckets at different horizons, especially if we compare single-managed funds with those managed by two individuals.

In what follows, we will investigate the impact of taxes on managers' investment decisions and the potential explanation the tax regime might have thereof. More specifically, we will examine whether tax considerations can absorb the effect that past experience has on portfolio weights and expectations formation.

4.2 Taxes

The differential treatment of short-term and long-term capital gains in terms of their taxation, together with the possibility to offset capital gains with capital losses, suggests that mutual funds will try to defer the realisation of gains and accelerate the realisation of losses. This implies that it is optimal from the point of view of minimising the tax bill for mutual funds to hold on to assets that performed well in the past and sell assets that had subpar performances[26]. This, in turn, will imply that the previous results could be simply driven by tax considerations. One way to solve the problem could be to model the optimal selling decision in the spirit of Barclay et al. (1998) or Sialm and Zhang (forthcoming) and check if the effect of experienced returns survives after we have taken tax considerations into account. However, in what follows, we are taking a reduced-form approach and make use of the large amount of data on managers who have managed different funds in their career. In particular, we will focus on the subsample of manager-stock pairs where the manager had positive holdings of the stock in the past while managing a different mutual fund compared to the one that he is currently managing. In this setting, tax considerations should be muted given that capital gain overhangs cannot be transferred from one fund to another.

Table 6 reports the results of a regression of expected returns on five buckets of past experience for only those managers that have changed funds, while Table 7 reports the results when we split the previous experience in ten buckets. The results are then summarised in Figure 6 where the upper panel reports the results for five buckets and the lower panel for ten buckets. While the number of observations is greatly reduced (from about 800,000 to slightly more than 110,000 observations if we do not include zero weights, and from about 2 million to approximately 225,000 if we do), the economic and statistical significance of the coefficients is virtually unchanged confirming the previous findings: experienced returns have a sizeable influence on expected excess returns, with the majority of the effect coming from the extreme buckets. If, for instance, we consider measure (1) we notice that the coefficient on the most recent bucket goes from 0.276 to 0.224, while on the earliest from 0.238 to 0.199. We infer, therefore, that no more than 20% of the effect might

[26]Bergstresser and Poterba (2002) show that inflows to mutual funds, and therefore managers' compensation, are affected by the amount of unrealised capital gains, implying that there might be a tension between postponing capital gains indefinitely to provide better after-tax returns for current investors and attracting new investors. Barclay et al. (1998) explicitly tackle this question, showing that indeed managers tend to realise gains early to attract new investors.

Table 6 : Managers Who Have Switched Funds - Five Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{5} \beta_q \bar{r}_{i,j,t\in\Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t\in\Delta T^q_{i,j,t}}$, $q \in \{1,2,3,4,5\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager must have in his current investment universe a stock that he has previously held in a different fund. A manager-stock pair must have at least 5 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.224***	0.272***	0.209***	0.240***	0.214***	0.250***
	(0.018)	(0.031)	(0.023)	(0.014)	(0.014)	(0.022)
β_2	0.133***	0.116***	0.112***	0.125***	0.117***	0.124***
	(0.015)	(0.016)	(0.015)	(0.009)	(0.009)	(0.011)
β_3	0.048***	0.046***	0.031***	0.063***	0.049***	0.066***
	(0.011)	(0.014)	(0.011)	(0.008)	(0.007)	(0.009)
β_4	0.066***	0.071***	0.051***	0.078***	0.065***	0.073***
	(0.010)	(0.010)	(0.011)	(0.007)	(0.006)	(0.007)
β_5	0.199***	0.199***	0.202***	0.216***	0.219***	0.211***
	(0.011)	(0.013)	(0.013)	(0.007)	(0.007)	(0.009)
N	110,037	110,037	110,037	225,676	225,676	225,676
R^2	0.892	0.885	0.889	0.843	0.834	0.842
Within-R^2	0.034	0.038	0.034	0.040	0.040	0.040
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

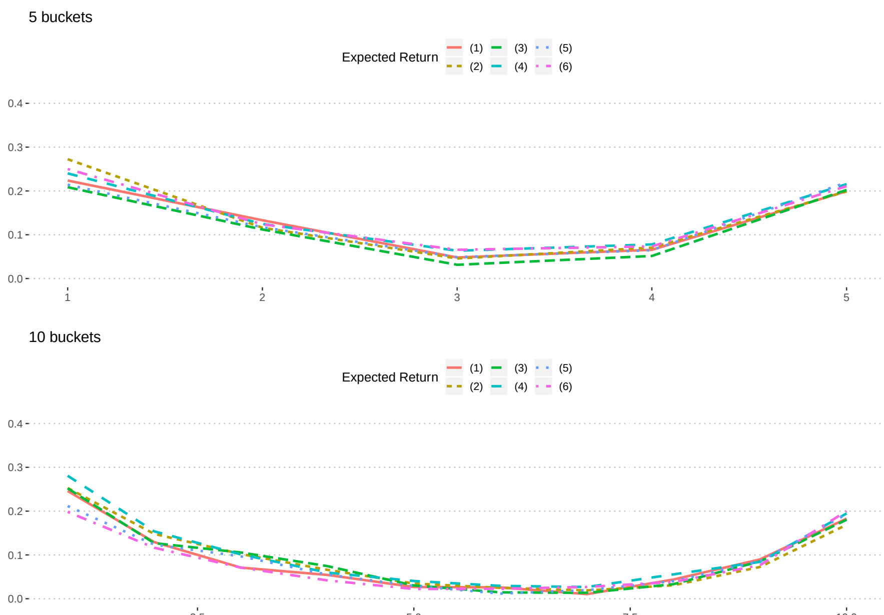
Table 7 : Managers Who Have Switched Funds - Ten Buckets

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{10} \beta_q \bar{r}_{i,j,t \in \Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t \in \Delta T^q_{i,j,t}}$, $q \in \{1, 2, ..., 10\}$, is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. To be included, a manager must have in his current investment universe a stock that he has previously held in a different fund. A manager-stock pair must have at least 10 quarters of experience. Standard errors are clustered at the same level of the fixed effects and are reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.246***	0.253***	0.252***	0.281***	0.212***	0.198***
	(0.026)	(0.026)	(0.027)	(0.055)	(0.017)	(0.021)
β_2	0.129***	0.148***	0.127***	0.154***	0.124***	0.116***
	(0.018)	(0.018)	(0.020)	(0.023)	(0.011)	(0.013)
β_3	0.071***	0.102***	0.106***	0.101***	0.097***	0.071***
	(0.015)	(0.014)	(0.018)	(0.011)	(0.010)	(0.010)
β_4	0.054***	0.066***	0.074***	0.060***	0.055***	0.042***
	(0.015)	(0.012)	(0.017)	(0.009)	(0.008)	(0.009)
β_5	0.027**	0.035***	0.030***	0.040***	0.029***	0.022***
	(0.013)	(0.011)	(0.012)	(0.008)	(0.007)	(0.008)
β_6	0.026**	0.025**	0.015	0.029***	0.012*	0.023***
	(0.011)	(0.011)	(0.013)	(0.007)	(0.007)	(0.006)
β_7	0.011	0.019*	0.013	0.027***	0.020***	0.027***
	(0.011)	(0.010)	(0.011)	(0.006)	(0.007)	(0.007)
β_8	0.044***	0.031***	0.033**	0.056***	0.040***	0.038***
	(0.012)	(0.011)	(0.013)	(0.007)	(0.006)	(0.007)
β_9	0.090***	0.073***	0.085***	0.084***	0.086***	0.077***
	(0.012)	(0.013)	(0.012)	(0.008)	(0.007)	(0.007)
β_{10}	0.183***	0.169***	0.180***	0.195***	0.193***	0.200***
	(0.014)	(0.014)	(0.016)	(0.010)	(0.009)	(0.009)
N	78,920	78,920	78,920	160,237	160,237	160,237
R^2	0.914	0.915	0.914	0.869	0.865	0.867
Within-R^2	0.038	0.037	0.039	0.044	0.040	0.039
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

Figure 6 : Weights on Past Experience - Managers Who Have Switched Funds
The figure reports the parameter estimates for β_q obtained from the following regression: $\mu_{i,j,t} - r_f = \sum_{q=1}^{Q} \beta_q \bar{r}_{i,j,t\in\Delta T^q_{i,j,t}} + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $\bar{r}_{i,j,t\in\Delta T^q_{i,j,t}}$ is the standardised average return in the q-th bucket, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. The upper panel reports the results for $Q = 5$, while the bottom panel for $Q = 10$. To be included, a manager must have in his current investment universe a stock that he has previously held in a different fund. In the upper panel, manager-stock pairs have at least 5 quarters of experience, while 10 quarters are needed for the bottom panel. Measures (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; measures (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Measures (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, measures (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and measures (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$.



be due to tax considerations and we confirm both the *recency* and the *first impression bias*.

Having presented the reduced-form results of our analysis, we now proceed to developing a structural model of learning and estimate its parameters.

5 Structural Results

The reduced-form evidence of the previous section taught us that: experience matters, i.e., average experienced returns are an important determinant of expected returns and; the effect of experience is neither constant nor monotone, in particular, earliest and most recent experience matters the most. However, as shown in Section 4, estimating the shape of the weighting function required us to drop a sizeable amount of observations and potentially lose significant information. For this reason we will posit a functional form for the learning weights and try to estimate its parameters. As Figures 4, 5 and 6 show, we need to allow for non-monotone weights if we want to accurately fit the data. Similarly to Section 4, we assume that the manager uses a weighted average of experienced returns in order to predict future returns. Recall the model in equation (17) where the weights $\frac{\delta_{i,j,t,k}}{T_{i,j,t}}$ captured the differential effect of returns experienced at different points in time. In this section, we will directly model these weights as follows:

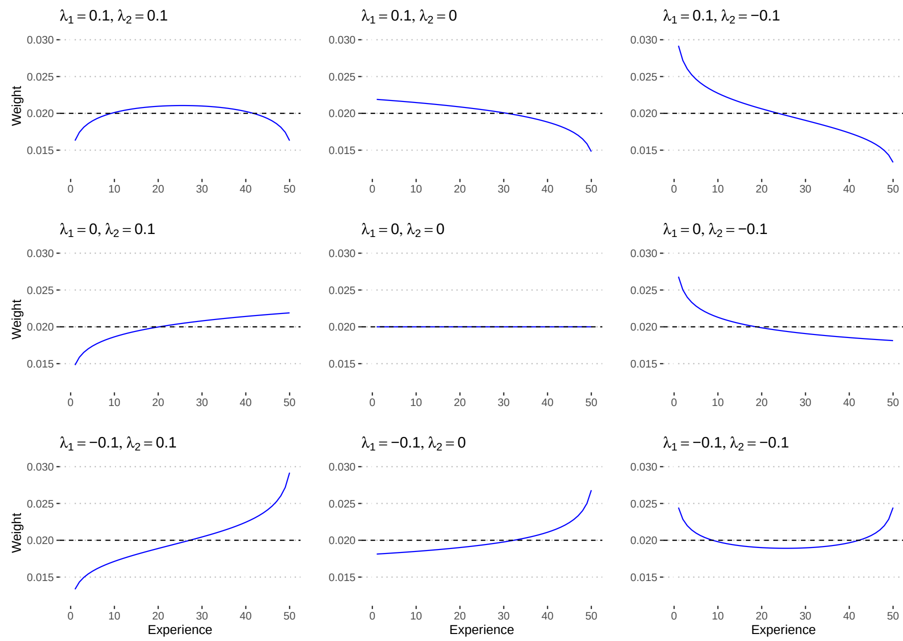
$$\omega_{i,j,t,k} = \frac{\delta_{i,j,t,k}}{T_{i,j,t}} = \frac{(T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}} \tag{21}$$

The functional form in equation (21) is similar to the one used by Malmendier and Nagel (2011) and Malmendier and Nagel (2016)[27]. The weighting function used by these papers depends only on $T_{i,j,t} - k$ and, as such, it confounds two separate effects: the *first impression bias* and the *recency bias*. On the other hand, our weighting function has the advantage of disentangling between these effects: the term $T_{i,j,t} - k$ measures the distance between the return observed at time $t + 1 - k$ and the beginning of a stock-specific experience, hence capturing the *first impression bias*, while k measures the distance from

[27]Our weighting scheme collapses to the one used by Malmendier and Nagel (2011) when $\lambda_2 = 0$.

the current date t, thus capturing the *recency bias*. Figure 7 shows how flexible the par-

Figure 7 : Weighting Functions - Various Examples



simonious parametrisation introduced in equation (21) is. We plot in blue the weighting function for a manager with $T_{ij,t} = 50$ quarters of experience for all the combinations of $\{\lambda_1, \lambda_2\} \in \{-0.1, 0, 0.1\} \times \{-0.1, 0, 0.1\}$[28] and compare it to the black dashed line representing the benchmark $\frac{1}{T_{i,j,t}}$ where the manager equally weights each observation that forms part of his experience. The first parameter, λ_1, governs the strength of the *first impression bias*: when it is negative, the manager is overweighting early experiences relative to the benchmark scenario. The second parameter, λ_2, controls the strength of the *recency bias*: when the sign of λ_2 is negative the manager overweights recent observations relative to the benchmark, and vice versa. As one can see from the examples in Figure 7, using only two parameters, we are able to capture a variety of shapes including linear, convex or concave, increasing or decreasing, monotone or non-monotone weighting schemes arising from the interplay of the *recency* and *first impression bias*. Given the evidence from the reduced-form regressions we expect λ_1 and λ_2 to be negative, implying

[28]Figure 1 in Appendix D plots the weighting function for $\{\lambda_1, \lambda_2\} \in \{-2, 0, 2\} \times \{-2, 0, 2\}$.

that the managers are subject to both effects.

Table 8 : Learning Parameters

The table reports the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \beta\left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{i,j,t+1-k}\right) + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t, $r_{i,j,t+1-k}$ is the realised return of stock j from time $t-k$ to $t+1-k$, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Weights are represented by the following functional form : $\omega_{i,j,t,k} = \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}$. Clustered standard errors are in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}_t^1$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}_t^2$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}_t^3$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.146***	0.139***	0.144***	0.205***	0.205***	0.207***
	0.005	0.005	0.005	0.005	0.005	0.005
λ_1	-1.901***	-1.838***	-1.873***	-1.663***	-1.700***	-1.683***
	0.068	0.064	0.064	0.034	0.038	0.035
λ_2	-1.659***	-1.487***	-1.563***	-1.574***	-1.610***	-1.590***
	0.108	0.116	0.108	0.053	0.061	0.053
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

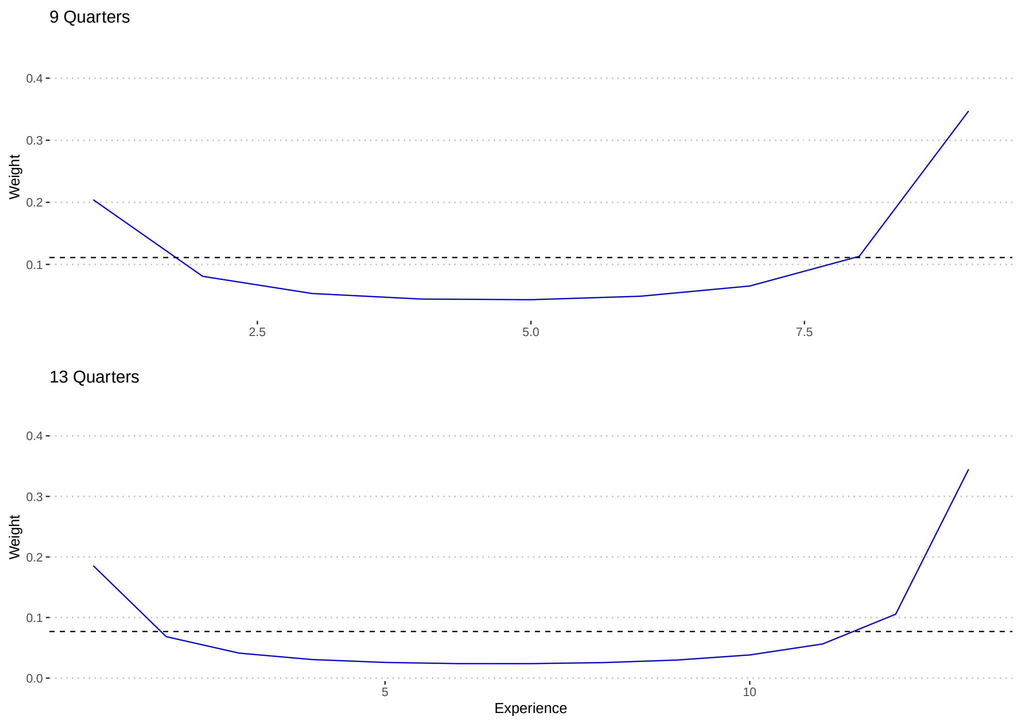
Similarly to the model in equation (19) we include manager-time and stock-time fixed effects to get rid of potentially time-varying unobservable characteristics shared across stocks and managers, respectively. Table 8 reports the NLS estimates of the following regression[29]:

$$\mu_{i,j,t} - r_f = \beta\left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{i,j,t+1-k}\right) + H_{i,t} + H_{j,t} + \epsilon_{i,j,t} \tag{22}$$

[29] Appendix C provides more details on the estimation procedure.

Figure 8 : Empirical Weighting Function

The figure plots the weights implied by the parameter estimates obtained from the following regression: $\mu_{i,j,t} - r_f = \beta \left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{i,j,t+1-k} \right) + H_{i,t} + H_{j,t} + \epsilon_{i,j,t}$, where $\mu_{i,j,t} - r_f$ is the recovered expected one-period ahead return of manager i for stock j at time t according to measure (1), $r_{i,j,t+1-k}$ is the realised return of stock j from time $t-k$ to $t+1-k$, $H_{i,t}$ is a manager-time fixed effect, and $H_{j,t}$ is a stock-time fixed effect. Weights are represented by the following functional form : $\omega_{i,j,t,k} = \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}$. The upper panel reports weights for a manager with stock-specific experience of 9 quarters and the lower for 13 quarters.



$$\omega_{i,j,t,k} = \frac{(T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}$$

Consistent with the reduced-form evidence, both λ_1 and λ_2 are negative and statistically significant across all specifications. The magnitude of the effects is illustrated in Figure 8 where we plot the weighting functions at median and average experience of $T_{i,j,t} = 9, 13$ quarters using the empirically estimated values for λ_1 and λ_2 under model (1). It is evident that the weighting function is always convex and non-monotone, implying that managers overweight the most recent and the earliest returns observed; for instance, a manager with an experience of nine quarters will assign a weight of 0.204 (0.347) to the most recent (earliest) observation, which is 1.84 (3.13) times the benchmark of $1/9$. On the contrary, he will only assign a weight of 0.043 to the middle observation which is 0.39 times the benchmark weight. The results display a slight asymmetry with λ_1 being always larger in magnitude than λ_2 implying that the *recency bias* is marginally weaker compared to the *first impression bias*. This is, however, not a robust feature of the data: Table 11 in Appendix D shows that λ_1 and λ_2 are almost identical once we include only manager-time and stock fixed effects, implying that a large fraction of the recency bias might be captured by stock-time fixed effects as we should expect. Pinning down the actual magnitude of the two biases is extremely difficult given that we have to get rid of a large fraction of the variation in expected returns to achieve identification. Finally, the parameter β in Table 8 measures the average impact of past experience on expected excess returns: the estimates range between 0.139 and 0.207. This is about 4 basis points larger than the baseline results in Table 2 where we did not allow for varying weights[30]. We therefore confirm that once we take into account the fact that recent and early returns might have a differential impact, we find an incremental effect of experience on expected returns. This concludes the main discussion of the paper. However, as explained in

[30]Note that all the results presented refer to standardised variables. In the case of the results in this section we estimate β and then scale its value by the standard deviation of $\left(\sum_{k=1}^{T_{i,j,t}} \omega_{i,j,t,k} r_{i,j,t+1-k}\right)$. This is to avoid directly scaling the weighted average which would affect the computation of the gradient of the right hand side of equation (22) needed to obtain standard errors.

Section 2.3, our methodology allows us to examine in more detail the preferences of investors, which will be the focus of the next section.

6 Risk Aversion

Recall equations (10)-(13); if we assume that subjective expected returns obey the law of iterated expectations, we are able to extract the risk aversion of managers by exploiting the cross-section of individual stock holdings. Running regressions of realised excess returns on scaled demands, as shown in equation (14), we can obtain an estimate for the risk aversion parameter γ and the bias (or residual hedging demand). We start this section by providing evidence from pooled regressions and then proceed by showing results pertaining to the distribution of γ_i obtained from multiple regressions. Table 9 reports the results of the following pooled regression:

$$r_{j,t+1} - r_f = \alpha + \gamma(\Sigma_t \boldsymbol{w}^*_{\boldsymbol{i},\boldsymbol{t}})_j + \epsilon_{i,j,t+1} \tag{23}$$

where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \boldsymbol{w}^*_{\boldsymbol{i},\boldsymbol{t}})_j$ is the expected excess return of stock j for manager i, at time t, obtained by using the optimal demand $\boldsymbol{w}^*_{i,t}$ of the same manager scaled by the conditional covariance matrix Σ_t. If we assume that preferences are constant across managers and time, we obtain a risk aversion coefficient close to unity (between 0.915 and 1.283 across specifications) for our *representative investor*. While the estimate is low compared to other measures obtained from asset prices (Mehra and Prescott, 1985; Kocherlakota, 1996), it is consistent with measures derived from labour choices (Chetty, 2006) and option prices (Martin, 2017). Our *representative investor* displays a quite large and statistically significant bias (or residual hedging demand) of about 1% per quarter.

The pooled results in Table 9 mask a sizeable amount of variability across managers. For this reason, we proceed in estimating separate regressions, one for each manager in

Table 9 : Risk Aversion - Pooled Regressions

The table reports the parameter estimates obtained from the following pooled regression: $r_{j,t+1} - r_f = \alpha + \gamma(\Sigma_t \boldsymbol{w^*_{i,t}})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w}^*_{\mathbf{i},\mathbf{t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. α is the pooled estimated bias across managers and time, γ is the pooled estimated risk aversion across managers and time. Standard errors are clustered at the manager-time and stock-time level and reported in parentheses. Columns (1)-(3) report results without $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; columns (4)-(6) include zero weights on stocks that belong to the manager's investment universe. Columns (1) and (4) use sample covariance matrices $\hat{\Sigma}^1_t$, columns (2) and (5) use Touloumis (2015) covariance matrices $\hat{\Sigma}^2_t$ and columns (3) and (6) use Ledoit and Wolf (2004) covariance matrices $\hat{\Sigma}^3_t$ in the computation of $\hat{\Sigma}_t \boldsymbol{w}_t$.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
α	0.011^{***}	0.011^{***}	0.011^{***}	0.010^{***}	0.010^{***}	0.010^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
γ	0.915^{***}	0.999^{***}	0.958^{***}	1.204^{***}	1.283^{***}	1.255^{***}
	(0.079)	(0.082)	(0.080)	(0.077)	(0.079)	(0.078)
N	5,383,850	5,383,850	5,383,850	12,545,295	12,545,295	12,545,295
R^2	0.004	0.004	0.004	0.006	0.006	0.006
$w_{ijt} = 0$	No	No	No	Yes	Yes	Yes
Covariance	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$	$\hat{\Sigma}^1_t$	$\hat{\Sigma}^2_t$	$\hat{\Sigma}^3_t$

Note: *p<0.1; **p<0.05; ***p<0.01

the sample:

$$r_{j,t+1} - r_f = \alpha_i + \gamma_i(\Sigma_t \boldsymbol{w^*_{i,t}})_j + \epsilon_{i,j,t+1} \tag{24}$$

Given that there seems to be limited difference resulting from the choice of the covariance matrix Σ_t, we report the results using the sample covariance $\hat{\Sigma}^1_t$. Table 10 reports summary statistics for elicited risk aversion and bias referring to measure (1)[31]. We obtain a median (average) relative risk aversion of 1.117 (1.236), in line with the pooled

[31] The results for measure (4) can be found in Appendix D.

Table 10 : Risk Aversion and Bias - Summary Statistics

The table reports the summary statistics of the parameter estimates $\hat{\alpha}_i$ and $\hat{\gamma}_i$ obtained by running one regression per manager with the following specification: $r_{j,t+1} - r_f = \alpha_i + \gamma_i(\Sigma_t \boldsymbol{w^*_{i,t}})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w^*_{i,t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. The reported results are obtained under measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights.

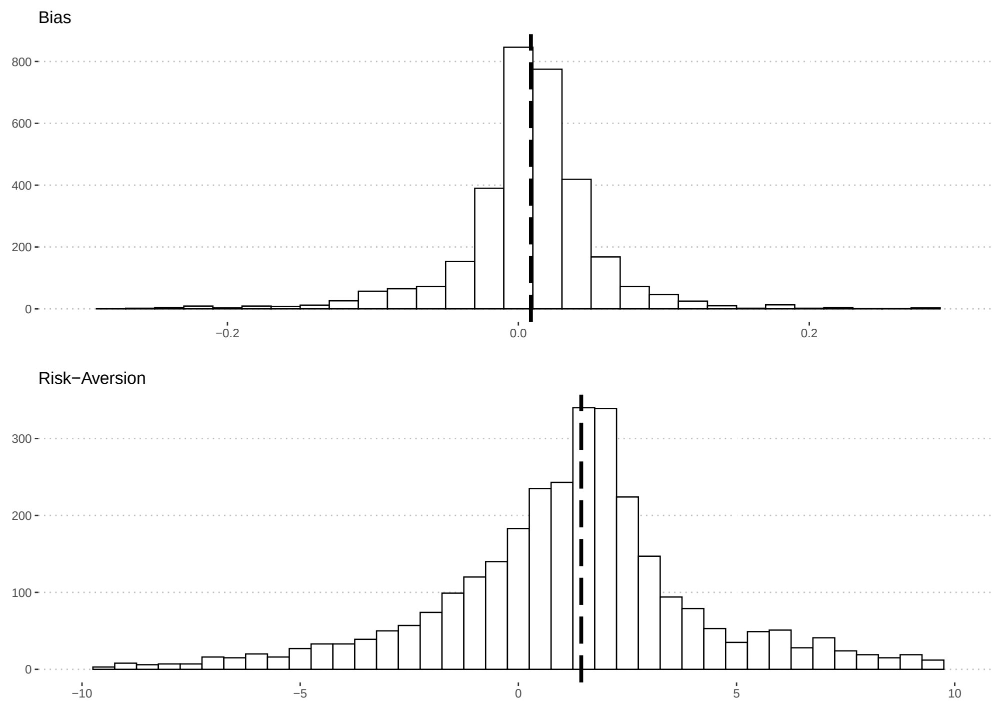
	$\hat{\alpha}_i$	$\hat{\gamma}_i$
mean	0.007	1.236
standard deviation	0.068	5.850
median	0.010	1.117
min	-0.676	-44.666
max	0.736	48.631
skewness	-0.626	1.075
kurtosis	27.395	13.200

results; however, there is a wide dispersion in the estimates with a standard deviation of 5.850. The estimates display positive skewness and are leptokurtic. When we allow for variation in preferences across managers, the bias is reduced on average: the mean bias is only 0.7% and the median bias is 1% per quarter. Figure 9 displays a histogram of the distribution of α_i and γ_i after we have removed outliers. Unfortunately our methodology does not prevent us from obtaining negative values for γ_i whenever the cross-section of revealed beliefs is negatively correlated with realised returns, conflating with the bias α_i. Most of the mass, however, seems to be represented by positive values of γ_i.

We then proceed to exploit the variation of preferences across managers and analyse whether tenure affects risk aversion and bias. Figure 10 displays the bias and the risk aversion as a function of tenure for measures (1) and (4). It is hard to detect a specific pattern in either of the measures; longer tenures seem to be dominated by noise, given that they make use of fewer estimations by construction. Finally, Figure 11 reports the results by date: also in this case it is hard to detect any conclusive evidence. Unfortu-

Figure 9 : Bias and Risk Aversion

The figure shows the empirical distribution of the parameter estimates $\hat{\alpha}_{i,t}$ and $\hat{\gamma}_{i,t}$ obtained by running one regression per manager with the following specification: $r_{j,t+1} - r_f = \alpha_i + \gamma_i(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w}^*_{\mathbf{i},\mathbf{t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. The dashed lines represent the median bias and risk aversion. The histogram is trimmed for outliers to visualise the centre of the distribution.



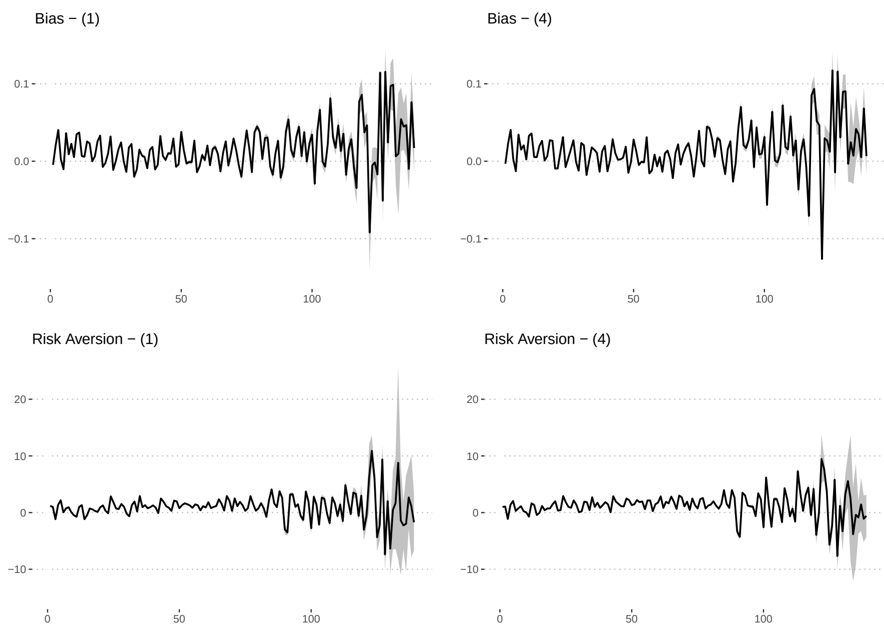
nately, our measures of risk aversion cannot be used to predict or explain future returns given they have been obtained from them: by construction they represent the best linear predictor of $r_{j,t+1} - r_f$ given the information contained in $(\Sigma_t \boldsymbol{w}^*_{i,t})_j$.

7 Conclusions

The paper contributes to the literature on the effect of personal experience on learning and expected returns by analysing a large sample of more than 3,000 professional investors (mutual fund managers) that have been tracked throughout their careers in the 35 years period between 1980 and 2015. Section 2.1 has shown that in a variety of cases it

Figure 10 : Bias and Risk Aversion by Tenure

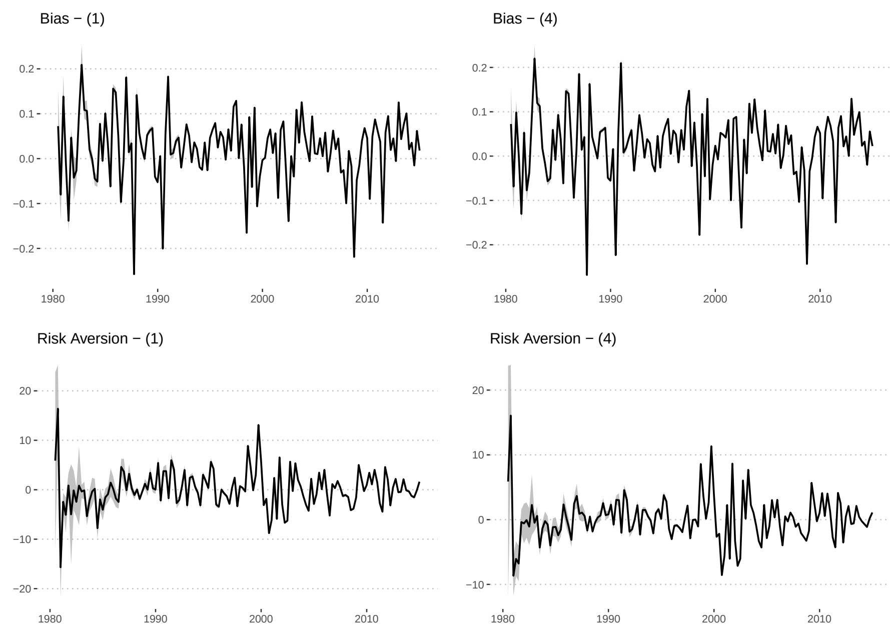
The figure plots the parameter estimates $\hat{\alpha}_\tau$ and $\hat{\gamma}_\tau$ obtained by running one regression per tenure τ with the following specification: $r_{j,t+1} - r_f = \alpha_\tau + \gamma_\tau(\Sigma_t \boldsymbol{w}^*_{i,t})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w}^*_{\mathbf{i},\mathbf{t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. Bias is the estimated parameter $\hat{\alpha}_\tau$, while Risk Aversion is the estimated parameter $\hat{\gamma}_\tau$. Tenure is measured in quarters since the first observation where we can identify the manager. The shaded grey area covers two standard deviations around the point estimate. The left panel reports results for measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; the right panel reports results for measure (4), using sample covariance matrices $\hat{\Sigma}^1_t$ and including zero weights on stocks that belong to the manager's investment universe.



is possible to invert the portfolio demands of our investors to obtain their subjective expected returns by using the identifying assumption that, while beliefs vary at the stock-investor-time level, risk aversion varies at the investor-time level, i.e., risk aversion is constant in the cross-section of stock holdings of a given manager. Similarly, we have been able to account for many cases in which demands display a hedging component by saturating the regressions with fixed effects. Indeed we have shown in Section 3 that almost ninety percent of recovered expected returns can be explained by manager-time and stock-time fixed effects. We have then provided reduced-form evidence showing that professional investors overweight experienced returns compared to other informa-

Figure 11 : Bias and Risk Aversion by Date

The figure plots the parameter estimates $\hat{\alpha}_{i,t}$ and $\hat{\gamma}_{i,t}$ obtained by running one regression per date with the following specification: $r_{j,t+1} - r_f = \alpha_t + \gamma_t(\Sigma_t \boldsymbol{w}^*_{\boldsymbol{i,t}})_j + \epsilon_{i,j,t+1}$, where $r_{j,t+1} - r_f$ is the realised excess return of stock j from time t to $t+1$, and $(\Sigma_t \mathbf{w}^*_{\mathbf{i,t}})_j$ is the demand of manager i for stock j at time t scaled by the conditional covariance matrix Σ_t. Bias is the estimated paramater $\hat{\alpha}_t$, while Risk Aversion is the estimated parameter $\hat{\gamma}_t$. The shaded grey area covers two standard deviations around the point estimate. The left panel reports results for measure (1), using sample covariance matrices $\hat{\Sigma}^1_t$ and no $w_{i,j,t} = 0$, namely including in the computations only strictly positive weights; the right panel reports results for measure (4), using sample covariance matrices $\hat{\Sigma}^1_t$ and including zero weights on stocks that belong to the manager's investment universe.



tion shared across stocks and individuals: having experienced one standard deviation increase in quarterly returns on average leads to an increased expected return of about 10-15 basis points per quarter. Various reduced-form specifications in Section 4 and the structural estimation in Section 5 confirm that the effect of experienced returns is neither constant nor monotone. We have shown that investors exhibit *recency* and *first impression bias*: an investor with a stock-specific experience of nine quarters overweights the most recently observed quarterly returns by 1.84 times and the first experienced return by 3.13 times relative to the constant weight benchmark. These results are most apparent for managers working alone, as opposed to in a team of two or more, suggesting that a significant fraction, though not the entirety, of the effect of personal experience cancels out once aggregated. By looking at managers who have switched funds, we have eliminated the possibility that these findings are purely driven by tax considerations: more than 80% of the effect remains unexplained by tax considerations. We finally turn to the issue of estimating risk aversion and find that a *representative investor* displays a coefficient of relative risk aversion around unity. The paper also finds that individual investors exhibit biases when forming expectations. Finally, when we look at more disaggregated measures, we find that there is large heterogeneity in biases and risk aversion across time and investors. The results in the paper can inform theorists willing to model the preferences and the learning behaviour of professional investors in a way that is consistent with the evidence obtained from portfolio holdings. Consistent with theory more than half of the variation in expected excess returns can be explained by a common time varying component. However, an incremental forty percent is due to investor-specific and stock-specific time-varying effects, hinting at the possibility of time variation in preferences and stock-specific factors shared across investors. Finally, if interested in modelling the idiosyncratic part of expected returns, one should pay particular attention to behavioural factors which play a prominent role as shown by the evidence provided in this paper.

References

Andonov, Aleksandar and Joshua D. Rauh, "The Return Expectations of Institutional Investors," *Working Paper*, November 2018.

Arrow, Kenneth J., "The Theory of Risk Aversion," in "Aspects of The Theory of Risk Bearing" 1965.

Asch, Solomon E, "Forming impressions of personality.," *The Journal of Abnormal and Social Psychology*, 1946, *41* (3), 258.

Back, Kerry E., *Asset Pricing and Portfolio Choice Theory*, Oxford University Press, 2017.

Barclay, Michael J., Neil D. Pearson, and Michael S. Weisbach, "Open-end mutual funds and capital-gains taxes," *Journal of Financial Economics*, 1998, *49* (1), 3 – 43.

Benos, Evangelos, Marek Jochec, and Victor Nyekel, "Can mutual funds time risk factors?," *The Quarterly Review of Economics and Finance*, 2010, *50* (4), 509 – 514.

Bergstresser, Daniel and James Poterba, "Do after-tax returns affect mutual fund inflows?," *Journal of Financial Economics*, 2002, *63* (3), 381 – 414.

Black, Fisher and Robert Litterman, "Global Portfolio Optimization," *Financial Analysts Journal*, 1992, *48* (05), 28–43.

Camerer, Colin and Teck Hua Ho, "Experience-weighted attraction learning in normal form games," *Econometrica*, 1999, *67* (4), 827–874.

Chernenko, Sergey, Samuel G. Hanson, and Adi Sunderam, "Who neglects risk? Investor experience and the credit boom," *Journal of Financial Economics*, 2016, *122* (2), 248–269.

Chetty, Raj, "A new method of estimating risk aversion," *American Economic Review*, 2006, *96* (5), 1821–1834.

Chiang, Yao-Min, David Hirshleifer, Yiming Qian, and Ann E Sherman, "Do investors learn from experience? Evidence from frequent IPO investors," *The Review of Financial Studies*, 2011, *24* (5), 1560–1589.

Choi, James J, David Laibson, Brigitte C Madrian, and Andrew Metrick, "Reinforcement learning and savings behavior," *The Journal of finance*, 2009, *64* (6), 2515–2534.

Cochrane, John H., "Macro-Finance," *Review of Finance*, 2017, *21* (3), 945–985.

Cohen, Randolph B, Christopher Polk, and Bernhard Silli, "Best ideas," LSE Research Online Documents on Economics 24471, London School of Economics and Political Science, LSE Library December 2008.

Deese, James and Roger A Kaufman, "Serial effects in recall of unorganized and sequentially organized verbal material.," *Journal of experimental psychology*, 1957, *54* (3), 180.

Ebbinghaus, Hermann, "Memory: A contribution to experimental psychology," 1913, p. 142.

Erev, Ido and Alvin E Roth, "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria," *American economic review*, 1998, pp. 848–881.

Evans, Richard B, "Mutual fund incubation," *The Journal of Finance*, 2010, *65* (4), 1581–1611.

Fazio, Russell, Mark Zanna, and Joel Cooper, "Direct Experience and Attitude-Behavior Consistency: An Information Processing Analysis," *Personality and Social Psychology Bulletin*, 01 1978, *4*, 48–51.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, "Five Facts About Beliefs and Portfolios," NBER Working Papers 25744, National Bureau of Economic Research, Inc 2019.

Greenwood, Robin and Andrei Shleifer, "Expectations of Returns and Expected Returns," *Review of Financial Studies*, 2014, *27* (3), 714–746.

_ and Stefan Nagel, "Inexperienced investors and bubbles," *Journal of Financial Economics*, August 2009, *93* (2), 239–258.

Harrison, J. Michael and David M. Kreps, "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic Theory*, June 1979, *20* (3), 381–408.

Hirshleifer, David A, Ben Lourie, Thomas Ruchti, and Phong Truong, "First Impressions and Analyst Forecast Bias," *Available at SSRN*, 2019.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, "Unobserved actions of mutual funds," *The Review of Financial Studies*, 2006, *21* (6), 2379–2416.

Kaustia, Markku and Samuli Knüpfer, "Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions," *Journal of Finance*, 2008, *63* (6), 2679–2702.

Klaus, Adam, Dmitry Matveev, and Stefan Nagel, "Do Surveys Expectations of Stock Returns Reflect Risk-Adjustments?," *Working Paper*, September 2018.

Knüpfer, Samuli, Elias Rantapuska, and Matti Sarvimäki, "Formative experiences and portfolio choice: Evidence from the Finnish great depression," *The Journal of Finance*, 2017, *72* (1), 133–166.

Kocherlakota, Narayana, "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature*, 1996, *34* (1), 42–71.

Koijen, Ralph S.J. and Motohiro Yogo, "A Demand System Approach to Asset Pricing," *Journal of Political Economy*, 2019, *127* (4), 1475–1515.

Ledoit, Olivier and Michael Wolf, "Honey, I Shrunk the Sample Covariance Matrix," *The Journal of Portfolio Management*, Summer 2004, *30*, 110–119.

Lochstoer, Lars A and Tyler Muir, "Volatility Expectations and Returns," *Available at SSRN 3473572*, 2019.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics*, 2011, *126* (1), 373–416.

_ and _ , "Learning from Inflation Experiences," *The Quarterly Journal of Economics*, 2016, *131* (1), 53–87.

_ , _ , and Zhen Yan, "The Making of Hawks and Doves: Inflation Experiences on the FOMC," Working Paper 23228, National Bureau of Economic Research March 2017.

Martin, Ian, "What is the Expected Return on the Market?," *The Quarterly Journal of Economics*, 2017, *132* (1), 367–433.

Mehra, Rajnish and Edward C Prescott, "The equity premium: A puzzle," *Journal of monetary Economics*, 1985, *15* (2), 145–161.

Merton, Robert C., "On estimating the expected return on the market : An exploratory investigation," *Journal of Financial Economics*, December 1980, *8* (4), 323–361.

Moreira, Alan and Tyler Muir, "Should Long-Term Investors Time Volatility?," *Journal of Financial Economics*, 2019, *131* (3), 507 – 527.

Murdock, Bennet B, *Human memory: Theory and data.*, Lawrence Erlbaum, 1974.

Murdock, Bennet B Jr, "The serial position effect of free recall.," *Journal of experimental psychology*, 1962, *64* (5), 482.

Nipher, FE, "On the distribution of errors in numbers written from memory," *Transactions of the Academy of Science of St. Louis*, 1878, *3*, 10–1.

Pratt, John W., "Risk Aversion in the small and in the large," *Econometrica*, 1964, *32*, 122–136.

Roth, Alvin E and Ido Erev, "Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term," *Games and economic behavior*, 1995, *8* (1), 164–212.

Sharpe, William F., "Imputing Expected Security Returns from Portfolio Composition," *Journal of Financial and Quantitative Analysis*, June 1974, *9* (03), 463–472.

Shumway, Tyler, Maciej J Szefler, and Kathy Yuan, "The information content of revealed beliefs in portfolio holdings," LSE Research Online Documents on Economics, London School of Economics and Political Science, LSE Library 2011.

Sialm, Clemens and Hanjiang Zhang, "Tax-Efficient Asset Management: Evidence from Equity Mutual Funds," *The Journal of Finance*, forthcoming.

Simonsohn, Uri, Niklas Karlsson, George Loewenstein, and Dan Ariely, "The tree of experience in the forest of information: Overweighing experienced relative to observed information," *Games and Economic Behavior*, 2008, *62* (1), 263–286.

Thorndike, Edward L, "Animal intelligence: An experimental study of the associative processes in animals.," *The Psychological Review: Monograph Supplements*, 1898, *2* (4), i.

Touloumis, Anestis, "Nonparametric Stein-type shrinkage covariance matrix estimators in high-dimensional settings," *Computational Statistics & Data Analysis*, 2015, *83* (C), 251–261.

Tversky, Amos and Daniel Kahneman, "Availability: A heuristic for judging frequency and probability," *Cognitive psychology*, 1973, *5* (2), 207–232.

Internet Appendix for "Revealed Expectations and Learning Biases: Evidence from the Mutual Fund Industry"

A Optimal Portfolio Choice

In what follows, we are going to provide four examples of optimal portfolio choice and describe how we can (or cannot) achieve identification of beliefs. We will first look at an investor facing borrowing constraints, second an investor facing short sale constraints, third we look at an investor worried about model misspecification and, finally, an investor who is tracking a benchmark. We show that we can identify beliefs in the first three cases, while the last one requires us to make additional assumptions.

A.1 Borrowing Constraint

We will follow the approach of Cvitanic and Karatzas (1992), Xu and Shreve (1992) and Tepla (2000). There exist a standard filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,\infty)}, \mathbb{P})$ where all the *usual* regularity conditions are satisfied. We assume that the investor maximises his expected utility over terminal wealth $\mathbb{E}_0[U(W_T)]$. Returns follow a geometric Brownian motion and the investor faces a borrowing constraint. He solves the following problem:

$$\sup_{\{\boldsymbol{w_s}\}_{s\in[0,T]}} \mathbb{E}_0\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] \quad \text{s.t.} \tag{1}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{2}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t} dt + \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{3}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t'}\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) \tag{4}$$

$$\boldsymbol{w_t'}\mathbf{1} \leq k \tag{5}$$

where B_t is the price of a risk-free bond, $\boldsymbol{S_t}$ is a vector of stock prices, $\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \left[\frac{dS_{1,t}}{S_{1,t}}, \dots, \frac{dS_{j,t}}{S_{j,t}}, \dots, \frac{dS_{N,t}}{S_{,N,t}}\right]'$, r_f is the instantaneous risk-free rate, $\boldsymbol{\mu_t}$ is the vector of stock return drifts, $\boldsymbol{w_t}$ is the vector of stock portfolio weights, $\Sigma^{\frac{1}{2}}$ is the matrix of instantaneous loadings on the Brownian motion processes $\boldsymbol{Z_t}$, $\mathbf{1}$ is a vector of ones and k is a real number.

Cvitanic and Karatzas (1992) show that the problem in (1)-(5) is equivalent to an unconstrained problem with modified drifts, i.e., where (2) and (3) are replaced by:

$$\frac{dB_t}{B_t} = (r_f + \delta(\boldsymbol{v_t}))dt \tag{6}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = (\boldsymbol{\mu_t} + \boldsymbol{v_t} + \delta(\boldsymbol{v_t})\mathbf{1})dt + \Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} \tag{7}$$

where the support function $\delta(\boldsymbol{x}) = \sup_{\boldsymbol{w}'\mathbf{1}\leq k}(-\boldsymbol{w}'\boldsymbol{x})$, $\boldsymbol{v_t}$ is such that $\delta(\boldsymbol{v_t}) < \infty$. Cvitanic and Karatzas (1992) show that the optimal $\boldsymbol{v_t^*}$ and portfolio weights $\boldsymbol{w_t^*}$ can be obtained by solving the 'dual' Hamilton-Jacobi-Bellman equation[1]. In particular, the optimal portfolio weights are:

$$\boldsymbol{w_t^*} = \frac{1}{\gamma}\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f\mathbf{1} - \boldsymbol{v_t^*}) \tag{8}$$

where $\boldsymbol{v_t^*} = \arg\min_{\{\boldsymbol{v} \text{ s.t. } \delta(\boldsymbol{v})<\infty\}}\left[||\boldsymbol{\theta_t} + \Sigma_t^{-\frac{1}{2}}\boldsymbol{v_t}||^2 + 2\gamma\delta(\boldsymbol{v_t})\right]$ and $\boldsymbol{\theta_t} = \Sigma_t^{-\frac{1}{2}}(\boldsymbol{\mu_t} - r_f\mathbf{1})$. Tepla (2000) shows that $\boldsymbol{v_t^*} = \bar{v}^*\mathbf{1}$ with $\bar{v}^* = \frac{\gamma(1-\gamma)-\mathbf{1}'\Sigma_t^{-1}(\boldsymbol{\mu_t}-r_f\mathbf{1})}{\mathbf{1}'\Sigma_t^{-1}\mathbf{1}}$ when the borrowing constraint binds, and zero otherwise. Notice that the above result implies that the solution to the constrained optimisation problem is equivalent to that of an unconstrained problem with a risk-free rate shifted by the scalar $\bar{v}^*$. Identification of beliefs is easily achieved in (8) by saturating the model with manager-time fixed effects in order to absorb any variation in manager-specific borrowing constraints. Specifically, for each manager i solving the above problem, the subjective beliefs can be expressed as:

$$\boldsymbol{\mu_{i,t}} - r_f\mathbf{1} = \gamma_i\Sigma_t\boldsymbol{w_{i,t}^*} + \boldsymbol{H_{i,t}} \tag{9}$$

where the manager-time fixed effect is equal to $H_{i,t} = \bar{v_i}^*\mathbf{1}$.

A.2 Short Sale Constraints

The managers solves the following problem[2]:

$$\sup_{\{\boldsymbol{w_s}\}_{s\in[0,T]}} \mathbb{E}_0\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right] \quad \text{s.t.} \tag{10}$$

[1] See Sections 12 and 15 of Cvitanic and Karatzas (1992). In particular, see equations (15.1), (15.2) and (15.10).

[2] The following problem is similar to the discrete problem analyzed by Koijen and Yogo (2019) as $\gamma \to 1$.

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{11}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t} dt + \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{12}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t'} \left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t} \mathbf{1} \right) \tag{13}$$

$$- w_{j,t} \leq 0 \quad \forall j = 1, 2, ..., N \tag{14}$$

The problem (10)-(14) can be solved by using Cvitanic and Karatzas (1992) and Xu and Shreve (1992) dual approach, similarly to the previous section. The support function now becomes $\delta(\boldsymbol{x}) = \sup_{\{-w_{j,t} \leq 0 \quad \forall j=1,2,...,N\}} (-\boldsymbol{w}'\boldsymbol{x})$. As before, we can find $\boldsymbol{v_t^*}$ by solving:

$$\min \left[||\boldsymbol{\theta_t} + \Sigma_t^{-\frac{1}{2}} \boldsymbol{v_t}||^2 + 2\gamma\delta(\boldsymbol{v_t}) \right] \text{ s.t.} \tag{15}$$

$$- \boldsymbol{v_t} \leq \mathbf{0} \tag{16}$$

Denote the vector of Lagrange multipliers on the the constraint in equation (16) by $\boldsymbol{\lambda_t} = [\lambda_{1,t}, ..., \lambda_{N,t}]'$. Taking first-order conditions of the above minimisation problem yields:

$$\Sigma_t^{-1} (\boldsymbol{\mu_t} - r_f \mathbf{1} + \boldsymbol{v_t^*}) + \boldsymbol{\lambda_t} = \mathbf{0} \tag{17}$$

Consider the following partitions: $\boldsymbol{v_t^*} = [\mathbf{0}' \quad \boldsymbol{v_t^{(2)*'}}]'$, $\boldsymbol{\lambda_t} = [\boldsymbol{\lambda_t^{(1)'}} \quad \mathbf{0}']'$, where we have divided between assets for which the short sale constraint does not bind and those for which it does. We can also partition the vector of expected excess returns and the covariance matrix: $\boldsymbol{\mu_t} - r_f \mathbf{1} = [(\boldsymbol{\mu_t^{(1)}} - r_f \mathbf{1})' \quad (\boldsymbol{\mu_t^{(2)}} - r_f \mathbf{1})']'$,

$$\Sigma_t = \begin{bmatrix} \Sigma_t^{(1,1)} & \Sigma_t^{(1,2)} \\ \Sigma_t^{(2,1)} & \Sigma_t^{(2,2)} \end{bmatrix},$$

Standard results imply that the inverse of the covariance matrix can be partitioned as:

$$\Sigma_t^{-1} = \begin{bmatrix} \Omega_t^{(1)} & -\Sigma_t^{(1,1)-1} \Sigma_t^{(1,2)} \Omega_t^{(2)} \\ -\Sigma_t^{(2,2)-1} \Sigma_t^{(2,1)} \Omega_t^{(1)} & \Omega_t^{(2)} \end{bmatrix}$$

where

$$\Omega_t^{(1)} = \left(\Sigma_t^{(1,1)} - \Sigma_t^{(1,2)}\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\right)^{-1}$$

$$\Omega_t^{(2)} = \left(\Sigma_t^{(2,2)} - \Sigma_t^{(2,1)}\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\right)^{-1}$$

Using the above, rewrite equation (17) as:

$$\mathbf{0} = \begin{bmatrix} \Omega_t^{(1)}(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}) - \Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\Omega_t^{(2)}\left(\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1} + \boldsymbol{v_t^{(2)*}}\right) + \boldsymbol{\lambda_t^{(1)}} \\ -\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\Omega_t^{(1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) + \Omega_t^{(2)}\left(\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1} + \boldsymbol{v_t^{(2)*}}\right) \end{bmatrix} \tag{18}$$

Multiplying the second row of (18) by $\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}$ and adding it to the first row allows us to solve for the Lagrange multipliers:

$$\boldsymbol{\lambda_t^{(1)}} = -\Sigma_t^{(1,1)-1}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) \tag{19}$$

Insert the multipliers into the first-order condition in equation (17) to obtain:

$$\boldsymbol{v_t^*} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{v_t^{(2)*}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \Sigma_t^{(1,1)-1}\Sigma_t^{(2,1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) - \left(\boldsymbol{\mu_t^{(2)}} - r_f\mathbf{1}\right) \end{bmatrix} \tag{20}$$

We can now substitute $\boldsymbol{v_t^*}$ into equation (8) and solve for the optimal weights:

$$\boldsymbol{w_t^*} = \begin{bmatrix} \boldsymbol{w_t^{(1)*}} \\ \mathbf{0} \end{bmatrix} = \frac{1}{\gamma}\begin{bmatrix} \Omega_t^{(1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) - \Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}\Omega_t^{(2)}\left(\Sigma_t^{(1,1)-1}\Sigma_t^{(2,1)}(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1})\right) \\ -\Sigma_t^{(2,2)-1}\Sigma_t^{(2,1)}\Omega_t^{(1)}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) + \Omega_t^{(2)}\left(\Sigma_t^{(1,1)-1}\Sigma_t^{(2,1)}(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1})\right) \end{bmatrix} \tag{21}$$

Multiplying the second row by $\Sigma_t^{(1,1)-1}\Sigma_t^{(1,2)}$ and adding the two rows together gives the optimal weights on the unconstrained assets:

$$\boldsymbol{w_t^{(1)*}} = \frac{1}{\gamma}\Sigma_t^{(1,1)-1}\left(\boldsymbol{\mu_t^{(1)}} - r_f\mathbf{1}\right) \tag{22}$$

Intuitively, the optimisation program of a short sale constrained investor results in an unconstrained portfolio allocation over the set of assets for which the constraint does not bind. For each manager i, identification of beliefs can be achieved by inverting equation (22):

$$\boldsymbol{\mu_{i,t}^{(1)}} - r_f \mathbf{1} = \gamma_i \Sigma_t^{(1,1)} \boldsymbol{w_{i,t}^{(1)*}} \tag{23}$$

A.3 Model Misspecification

In this section we follow the approach of Maenhout (2004) and analyse the behaviour of an investor worried about model misspecification. The investor solves the following problem:

$$J_0 = \sup_{\{\boldsymbol{w_s}, C_s\}} \mathbb{E}_0 \left[\int_0^\infty f(c_s, J_s) ds \right] \quad \text{s.t.} \tag{24}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{25}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t} dt + \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} \tag{26}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t'} \left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t} \mathbf{1} \right) - \frac{C_t}{W_t} dt \tag{27}$$

where we have remained vague on the functional form of the value function. Standard dynamic optimisation arguments yield the following HJB equation:

$$0 = \sup_{\{\boldsymbol{w_t}, C_t\}} \{ f(c_t, J_t) dt + \mathbb{E}_t [dJ_t] \} \tag{28}$$

Equation (28) assumes that the investor is certain about the value of $\mathbb{E}_t[dJ_t]$ and chooses his portfolio accordingly. An investor worried about model misspecification will choose the optimal allocation given the worst-case scenario. Following Anderson et al. (2003), Maenhout (2004) shows that the wealth of the investor under the endogenously chosen model for $u(W_t)$ will evolve according to:

$$dW_t = W_t \left(r_f + \boldsymbol{w_t'}(\boldsymbol{\mu_t} - r_f \mathbf{1}) - \frac{C_t}{W_t} \right) dt + W_t \boldsymbol{w_t'} \Sigma_t^{\frac{1}{2}} \boldsymbol{dZ_t} + W_t^2 \boldsymbol{w_t'} \Sigma_t \boldsymbol{w_t} u(W_t) dt \tag{29}$$

where $u(W_t)$ is a drift term chosen by the investor to minimise the following expression:

$$u^*(W_t) = \inf_{u_t} \left\{ \mathbb{E}_t[dJ_t | u_t] + \frac{1}{2\Psi} u_t^2 W_t^2 \boldsymbol{w_t'} \Sigma_t \boldsymbol{w_t} dt \right\} \tag{30}$$

where $\mathbb{E}_t[dJ_t|u_t]$ is computed under the law of motion in equation (29). Among all the models for $u(W_t)$ the investor chooses the least favourable one in terms of its effect on $\mathbb{E}_t[dJ_t|u_t]$, subject to the entropy constraint $\frac{1}{2\Psi}u_t^2W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t}dt$. The HJB equation thus becomes:

$$\begin{aligned}0 = \sup_{\{\boldsymbol{w_t},C_t\}} \inf_{u_t} \quad & f(c_t,J_t) + \frac{\partial J_t}{\partial t} + J_{W_t}W_t\left(r_f + \boldsymbol{w_t'}(\boldsymbol{\mu_t} - r_f\mathbf{1}) - \frac{C_t}{W_t}\right) + \\ & J_{W_t}W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t}u_t + \frac{1}{2\Psi}u_t^2W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t} + \frac{1}{2}J_{W_tW_t}W_t^2\boldsymbol{w_t'}\Sigma_t\boldsymbol{w_t}\end{aligned} \tag{31}$$

The agent will choose $u(W_t)^* = -J_{W_t}\Psi$. The optimal portfolio, therefore, will be:

$$\boldsymbol{w_t^*} = -\frac{J_{W_t}}{[J_{W_tW_t} - J_{W_t}^2\Psi]W_t}\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f\mathbf{1}) \tag{32}$$

An investor concerned about model misspecification will behave like an otherwise identical investor with relative risk aversion of $\gamma_{i,t} = -\frac{[J_{W_tW_t} - J_{W_t}^2\Psi]W_t}{J_{W_t}}$. In this case, identification follows in a way similar to the standard model presented in the main text.

A.4 Benchmarking

In the spirit of van Binsbergen et al. (2008), consider an investor who has his objective function defined over his terminal wealth W_T relative to a benchmark portfolio M_T. He will solve the following problem:

$$J_0 = \sup_{\{\boldsymbol{w_s}\}} \mathbb{E}_0\left[f\left(\frac{W_T}{M_T^{\beta}}\right)\right] \quad \text{s.t.} \tag{33}$$

$$\frac{dB_t}{B_t} = r_f dt, \quad B_0 = 1 \tag{34}$$

$$\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} = \boldsymbol{\mu_t}dt + \Sigma_t^{\frac{1}{2}}\boldsymbol{dZ_t} \tag{35}$$

$$\frac{dW_t}{W_t} = \frac{dB_t}{B_t} + \boldsymbol{w_t'}\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) \tag{36}$$

The benchmark will have weights $\boldsymbol{\theta_t}$ in the N risky assets and will therefore evolve according to:

$$\frac{dM_t}{M_t} = \frac{dB_t}{B_t} + \boldsymbol{\theta_t'}\left(\frac{\boldsymbol{dS_t}}{\boldsymbol{S_t}} - \frac{dB_t}{B_t}\mathbf{1}\right) \tag{37}$$

The problem can be recast in terms of the state variable $X_t = \frac{W_t}{M_t^\beta}$ with the following law of motion:

$$\begin{aligned}\frac{dX_t}{X_t} =&((1-\beta)r_f + (\boldsymbol{w_t} - \beta\boldsymbol{\theta_t})'(\boldsymbol{\mu_t} - r_f\mathbf{1}))dt - \frac{1}{2}\beta(\beta-1)\boldsymbol{\theta_t'}\Sigma_t\boldsymbol{\theta_t}dt + \\ &(\boldsymbol{w_t} - \beta\boldsymbol{\theta_t})'\Sigma_t^{\frac{1}{2}}dZ_t - (\boldsymbol{w_t} - \beta\boldsymbol{\theta_t})'\Sigma_t\beta\boldsymbol{\theta_t}dt\end{aligned} \tag{38}$$

If we set up the HJB equation and take first-order conditions, we obtain the optimal weights:

$$\boldsymbol{w_t^*} = -\frac{J_{X_t}}{J_{X_tX_t}X_t}\Sigma_t^{-1}(\boldsymbol{\mu_t} - r_f\mathbf{1}) + \beta\boldsymbol{\theta_t}\left(1 + \frac{J_{X_t}}{J_{X_tX_t}X_t}\right) \tag{39}$$

In this case, it is not obvious that we can identify beliefs. However, if there is no variation in the objective function in the cross-section of managers adopting the same benchmark portfolio $\boldsymbol{\theta_t}$, stock-time fixed effects would suffice to recover expectations. Although the above model requires an additional assumption to achieve identification, this is consistent with the common practice of evaluating managers using summary statistics such as CAPM alphas (Berk and Van Binsbergen, 2016; Barber et al., 2016). For instance, set $f\left(\frac{W_T}{M_T^\beta}\right) = \frac{1}{1-\gamma}\left(\frac{W_T/W_0}{(M_T/M_0)^\beta}\right)^{1-\gamma} = \frac{1}{1-\gamma}\left(\frac{R_{W,T}}{R_{M,T}^\beta}\right)^{1-\gamma}$. That would be equivalent to solving:

$$\sup_{\{\boldsymbol{w_s}\}} \quad \mathbb{E}_0[r_{W,T}] - \beta\mathbb{E}_0[r_{M,T}] - \frac{(\gamma-1)}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T}) \tag{40}$$

where $r_{W,T} = \log W_T/W_0$ and $r_{M,T} = \log M_T/M_0$ are log-returns. The manager is maximising $\alpha = \mathbb{E}_0[r_{W,T}] - \beta\mathbb{E}_0[r_{M,T}]$ subject to the tracking error penalisation $\frac{(\gamma-1)}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T})$[3]. In this case $-\frac{J_{X_t}}{J_{X_tX_t}X_t} = 1/\gamma$ and we could recover beliefs using:

$$\boldsymbol{\mu_{i,t}} - r_f\mathbf{1} = \gamma\Sigma_t\boldsymbol{w_{i,t}^*} + \boldsymbol{H_{j,t}} \tag{41}$$

where the stock-time fixed effect is $\boldsymbol{H_{j,t}} = (1-\gamma)\beta\Sigma_t\boldsymbol{\theta_t}$.

[3] As it is well known, the agent penalises tracking error for any value of $\gamma > 0$, even for $0 \leq \gamma \leq 1$. To see this, notice that we can substitute $\mathbb{E}_t[r_{W,T} - \beta r_{M,T}] = \log\mathbb{E}_t\left[\frac{R_{W,T}}{R_{M,T}^\beta}\right] - \frac{1}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T})$ and obtain the following objective:

$$\sup_{\{\boldsymbol{w_s}\}} \quad \log\mathbb{E}_t\left[\frac{R_{W,T}}{R_{M,T}^\beta}\right] - \frac{\gamma}{2}\mathbb{V}ar_0(r_{W,T} - \beta r_{M,T})$$

B Data Construction

In this section we provide details on the construction of the data that have been used in the paper. We start with the universe of mutual funds in the CRSP database. We remove funds whose manager name clearly does not refer to a person[4]. After having obtained a list of names of managers, we look for cases in which the same manager is spelled differently, e.g. ”John Smith”, ”J. Smith”, ”J Smith” or just ”Smith”. To be sure that the pairing is done correctly we proceed in the following way: first, we compute a matrix of distances between names using cosine, Jaccard and Jaro-Winkler methods. We then keep pairs that have a distance below a distance-specific threshold (0.10, 0.17, 0.10 for the cosine, Jaccard and Jaro-Wrinkler methods, respectively) that is set to make sure that we avoid false negatives. We then proceed to manually check over 15,000 pairs to guarantee proper matching with the help of online resources and common sense. After having obtained a list of managers with the dates in which they manage a specific fund, we follow Evans (2010) and Benos et al. (2010) to screen for equity mutual funds. First, if available, we keep funds with the following Lipper class: EIEI, G, LCCE, LCGE, LCVE, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, SCCE, SCGE, SCVE. We then keep the funds with missing Lipper class and the following Strategic Insight Objective Code: AGG, GMC, GRI, GRO, ING, SCG. If neither of the previous are available, we use the following Wiesenberger Fund Type Codes: G, G-I, AGG, GCI, GRI, GRO, LTG, MCG, and SCG. We then keep all the funds with policy equal to CS. Finally, we remove funds with less than 80% of holdings in common equity, similarly to Kacperczyk et al. (2006). To check for possible mistakes we keep funds with CRSP objective code starting with E and M and remove those starting with EF. This provides us with a manager-by-manager history of the funds managed that we subsequently match with the S12 type1 file from the Thomson-Reuters Institutional Holdings database, using Russ Wermer’s MFLinks tables. We then proceed by joining with the S12 type2 and type3 files to obtain a history of holdings.

We continue by adding stock return and balance sheet data using CRSP and Compustat, respectively. From the CRSP Compustat Merged Database we select LinkTypes LU and LC and LinkPrim P and C for stocks with share codes of 10 and 11. After we have merged

[4]We use various automatic screens like “advisors”, “ltd”, “limited”, etc..., paired with manual inspection.

the two datasets, we compute dividends using CRSP returns and returns not including distributions, similarly to Koijen and Yogo (2019). From Compustat we compute the following quantities: *me* as market equity, *beme* as the book to market equity ratio, *dp* as the ratio between dividends and market prices, *profitability* as the ratio between operating profits and book equity and *investment* as the growth rate of assets similarly to Fama and French (2015).

We then proceed with the construction of the scaled demands $\hat{\Sigma}_t \boldsymbol{w_t}$. We start from CRSP daily return data and compute covariance matrices using the previous year. We compute three daily covariance matrices: $\hat{\Sigma}_t^{d,1}$ which is the sample covariance matrix, and two Bayesian shrinkage estimates. The first one follows Touloumis (2015) and shrinks the daily sample covariance towards a target diagonal matrix with the sample variances on the diagonal, i.e., the resulting estimator is $\hat{\Sigma}_t^{d,2} = \lambda \hat{\Sigma}_t^{d,1} + (1-\lambda)\Sigma_t^{target}$, with $\Sigma_t^{target} = I_N * \hat{\Sigma}_t^{d,1}$, where $*$ denotes the Hadamard product and I_N is an $N \times N$ identity matrix with N being the number of stocks. The third covariance estimator follows Ledoit and Wolf (2004) and shrinks the daily covariance matrix towards a diagonal matrix with the average variance on the diagonal, i.e., $\hat{\Sigma}_t^{d,3} = \lambda \hat{\Sigma}_t^{d,1} + (1-\lambda)\tilde{\Sigma}_t^{target}$, where $\tilde{\Sigma}_t^{target} = \frac{tr(\hat{\Sigma}_t^{d,1})}{N} I_N$, where $tr(\hat{\Sigma}_t^{d,1})$ is the trace of the daily sample covariance matrix. The shrinkage intensity λ is chosen similarly to Touloumis (2015) to minimize the risk function $\mathbb{E}[||\hat{\Sigma}_t^{d,k} - \Sigma_t^d||_F^2]$ where $||S||_F^2 = \frac{tr(S'S)}{dim(S)}$ denotes the Frobenius norm of matrix S, which results in $\lambda = \frac{Y_{2,T} + Y_{1,T}^2}{TY_{2,T} + \frac{N-T+1}{N} Y_{1,T}^2}$, where $Y_{1,T} = \frac{1}{T}\sum_{s=1}^T X_s' X_s - \frac{1}{P_2^T}\sum_{s \neq h} X_h' X_s$, $Y_{2,T} = \frac{1}{P_2^T}\sum_{s\neq h}(X_h' X_s)^2 - 2\frac{1}{P_3^T}\sum_{s\neq h\neq k} X_s' X_h X_s' X_k + \frac{1}{P_4^T}\sum_{s\neq h\neq k\neq w} X_s X_h' X_k X_w'$ with X_j being the vector of stock returns for which we have T observations and $P_a^b = \frac{b!}{(b-a)!}$. Finally, we can scale the matrices $\hat{\Sigma}_t^{d,k}$ by the average number of trading days in a quarter, which in our sample is equal to $\frac{num.obs}{num.quarters} = 63.07$ to obtain our quarterly estimators $\hat{\Sigma}_t^k = \frac{num.obs}{num.quarters} \times \hat{\Sigma}_t^{d,k}$. We can then proceed to compute scaled demands as $\hat{\Sigma}_t^k \boldsymbol{w_t}$. We compute two vectors of scaled demands for each estimator: one that does not include stocks that currently have zero weights, but belong to the investment opportunity set of the manager, and one that does, i.e., in the first case all the $w_{j,t}$ in $\boldsymbol{w_t}$ are different from zero, while in the second $\boldsymbol{w_t}$ has some zero elements. The investment opportunity set is constructed similarly to Koijen and Yogo (2019) and includes all stocks that are currently held or have ever been held by the manager in the past 11 quarters.

C Structural Estimation

As described in Section 5, we estimate the model in equation (22) via non-linear least squares (NLS). In particular we obtain the coefficients $\hat{\boldsymbol{\theta}} = (\hat{\beta}, \hat{\lambda}_1, \hat{\lambda}_2)'$ by minimising the sum of squared errors:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_i \sum_j \sum_t \left(\mu_{i,j,t} - r_f - \beta \left(\sum_{k=1}^{T_{i,j,t}} \frac{(T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t} - k)^{\lambda_1} k^{\lambda_2}} r_{i,j,t+1-k} \right) - H_{i,t} - H_{j,t} \right)^2 \tag{42}$$

We perform the minimisation with $(\hat{\lambda}_1, \hat{\lambda}_2) \in [-5, 5] \times [-5, 5]$ via Simulated Annealing and limited-memory BFGS[5]. Fixed effects are partialled out by demeaning $\mu_{i,j,t} - r_f$ and $\left(\sum_{k=1}^{T_{i,j,t}} \frac{(T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}}{\sum_{k=1}^{T_{i,j,t}} (T_{i,j,t}-k)^{\lambda_1} k^{\lambda_2}} r_{i,j,t+1-k} \right)$. To compute standard errors, we can rewrite (42) as:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \frac{1}{2} \sum_{p=1}^{P} (y_p - \varphi(\boldsymbol{x_p}; \boldsymbol{\theta}))^2 \tag{43}$$

where the index p is a short-hand for all the P combinations of i, j, t. We next follow the approach of Davidson and MacKinnon (2001) and recover standard errors using Gauss-Newton Regressions. Consider the 3×1 gradient vector $\Psi(\boldsymbol{x_p}; \boldsymbol{\theta}) = \frac{\partial \varphi(\boldsymbol{x_p}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ and the following regression:

$$y_p - \varphi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}}) = \Psi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}})' \boldsymbol{b} + u_p \tag{44}$$

where we regress the residuals $y_p - \varphi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}})$ on the estimated gradient $\Psi(\boldsymbol{x_p}; \hat{\boldsymbol{\theta}})$[6]. Denote the $P \times 3$ matrix of gradient observations as $\hat{\Psi} = [\Psi(\boldsymbol{x_1}; \hat{\boldsymbol{\theta}}), ..., \Psi(\boldsymbol{x_P}; \hat{\boldsymbol{\theta}})]'$, then we can estimate the covariance matrix of the coefficients $\boldsymbol{b}$ using the standard clustered "sand-wich" estimator:

$$S(\hat{\boldsymbol{b}}) = (\hat{\Psi}'\hat{\Psi})^{-1} \hat{\Psi}' \hat{\Omega} \hat{\Psi} (\hat{\Psi}'\hat{\Psi})^{-1} \tag{45}$$

Davidson and MacKinnon (2001) show that the covariance matrix of $\boldsymbol{b}$ in (45) is a consistent estimator for the covariance of $\boldsymbol{\theta}$.

[5] Notice that, conditional on λ_1 and λ_2, β can be estimated via OLS and, therefore, is left unconstrained.

[6] For expositional reasons we exclude the estimated fixed effects from $\boldsymbol{\theta}$. Given that they enter linearly in $\psi(\boldsymbol{x_p}; \boldsymbol{\theta})$, their gradients are identical to the matrix containing the full set of dummies and, therefore, can be taken care of by including dummies on the right hand side of (44) or by demeaning.

D Additional Results

D.1 Tables

Table 1 : The Effect of Average Experienced Returns

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.148***	0.140***	0.146***	0.188***	0.179***	0.189***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
profitability	-0.002	-0.0010	-0.002	-0.003	-0.002	-0.004
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
investment	0.040***	0.032***	0.035***	0.051***	0.039***	0.042***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
BE/ME	0.012	0.020***	0.012*	0.016*	0.019**	0.017**
	(0.008)	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)
ME	0.011	0.009	0.012	0.009	0.0009	0.008
	(0.015)	(0.012)	(0.013)	(0.018)	(0.016)	(0.017)
D/P	-0.019***	-0.017***	-0.018***	-0.005	-0.006	-0.005
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
N	1, 153, 333	1, 153, 333	1, 153, 333	2, 596, 853	2, 596, 853	2, 596, 853
R^2	0.591	0.583	0.588	0.546	0.538	0.536
Within-R^2	0.016	0.014	0.015	0.021	0.019	0.021
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2 : The Effect of Experienced Returns - Five Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.297***	0.283***	0.286***	0.274***	0.264***	0.277***
	(0.009)	(0.010)	(0.008)	(0.006)	(0.007)	(0.006)
β_2	0.137***	0.129***	0.138***	0.125***	0.115***	0.121***
	(0.009)	(0.008)	(0.008)	(0.005)	(0.005)	(0.005)
β_3	0.061***	0.054***	0.057***	0.055***	0.048***	0.054***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.004)	(0.005)
β_4	0.084***	0.083***	0.088***	0.085***	0.078***	0.084***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
β_5	0.266***	0.259***	0.262***	0.267***	0.258***	0.261***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
profitability	-0.005*	0.0009	-0.005	-0.010**	-0.006	-0.008*
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
investment	0.006	0.003	0.002	0.019***	0.011*	0.013**
	(0.008)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)
BE/ME	0.066***	0.072***	0.062***	0.053***	0.056***	0.054***
	(0.014)	(0.015)	(0.016)	(0.010)	(0.010)	(0.010)
ME	-0.009	-0.007	-0.005	-0.012	-0.017	-0.013
	(0.015)	(0.012)	(0.013)	(0.020)	(0.019)	(0.019)
D/P	-0.008	-0.003	-0.006	0.005	0.003	0.004
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)
N	724,999	724,999	724,999	1,783,648	1,783,648	1,783,648
R^2	0.594	0.587	0.591	0.556	0.547	0.545
Within-R^2	0.066	0.064	0.065	0.070	0.067	0.069
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3 : The Effect of Experienced Returns - Ten Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.259***	0.224***	0.237***	0.235***	0.229***	0.242***
	(0.011)	(0.008)	(0.008)	(0.007)	(0.006)	(0.007)
β_2	0.123***	0.116***	0.131***	0.124***	0.115***	0.120***
	(0.007)	(0.006)	(0.007)	(0.005)	(0.005)	(0.005)
β_3	0.098***	0.094***	0.102***	0.088***	0.083***	0.084***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.004)	(0.004)
β_4	0.078***	0.064***	0.073***	0.069***	0.063***	0.065***
	(0.006)	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)
β_5	0.059***	0.048***	0.049***	0.047***	0.041***	0.043***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
β_6	0.061***	0.053***	0.055***	0.053***	0.053***	0.052***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
β_7	0.067***	0.066***	0.066***	0.066***	0.057***	0.063***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)	(0.004)
β_8	0.074***	0.063***	0.067***	0.071***	0.070***	0.074***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
β_9	0.107***	0.109***	0.113***	0.120***	0.114***	0.112***
	(0.005)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
β_{10}	0.243***	0.239***	0.239***	0.243***	0.247***	0.246***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)
profitability	-0.005	-0.003	-0.007	-0.013**	-0.009	-0.011*
	(0.004)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
investment	-0.015*	-0.011	-0.015*	-0.005	-0.011*	-0.007
	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)	(0.007)
BE/ME	0.076***	0.078***	0.065***	0.069***	0.071***	0.066***
	(0.018)	(0.019)	(0.024)	(0.014)	(0.013)	(0.013)
ME	-0.019	-0.014	-0.011	-0.022	-0.028	-0.021
	(0.016)	(0.015)	(0.015)	(0.023)	(0.021)	(0.022)
D/P	-0.001	-0.003	-0.006	0.008	0.006	0.010
	(0.010)	(0.010)	(0.009)	(0.008)	(0.008)	(0.008)
N	403, 968	403, 968	403, 968	980, 175	980, 175	980, 175
R^2	0.598	0.588	0.596	0.567	0.557	0.555
Within-R^2	0.065	0.061	0.063	0.070	0.070	0.071
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4 : The Effect of Experienced Returns - Three Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.283***	0.294***	0.288***	0.284***	0.288***	0.288***
	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)	(0.004)
β_2	0.077***	0.082***	0.078***	0.078***	0.079***	0.080***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
β_3	0.229***	0.231***	0.232***	0.231***	0.231***	0.233***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)
N	1,031,564	1,031,564	1,031,564	2,483,275	2,483,275	2,483,275
R^2	0.777	0.762	0.769	0.704	0.688	0.690
Within-R^2	0.039	0.041	0.040	0.040	0.040	0.040
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5 : The Effect of Experienced Returns - Three Buckets

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.280***	0.273***	0.277***	0.273***	0.267***	0.277***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)
β_2	0.073***	0.069***	0.071***	0.066***	0.060***	0.066***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.004)	(0.005)
β_3	0.236***	0.230***	0.233***	0.238***	0.229***	0.235***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
profitability	-0.001	-0.000	-0.002*	-0.003	-0.002	-0.004
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
investment	0.019***	0.013*	0.014**	0.032***	0.021***	0.024***
	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)
BE/ME	0.048***	0.056***	0.048***	0.043***	0.044***	0.044***
	(0.010)	(0.011)	(0.011)	(0.009)	(0.009)	(0.008)
ME	-0.002	-0.003	0.000	-0.006	-0.014	-0.007
	(0.014)	(0.012)	(0.012)	(0.019)	(0.017)	(0.017)
D/P	-0.009	-0.007	-0.008	0.004	0.002	0.003
	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)
N	937, 382	937, 382	937, 382	2, 258, 925	2, 258, 925	2, 258, 925
R^2	0.582	0.573	0.578	0.545	0.536	0.535
Within-R^2	0.056	0.055	0.056	0.058	0.056	0.059
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6 : The Effect of Experienced Returns - Three Buckets and $k = 4$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_2	0.016***	0.015***	0.022***	0.008***	0.008***	0.014***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
β_3	0.161***	0.159***	0.160***	0.166***	0.168***	0.165***
	(0.004)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
N	618,451	618,451	618,451	1,499,594	1,499,594	1,499,594
R^2	0.812	0.799	0.807	0.744	0.729	0.733
Within-R^2	0.021	0.021	0.021	0.021	0.021	0.020
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 7 : The Effect of Experienced Returns - Three Buckets and $k = 4$ Quarters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.208***	0.189***	0.198***	0.206***	0.194***	0.205***
	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
β_2	0.093***	0.083***	0.089***	0.077***	0.070***	0.076***
	(0.008)	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)
β_3	0.215***	0.208***	0.209***	0.229***	0.225***	0.223***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
profitability	-0.005	-0.003	-0.007	-0.013**	-0.010	-0.011*
	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.006)
investment	-0.002	-0.003	-0.006	0.004	-0.001	0.001
	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)	(0.007)
BE/ME	0.056***	0.059***	0.048***	0.052***	0.051***	0.049***
	(0.013)	(0.015)	(0.018)	(0.012)	(0.011)	(0.011)
ME	-0.006	-0.002	0.001	-0.009	-0.012	-0.007
	(0.016)	(0.014)	(0.015)	(0.022)	(0.020)	(0.021)
D/P	-0.008	-0.009	-0.012	0.004	0.002	0.004
	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)
N	564,287	564,287	564,287	1,367,732	1,367,732	1,367,732
R^2	0.598	0.590	0.597	0.570	0.560	0.558
Within-R^2	0.042	0.039	0.040	0.046	0.044	0.045
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8 : The Effect of Experienced Returns - Three Buckets and $k = 8$ Quarters.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_2	0.020***	0.017***	0.026***	0.014***	0.020***	0.019***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.002)
β_3	0.137***	0.131***	0.135***	0.136***	0.144***	0.141***
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
N	343,058	343,058	343,058	753,526	753,526	753,526
R^2	0.870	0.864	0.866	0.834	0.821	0.824
Within-R^2	0.021	0.020	0.021	0.020	0.022	0.021
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time	Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 9 : The Effect of Experienced Returns - Three Buckets and $k = 8$ Quarters.

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1	0.168***	0.149***	0.157***	0.165***	0.152***	0.160***
	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.008)
β_2	0.067***	0.058***	0.065***	0.066***	0.063***	0.065***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
β_3	0.179***	0.173***	0.178***	0.183***	0.187***	0.186***
	(0.006)	(0.007)	(0.005)	(0.005)	(0.004)	(0.004)
profitability	-0.003	-0.003	-0.007	-0.016*	-0.014	-0.014
	(0.004)	(0.005)	(0.005)	(0.008)	(0.009)	(0.009)
investment	-0.029***	-0.027***	-0.029***	-0.029***	-0.035***	-0.032***
	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
BE/ME	0.093***	0.100***	0.092***	0.077***	0.074***	0.069***
	(0.027)	(0.027)	(0.026)	(0.019)	(0.019)	(0.018)
ME	-0.015	-0.007	-0.008	-0.023	-0.027	-0.021
	(0.019)	(0.017)	(0.018)	(0.027)	(0.025)	(0.026)
D/P	-0.001	0.004	-0.007	0.014	0.009	0.015
	(0.010)	(0.011)	(0.010)	(0.011)	(0.011)	(0.011)
N	314, 557	314, 557	314, 557	691, 634	691, 634	691, 634
R^2	0.671	0.661	0.666	0.655	0.644	0.644
Within-R^2	0.034	0.031	0.033	0.036	0.036	0.036
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 10 : The Effect of Experienced Returns by Number of Managers

	Expected Returns							
	(1)				(4)			
Nr. Managers	1	2	3	≥ 4	1	2	3	≥ 4
β_1	0.280***	0.164***	-0.001	0.014***	0.275***	0.160***	-0.001	0.008***
	(0.012)	(0.010)	(0.003)	(0.004)	(0.010)	(0.008)	(0.002)	(0.003)
β_2	0.149***	0.025***	-0.002	0.008**	0.151***	0.031***	-0.002	-0.000
	(0.008)	(0.007)	(0.002)	(0.004)	(0.006)	(0.004)	(0.002)	(0.002)
β_3	0.101***	0.028***	-0.002	0.005	0.101***	0.027***	-0.001	-0.001
	(0.006)	(0.004)	(0.002)	(0.003)	(0.004)	(0.003)	(0.001)	(0.002)
β_4	0.060***	0.017***	-0.004**	0.001	0.059***	0.017***	-0.002**	-0.006***
	(0.006)	(0.003)	(0.002)	(0.003)	(0.004)	(0.002)	(0.001)	(0.002)
β_5	0.028***	-0.001	-0.003**	0.001	0.029***	-0.001	-0.000	-0.003**
	(0.005)	(0.002)	(0.001)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)
β_6	0.022***	0.003	-0.001	-0.002	0.021***	0.002	-0.001	-0.001
	(0.004)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.001)	(0.001)
β_7	0.020***	-0.002	-0.002*	-0.003*	0.024***	0.000	0.000	-0.001
	(0.004)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
β_8	0.041***	0.008***	-0.002	-0.002	0.044***	0.008***	-0.000	0.000
	(0.004)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
β_9	0.077***	0.004*	0.000	-0.002*	0.083***	0.005***	-0.000	-0.001
	(0.006)	(0.002)	(0.001)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)
β_{10}	0.203***	0.017***	-0.002***	-0.002	0.204***	0.016***	-0.001	0.001
	(0.005)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
N	442, 353	579, 965	558, 722	428, 591	1, 073, 779	1, 454, 292	1, 524, 108	1, 158, 163
R^2	0.824	0.912	0.993	0.991	0.750	0.867	0.988	0.982
Within-R^2	0.039	0.010	0.000	0.001	0.039	0.012	0.000	0.001
$w_{i,j,t} = 0$	No	No	No	No	Yes	Yes	Yes	Yes
FE	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time	Mgr×Time Stock×Time
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^1$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 11 : Learning Parameters

	Expected Returns					
	(1)	(2)	(3)	(4)	(5)	(6)
β	0.203***	0.190***	0.198***	0.246***	0.241***	0.251***
	0.007	0.007	0.007	0.006	0.006	0.006
λ_1	-2.225***	-2.223***	-2.157***	-1.800***	-1.929***	-1.854***
	0.128	0.119	0.114	0.066	0.068	0.065
λ_2	-2.362***	-2.313***	-2.265***	-1.881***	-2.012***	-1.951***
	0.137	0.126	0.121	0.073	0.074	0.072
$w_{i,j,t} = 0$	No	No	No	Yes	Yes	Yes
FE	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time	Mgr×Time
	Stock	Stock	Stock	Stock	Stock	Stock
Covariance	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$	$\hat{\Sigma}_t^1$	$\hat{\Sigma}_t^2$	$\hat{\Sigma}_t^3$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 12 : Risk Aversion and Bias Including Zero Weights - Summary Statistics

	$\hat{\alpha}_i$	$\hat{\gamma}_i$
mean	0.006	1.501
standard deviation	0.056	5.266
median	0.009	1.441
min	-0.431	-43.532
max	0.398	42.658
skewness	-1.111	0.954
kurtosis	13.375	13.727

D.2 Plots

Figure 1 : Weighting Functions - Various Examples

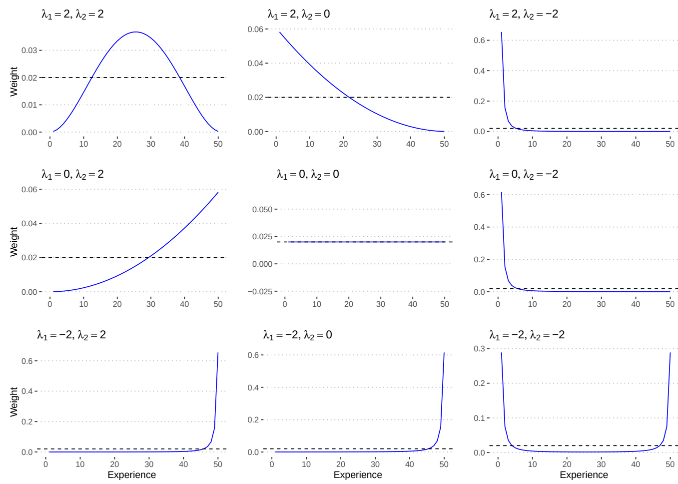
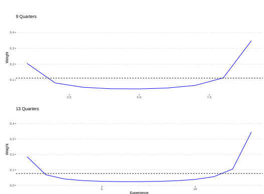
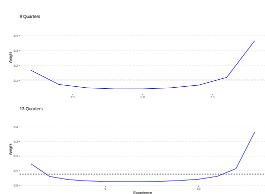


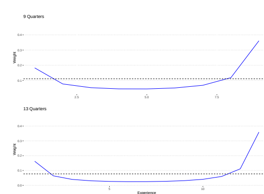
Figure 2 : Estimated Weighting Functions - Manager-Time, Stock-Time Fixed Effects



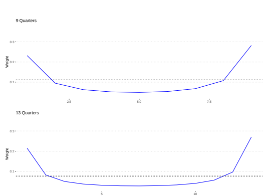
(a) $\hat{\Sigma}_t^1 \boldsymbol{w}^*_{i,t}, w_{i,j,t} > 0$



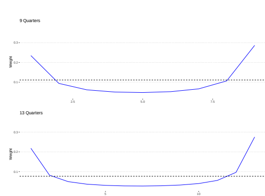
(b) $\hat{\Sigma}_t^2 \boldsymbol{w}^*_{i,t}, w_{i,j,t} > 0$



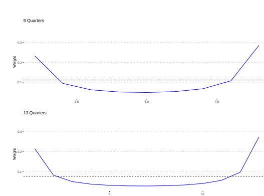
(c) $\hat{\Sigma}_t^3 \boldsymbol{w}^*_{i,t}, w_{i,j,t} > 0$



(d) $\hat{\Sigma}_t^1 \boldsymbol{w}^*_{i,t}, w_{i,j,t} \geq 0$

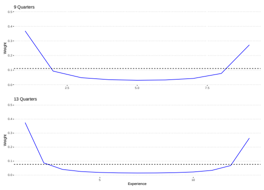


(e) $\hat{\Sigma}_t^2 \boldsymbol{w}^*_{i,t}, w_{i,j,t} \geq 0$

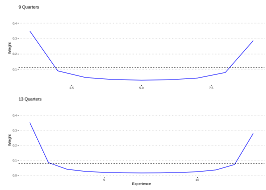


(f) $\hat{\Sigma}_t^3 \boldsymbol{w}^*_{i,t}, w_{i,j,t} \geq 0$

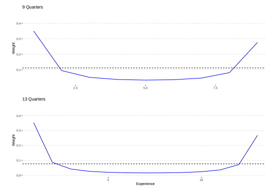
Figure 3 : Estimated Weighting Functions - Manager-Time, Stock Fixed Effects



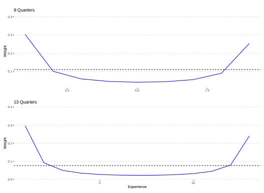
(a) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} > 0$



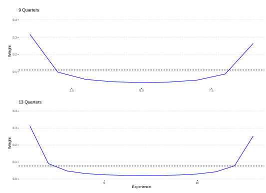
(b) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} > 0$



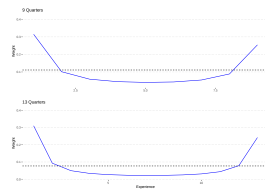
(c) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} > 0$



(d) $\hat{\Sigma}_t^1 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} \geq 0$

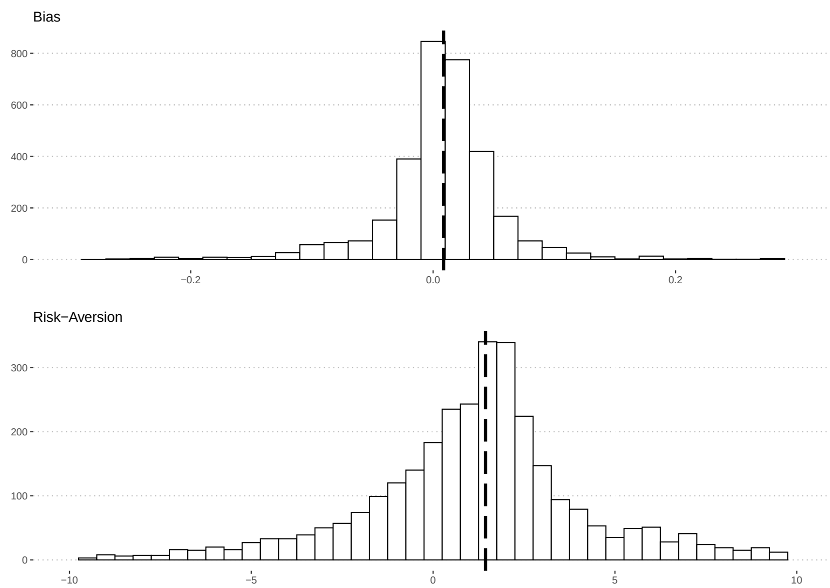


(e) $\hat{\Sigma}_t^2 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} \geq 0$



(f) $\hat{\Sigma}_t^3 \boldsymbol{w}_{i,t}^*$, $w_{i,j,t} \geq 0$

Figure 4 : Risk Aversion and Bias Including Zero Weights



References

Anderson, Evan W, Lars Peter Hansen, and Thomas J Sargent, "A quartet of semigroups for model specification, robustness, prices of risk, and model detection," *Journal of the European Economic Association*, 2003, *1* (1), 68–123.

Barber, Brad M, Xing Huang, and Terrance Odean, "Which factors matter to investors? Evidence from mutual fund flows," *The Review of Financial Studies*, 2016, *29* (10), 2600–2642.

Benos, Evangelos, Marek Jochec, and Victor Nyekel, "Can mutual funds time risk factors?," *The Quarterly Review of Economics and Finance*, 2010, *50* (4), 509 – 514.

Berk, Jonathan B and Jules H Van Binsbergen, "Assessing asset pricing models using revealed preference," *Journal of Financial Economics*, 2016, *119* (1), 1–23.

Cvitanic, Jaksa and Ioannis Karatzas, "Convex duality in constrained portfolio optimization," *The Annals of Applied Probability*, 1992, *2* (4), 767–818.

Davidson, Russell and James G MacKinnon, "Artificial regressions," *A companion to theoretical econometrics*, 2001, pp. 16–37.

Evans, Richard B, "Mutual fund incubation," *The Journal of Finance*, 2010, *65* (4), 1581–1611.

Fama, Eugene F and Kenneth R French, "A five-factor asset pricing model," *Journal of financial economics*, 2015, *116* (1), 1–22.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, "Unobserved actions of mutual funds," *The Review of Financial Studies*, 2006, *21* (6), 2379–2416.

Koijen, Ralph S.J. and Motohiro Yogo, "A Demand System Approach to Asset Pricing," *Journal of Political Economy*, 2019, *127* (4), 1475–1515.

Ledoit, Olivier and Michael Wolf, "Honey, I Shrunk the Sample Covariance Matrix," *The Journal of Portfolio Management*, Summer 2004, *30*, 110–119.

Maenhout, Pascal J, "Robust portfolio rules and asset pricing," *Review of financial studies*, 2004, *17* (4), 951–983.

Tepla, Lucie, "Optimal portfolio policies with borrowing and shortsale constraints," *Journal of Economic Dynamics and Control*, 2000, *24* (11-12), 1623–1639.

Touloumis, Anestis, "Nonparametric Stein-type shrinkage covariance matrix estimators in high-dimensional settings," *Computational Statistics & Data Analysis*, 2015, *83* (C), 251–261.

van Binsbergen, Jules H., Michael W. Brandt, and Ralph S.J. Koijen, "Optimal Decentralized Investment Management," *Journal of Finance*, August 2008, *63* (4), 1849–1895.

Xu, Gan-Lin and Steven E. Shreve, "A Duality Method for Optimal Consumption and Investment Under Short- Selling Prohibition. I. General Market Coefficients," *Ann. Appl. Probab.*, 02 1992, *2* (1), 87–112.