Rational Attention Allocation Over the Business Cycle

Marcin Kacperczyk* **Stijn Van Nieuwerburgh†** **Laura Veldkamp‡**

August 14, 2012§

*Department of Finance Stern School of Business and NBER, New York University, 44 W. 4th Street, New York, NY 10012; mkacperc@stern.nyu.edu; http://www.stern.nyu.edu/∼mkacperc.

†Department of Finance Stern School of Business, NBER, and CEPR, New York University, 44 W. 4th Street, New York, NY 10012; svnieuwe@stern.nyu.edu; http://www.stern.nyu.edu/∼svnieuwe.

‡Department of Economics Stern School of Business, NBER, and CEPR, New York University, 44 W. 4th Street, New York, NY 10012; lveldkam@stern.nyu.edu; http://www.stern.nyu.edu/∼lveldkam.

§We thank John Campbell, Joseph Chen, Xavier Gabaix, Vincent Glode, Ralph Koijen, Jeremy Stein, Matthijs van Dijk, Robert Whitelaw, and seminar participants at NYU Stern (economics and finance), Harvard Business School, Chicago Booth, MIT Sloan, Yale SOM, Stanford University (economics and finance), University of California at Berkeley (economics and finance), UCLA economics, Duke economics, University of Toulouse, University of Vienna, Australian National University, University of Melbourne, University of New South Wales, University of Sydney, University of Technology Sydney, Erasmus University, University of Mannheim, University of Alberta, Concordia, Lugano, the Amsterdam Asset Pricing Retreat, the Society for Economic Dynamics meetings in Istanbul, CEPR Financial Markets conference in Gerzensee, UBC Summer Finance conference, and Econometric Society meetings in Atlanta for useful comments and suggestions. Thank you to Isaac Baley for outstanding research assistance. Finally, we thank the Q-group for their generous financial support.

Abstract

The literature assessing whether mutual fund managers have skill typically regards market timing or stock picking skills as immutable attributes of a manager or fund. Yet, measures of these skills appear to vary over the business cycle. This paper offers a rational explanation, arguing that timing and picking are tasks. A skilled manager can choose how much of each task to attend to. Using tools from the rational inattention literature, we show that in booms, a manger should pick stocks and in recessions, he should pay more attention to his market timing. The model predicts equilibrium outcomes in a world where a fraction of managers have skill and invest alongside unskilled investors. The predictions about funds' covariance with payoff shocks, cross-fund dispersion, and their excess returns are all supported by the data. In turn, these findings offer new evidence to support two broader ideas: that some investment managers have skill and that attention is allocated rationally.

> "What information consumes is rather obvious: It consumes the attention of its recipients. Hence a wealth of information creates a poverty of attention, and a need to allocate that attention efficiently among the overabundance of information sources that might consume it." Simon (1971)

The literature that evaluates skills of mutual fund managers typically regards skill as an immutable attribute of the manager or the fund.[1] Yet, many skill measures vary over the business cycle, such as returns, alphas (Glode 2011), and return-based measures of stock picking and market timing (Kacperczyk, Van Nieuwerburgh, and Veldkamp 2012). Because time-varying ability seems far-fetched, these results call into question the existence of skill itself. This paper examines a rational explanation for time-varying skill, where skill is a general cognitive ability that can be applied to different tasks, such as picking stocks or market timing, at different points in time. Each period, skilled managers choose how much of their time or cognitive ability (call that "attention") to allocate to each task. When the economic environment changes, the relative payoffs of paying attention to market timing and stock selection shift. The resulting fluctuations in attention allocation look like time-varying skill. While this story might sound plausible, it leaves open three questions. First, why would a manager want his attention allocation to depend on the state of the business cycle? Second, do the managers' attention choices exhibit the same pattern as the time-varying skill observed in the data? If managers want to allocate more attention to stock picking in booms, do we see better stock picking in booms? Third, if there are many skilled and unskilled managers in an asset market, would the time-series and cross-sectional portfolio and return patterns resemble those in the data? This paper builds a simple theory of attention allocation and portfolio choice and subjects it to these three tests.

The model uses tools from the rational inattention literature (Sims 2003) to analyze the trade-off between allocating attention to each task. In recessions, the abundance of aggregate risk and its high price both work in the same direction to make market timing more valuable. The model generates indirect predictions for the dispersion and returns of fund portfolios that distinguish this explanation from other potential explanations for time-varying skill. It reveals that when skilled managers devote more time to market timing, portfolio dispersion is higher, both among skilled managers and between skilled and unskilled

[1] For theoretical models, see e.g., Mamaysky and Spiegel (2002), Berk and Green (2004), Kaniel and Kondor (2012), Cuoco and Kaniel (2011), Vayanos and Woolley (2010), Chien, Cole, and Lustig (2011), Chapman, Evans, and Xu (2010), and Pástor and Stambaugh (2012). A number of recent papers in the empirical mutual fund literature also find that some managers have skill, e.g., Kacperczyk, Sialm, and Zheng (2005, 2008), Kacperczyk and Seru (2007), Cremers and Petajisto (2009), Huang, Sialm, and Zhang (2011), Koijen (2012), Baker, Litov, Wachter, and Wurgler (2010).

managers. It predicts that recessions are times when skilled managers outperform others by a larger margin. Finally, it predicts that volatility and recessions should each have an independent effect on attention, dispersion, and performance. All of these predictions are borne out in the mutual fund data.

These findings offer useful evidence to support a variety of theories that use rational attention allocation to explain phenomena in many economic environments. Recent work has shown that introducing attention constraints into decision problems can help explain observed consumption, price-setting, and investment patterns as well as the timing of government announcements and the propensity for governments to be unprepared for rare events.[2] An obstacle to the progress of this line of work is that information is not directly observable, precluding a direct test of whether decision makers actually allocate their attention in a value-maximizing way.[3]

To surmount the problem that attention is unobservable, our model uses an observable variable – the state of the business cycle – to predict attention allocation. Attention, in turn, predicts aggregate investment patterns. Because the theory begins and ends with observable variables, it becomes testable. To carry out these tests, we use data on actively managed equity mutual funds. A wealth of detailed data on portfolio holdings and returns makes this industry an ideal setting in which to test whether decision makers allocate attention optimally.

To explore whether a rational attention allocation can explain the behavior of mutual fund managers, we build a general equilibrium model in which a fraction of investment managers have skill. These skilled managers can observe a fixed number of signals about asset payoffs and choose what fraction of those signals will contain aggregate versus stock-specific information. We think of aggregate signals as macroeconomic data that affect future cash flows of all firms, and of stock-specific signals as firm-level data that forecast the part of firms' future cash flows that is independent of the aggregate shocks. Based on their signals, skilled managers form portfolios, choosing larger portfolio weights for assets that are more likely to have high returns.

[2]See, for example, Sims (2003) on consumption, Maćkowiak and Wiederholt (2010, 2009), and Matejka (2011) on price setting, and Van Nieuwerburgh and Veldkamp (2009, 2010) and Kondor (2012) on financial investment. Reis (2011) considers the optimal timing of government announcements and Maćkowiak and Wiederholt (2012) use rational inattention constraint to model the allocation of cognitive energy to planning for rare events. A related attention constraint called inattentiveness is explored in Reis (2006). Veldkamp (2011) provides a survey of this literature.

[3]While papers such as Klenow and Willis (2007), Mondria, Wu, and Zhang (2010) and Maćkowiak, Moench, and Wiederholt (2009) have also tested predictions of rational inattention models, none has looked for evidence that attention is reallocated, arguably a more stringent test of the theory.

The model produces four main predictions. The first prediction is that attention should be reallocated over the business cycle. In the data, recessions are times when unexpected returns are low, aggregate volatility rises, and the price of risk surges. When we embed these three forces in our model, the first has little effect on attention allocation, but the second and third forces both draw attention to aggregate shocks in recessions. The increased volatility of aggregate shocks makes it optimal to allocate more attention to them, because it is more valuable to pay attention to more uncertain outcomes. The elevated price of risk amplifies this reallocation: Since aggregate shocks affect a large fraction of the portfolio's value, paying attention to aggregate shocks resolves more portfolio risk than learning about stock-specific risks. When the price of risk is high, such risk-minimizing attention choices become more valuable. While the idea that it is more valuable to shift attention to more volatile shocks may not be all that surprising, whether changes in the price of risk would amplify or counteract this effect is not obvious.

The second and third predictions do not come from the reallocation of attention. Rather, they help to distinguish this theory from non-informational alternatives and support the idea that at least some portfolio managers are engaging in value-maximizing behavior. The second prediction is counter-cyclical dispersion in portfolio holdings and profits. In recessions, when aggregate shocks to asset payoffs are larger in magnitude, asset payoffs exhibit more comovement. Thus, any passive portfolio strategies that put exogenously fixed weights on assets would have returns that also comove more in recessions, which would imply less dispersion. In contrast, when investment managers learn about asset payoffs and manage their portfolios according to what they learn, fund returns comove *less* and dispersion increases in recessions. The reason is that when aggregate shocks become more volatile, managers who learn about aggregate shocks put less weight on their common prior beliefs, which have less predictive power, and more weight on their heterogeneous signals. This generates more heterogeneous beliefs in recessions and therefore more heterogeneous investment strategies and fund returns.

Third, the model predicts time variation in fund performance. Since the average fund can only outperform the market if there are other, non-fund investors who underperform, the model also includes unskilled non-fund investors. Because asset payoffs are more uncertain, recessions are times when information is more valuable. Therefore, the informational advantage of the skilled over the unskilled increases and generates higher returns for informed managers. The average fund's outperformance rises.

The fourth prediction is that all three of the above effects of recessions come in part from high aggregate volatility, and in part from the high price of risk. Therefore, periods of high aggregate volatility should be periods in which attention is allocated to aggregate shocks. Furthermore, these high-volatility periods should also be times when portfolio dispersion is high and skilled funds outperform. Then, after controlling for volatility and as long as prices are sufficiently noisy, there should also be an additional positive effect of recessions on all three measures. This additional effect comes from the fact that recessions are also times when the price of risk is high. In other words, both volatility and the price of risk have separate effects on skill, dispersion, and performance.

We test the model's four main predictions on the universe of actively managed U.S. mutual funds. To test the first prediction, a key insight is that managers can only choose portfolios that covary with shocks they pay attention to. Thus, to detect cyclical changes in attention, we should look for changes in covariances. We estimate the covariance of each fund's portfolio holdings with the aggregate payoff shock, proxied by innovations in industrial production growth. This covariance measures a manager's ability to time the market by increasing (decreasing) her portfolio positions in anticipation of good (bad) macroeconomic news. This timing covariance rises in recessions. We also calculate the covariance of a fund's portfolio holdings with asset-specific shocks, proxied by innovations in firms' earnings. This covariance measures managers' ability to pick stocks that subsequently experience unexpectedly high earnings. Consistent with the theory, this stock-picking covariance increases in expansions.

Second, we test for cyclical changes in portfolio dispersion. We find that, in recessions, funds hold portfolios that differ more from one another. As a result, their cross-sectional return dispersion increases, consistent with the theory. In the model, much of this dispersion comes from taking different bets on market outcomes, which should show up as dispersion in CAPM betas. We find evidence in the data for higher beta dispersion in recessions.

Third, we document fund outperformance in recessions, extending earlier results in the literature. Risk-adjusted excess fund returns (alphas) are around 1.6 to 4.6% per year higher in recessions, depending on the specification. Gross alphas (before fees) are not statistically different from zero in expansions, but they are significantly positive in recessions.[4] These cyclical differences are statistically and economically significant.

Fourth, we document an effect of recessions on covariance, dispersion, and performance,

[4]Net alphas (after fees) are negative in expansions (-0.6%) and positive (1.0%) in recessions for our most conservative metric. Gross alphas are higher by about 1% point per year. Since funds do not set fees in our model, we have no predictions about after-fee alphas. For a theory about why we should expect net alphas to be zero, see Berk and Green (2004). For recent empirical work, see Berk and van Binsbergen (2012).

above and beyond that which comes from volatility alone. When we use both a recession indicator and aggregate volatility as explanatory variables, we find that both contribute about equally to our three main results. Showing that these results are truly business-cycle phenomena – as opposed to merely high volatility phenomena – is interesting because it connects these results with the existing macroeconomics literature on rational inattention, e.g., Maćkowiak and Wiederholt (2010, 2009).

The rest of the paper is organized as follows. Section 1 lays out our model. After describing the setup, we characterize the optimal information and investment choices of skilled and unskilled investors. We show how equilibrium asset prices are formed. We derive theoretical predictions for funds' attention allocation, portfolio dispersion, and performance. Section 2 explains how we bring the model to the data. Section 3 tests the model's predictions using the context of actively managed mutual funds. Section 4 concludes.

1 Model

We develop a model whose purpose is to understand how the optimal attention allocation of investment managers depends on the business cycle, and how attention affects asset holdings and asset prices. Most of the complexity of the model comes from the fact that it is an equilibrium model. But in order to study the effects of attention on asset holdings, asset prices and fund performance, having an equilibrium model is a necessity. In particular, an equilibrium model ensures that for every investor that outperforms, there is someone who under-performs.

1.1 Setup

The model has three periods. At time 1, skilled investment managers choose how to allocate their attention across different assets. At time 2, all investors choose their portfolios of risky and riskless assets. At time 3, asset payoffs and utility are realized.

Assets The model features 1 riskless and n risky assets. The price of the riskless asset is normalized to 1 and it pays off r at time 3. Risky assets $i \in \{1, ..., n-1\}$ have random payoffs f_i with respective loadings $b_i, ..., b_{n-1}$ on an aggregate shock z_a, and face stock-specific shocks $z_1, ..., z_{n-1}$. The n-th asset, is a composite asset whose payoff has no stock-specific shock and a loading of one on the aggregate shock. We use this composite asset as a stand-in for all other assets. Formally,

$$
\begin{aligned}
f_i &= \mu_i + b_i z_a + z_i, \qquad i \in \{1, \dots, n-1\} &(1)\\
f_n &= \mu_n + z_a &(2)
\end{aligned}
$$

where the risk factors $z_a \sim N(0, \sigma_a)$ and $z_i \sim N(0, \sigma_i)$, are mutually independent for $i \in \{1, \dots, n-1\}$. We define the $n \times 1$ vector $f = [f_1, f_2, ..f_n]'$.

Risk factors The vector of risk factor shocks, $z = [z_1, z_2, ..., z_{n-1}, z_a]'$, is normally distributed as: $z \sim \mathcal{N}(0, \Sigma)$ where Σ is a diagonal matrix. Stacking the equations above, we can write $f = \mu + \Gamma z$, where Γ is a $n \times n$ invertible matrix of loadings that map risk factors, z, into the mean-zero payoffs $(f - \mu)$. We define the payoff of the risk factors as $\tilde{f} \equiv \Gamma^{-1} f = \Gamma^{-1}\mu + z$. Thus, payoffs of risk factors are linear combinations of payoffs of the underlying assets. In other words, they are a payoff to a particular portfolio of assets. Working with risk factor payoffs and prices (denoted with tildes) allows us to solve the model in a tractable way.[5]

Each risk factor has a stochastic supply given by $\bar{x}_i + x_i$, where noise x_i is normally-distributed, with mean zero, variance σ_x, and no correlation with other noises: $x \sim \mathcal{N}(0, \sigma_x I)$. The vector of noisy asset supplies is $\Gamma(\bar{x} + x)$. As in the standard noisy rational expectations equilibrium model, the supply is random to prevent the price from fully revealing the information of informed investors.

Portfolio Choice Problem There is a continuum of atomless investors. Investors are each endowed with initial wealth, W_0. They have mean-variance preferences over time-3 wealth, with a risk-aversion coefficient, ρ. Let E_j and V_j denote j's expectations and variances conditioned on all information known at time 2, which includes prices and signals. Thus, investor j chooses how many shares of each asset to hold q_j to maximize time-2 expected utility, U_{2j}:

$$U_{2j} = \rho E_j[W_j] - \frac{\rho^2}{2} V_j[W_j] \tag{3}$$

subject to the budget constraint: $W_j = rW_0 + q_j'(f - pr)$, where q_j, p are $n \times 1$ vectors of prices and quantities of each asset held by investor j. Since there are no wealth effects with

[5] The existence of the composite asset ensures that the assets span the shocks, which allows Γ to be invertible. An invertible mapping Γ allows us to solve for prices and quantities of risk factors z and then map them back into asset prices and quantities. This solution approach is what keeps the model analytically tractable.

mean-variance utility, we normalize W_0 to zero for the theoretical results. Furthermore, we can rewrite the budget constraint in terms of risk factor prices and quantities by defining $\tilde{p}_j \equiv \Gamma^{-1} p_j$, $\tilde{q}_j \equiv \Gamma' q_j$ and substituting $f = \Gamma \tilde{f}$ to get

$$W_j = \tilde{q}_j'(\tilde{f} - \tilde{p}r). \tag{4}$$

Prices Equilibrium prices are determined by market clearing:

$$\int \tilde{q}_j dj = \bar{x} + x, \tag{5}$$

where the left-hand side of the equation is the vector of aggregate demand and the right-hand side is the vector of aggregate supply of each risk factor.

Information, updating, and attention allocation At time 1, a skilled investment manager j chooses the precisions of signals that she will receive at time 2. Allocating attention to a risk factor means that a manager gets a more precise signal about that risky outcome. Mathematically, manager j's vector of signals is $\eta_j = z + \varepsilon_j$, where the vector of signal noise is distributed as $\varepsilon_j \sim \mathcal{N}(0, \Sigma_{\eta j})$. The variance matrix $\Sigma_{\eta j}$ is diagonal with ith diagonal element K_{ij}^{-1}. Thus, K_{ij} is the precision of investor j's signal about risk i. Private signal noise is independent across risks i and managers j. Managers combine signal realizations with priors and information extracted from asset prices to update their beliefs, using Bayes' law.

Signal precision choices $\{K_{ij}\}$ maximize time-1 expected utility, U_{1j}, of the fund's terminal wealth W_j:

$$U_{1j} = E_1 \left[\rho E_2[W_j] - \frac{\rho^2}{2} V_2[W_j] \right], \tag{6}$$

subject to two constraints.[6]

The first constraint is the *information capacity constraint*. It states that the sum of the signal precisions must not exceed the information capacity:

$$K_{aj} + \sum_{i=1}^{n-1} K_{ij} \leq K. \tag{7}$$

[6]See Veldkamp (2011) for a discussion of the use of expected mean-variance utility in information choice problems. The supplementary appendix proves versions of the main propositions for the expected exponential utility model.

In Bayesian updating with normal variables, observing one signal with precision K_i or two signals, each with precision $K_i/2$, is equivalent. Therefore, one interpretation of the capacity constraint is that it allows the manager to observe N signal draws, each with precision K_i/N, for large N. The investment manager then chooses how many of those N signals will be about each shock.[7] Note that our model holds each manager's total attention fixed and studies its allocation in recessions and expansions. In Section 1.7, we consider a manager who chooses how much capacity for attention to acquire.

The second constraint is the *no-forgetting constraint*, which ensures that the chosen precisions are non-negative:

$$K_{ij} \geq 0 \quad i \in \{1, ..., n-1, a\} \tag{8}$$

It prevents the manager from erasing any prior information, to make room to gather new information about another shock.

Skilled and Unskilled Investors The only ex-ante difference between investors is that a fraction χ of them have *skill*, meaning that they can choose to observe a set of informative *private* signals about the risk factor shocks z_i. Unskilled investors are ones with zero signal precision: $\Sigma_{\eta j}^{-1} = 0$, or equivalently, $K_{ij} = 0,\ \forall i$. Both unskilled and skilled investors observe the information in prices, which are public signals, costlessly.

When we bring the model to the data, we will call all skilled investors mutual funds. Furthermore, we will distinguish between two types of unskilled investors: unskilled mutual funds and non-fund investors.[8] In the model, these two types are identical. The reason for modeling non-fund investors is that without them, we cannot talk about average fund performance. The sum of all funds' holdings would have to equal the market (market clearing) and therefore, the average fund return would have to equal the market return. There could be no excess return in expansions or recessions.

[7]The results are not sensitive to the additive nature of the information capacity constraint. They also hold, for example, for a product constraint on precisions. The entropy constraints often used in information theory take this multiplicative form. Results available upon request. See also Van Nieuwerburgh and Veldkamp (2010).

[8]For our results to hold, it is sufficient to assume that the fraction of non-fund investors that are unskilled is higher than that for the mutual funds.

Modeling recessions Since this is a static model, the investment world is either in the recession (R) or in the expansion state (E).[9] The asset pricing literature identifies three principal effects of recessions: (1) returns are unexpectedly low, (2) returns are more volatile, and (3) the price of risk is high. Section 3 discusses the empirical evidence supporting the latter two effects. To capture the return volatility effect (2) in the model, we assume that the prior variance of the aggregate shock in recessions (R) is higher than the one in expansions (E): $\sigma_a(R) > \sigma_a(E)$. To capture the varying price of risk (3), we vary the parameter that governs the price of risk, which is risk aversion. We assume $\rho(R) > \rho(E)$. We continue to use σ_a and ρ to denote aggregate shock variance and risk aversion in the current business cycle state.

The first effect of recessions, unexpectedly low returns, cannot affect attention allocation because attention must be allocated before returns are observed. Yet, unexpected returns could affect managers' return covariances. The difficulty in analyzing this effect is that since agents in our model always know the current state of the business cycle, they cannot be systematically surprised by low asset payoffs in recessions. When low payoffs are expected, asset prices fall right away, leaving subsequent returns unaffected. Therefore, exploring (1) requires a slightly modified model that relaxes rational expectations. The Supplementary Appendix explores this model numerically and shows that the unexpectedly low returns have little effect on the results. The main body of the paper explores the volatility and price of risk effects.

1.2 Model Solution

We begin by working through the mechanics of Bayesian updating. There are three types of information that are aggregated in time-2 posteriors beliefs: prior beliefs, price information and (private) signals. We conjecture and later verify that a transformation of prices $\tilde{p}$ generates an unbiased signal about the risk factor payoffs, $\eta_p = z + \epsilon_p$, where $\epsilon_p \sim N(0, \Sigma_p)$, for some diagonal variance matrix Σ_p. Then, by Bayes' law, the posterior beliefs about z are normally-distributed with mean $\hat{z}_j = \hat{\Sigma}_j(\Sigma_{\eta j}^{-1}\eta_j + \Sigma_p^{-1}\eta_p)$ and posterior precision $\hat{\Sigma}_j^{-1} = \Sigma^{-1} + \Sigma_p^{-1} + \Sigma_{\eta j}^{-1}$. Using the definition $\tilde{f} = \Gamma^{-1}\mu + z$, we find that posterior beliefs about

[9] We do not consider transitions between recessions and expansions, although such an extension would be easy in our setting because assets are short lived and their payoffs are realized and known to all investors at the end of each period. Thus, a dynamic model would amount to a succession of static models that are either in the expansion or in the recession state.

risk factor payoffs are $\tilde{f} \sim N(E_j[\tilde{f}], \hat{\Sigma}_j^{-1})$ where

$$E_j[\tilde{f}] = \Gamma^{-1}\mu + \hat{\Sigma}_j(\Sigma_{\eta j}^{-1}\eta_j + \Sigma_p^{-1}\eta_p). \tag{9}$$

Next, we solve the model in four steps.

Step 1: Solve for the optimal portfolios, given information sets. Substituting the budget constraint (4) into the objective function (3) and taking the first-order condition with respect to $\tilde{q}_j$ reveals that optimal holdings are increasing in the investor's risk tolerance, precision of beliefs, and expected return:

$$\tilde{q}_j = \frac{1}{\rho}\hat{\Sigma}_j^{-1}(E_j[\tilde{f}] - \tilde{p}r). \tag{10}$$

Step 2: Clear the asset market. Substitute each agent j's optimal portfolio (10) into the market-clearing condition (5). Collecting terms and simplifying reveals that equilibrium asset prices are linear in payoff risk shocks and in supply shocks:

Lemma 1. $\tilde{p} = \frac{1}{r}(A + Bz + Cx)$

A detailed derivation of coefficients A, B and C, expected utility, and the proofs of this and all further propositions are in the appendix.

In this model, agents learn from prices because prices are informative about the payoff shocks z. Next, we deduce what information is implied by Lemma 1. Price information is the signal about z contained in prices. The transformation of the price vector $\tilde{p}$ that yields an unbiased signal about z is $\eta_p \equiv B^{-1}(\tilde{p}r - A)$. Note that applying the formula for η_p to Lemma 1 reveals that $\eta_p = z + \varepsilon_p$, where the signal noise in prices is $\varepsilon_p = B^{-1}Cx$. Since we assumed $x \sim N(0, \sigma_x I)$, the price noise is distributed $\varepsilon_p \sim N(0, \Sigma_p)$, where $\Sigma_p \equiv \sigma_x B^{-1}CC'B^{-1'}$. Substituting in the coefficients B and C from the proof of Lemma 1 shows that public signal precision Σ_p^{-1} is a diagonal matrix with ith diagonal element $\sigma_{pi}^{-1} = \frac{\bar{K}_i^2}{\rho^2\sigma_x}$, where $\bar{K}_i \equiv \int K_{ij}dj$ is the average capacity allocated to shock i.

Step 3: Compute ex-ante expected utility. Substitute optimal risky asset holdings from equation (10) into the first-period objective function (6) to get: $U_{1j} = \frac{1}{2}E_1\left[(E_j[\tilde{f}] - \tilde{p}r)\hat{\Sigma}_j^{-1}(E_j[\tilde{f}] - \tilde{p}r)\right]$. Note that the expected excess returns $(E_j[\tilde{f}] - \tilde{p}r)$ depends on signals and prices, both of which are unknown at time 1. Because asset prices are linear functions of normally distributed shocks, $E_j[\tilde{f}] - \tilde{p}r$, is normally distributed as well.

Thus, $(E_j[\tilde{f}] - \tilde{p}r)\hat{\Sigma}_j^{-1}(E_j[\tilde{f}] - \tilde{p}r)$ is a non-central χ^2-distributed variable. Computing its mean yields:

$$U_{1j} = \frac{1}{2} trace(\hat{\Sigma}_j^{-1} V_1[E_j[\tilde{f}] - \tilde{p}r]) + \frac{1}{2} E_1[E_j[\tilde{f}] - \tilde{p}r]' \hat{\Sigma}_j^{-1} E_1[E_j[\tilde{f}] - \tilde{p}r]. \tag{11}$$

Step 4: Solve for information choices. Note that in expected utility (11), the choice variables K_{ij} enter only through the posterior variance $\hat{\Sigma}_j$ and through $V_1[E_j[\tilde{f}] - \tilde{p}r] = V_1[\tilde{f} - \tilde{p}r] - \hat{\Sigma}_j$. Since there is a continuum of investors, and since $V_1[\tilde{f} - \tilde{p}r]$ and $E_1[E_j[\tilde{f}] - \tilde{p}r]$ depend only on parameters and on aggregate information choices, each investor takes them as given.

Since $\hat{\Sigma}_j^{-1}$ and $V_1[E_j[\tilde{f}] - \tilde{p}r]$ are both diagonal matrices and $E_1[E_j[\tilde{f}] - \tilde{p}r]$ is a vector, both terms of (11) are weighted sums of the diagonal elements of $\hat{\Sigma}_j^{-1}$. Thus, (11) can be rewritten as $U_{1j} = \sum_i \lambda_i \hat{\Sigma}_j^{-1}(i,i) - N/2$, for positive coefficients λ_i. Since $\hat{\Sigma}_j^{-1}(i,i) = \Sigma^{-1}(i,i) + \Sigma_p^{-1}(i,i) + K_{ij}$, we can write the information choice problem as:

$$\max_{K_{1j},\ldots,K_{(n-1)j},K_{aj}} \sum_i \lambda_i K_{ij} + \text{constant} \tag{12}$$

$$s.t. \quad K_{aj} + \sum_{i=1}^{n-1} K_{ij} \leq K \tag{13}$$

$$\text{where} \quad \lambda_i = \bar{\sigma}_i[1 + (\rho^2 \sigma_x + \bar{K}_i)\bar{\sigma}_i] + \rho^2 \bar{x}_i^2 \bar{\sigma}_i^2, \tag{14}$$

$\bar{K}_i = \int K_{ij} dj$ is the average signal precision across investors and $\bar{\sigma}_i^{-1} = \int \hat{\Sigma}_j^{-1}(i,i) dj$ is the average precision of posterior beliefs about risk i. Its inverse, average variance $\bar{\sigma}_i$ is decreasing in $\bar{K}_i$. Equation (14) is derived in the appendix.

To maximize a weighted sum subject to an unweighted sum, the manager optimally assigns all capacity to the risk(s) with the highest weight. If there is a unique $i^* = \text{argmax}_i \lambda_i$, then the solution is to set $K_{i^*j} = K$ and $K_{lj} = 0$, $\forall l \neq i^*$.

In many cases, there will be multiple risks with identical λ weights. That is because λ_i is decreasing in $\bar{K}_i$, the average investor's signal precision. This is the same strategic substitutability effect first noted by Grossman and Stiglitz (1980). The more others learn about a risk, the more informative prices are and the less valuable it is for others to learn about the same risk. When there exist risks i, l s.t. $\lambda_i = \lambda_l$, then investors are indifferent about which risk to learn about. For simplicity, we restrict attention to the unique symmetric equilibrium where all skilled investors choose the same information precisions. However, none of the propositions depend on this restriction.

The following sections explain the model's key predictions: attention allocation, dispersion in investors' portfolios, average performance, and the effect of recessions on these objects beyond that of aggregate volatility. For each prediction, we state and prove a hypothesis. The next section explains how we test the hypothesis in the data.

1.3 Prediction 1: Cyclical Attention Reallocation

First, we derive from the model the prediction that the optimal attention allocation in expansions differs from that in recessions. Specifically, there should be more attention paid to aggregate shocks in recessions and more attention paid to stock-specific shocks in expansions. Recessions involve changes in the volatility of aggregate shocks and changes in the price of risk. In order to see the effect of each aspect of a recession, we consider each separately, beginning with the rise in volatility.

Proposition 1. *For a given investor j, the optimal choice of attention allocation to risk i is weakly increasing in its variance σ_i: $\partial K_{ij}/\partial \sigma_i \geq 0$.*

Intuitively, investors prefer to learn about large shocks that are an important component of the overall asset supply, and volatile shocks that have high prior payoff variance. Aggregate shocks are larger in scale, but are less volatile than stock-specific shocks. Recessions are times when aggregate volatility increases, which makes aggregate shocks more valuable to learn about. The converse is true in expansions.

Note that this is a partial derivative result. It holds information choices fixed. In any interior equilibrium, attention will be reallocated until the marginal utility of learning about aggregate and stock-specific shocks is equalized.[10] But it is the initial increase in marginal utility which drives this reallocation.

Next, we consider the effect of an increase in the price of risk. An increase in the price of risk induces managers to allocate even more attention to the aggregate shock in recessions. The additional price of risk effect should show up as an effect of recessions, above and beyond what aggregate volatility alone can explain. The parameter that governs the price of risk in our model is risk aversion. The following result shows that an increase in the price of risk (risk aversion) in recessions is an independent force driving the reallocation of attention from stock-specific to aggregate shocks.

[10] Mathematically, $\lambda_i = \lambda_a, \forall i \in \{1, \cdots, n-1\}$. Full equalization across all risk factors only takes place provided that there is enough aggregate capacity (χK exceeds some threshold). If not, equalization takes place among a subset of risk factors.

Proposition 2. *For a given investor j, an increase in risk aversion ρ weakly increases the attention allocated to risk i if its supply $\bar{x}_i$ is sufficiently high: $\partial K_{ij}/\partial\rho \geq 0$ if $\bar{x}_i \geq x^*$.*

The intuition for this result rests on the fact that we defined the aggregate shock to be the shock in the greatest supply. Therefore, it affects a large fraction of the value of one's portfolio. Therefore, a marginal reduction in the uncertainty about an aggregate shock reduces total portfolio risk by more than the same-sized reduction in the uncertainty about a stock-specific shock. In other words, learning about the aggregate shock is the most efficient way to reduce portfolio risk. The more risk averse an agent is, the more attractive aggregate attention allocation becomes.

Investors' optimal attention allocation decisions are reflected in their portfolio holdings. In recessions, skilled investors predominantly allocate attention to the aggregate payoff shock, z_a. They use the information they observe to form a portfolio that covaries with z_a. In times when they learn that z_a will be high, they hold more risky assets whose returns are increasing in z_a. This positive covariance can be seen from equation (10) in which $\tilde{q}$ is increasing in $E_j[\tilde{f}]$ and from equation (9) in which $E_j[\tilde{f}]$ is increasing in η_j, which is further increasing in z_a. The positive covariances between the aggregate shock and funds' portfolio holdings in recessions, on the one hand, and between stock-specific shocks and the portfolio holdings in expansions, on the other hand, directly follow from optimal attention allocation decisions switching over the business cycle. As such, these covariances are the key moments that enable us to test the attention allocation predictions of the model. We define the empirical counterparts to these covariances in Section 2.

1.4 Prediction 2: Dispersion

Since many studies detect no skill, perhaps the most controversial implication of the attention reallocation result is that investment managers are processing information at all. Our second and third predictions speak directly to that implication. They do not identify changes in attention allocation, but they help to distinguish our theory from non-information-based alternatives.

In recessions, as aggregate shocks become more volatile, the firm-specific shocks to assets' payoffs account for less of the variation, and the comovement in stock payoffs rises. Since asset payoffs comove more, the payoffs to all passive investment strategies that put fixed weights on assets should also comove more. Dispersion across investor portfolios and portfolio returns should fall. But when investment managers are processing information, this prediction is reversed. To see why, consider a simple example where there is no learning from prices. A skilled agent is updating beliefs about a random variable $\tilde{f} \sim N(\mu, \Sigma)$, using a signal $\eta_j | \tilde{f} \sim N(\tilde{f}, \Sigma_\eta)$. Bayes' law says that the posterior mean is a weighted average of the prior mean μ and the signal, where each is weighted by their relative precision:

$$E[\tilde{f}|\eta_j] = \left(\Sigma^{-1} + \Sigma_\eta^{-1}\right)^{-1} \left(\Sigma^{-1}\mu + \Sigma_\eta^{-1}\eta_j\right) \tag{15}$$

If in recessions, aggregate shock variance σ_a rises, then the prior beliefs about asset payoffs become more uncertain: Σ rises and Σ^{-1} falls. This makes the weight on prior beliefs μ decrease and the weight on the signal η_j increase. The prior μ is common across agents, while the signal realization η_j is heterogeneous. When informed managers weigh their heterogeneous signals more, their resulting posterior beliefs become more different from each other and more different from the beliefs of uninformed managers or investors. More disagreement about asset payoffs results in more heterogeneous portfolios and portfolio returns.

Thus, the model's second prediction is that in recessions, the cross-sectional dispersion in funds' investment strategies and returns should rise. The following Proposition shows that funds' portfolio holdings, q, and portfolio excess returns, $q_j'(f - pr)$, display higher cross-sectional dispersion when risk is higher, in recessions.

Proposition 3. *For given precisions of an investor j, an increase in variance σ_i: a) increases the dispersion of fund portfolios $E[(q_j - \bar{q})'(q_j - \bar{q})]$ and b) increases the dispersion of portfolio excess returns $E[((q_j - \bar{q})'(f - pr))^2]$.*

This result holds generally, when any shock becomes more volatile, not only when aggregate risk rises. However, the effect is particularly large for the aggregate shock because it affects every asset and therefore is in abundant supply. This shows up in the proof (see appendix) as a high $\bar{x}_a$, which amplifies the effect of σ_a on portfolio and return dispersion.

Next, we consider the second effect of recessions: an increase in the price of risk. The following result shows that, when prices are sufficiently noisy, an increase in the price of risk increases the dispersion of portfolio returns.

Proposition 4. *If σ_x is sufficiently large, then an increase in risk aversion ρ increases the dispersion of portfolio excess returns $E[((q_j - \bar{q})'(f - pr))^2]$.*

When risk aversion rises, skilled investors make smaller bets on their information. These smaller deviations from the market portfolio affect prices less and make prices less informative. The reduced precision of price information implies that agents weight prices less and private signals more in their posterior beliefs. Just like priors, information in prices is common. Thus, weighting common signals less and heterogenous private signals more leads to higher dispersion in beliefs and therefore in portfolio returns as well.

But this effect has to offset a counter-acting force. Recall that the optimal portfolio for investor j takes the form $q = (1/\rho)\hat{\Sigma}_j^{-1}(f - pr)$. If ρ increases, investors scale down their risky asset positions and q falls. The increase in returns needs to increase dispersion enough to offset the decrease in dispersion coming from the effect of $1/\rho$ reducing q. The proof of the proposition in the appendix shows that a sufficient condition for the first effect to dominate is that the elasticity of $V_1[\tilde{f} - \tilde{p}r]$ with respect to ρ is greater than 1, which requires a large enough asset supply variance.

1.5 Prediction 3: Performance

The third prediction of the model is that the average performance of investment managers is higher in recessions than it is in expansions. To measure performance, we want to measure the portfolio return, adjusted for risk. One risk adjustment that is both analytically tractable in our model and often used in empirical work is the certainty equivalent return, which is also an investor's objective (6). The following Proposition shows that risk-unadjusted abnormal portfolio returns, defined as the fund's portfolio return, $q_j'(f - pr)$, minus the market return, $\bar{q}'(f - pr)$ for skilled funds exceeds that of unskilled funds and non-fund investors by more when volatility is higher, that is, in recessions.

Proposition 5. *An increase in the variance of any shock σ_i increases the portfolio excess return of an informed fund, $E[(q_j - \bar{q})'(z - pr)]$.*

Because asset payoffs are more uncertain, recessions are times when information is more valuable. In principle, the result is general and holds when any shock becomes more volatile. But we show that the empirically relevant case is that only aggregate shock volatility rises in recessions. Furthermore, the return effect is larger for the aggregate shock because it depends on how abundant the risk is ($\bar{x}_a$) and the aggregate shock is naturally the most abundant one.

Therefore, the advantage of the skilled over the unskilled investors increases in recessions. This informational advantage generates higher returns for informed managers. In equilibrium, market clearing dictates that abnormal returns average to zero across all investors. However, because the data only include mutual funds, our model calculations must similarly exclude non-fund investors. Since investment managers are skilled or unskilled, while other investors are only unskilled, an increase in the skill premium implies that the *average* mutual fund's abnormal return rises in recessions.

Next, we consider the effect of an increase in the price of risk on performance given in the second part of the previous proposition.

Proposition 6. *If σ_x is sufficiently large, an increase in risk aversion ρ increases the portfolio excess return of an informed fund, $E[(q_j - \bar{q})'(z - pr)]$.*

The reason that a higher price of risk leads to higher performance is that information can resolve risk. Therefore, informed managers are compensated for risk that they do not bear because the information has resolved some of their uncertainty about random asset payoffs. When the price of risk rises, the value of being able to resolve this risk rises as well. Put differently, informed funds take larger positions in risky assets because they are less uncertain about their returns. These larger positions pay off more on average when the price of risk is high.

1.6 Who Underperforms?

The requirement that markets clear implies that not all investors can be successful stock-pickers or market-timers. In each period, someone must make poor stock-picking or market-timing decisions if someone else makes profitable decisions. We explain now why rational, unskilled investors underperform in equilibrium.

Unskilled investors have negative timing ability in recessions. When the aggregate state z_a is low, most skilled investors sell, pushing down asset prices, p, and making prior expected returns, $E[f - pr]$, high. Equation (10) shows that uninformed investors' asset holdings increase in $(\mu - pr)$. The uninformed misinterpret the low prices as a buying opportunity and earn low returns. In terms of our measure of market timing, the informed (uninformed) investors' holdings covary positively (negatively) with aggregate payoffs, making their market timing measure positive (negative). Since no investors learn about the aggregate shock in expansions, prices do not fall when unexpected aggregate shocks are negative. Since the price mechanism is shut down, the market timing measure is close to zero for both skilled and unskilled in expansions. Taken together, the average fund exhibits some ability to time the market and exploits that ability at the expense of the uninformed investors, in recessions.

Likewise, unskilled investors will show negative stock-picking ability in expansions. When the stock-specific shock z_i is low, and some investors know that it will be low, they will sell and depress the price of asset i. A low price raises the expected return on the asset $(\mu_i - p_i r)$

for uninformed investors. The high expected return induces them to buy more of the asset. Since they buy more of assets that subsequently have negative asset-specific payoff shocks, these uninformed investors display negative stock-picking ability.

We note that when there is a positive aggregate supply shock, prices will be lower (Lemma 1), and assets will look more attractive to both uninformed and informed agents, all else equal. In that case, both informed and uninformed can trade in the same direction because of the additional asset supply.

1.7 Endogenous Capacity Choice

So far, we have assumed that skilled investment managers choose how to allocate a fixed information-processing capacity, K. We now extend the model to allow for skilled managers to add capacity at a cost $\mathcal{C}(K)$.[11] We draw three main conclusions. First, the proofs of Propositions 1 and 2 hold for any chosen level of capacity K, below an upper bound, no matter the functional form of $\mathcal{C}$. The other propositions also continue to hold because they hold for any level of capacity. Endogenous capacity only has quantitative, not qualitative implications. Second, because the marginal utility of learning about the aggregate shock is increasing in its prior variance (Proposition 1), skilled managers choose to acquire higher capacity in recessions. This extensive-margin effect amplifies our benchmark, intensive-margin result. Third, the degree of amplification depends on the convexity of the cost function, $\mathcal{C}(K)$. The convexity determines how elastic equilibrium capacity choice is to the cyclical changes in the marginal benefit of learning. The supplementary appendix discusses numerical simulation results from an endogenous-K model; they are similar to our benchmark exogenous-K results.

A second margin along which one could imagine extending the results is to let investors choose whether to become informed at a fixed capacity level K or not. Under this discrete choice, an asymmetric equilibrium can arise where some investors become informed and other not and where all investors are ex-ante indifferent about whether or not to become informed. When capacity is more valuable, in recessions, more investors would become informed. This would dampen the market timing, stock picking and outperformance metrics for the informed because the difference between informed and uninformed investors would weaken. However, it cannot overturn our results. If more information acquisition in recessions were to result

[11]We model this cost as a utility penalty, akin to the disutility from labor in business cycle models. Since there are no wealth effects in our setting, it would be equivalent to modeling a cost of capacity through the budget constraint. For a richer treatment of information production modeling, see Veldkamp (2006).

in worse performance (or market timing) in recessions than in booms, then the returns to information would be lower than in booms, which would imply that there should be less information acquisition in recessions, a contradiction. The supplementary appendix discusses numerical simulation results that vary exogenously the fraction of informed investors χ. With more informed fund managers, each have a smaller informational advantage than with fewer informed managers, resulting in lower market-timing (stock-picking) skill in recessions (expansions) and lower outperformance. But because there are now more informed and fewer uninformed funds in the economy, the *average* fund's market-timing (stock-picking) skill rises in recessions (expansions) and so does its outperformance.

2 Bringing Model to Data

This section introduces the empirical measures that we use in Section 3 to test the theory of Section 1. It argues that they have the same comparative statics as their theoretical counterparts and are themselves well-defined objects.

2.1 Market Timing and Stock Picking Measures

We define a fund's *fundamentals-based timing ability*, $Ftiming$, as the covariance between its portfolio weights in deviation from the market portfolio weights, $w_i^j - w_i^m$, and the aggregate payoff shock, z_a, over a T-period horizon, averaged across assets:

$$Ftiming_t^j = \frac{1}{TN^j} \sum_{i=1}^{N^j} \sum_{\tau=0}^{T-1} (w_{it+\tau}^j - w_{it+\tau}^m)(z_{a(t+\tau+1)}), \tag{16}$$

where N^j is the number of individual assets held by fund j. The subscript t on the portfolio weights and the subscript $t+1$ on the aggregate shock signify that the aggregate shock is unknown at the time of portfolio formation. Relative to the market, a fund with a high $Ftiming$ overweights assets that have high (low) sensitivity to the aggregate shock in anticipation of a positive (negative) aggregate shock realization and underweights assets with a low (high) sensitivity.

When skilled investment managers allocate attention to stock-specific payoff shocks, z_i, information about z_i allows them to choose portfolios that covary with z_i. *Fundamentals-based stock picking ability*, $Fpicking$, measures the covariance of a fund's portfolio weights of each stock, relative to the market, with the stock-specific shock, z_i:

$$Fpicking_t^j = \frac{1}{N^j} \sum_{i=1}^{N^j} (w_{it}^j - w_{it}^m)(z_{it+1}). \tag{17}$$

How well the manager can choose portfolio weights in anticipation of future asset-specific payoff shocks is closely linked to her stock-picking ability.

Ftiming and *Fpicking* are closely related to commonly-used market-timing and stock-picking variables, which measure how a fund's holdings of each asset, relative to the market, covary with the systematic and idiosyncratic components of the *stock return*. The key difference is that we measure how a portfolio covaries with aggregate and firm-specific *fundamentals*. We use the fundamentals-based measures because they correspond more closely to the idea in the model that funds are learning about fundamentals and using signals about those fundamentals to time the market and pick stocks.

2.2 Dispersion and Outperformance Measures

To connect the propositions to the data, we measure portfolio dispersion as the sum of squared deviations of fund j's portfolio weight in asset i at time t, w_{it}^j, from the average fund's portfolio weight in asset i at time t, w_{it}^m, summed over all assets held by fund j, N^j:

$$Portfolio\ Dispersion_t^j = \sum_{i=1}^{N^j} \left(w_{it}^j - w_{it}^m\right)^2 \tag{18}$$

This measure is similar to the portfolio concentration measure in Kacperczyk, Sialm, and Zheng (2005) and the active share measure in Cremers and Petajisto (2009). It is the same quantity as in Proposition 3, except that the number of shares q is replaced with portfolio weights w. In recessions, the portfolios of the informed managers differ more from each other and more from those of the uninformed investors. Part of this difference comes from a change in the composition of the risky asset portfolio and part comes from differences in the fraction of assets held in riskless securities. Fund j's portfolio weight w_{it}^j is a fraction of the fund's assets, including both risky and riskless, held in asset i. Thus, when one informed fund gets a bearish signal about the market, its w_{it}^j for all risky assets i falls. Dispersion can rise when funds take different positions in the risk-free asset, even if the fractional allocation among the risky assets remains identical.

The higher dispersion across funds' portfolio strategies translates into a higher cross-sectional dispersion in fund abnormal returns ($R^j - R^m$). To facilitate comparison with the data, we define the dispersion of variable X as $|X^j - \bar{X}|$. The notation $\bar{X}$ denotes the equally weighted cross-sectional average across all investment managers (excluding non-fund investors).

When funds get signals about the aggregate state z_a that are heterogenous, they take different directional bets on the market. Some funds tilt their portfolios to high-beta assets and other funds to low-beta assets, thus creating dispersion in fund betas. To look for evidence of this mechanism, we form a CAPM regression for fund j

$$R_t^j = \alpha^j + \beta^j R_t^m + \sigma_\varepsilon^j \varepsilon_t^j \tag{19}$$

and test for an increase in the beta dispersion in recessions as well.

We measure outperformance by looking at abnormal fund returns, measured as the fund's return minus the market return, and several risk-adjusted returns. One way to do that risk adjustment is to estimate (19) for each fund and look at the α of that equation. We also compute αs for similar models with multiple risk factors that are common in the empirical literature. The three-factor alpha controls for a size and a value factor, while the four-factor additionally controls for a momentum factor.

2.3 Do the Theoretical Measures and Empirical Measures Have the Same Properties?

The theoretical propositions refer to payoffs and quantities that have analytical expressions in a model with CARA preferences and normally distributed asset payoffs. But they do not correspond to the returns and portfolio weights that are commonly used in the empirical literature. The commonly used empirical measures, however, are not tractable analytically. This raises the concern that, if we construct $Ftiming$ and $Fpicking$ inside the model, allocating attention to aggregate shocks might not manifest itself as high $Ftiming$ and allocating attention to stock-specific risks might not be captured by high $Fpicking$. Similar concerns arise for the dispersion and outperformance results. In addition, these empirical measures may not be well-defined theoretical objects.

We start by noting that the empirical measures we use in the next section to test our theory are well-defined objects. Our empirical measures use conventional definitions of fund portfolio returns and portfolio weights. The weight fund j puts on asset i in its portfolio is the time-2 market value held in asset i, $p_i q_i$, divided by the value of all assets held by the fund. By the budget constraint, the value of all assets held by the fun is initial wealth W_0. Thus, $w_i^j \equiv \frac{p_i q_i^j}{W_0}$. Similarly, the weight of asset i in the market is the value of asset i $p_i(\bar{x}_i + x_i)$, divided by the value of all the risky assets in the market. If we assume that riskless assets are in zero net supply, then the value of all risky assets is the sum of all initial wealth. Thus, $w_i^m \equiv \frac{p_i q_i}{\int W_{0j} dj}$ and $q_i \equiv \int_j q_i^j$ is the total demand for asset j. Likewise, a fund j's return is $R^j \equiv \sum_{i \in \{0,1,2,c\}} w_i^j R^i$, where. Rearranging reveals that this is equal to one plus the fund's end-of-period wealth W_j, divided by their initial wealth W_0: $R^j = W_0^j / W_j - 1$. All these measures have only initial wealth in their denominators. Since initial wealth is known at the start of period 1, each one has finite moments in the model.

Second, to further allay the concern, we choose parameters and simulate our model in which each fund manager allocates attention and chooses his portfolio optimally. Then, we compute equilibrium prices and portfolio weights and estimate the same regressions on the model-generated data as we do in the real data. This exercise verifies that the empirical and theoretical measures have the same comparative statics. The supplementary appendix explains how parameters are chosen to match moments of the aggregate and individual stock returns in expansions and recessions, and it documents a complete set of results. In the simulation, we set the number of assets $n = 3$. For brevity, we only discuss the key comparative statics here.

For our benchmark parameter values, all skilled managers exclusively allocate attention to stock-specific shocks in expansions. In contrast, the bulk of skilled managers learn about the aggregate shock in recessions (87%, with the remaining 13% equally split between shocks 1 and 2). Thus, managers reallocate their attention over the business cycle. Such large swings in attention allocation occur for a wide range of parameters.

This shift in attention allocation is clearly reflected in the fluctuations in $Ftiming$ and $Fpicking$. The simulation results show that skilled investors' $Ftiming$ in recessions is orders of magnitude higher than in expansions. Similarly, we find that skilled funds have positive $Fpicking$ ability in expansions, when they allocate their attention to stock-specific information. Our numerical results also confirm that there is a higher dispersion in the funds' betas, and in their abnormal returns, in recessions. Lastly, the simulations confirm that abnormal returns and alphas, defined as in the empirical literature, and averaged over all funds, are higher in recessions than in expansions. Skilled investment managers have positive excess returns, while the uninformed ones have negative excess returns. Aggregating returns across skilled and unskilled funds results in higher average alphas in recessions.

3 Evidence from Equity Mutual Funds

Our model studies attention allocation over the business cycle, and its consequences for investors' strategies. We now turn to a specific set of investors, active U.S. mutual fund managers, to test the predictions of the model. The richness of the data makes the mutual fund industry a great laboratory for these tests. In principle, similar tests could be conducted for hedge funds, real estate investment trusts, other professional investment managers, or even individual investors.

3.1 Data

Our sample builds upon several data sets. We begin with the Center for Research on Security Prices (CRSP) survivorship bias-free mutual fund database. The CRSP database provides comprehensive information about fund returns and a host of other fund characteristics, such as size (total net assets), age, expense ratio, turnover, and load. Given the nature of our tests and data availability, we focus on actively managed open-end U.S. equity mutual funds. We further merge the CRSP data with fund holdings data from Thomson Financial. The total number of funds in our merged sample is 3,477. We also use the CRSP/Compustat stock-level database, which is a source of information on individual stocks' returns, market capitalizations, book-to-market ratios, momentum, liquidity, and standardized unexpected earnings (SUE). The aggregate stock market return is the value-weighted average return of all stocks in the CRSP universe.

We use innovations in monthly seasonally-adjusted industrial production, obtained from the Federal Reserve Statistical Release, as a proxy for aggregate shocks. We measure recessions using the definition of the National Bureau of Economic Research (NBER) business cycle dating committee. The start of the recession is the peak of economic activity and its end is the trough. Our aggregate sample spans 312 months of data from January 1980 until December 2005, among which 38 are NBER recession months (12%). We consider several alternative recession indicators and find our results to be robust.[12]

[12]Results are omitted for brevity but are available from the authors upon request.

3.2 Motivating Fact: Aggregate Risk and Prices of Risk Rise in Recessions

At the outset, we present empirical evidence for the main assumption in our model: Recessions are periods in which individual *stocks* contain more aggregate risk and when prices of risk are higher.

Table 1 shows that an average stock's aggregate risk increases substantially in recessions whereas the change in idiosyncratic risk is not statistically different from zero. The table uses monthly returns for all stocks in the CRSP universe. For each stock and each month, we estimate a CAPM equation based on a twelve-month rolling-window regression, delivering the stock's beta, β_t^i, and its residual standard deviation, $\sigma_{\varepsilon t}^i$. We define the aggregate risk of stock i in month t as $|\beta_t^i \sigma_t^m|$ and its idiosyncratic risk as $\sigma_{\varepsilon t}^i$, where σ_t^m is formed monthly as the realized volatility from daily return observations. Panel A reports the results from a time-series regression of the aggregate risk (Columns 1 and 2), the idiosyncratic risk (Columns 3 and 4), and the ratio of aggregate to idiosyncratic risk (Columns 5 and 6), all averaged across stocks, on the NBER recession indicator variable.[13] The aggregate risk is twenty percent higher in recessions than it is in expansions (6.69% versus 8.04% per month), an economically and statistically significant difference. In contrast, the stock's idiosyncratic risk is essentially identical in expansions and in recessions. As a result, the ratio of aggregate to idiosyncratic risk increases from 0.508 in expansions to 0.606 in recessions, and this cyclicality is driven exclusively by the numerator. The results are similar whether one controls for other aggregate risk factors (Columns 2, 4, and 6) or not (Columns 1, 3, and 5).

Panel B reports estimates from pooled (panel) regressions of each stock's aggregate risk (Columns 1 and 2), idiosyncratic risk (Columns 3 and 4), or the ratio of aggregate to idiosyncratic risk (Columns 5 and 6) on the recession indicator variable, *Recession*, and additional stock-specific control variables including size, book-to-market ratio, and leverage. The panel results confirm the time-series findings.

A large literature in economics and finance presents evidence supporting the results in Table 1. First, Ang and Chen (2002), Ribeiro and Veronesi (2002), and Forbes and Rigobon (2002) document that stocks exhibit more comovement in recessions, consistent with stocks carrying higher systematic risk in recessions. Second, Schwert (1989, 2011), Hamilton and Lin (1996), Campbell, Lettau, Malkiel, and Xu (2001), and Engle and Rangel (2008) show that aggregate stock market return volatility is much higher during periods of low economic

[13]The reported results are for equally weighted averages. Unreported results confirm that value-weighted averaging across stocks delivers the same conclusion.

activity. Diebold and Yilmaz (2008) find a robust cross-country link between volatile stock markets and volatile fundamentals. Third, Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2012) find that the volatilities of GDP and industrial production growth, obtained from GARCH estimation, and the volatility implied by stock options are much higher during recessions. The same result holds for the uncertainty in several establishment-, firm- and industry-level payoff measures they consider.

Our second assumption, that the price of risk rises in recessions, is supported by four pieces of evidence. First is an empirical literature that documents the counter-cyclical nature of risk premia and Sharpe ratios on equity, bonds, options, and currencies.[14] Second, a large theoretical literature has developed that generates such counter-cyclical market prices of risk, e.g., the external habit model of Campbell and Cochrane (1999), the consumption commitments model of Chetty and Szeidl (2007), the variable rare disasters model of Gabaix (2012), or heterogeneous-agent models where agents have different risk aversion parameters. Third, a recent paper by Dew-Becker (2012) combines habit with Epstein-Zin preferences in an asset pricing model with production. He uses the structure of the model to construct an empirical proxy for risk aversion, which rises in recessions. Fourth, a handful of recent papers show that aggregate risk aversion rises in recessions because of properties of aggregation.[15] In these models, heterogeneous agents with the same preferences but different risk aversion parameters aggregate into a representative agent who has wealth-weighted functions of the individual agent's parameters. Because more risk-averse agents are more conservative, their relative wealth rises in recessions, making aggregate risk aversion counter-cyclical.

3.3 Testing Prediction 1: Time-Varying Skill

Turning to our main model predictions, we first test whether skilled investment managers reallocate their attention over the business cycle in a way that is consistent with measures of time-varying skill. Learning about the aggregate payoff shock in recessions makes managers choose portfolio holdings that covary more with the aggregate shock. Conversely, in expansions, their holdings covary more with stock-specific information.

To estimate time-varying skill, we need measures of $Ftiming$ and $Fpicking$ for each fund j in each month t. We proxy for the aggregate payoff shock with the innovation in log industrial

[14] Among many others, Fama and French (1989), Cochrane (2006), Ludvigson and Ng (2009), Lettau and Ludvigson (2010), Lustig, Roussanov, and Verdelhan (2012), and the references therein. A related fact consistent with counter-cyclical market prices of risk is high corporate bond yields in recessions despite only modestly higher default rates, see Chen (2010).

[15] See Dumas (1989), Chan and Kogan (2002), and Gârleanu and Panageas (2010), among others.

production growth, estimated from an AR(1).[16] A time series of $Ftiming_t^j$ is obtained by computing the covariance of the innovations and each fund j's portfolio weights (as in equation (16)), using twelve-month rolling windows. Following equation (17), *Fpicking* is computed in each month t as a cross-sectional covariance across the assets between the fund's portfolio weights and firm-specific earnings shocks (SUE). We then estimate the following two equations using pooled (panel) regression model and calculating standard errors by clustering at the fund and time dimensions.

$$
\begin{aligned}
Fpicking_t^j &= a_0 + a_1 Recession_t + \mathbf{a_2 X_t^j} + \epsilon_t^j, \qquad (20)\\
Ftiming_t^j &= a_3 + a_4 Recession_t + \mathbf{a_5 X_t^j} + \varepsilon_t^j, \qquad (21)
\end{aligned}
$$

$Recession_t$ is an indicator variable equal to one if the economy in month t is in recession, as defined by the NBER, and zero otherwise. X is a vector of fund-specific control variables, including the fund age, the fund size, the average fund expense ratio, the turnover rate, the percentage flow of new funds, the fund load, the volatility of fund flows, and the fund style characteristics along the size, value, and momentum dimensions.

Our model predicts that *Ftiming* should be higher in recessions, which means that the coefficient on *Recession*, a_4, should be positive. Conversely, the fund's portfolio holdings and its returns covary more with subsequent firm-specific shocks in expansions. Therefore, our hypothesis is that *Fpicking* should fall in recessions, or that a_1 should be negative.

The parameter estimates appear in columns 1, 2, 4 and 5 of Table 2. Column 1 shows the results for a univariate regression model. In expansions, *Ftiming* is not different from zero, implying that funds' portfolios do not comove with future macroeconomic information in those periods. In recessions, *Ftiming* increases. The increase amounts to ten percent of a standard deviation of *Ftiming*. It is measured precisely, with a t-statistic of 3. To remedy the possibility of a bias in the coefficient due to omitted fund characteristics correlated with recession times, we turn to a multivariate regression. Our findings, in Column 2, remain largely unaffected by the inclusion of the control variables. Columns 4 and 5 of Table 2 show that the average *Fpicking* across funds is positive in expansions and substantially lower in recessions. The effect is statistically significant at the 1% level. It is also economically significant: Fpicking decreases by approximately ten percent of one standard deviation.[17] In sum,

[16]Our results are robust to using industrial productions growth itself. Our results are also robust to measuring aggregate shocks to fundamentals as innovations in non-farm employment growth

[17]We note that the R^2 statistics in this and all following tables are quite low. We share this feature with

the data support both main predictions of the theory: Portfolio holdings are more sensitive to aggregate shocks in recessions and more sensitive to firm-specific shocks in expansions.

Kacperczyk, Van Nieuwerburgh, and Veldkamp (2012) show that these results are robust to alternative, return-based measures of picking and timing, to alternative recession indicator variables, and they investigate in more detail the strategies funds use to time the market.

Testing for Separate Effects of Volatility and Recessions. To identify a more nuanced prediction of the model, we can split the recession effect into that which comes from aggregate volatility and that which comes from an increased price of risk. Proposition 1 predicts that an increase in aggregate volatility alone should cause managers to reallocate attention to aggregate shocks. Furthermore, there should be an additional effect of recessions, after controlling for aggregate volatility, that comes from the increase in the price of risk (Proposition 2). To test for these two separate effects, we re-estimate the previous results with both an indicator for recessions and an indicator for high aggregate payoff volatility. The high-volatility indicator variable equals one in months with the highest volatility of aggregate earnings growth, where aggregate volatility is estimated from Shiller's S&P 500 earnings growth data.[18] We include both NBER recession and high aggregate payoff volatility indicators as explanatory variables in an empirical horse race.

Columns 3 and 6 of Table 2 show that both recession and volatility contribute to a lower $Fpicking$ in expansions and a higher $Ftiming$ in recessions. For the $Ftiming$ result, the recession effect is much stronger, while for the $Fpicking$ result both recession and volatility contribute about equally. Clearly, there is an effect of recessions beyond the one coming through volatility. This is consistent with the predictions of our model, where recessions are characterized both by an increase in aggregate payoff volatility and an increase in the price of risk.

3.4 Testing Prediction 2: Dispersion

The second main prediction of the model states that heterogeneity in fund investment strategies and portfolio returns rises in recessions. To test this hypothesis, we estimate the fol-

the entire literature that analyzes mutual fund data. The reason is that mutual fund-level data display a large amount of heterogeneity that is hard to account for.

[18] We calculate the twelve-month rolling-window standard deviation of aggregate earnings growth. The volatility cutoff selects 6% of months. Of the high-volatility periods, 28% are recessions. Of all other periods (when high-volatility indicator is 0), 10.6% are recessions. Conversely, 14% of recessions are also high-volatility periods whereas only 4.8% of expansions are high-volatility periods.

lowing regression specification, using various return and investment heterogeneity measures, generically denoted as $Dispersion_t^j$, the dispersion of fund j at month t.

$$Dispersion_t^j = b_0 + b_1 Recession_t + \mathbf{b_2 X_t^j} + \epsilon_t^j, \tag{22}$$

The definitions of *Recession* and other controls mirror those in regression (20). Our coefficient of interest is b_1.

The first dispersion measure we examine is $Portfolio\ Dispersion$, defined in equation (18). It measures a deviation of a fund's investment strategy from a passive market strategy, and hence, in equilibrium, from the strategies of other investors. The results in Columns 1 and 2 of Table 3 indicate an increase in average $Portfolio\ Dispersion$ across funds in recessions. The increase is statistically significant at the 1% level. It is also economically significant: The value of portfolio dispersion in recessions goes up by about 15% of a standard deviation.

Since dispersion in fund strategies should generate dispersion in fund returns, we next look for evidence of higher return dispersion in recessions. To measure dispersion, we use the absolute deviation between fund j's return and the equally weighted cross-sectional average, $|return_t^j - \overline{return}_t|$, as the dependent variable in (22). Columns 5 and 6 of Table 3 show that return dispersion increases by 17% in recessions. Finally, portfolio and return dispersion in recessions should come from different directional bets on the market. This should show up as an increase in the dispersion of portfolio betas. Columns 3 and 4 show that the CAPM-beta dispersion increases by 36% in recessions, all consistent with the predictions of our model.

These findings are robust. Counter-cyclical dispersion in funds' portfolio strategies is also found in measures of fund style shifting and sectoral asset allocation. The dispersion in returns is also found for abnormal returns and fund alphas. Results are available on request.

Testing for Separate Effects of Volatility and Recessions. Propositions 3 and 4 tell us that return dispersion increases in recessions for two reasons. One is that the volatility of aggregate shocks increases and the other reason is that the price of risk increases. We can disentangle these two effects by regressing return dispersion on volatility and recession simultaneously. The model would predict that volatility should be a significant determinant of dispersion and that after controlling for volatility, there should be some additional explanatory power of recessions that comes from the price of risk effect.

Column 7 of Table 3 shows that both the recession and the volatility effects are present in the data. Both are associated with a significant increase in the dispersion of returns. After including the volatility variable, the magnitude of the coefficient on *Recession* falls by 25%, but the recession variable retains its statistical significance. The volatility and price of risk fluctuations both have significant effects on portfolio dispersion, with the effect of volatility being somewhat larger.

3.5 Testing Prediction 3: Performance

The third prediction of our model is that recessions are times when information allows funds to earn higher average risk-adjusted returns. Empirical work by Moskowitz (2000), Kosowski (2011), de Souza and Lynch (2012), and Glode (2011) also documents such evidence, but their focus is solely on performance, not on managers' attention allocation nor their investment strategies. Their results are based on time-series analysis, while we account for differences in fund size, age, turnover, flows, loads, style and flow volatility. Furthermore, these studies are silent on the specific mechanism that drives the outperformance result, which is one of the main contributions of our paper.

We evaluate this hypothesis using the following regression specification:

$$Performance_t^j = c_0 + c_1 Recession_t + \mathbf{c_2 X_t^j} + \epsilon_t^j \tag{23}$$

where $Performance_t^j$ denotes fund j's performance in month t, measured as fund abnormal returns, or CAPM, three-factor, or four-factor alphas. All returns are net of management fees. The coefficient of interest is c_1.

Column 1 of Table 4 shows that the average fund's net return is 3bp per month lower than the market return in expansions, which is statistically indistinguishable from zero. But the coefficient of *Recession* is is 38bp per month, implying that the average mutual fund's abnormal return is 4.6% per year higher in recessions. This difference is highly statistically significant and increases further after we control for fund characteristics (Column 2). Similar results (Columns 3 and 4) obtain when we use the CAPM alpha as a measure of fund performance, except that the net alpha in expansions is now statistically significantly negative. The 34bp per month higher net alpha in recessions corresponds to 4% per year. When we use alphas based on the three- and four-factor models, the recession return premium diminishes (Columns 5-8). But in recessions, the four-factor alpha still represents a non-trivial 1% per year risk-adjusted excess return, 1.6% higher (significant at the 1% level) than the -0.6% recorded in expansions.

The advantage of this cross-sectional regression model is that it allows us to include a host of fund-specific control variables. The disadvantage is that performance is measured using past twelve-month rolling-window regressions. Thus, a given observation can be classified as a recession when some or even all of the remaining eleven months of the window are expansions. To verify the robustness of our cross-sectional results, we also employ a time-series approach. In each month, we form the equally weighted portfolio of funds and calculate its net return, in excess of the risk-free rate. We then regress this time series of fund portfolio returns on *Recession* and common risk factors, adjusting standard errors for heteroscedasticity and autocorrelation. We find similar outperformance in recessions. Our results are also robust to alternative performance measures, such as *gross* fund returns, gross alphas, or the information ratio (the ratio of the CAPM alpha to the CAPM residual volatility). All increase sharply in recessions. Finally, we find similar results when we lead alpha on the left-hand side by one month instead of using a contemporaneous alpha. All results point in the same direction: Outperformance increases in recessions.

Testing for Separate Effects of Volatility and Recessions. As before, two forces increase the performance of funds relative to non-funds in recessions: the increase in volatility and the increase in the price of risk (propositions 5 and 6). Column 9 of Table 4 shows that the data are consistent with each force having a distinct effect on fund outperformance. We use the 4-factor alpha as the dependent variable for this exercise because we want to avoid conflating more risk taking in recessions with greater fund outperformance in recessions. When we regress each fund's 4-factor alpha on a recession indicator and a volatility measure, both have positive, significant coefficients. Adding the volatility variable reduces the size of the recession effect by 28%. This suggests that fund outperformance in recessions is due mostly to the increased price of risk and is due to a lesser extent to the higher volatility of aggregate shocks. But the fact that both variables have a significant relationship with fund outperformance, dispersion, and attention, in the direction predicted by the theory offers solid support for the model.

3.6 Alternative Explanations

The existing literature has not yet advanced any alternative explanations for time-varying skill, as far as we know. In Kacperczyk, Van Nieuwerburgh, and Veldkamp (2012), we consider six possible alternative explanations in detail: 1) Measures of fund skill change because fund managers change; 2) the results are driven by sample selection, because highly successful managers are hired by hedge funds; 3) the convex flow-performance relationship changes incentives over the business cycle; 4) young managers with career concerns have an incentive to herd that might vary over the cycle; 5) investors in mutual funds have time-varying marginal utility; and 6) the effects arise mechanically from the properties of asset returns. Ultimately, we conclude that while some of these hypotheses can account for some of the facts, they do not account for all facts *jointly*. In particular, none explains the main fact, that market timing is better in recessions and stock picking is more successful in booms.

4 Conclusion

Do investment managers add value for their clients? The answer to this question matters for issues ranging from the discussion of market efficiency to practical portfolio advice for households. The large amount of randomness in financial asset returns makes it a difficult question to answer. The multi-billion investment management business is first and foremost an information-processing business. We model investment managers not only as agents making optimal portfolio decisions, but also as human beings with finite mental capacity (attention), who optimally allocate that scarce capacity to process information at each point in time. Since the optimal attention allocation varies with the state of the economy, so do investment strategies and fund returns. As long as a subset of skilled investment managers can process information about future asset payoffs, the model predicts a higher covariance of portfolio holdings with aggregate asset payoff shocks, more cross-sectional dispersion in portfolio investment strategies and returns across funds, and a higher average outperformance in recessions. We observe these patterns in investments and returns of actively managed U.S. mutual funds. Hence, the data are consistent with a world in which some investment managers have skill.

Beyond the mutual fund industry, a sizeable fraction of GDP currently comes from industries that produce and process information (consulting, business management, product design, marketing analysis, accounting, rating agencies, equity analysts, etc.). Ever increasing access to information has made the problem of how to best allocate a limited amount of information-processing capacity ever more relevant. While information choices have consequences for real outcomes, they are often poorly understood because they are difficult to measure. By predicting how information choices are linked to observable variables (such as the state of the economy) and by tying information choices to real outcomes (such as portfolio investment), we show how models of information choices can be brought to the data. This information-choice-based approach could be useful in examining other information-processing sectors of the economy.

References

Ang, A., and J. Chen (2002): "Asymmetric correlations of equity portfolios," *Journal of Financial Economics*, 63 (3), 443–494.

Baker, M., L. Litov, J. Wachter, and J. Wurgler (2010): "Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements," *Journal of Financial and Quantitative Analysis*, 45 (5), 1111–1131.

Berk, J., and R. C. Green (2004): "Mutual fund flows and performance in rational markets," *Journal of Political Economy*, 112, 1269–1295.

Berk, J., and J. H. van Binsbergen (2012): "Measuring managerial skill in the mutual fund industry," Working Paper, Stanford University.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2012): "Really uncertain business cycles," NBER Working Paper No 18245.

Campbell, J., M. Lettau, B. Malkiel, and Y. Xu (2001): "Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk," *Journal of Finance*, 56 (1), 1–44.

Campbell, J. Y., and J. H. Cochrane (1999): "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of Political Economy*, 107(2), 205–251.

Chan, Y. L., and L. Kogan (2002): "Catching up with the Joneses: Heterogeneous preferences and the dynamics of asset prices," *Journal of Political Economy*, 110 (6), 1255–1285.

Chapman, D., R. Evans, and Z. Xu (2010): "The portfolio choices of young and old active mutual fund managers," Working Paper, Boston College.

Chen, H. (2010): "Macroeconomic conditions and the puzzles of credit spreads and capital structure," *Journal of Finance*, 65 (6), 2171–2212.

Chetty, R., and A. Szeidl (2007): "Consumption commitments and risk preferences," *Quarterly Journal of Economics*, 122 (2), 831–877.

Chien, Y., H. Cole, and H. Lustig (2011): "A multiplier approach to understanding the macro implications of household finance," *Review of Economic Studies*, 78 (1), 199–234.

Cochrane, J. H. (2006): "Financial markets and the real economy," in *International Library of Critical Writings in Financial Economics*, vol. 18. London: Edward Elgar.

Cremers, M., and A. Petajisto (2009): "How active is your fund manager? A new measure that predicts performance," *Review of Financial Studies*, 22, 3329–3365.

Cuoco, D., and R. Kaniel (2011): "Equilibrium prices in the presence of delegated portfolio management," *Journal of Financial Economics*, 101 (2), 264–296.

de Souza, A., and A. W. Lynch (2012): "Does mutual fund performance vary over the business cycle?," Working Paper, Fordham University and New York University.

Dew-Becker, I. (2012): "A model of time-varying risk premia with habits and production," Working Paper, Harvard.

Diebold, F., and K. Yilmaz (2008): "Macroeconomic volatility and stock market volatility world-wide," in *Volatility and time-series econometrics: Essays in honor of Robert F. Engle*, ed. by T. Bollerslev, J. Russell, and M. Watson. Oxford University Press.

Dumas, B. (1989): "Two-person dynamic equilibrium in the capital market," *Review of Financial Studies*, 2 (2), 157–188.

Engle, R., and J. Rangel (2008): "The spline-GARCH model for low-frequency volatility and its global macroeconomic causes," *Review of Financial Studies*, 21 (3), 1187–1222.

Fama, E. F., and K. French (1997): "Industry costs of equity," *Journal of Financial Economics*, 43, 153–193.

Fama, E. F., and K. R. French (1989): "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics*, 25 (1), 23–49.

Forbes, K., and R. Rigobon (2002): "No contagion, only interdependence: Measuring stock co-movements," *Journal of Finance*, 57, 2223–2261.

Gabaix, X. (2012): "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance," *Quarterly Journal of Economics*, 127 (2), 645–700.

Gârleanu, N., and S. Panageas (2010): "Young, old, conservarive and bold. The implications of finite lives and heterogeneity for asset pricing," Working Paper, University of California, Berkeley and University of Chicago.

Glode, V. (2011): "Why mutual funds underperform?," *Journal of Financial Economics*, 99 (3), 546–559.

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

Hamilton, J. D., and G. Lin (1996): "Stock market volatility and the business cycle," *Journal of Applied Econometrics*, 11 (5), 573–593.

HUANG, J., C. SIALM, AND H. ZHANG (2011): "Risk shifting and mutual fund performance," *Review of Financial Studies*, 24 (8), 2575–2616.

KACPERCZYK, M., AND A. SERU (2007): "Fund manager use of public information: New evidence on managerial skills," *Journal of Finance*, 62, 485–528.

KACPERCZYK, M., C. SIALM, AND L. ZHENG (2005): "On the industry concentration of actively managed equity mutual funds," *Journal of Finance*, 60, 1983–2012.

——— (2008): "Unobserved actions of mutual funds," *Review of Financial Studies*, 21, 2379–2416.

KACPERCZYK, M., S. VAN NIEUWERBURGH, AND L. VELDKAMP (2012): "Time-varying fund manager skill," Working Paper, New York University.

KANIEL, R., AND P. KONDOR (2012): "The delegated Lucas tree," *Review of Financial Studies*, forthcoming.

KLENOW, P. J., AND J. L. WILLIS (2007): "Sticky information and sticky prices," *Journal of Monetary Economics*, 54, 79–99.

KOIJEN, R. (2012): "The cross-section of managerial ability, incentives, and risk preferences," *Journal of Finance*, f orthcoming.

KONDOR, P. (2012): "The more we know about the fundamental, the less we agree," *Review of Economic Studies*, 79(3), 1175–1207.

KOSOWSKI, R. (2011): "Do mutual funds perform when it matters most to investors? US mutual fund performance and risk in recessions and expansions," *Quarterly Journal of Finance*, 1 (3), 607–664.

LETTAU, M., AND S. LUDVIGSON (2010): "Measuring and modeling variation in the risk-return tradeoff," in *Handbook of Financial Econometrics*, ed. by Y. Ait-Sahalia, and L. P. Hansen, vol. 1. Elsevier Science B.V., North Holland.

LUDVIGSON, S., AND S. NG (2009): "Macro factors in bond risk premia," *Review of Financial Studies*, 22, 5027–5067.

LUSTIG, H., N. ROUSSANOV, AND A. VERDELHAN (2012): "Countercyclical currency risk premia," Working Paper UCLA Anderson, Wharton School, and MIT Sloan.

MAĆKOWIAK, B., E. MOENCH, AND M. WIEDERHOLT (2009): "Sectoral price data and models of price setting," *Journal of Monetary Economics*, 56S, 78–99.

MAĆKOWIAK, B., AND M. WIEDERHOLT (2009): "Optimal sticky prices under rational inattention," *American Economic Review*, 99 (3), 769–803.

——— (2010): "Business cycle dynamics under rational inattention," Working Paper, Northwestern University.

——— (2012): "Inattention to rare events," Working Paper, Northwestern University.

Mamaysky, H., and M. Spiegel (2002): "A theory of mutual funds: Optimal fund objectives and industry organization," Working Paper Yale School of Management.

Matejka, F. (2011): "Rigid pricing and rationally inattentive consumers," Working Paper CERGE.

Mondria, J., T. Wu, and Y. Zhang (2010): "The determinants of international investment and attention allocation: Using internet search query data," *Journal of International Economics*, 82 (1), 85–95.

Moskowitz, T. J. (2000): "Discussion of mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses," *Journal of Finance*, 55, 1695–1704.

Pástor, L., and R. F. Stambaugh (2012): "On the size of the active management industry," *Journal of Political Economy*, forthcoming.

Reis, R. (2006): "Inattentive producers," *Review of Economic Studies*, 73(3), 793–821.

——— (2011): "When should policymakers make announcements?," Working Paper, Columbia University.

Ribeiro, R., and P. Veronesi (2002): "Excess co-movement of international stock markets in bad times: A rational expectations equilibrium model," Working Paper, University of Chicago.

Schwert, G. W. (1989): "Why does stock market volatility change over time?," *Journal of Finance*, 44(5), 1115–1153.

——— (2011): "Stock volatility during the recent financial crisis," *European Financial Management*, 17 (5), 789–805.

Simon, H. (1971): *Computers, communications, and the public interest*chap. Designing organizations for an information-rich world. The Johns Hopkins Press.

Sims, C. (2003): "Implications of rational inattention," *Journal of Monetary Economics*, 50(3), 665–90.

Van Nieuwerburgh, S., and L. Veldkamp (2009): "Information immobility and the home bias puzzle," *Journal of Finance*, 64 (3), 1187–1215.

——— (2010): "Information acquisition and under-diversification," *Review of Economic Studies*, 77 (2), 779–805.

Vayanos, D., and P. Woolley (2010): "Fund flows and asset prices: A baseline model," Working Paper, London School of Economics.

Veldkamp, L. (2006): "Media frenzies in markets for financial information," *American Economic Review*, 53 (4), 577–601.

——— (2011): *Information choice in macroeconomics and finance.* Princeton University Press.

Table 1: **Individual Stocks Have More Aggregate Risk in Recessions**

For each stock i and month t, we estimate a CAPM equation based on twelve months of data (a twelve-month rolling-window regression). This estimation delivers the stock's beta, β_t^i, and its residual standard deviation, $\sigma_{\varepsilon t}^i$. We define stock i's aggregate risk in month t as $\left|\beta_t^i \sigma_t^m\right|$ and its idiosyncratic risk as $\sigma_{\varepsilon t}^i$, where σ_t^m is the realized volatility from daily market return observations. Panel A reports results from a time-series regression of the average stock's aggregate risk, $\frac{1}{N}\sum_{i=1}^N \left|\beta_t^i \sigma_t^m\right|$, in Columns 1 and 2, of the average idiosyncratic risk, $\frac{1}{N}\sum_{i=1}^N \sigma_{\varepsilon t}^i$, in Columns 3 and 4, and of the ratio of aggregate to average idiosyncratic risk, in Columns 5 and 6, on *Recession*. *Recession* is an indicator variable equal to one for every month the economy is in a recession according to the NBER, and zero otherwise. In Columns 2, 4, and 6 we include several aggregate control variables: the market excess return (MKTPREM), the return on the small-minus-big portfolio (SMB), the return on the high-minus-low book-to-market portfolio (HML), the return on the up-minus-down momentum portfolio (UMD). The data are monthly from 1980-2005 (309 months). Standard errors (in parentheses) are corrected for autocorrelation and heteroscedasticity. Panel B reports results of panel regressions of each stock's aggregate risk, $\left|\beta_t^i \sigma_t^m\right|$, in Columns 1 and 2 and of its idiosyncratic risk, $\sigma_{\varepsilon t}^i$, in Columns 3 and 4, and of the ratio of a stock's aggregate to idiosyncratic risk, in Columns 5 and 6, on *Recession*. In Columns 2, 4, and 6 we include several firm-specific control variables: the log market capitalization of the stock, $log(Size)$, the ratio of book equity to market equity, $B-M$, the average return over the past year, $Momentum$, the stock's ratio of book debt to book debt plus book equity, $Leverage$, and an indicator variable, $NASDAQ$, equal to one if the stock is traded on NASDAQ. All control variables are lagged one month. The data are monthly and cover all stocks in the CRSP universe for 1980-2005. Standard errors (in parentheses) are clustered at the stock and time dimensions.

	(1)	(2)	(3)	(4)	(5)	(6)
	Aggregate Risk		Idiosyncratic Risk		Aggregate/Idiosyncratic Risk	
	Panel A: Time-Series Regression					
Recession	1.348	1.308	0.058	0.016	0.098	0.097
	(0.693)	(0.678)	(1.018)	(1.016)	(0.027)	(0.027)
MKTPREM		-4.034		-1.865		-0.215
		(3.055)		(3.043)		(0.226)
SMB		8.110		12.045		0.167
		(3.780)		(4.923)		(0.199)
HML		0.292		9.664		-0.308
		(5.458)		(8.150)		(0.302)
UMD		-4.279		-1.112		-0.270
		(2.349)		(3.888)		(0.178)
Constant	6.694	6.748	13.229	13.196	0.508	0.513
	(0.204)	(0.212)	(0.286)	(0.276)	(0.013)	(0.014)
Observations	309	309	309	309	309	309
R-squared	6.85	9.70	0.10	3.33	8.58	10.52
	Panel B: Pooled Regression					
Recession	1.203	1.419	0.064	0.510	0.096	0.104
	(0.242)	(0.238)	(0.493)	(0.580)	(0.021)	(0.024)
Log(Size)		-0.145		-1.544		0.043
		(0.021)		(0.037)		(0.002)
B-M Ratio		-0.934		-2.691		0.008
		(0.056)		(0.086)		(0.004)
Momentum		0.097		2.059		-0.040
		(0.101)		(0.177)		(0.005)
Leverage		-0.600		-1.006		-0.010
		(0.074)		(0.119)		(0.003)
NASDAQ		0.600		1.937		-0.043
		(0.075)		(0.105)		(0.005)
Constant	4.924	4.902	12.641	12.592	0.450	0.450
	(0.092)	(0.095)	(0.122)	(0.144)	(0.009)	(0.009)
Observations	1,312,216	1,312,216	1,312,216	1,312,216	1,312,216	1,312,216
R-squared	0.62	2.90	0.000	19.33	0.58	7.56

Table 2: **Attention Allocation is Cyclical**

Dependent variables: Fund j's $Ftiming_t^j$ is defined in equation (16), where the rolling window T is 12 months and the aggregate shock a_{t+1} is the change in industrial production growth between t and $t+1$. A fund j's $Fpicking_t^j$ is defined as in equation (17), where s_{it+1} is the change in asset i's earnings growth between t and $t+1$. All are multiplied by 10,000 for readability.
Independent variables: *Recession* is an indicator variable equal to one for every month the economy is in a recession according to the NBER, and zero otherwise. *Log(Age)* is the natural logarithm of fund age in years. $Log(TNA)$ is the natural logarithm of a fund total net assets. *Expenses* is the fund expense ratio. *Turnover* is the fund turnover ratio. *Flow* is the percentage growth in a fund's new money. *Load* is the total fund load. *Flowvol* is the volatility of fund flows, measures from the last twelve months of fund flows. The last three control variables measure the style of a fund along the size, value, and momentum dimensions, calculated from the scores of the stocks in their portfolio in that month. They are omitted for brevity. All control variables are demeaned. Flow and Turnover are winsorized at the 1% level. $Volatility$ is an indicator variable for periods of volatile earnings. We calculate the twelve-month rolling-window standard deviation of the year-to-year log change in the earnings of S&P 500 index constituents; the earnings data are from Robert Shiller for 1926-2008. Volatility equals one if this standard deviation is in the highest 10% of months in the 1926-2008 sample. During 1985-2005, 12% of months are such high volatility months. The data are monthly and cover the period 1980 to 2005. Standard errors (in parentheses) are clustered by fund and time.

	(1)	(2)	(3)	(4)	(5)	(6)
		Ftiming			Fpicking	
Recession	0.011	0.012	0.011	-0.742	-0.680	-0.617
	(0.003)	(0.003)	(0.004)	(0.138)	(0.126)	(0.127)
Volatility			0.000			-0.415
			(0.003)			(0.099)
Log(Age)		-0.001	-0.001		0.447	0.445
		(0.001)	(0.001)		(0.061)	(0.061)
Log(TNA)		-0.001	-0.001		-0.130	-0.129
		(0.000)	(0.000)		(0.029)	(0.029)
Expenses		-0.208	-0.227		96.748	96.205
		(0.219)	(0.217)		(11.200)	(11.007)
Turnover		-0.004	-0.004		-0.260	-0.262
		(0.001)	(0.001)		(0.063)	(0.063)
Flow		-0.010	-0.010		0.637	0.631
		(0.011)	(0.011)		(0.652)	(0.652)
Load		0.006	0.009		-9.851	-9.851
		(0.022)	(0.022)		(1.951)	(1.935)
Flow vol		-0.006	-0.004		6.684	6.711
		(0.017)	(0.017)		(1.042)	(1.032)
Constant	-0.001	0.000	-0.001	3.082	3.238	3.119
	(0.001)	(0.002)	(0.001)	(0.069)	(0.107)	(0.072)
Observations	221,488	221,488	221,488	165,029	165,029	165,029
R-squared	0.03	0.09	0.08	0.03	0.25	0.25

Table 3: **Portfolio and Return Dispersion Rises in Recessions**

Dependent variables: Portfolio dispersion is the Herfindahl index of portfolio weights in stocks $i \in \{1, \cdots, N\}$ in deviation from the market portfolio weights $\sum_{i=1}^{N}(w_{it}^{j} - w_{it}^{m})^2 \times 100$. Return dispersion is $|return_t^j - \overline{return}_t|$, where $\overline{return}$ denotes the (equally weighted) cross-sectional average. The CAPM beta comes from twelve-month rolling-window regressions of fund-level excess returns on excess market returns (and returns on SMB, HML, and MOM). Beta dispersion is constructed analogously to return dispersion. The right-hand side variables, the sample period, and the standard error calculation are the same as in Table 2.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Portfolio Dispersion		Beta Dispersion		Return Dispersion		
Recession	0.204	0.118	0.083	0.088	0.316	0.380	0.286
	(0.027)	(0.025)	(0.015)	(0.014)	(0.147)	(0.146)	(0.142)
Volatility							0.593
							(0.220)
Log(Age)		0.210		-0.005		-0.121	-0.109
		(0.028)		(0.002)		(0.017)	(0.018)
Log(TNA)		-0.165		0.004		0.043	0.036
		(0.014)		(0.001)		(0.009)	(0.010)
Expenses		31.986		4.162		28.330	25.602
		(4.867)		(0.212)		(2.621)	(2.564)
Turnover		-0.113		0.013		0.090	0.075
		(0.026)		(0.001)		(0.013)	(0.014)
Flow		-0.230		-0.004		-0.230	-0.268
		(0.108)		(0.018)		(0.223)	(0.217)
Load		-1.658		-0.318		-4.071	-3.548
		(0.900)		(0.041)		(0.509)	(0.527)
Flow vol		2.379		0.075		1.570	1.852
		(0.304)		(0.027)		(0.240)	(0.261)
Constant	1.525	1.524	0.228	0.228	1.904	1.899	1.843
	(0.024)	(0.022)	(0.006)	(0.006)	(0.084)	(0.077)	(0.078)
Observations	227,141	227,141	224,130	224,130	227,141	227,141	227,141
R-squared	0.10	4.80	1.35	8.10	0.19	7.00	7.89

Table 4: **Fund Performance Improves in Recessions**

Dependent variables: *Abnormal Return* is the fund return minus the market return. The alphas come from twelve-month rolling-window regressions of fund-level excess returns on excess market returns for the CAPM alpha, additionally on the SMB and the HML factors for the three-factor alpha, and additionally on the UMD factor for the four-factor alpha. The independent variables, the sample period, and the standard error calculations are the same as in Table 2.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Abnormal Return		CAPM Alpha		3-Factor Alpha		4-Factor Alpha		
Recession	0.384	0.433	0.339	0.399	0.043	0.062	0.108	0.131	0.086
	(0.056)	(0.059)	(0.048)	(0.050)	(0.034)	(0.026)	(0.041)	(0.033)	(0.035)
Volatility									0.138
									(0.064)
Log(Age)		-0.015		-0.032		-0.023		-0.035	-0.032
		(0.021)		(0.008)		(0.006)		(0.006)	(0.006)
Log(TNA)		0.023		0.040		0.018		0.019	0.018
		(0.013)		(0.004)		(0.003)		(0.003)	(0.003)
Expenses		-5.120		-0.929		-5.793		-5.970	-6.590
		(2.817)		(0.892)		(0.720)		(0.677)	(0.715)
Turnover		0.021		-0.054		-0.087		-0.076	-0.080
		(0.039)		(0.010)		(0.010)		(0.008)	(0.008)
Flow		2.127		2.308		1.510		1.386	1.378
		(0.672)		(0.172)		(0.096)		(0.096)	(0.096)
Load		-0.698		-0.810		-0.143		-0.371	-0.249
		(0.457)		(0.174)		(0.129)		(0.139)	(0.139)
Flow vol		-0.106		1.025		1.461		1.210	1.278
		(0.588)		(0.137)		(0.109)		(0.104)	(0.106)
Constant	-0.032	-0.036	-0.060	-0.065	-0.059	-0.061	-0.051	-0.053	-0.066
	(0.064)	(0.063)	(0.025)	(0.024)	(0.020)	(0.018)	(0.023)	(0.021)	(0.021)
Observations	224,130	224,130	224,130	224,130	224,130	224,130	224,130	224,130	224,130
R-squared	0.01	0.57	1.15	10.70	0.03	6.20	0.16	5.50	5.82

A Technical Appendix

A.1 Useful notation, matrices and derivatives

All the following matrices are diagonal with ii entry given by:

1. Posterior precision of shock i for an investor j is $\hat{\sigma}_{ij}^{-1}$, which is equivalent to

$$(\hat{\Sigma}_j^{-1})_{ii} = (\Sigma^{-1} + \Sigma_{\eta j}^{-1} + \Sigma_p^{-1})_{ii} = \sigma_i^{-1} + K_{ij} + \frac{\bar{K}_i^2}{\rho^2 \sigma_x} = \hat{\sigma}_i^{-1} \tag{24}$$

2. Average signal precision: $(\bar{\Sigma}_\eta^{-1})_{ii} = \bar{K}_i$, where $\bar{K}_i \equiv \int_j K_{ij}$

3. Precision of the information convey by prices about shock i: $(\Sigma_p^{-1})_{ii} = \frac{1}{\rho^2 \sigma_x}(\bar{\Sigma}_\eta^{-1})_{ii}^2 = \frac{\bar{K}_i^2}{\rho^2 \sigma_x} = \sigma_{ip}^{-1}$

4. Average posterior precision of shock i: $\bar{\sigma}_i^{-1} \equiv \sigma_i^{-1} + \bar{K}_i + \frac{\bar{K}_i^2}{\rho^2 \sigma_x}$. The average variance is therefore $(\bar{\Sigma})_{ii} = \frac{1}{\left(\sigma_i^{-1} + \bar{K}_i + \frac{\bar{K}_i^2}{\rho^2 \sigma_x}\right)} = \bar{\sigma}_i$, with derivatives:

$$\frac{\partial (\bar{\Sigma})_{ii}}{\partial \sigma_i} = \left(\frac{\bar{\sigma}_i}{\sigma_i}\right)^2 > 0 \tag{25}$$

$$\frac{\partial (\bar{\Sigma})_{ii}}{\partial \rho} = \frac{2}{\rho} \frac{\bar{\sigma}_i^2}{\sigma_{ip}} > 0 \tag{26}$$

Applying the chain rule yields

$$\frac{\partial (\bar{\Sigma}\bar{\Sigma})_{ii}}{\partial \rho} = \frac{4\bar{\sigma}_i}{\rho} \frac{\bar{\sigma}_i^2}{\sigma_{ip}} > 0. \tag{27}$$

5. Difference from average posterior beliefs: Recall that $\bar{\Sigma}_\eta^{-1} \equiv \int_j \Sigma_{\eta j}^{-1} dj$ is the average private signal precision and that $\bar{\Sigma}^{-1} \equiv \int_j \hat{\Sigma}_j^{-1} dj = \Sigma^{-1} + \Sigma_p^{-1} + \bar{\Sigma}_\eta^{-1}$ is the average posterior precision. Now define Δ as the difference between the precision of an informed investor's posterior beliefs and the average posterior precision. Since the $\Sigma^{-1} + \Sigma_p^{-1}$ terms are equal for all investors, this quantity is also equal to the difference between the precision of an informed investor's private signals and the average private signal precision:

$$\Delta \equiv \hat{\Sigma}_j^{-1} - \bar{\Sigma}^{-1} = \Sigma_{\eta j}^{-1} - \bar{\Sigma}_\eta^{-1}. \tag{28}$$

6. Ex-ante mean and variance of returns:

   Using lemma 1 and the coefficients given by (44), we can write the risk factor return as

$$\begin{aligned} \tilde{f} - \tilde{p}r &= (I - B)z - Cx - A \\ &= \bar{\Sigma}\left[\Sigma^{-1} z + \rho \left(I + \frac{1}{\rho^2 \sigma_x} \bar{\Sigma}_\eta^{-1\prime}\right) x\right] + \rho \bar{\Sigma} \bar{x} + \Gamma^{-1} \mu. \end{aligned}$$

   This expression is a constant plus a linear combination of two normal variables, which is also a normal variable. Therefore, we can write

$$\tilde{f} - \tilde{p}r = V^{1/2} u + w \tag{29}$$

where $u \sim N(0, I)$,

$$w \equiv \rho\bar{\Sigma}\bar{x} + \Gamma^{-1}\mu \tag{30}$$

and

$$\begin{aligned}
V &\equiv \bar{\Sigma}\left[\Sigma^{-1} + \rho^2\sigma_x\left(I + \frac{1}{\rho^2\sigma_x}\bar{\Sigma}_\eta^{-1'}\right)\left(I + \frac{1}{\rho^2\sigma_x}\bar{\Sigma}_\eta^{-1'}\right)'\right]\bar{\Sigma} \\
&= \bar{\Sigma}\left[\Sigma^{-1} + \rho^2\sigma_x\left(I + \frac{1}{\rho^2\sigma_x}(\bar{\Sigma}_\eta^{-1'} + \bar{\Sigma}_\eta^{-1}) + \frac{1}{\rho^4\sigma_x^2}\bar{\Sigma}_\eta^{-1'}\bar{\Sigma}_\eta^{-1}\right)\right]\bar{\Sigma} \\
&= \bar{\Sigma}\left[\Sigma^{-1} + \rho^2\sigma_x I + (\bar{\Sigma}_\eta^{-1'} + \bar{\Sigma}_\eta^{-1}) + \frac{1}{\rho^2\sigma_x}\bar{\Sigma}_\eta^{-1'}\bar{\Sigma}_\eta^{-1}\right]\bar{\Sigma} \\
&= \bar{\Sigma}\left[\rho^2\sigma_x I + \bar{\Sigma}_\eta^{-1'} + \Sigma^{-1} + \bar{\Sigma}_\eta^{-1} + \Sigma_p^{-1}\right]\bar{\Sigma} \\
&= \bar{\Sigma}\left[\rho^2\sigma_x I + \bar{\Sigma}_\eta^{-1'} + \bar{\Sigma}^{-1}\right]\bar{\Sigma}
\end{aligned} \tag{31}$$

The first line uses $E[xx'] = \sigma_x I$ and $E[zz'] = \Sigma$, the fourth line uses (45) and the fifth line uses $\bar{\Sigma}^{-1} = \Sigma^{-1} + \Sigma_p^{-1} + \bar{\Sigma}_\eta^{-1}$.

This variance matrix V is a diagonal matrix. It diagonal elements are

$$\begin{aligned}
(V)_{ii} &= (\bar{\Sigma}\left[\rho^2\sigma_x I + \bar{\Sigma}_\eta^{-1} + \bar{\Sigma}^{-1}\right]\bar{\Sigma})_{ii} \\
&= \bar{\sigma}_i[1 + (\rho^2\sigma_x + \bar{K}_i)\bar{\sigma}_i]
\end{aligned} \tag{32}$$

Diagonals of V have the following derivatives:

$$\frac{\partial V_{ii}}{\partial \sigma_i} = \left(\frac{\bar{\sigma}_i}{\sigma_i}\right)^2\left(1 + 2(\rho^2\sigma_x + \bar{K}_i)\bar{\sigma}_i\right) > 0 \tag{33}$$

$$\frac{\partial V_{ii}}{\partial \rho} = 2\rho\sigma_x\bar{\sigma}_i^2\left[1 + \frac{1}{\rho^2\sigma_{lp}}(1 + 2(\rho^2\sigma_x + \bar{K}_i)\bar{\sigma}_i)\right] > 0 \tag{34}$$

The elasticity of V_{ii} with respect ρ is given by:

$$\begin{aligned}
\frac{\partial V_{ii}}{\partial \rho}\frac{\rho}{V_{ii}} &= 2\rho\sigma_x\bar{\sigma}_i^2\left[1 + \frac{1}{\rho^2\sigma_{lp}}(1 + 2(\rho^2\sigma_x + \bar{K}_i)\bar{\sigma}_i)\right]\frac{\rho}{\bar{\sigma}_i[1 + (\rho^2\sigma_x + \bar{K}_i)\bar{\sigma}_i]} \\
&= \frac{2\rho^2\sigma_x\bar{\sigma}_i}{[1 + (\rho^2\sigma_x + \bar{K}_i)\bar{\sigma}_i]}\left[1 + \frac{1}{\rho^2\sigma_{lp}}(1 + 2(\rho^2\sigma_x + \bar{K}_i)\bar{\sigma}_i)\right]
\end{aligned}$$

The second term is always larger than one. We look for a sufficient condition that makes the first term larger than one too:

$$\begin{aligned}
2\rho^2\sigma_x\bar{\sigma}_i &> 1 + (\rho^2\sigma_x + \bar{K}_i)\bar{\sigma}_i \\
\rho^2\sigma_x &> \bar{\sigma}_i^{-1} + \bar{K}_i \\
\rho^2\sigma_x &> \sigma_i^{-1} + 2\bar{K}_i + \frac{\bar{K}_i^2}{\rho^2\sigma_x}
\end{aligned} \tag{35}$$

Since the LHS is increasing in σ_x and the RHS is decreasing in σ_x, if σ_x is sufficiently high, the elasticity of V_{ii} with respect to ρ becomes larger than one.

A.2 Solving the Model

Step 1: Portfolio Choices From the FOC, the optimal portfolio of risks factors chosen by investor j is

$$\tilde{q}_j = \frac{1}{\rho}\hat{\Sigma}_j^{-1}(E_j[\tilde{f}] - \tilde{p}r) \tag{36}$$

where $E_j[\tilde{f}]$ and $\hat{\Sigma}_j$ depend on the skill of the investor.

Next, we compute the risk factor portfolio of the average investor.

$$\begin{aligned}
\int_j \tilde{q}_j &= \frac{1}{\rho}\int_j \hat{\Sigma}_j^{-1}(E_j[\tilde{f}] - \tilde{p}r)dj \\
&= \frac{1}{\rho}\left(\int_j \hat{\Sigma}_j^{-1}(\Gamma^{-1}\mu + E_j[\tilde{f}])dj - \bar{\Sigma}^{-1}\tilde{p}r\right) \\
&= \frac{1}{\rho}\left(\int_j \Sigma_{\eta j}^{-1}\eta_j dj + \Sigma_p^{-1}\eta_p + \bar{\Sigma}^{-1}(\Gamma^{-1}\mu - \tilde{p}r)\right) \\
&= \frac{1}{\rho}\left(\bar{\Sigma}_\eta^{-1}z + \Sigma_p^{-1}\eta_p + \bar{\Sigma}^{-1}(\Gamma^{-1}\mu - \tilde{p}r)\right) \qquad (37)
\end{aligned}$$

where the fourth equality uses the fact that average noise of private signals is zero. Using the portfolio expressions (36) and (37), we compute the difference between the portfolio of investor j and the market portfolio:

$$\begin{aligned}
\tilde{q}_j - \int_j \tilde{q}_j &= \frac{1}{\rho}\left(\hat{\Sigma}_j^{-1}(E_j[\tilde{f}] - \tilde{p}r) - (\bar{\Sigma}_\eta^{-1} + \Sigma_p^{-1})z - \Sigma_p^{-1}\varepsilon_p - \bar{\Sigma}^{-1}(\Gamma^{-1}\mu - \tilde{p}r)\right) \\
&= \frac{1}{\rho}\left((\Sigma_{\eta j}^{-1}\eta_j + \Sigma_p^{-1}\eta_p) - \bar{\Sigma}_\eta^{-1}z - \Sigma_p^{-1}\eta_p + (\hat{\Sigma}_j^{-1} - \bar{\Sigma}^{-1})(\Gamma^{-1}\mu - \tilde{p}r)\right) \\
&= \frac{1}{\rho}\left((\Sigma_{\eta j}^{-1} - \bar{\Sigma}_\eta^{-1})z + \Sigma_{\eta j}^{-1}\varepsilon_j + (\hat{\Sigma}_j^{-1} - \bar{\Sigma}^{-1})(\Gamma^{-1}\mu - \tilde{p}r)\right) \\
&= \frac{1}{\rho}\left(\Delta(\tilde{f} - \tilde{p}r) + \Sigma_{\eta j}^{-1}\varepsilon_j\right) \qquad (38) \\
&= \frac{1}{\rho}\left[\Sigma_{\eta j}^{-1}\varepsilon_j + \Delta(V^{1/2}u + w)\right] \qquad (39)
\end{aligned}$$

where the third equality uses $\eta_j = z + \varepsilon_j$, the fourth equality uses (28) and the definition $\tilde{f} = \Gamma^{-1}\mu + z$ and the last line uses (29).

Step 2: Clearing the asset market and computing expected excess return Lemma 1 describes the solution to the market-clearing problem and derives the coefficients A, B and C in the pricing equation. The equilibrium price, along with the random signal realizations determines the time-2 expected return $(E_j[\tilde{f}] - \tilde{p}r)$. But at time 1, the equilibrium price and one's realized signals are not known. To compute period-1 utility, we need to know the time-1 expectation and variance of this time-2 expected return.

The time-2 expected excess return can be written as: $E_j[\tilde{f}] - \tilde{p}r = E_j[\tilde{f}] - \tilde{f} + \tilde{f} - \tilde{p}r$ and therefore its variance is

$$V_1[E_j[\tilde{f}] - \tilde{p}r] = V_1[E_j[\tilde{f}] - \tilde{f}] + V_1[\tilde{f} - \tilde{p}r] + 2Cov_1[E_j[\tilde{f}] - \tilde{f}, \tilde{f} - \tilde{p}r] \tag{40}$$

Combining (9) with the definitions $\eta_j = z + \varepsilon_j$ and $\eta_p = z + \varepsilon_p$, we can compute expectation errors:

$$\begin{aligned} E_j[\tilde{f}] - \tilde{f} &= \hat{\Sigma}_j \left[(\Sigma_{\eta j}^{-1} + \Sigma_p^{-1} - \hat{\Sigma}_j^{-1}) z + \Sigma_{\eta j}^{-1} \varepsilon_j + \Sigma_p^{-1} \varepsilon_p \right] \\ &= \hat{\Sigma}_j \left[-\Sigma^{-1} z + \Sigma_{\eta j}^{-1} \varepsilon_j + \Sigma_p^{-1} \varepsilon_p \right] \end{aligned}$$

Since this is a sum of mean-zero variables, its expectation is $E_1[E_j[\tilde{f}] - \tilde{f}] = 0$ and its variance is $V_1[E_j[\tilde{f}] - \tilde{f}] = \hat{\Sigma}_j \left[\Sigma^{-1} + \Sigma_{\eta j}^{-1} + \Sigma_p^{-1} \right] \hat{\Sigma}_j' = \hat{\Sigma}_j$.

From (29) we know that $V_1[\tilde{f} - \tilde{p}r] = V$. To compute the covariance term, we can use the definition $\tilde{f} = \Gamma^{-1}\mu + z$ and rearrange the definition of η_p to get $\tilde{p}r = B\eta_p + A$ and $\eta_p = z + \varepsilon_p$ to write

$$\begin{aligned} \tilde{f} - \tilde{p}r &= \Gamma^{-1}\mu + (I - B)z - A - B\varepsilon_p & (41) \\ &= \rho\bar{\Sigma}\bar{x} + \bar{\Sigma}\Sigma^{-1} z - (I - \bar{\Sigma}\Sigma^{-1})\varepsilon_p & (42) \end{aligned}$$

where the second line comes from substituting the coefficients A and B from lemma 1. Since the constant $\rho\bar{\Sigma}\bar{x}$ does not affect the covariance, we can write

$$\begin{aligned} Cov_1[E_j[\tilde{f}] - \tilde{f}, \tilde{f} - \tilde{p}r] &= Cov[-\hat{\Sigma}_j \Sigma^{-1} z + \hat{\Sigma}_j \Sigma_p^{-1} \varepsilon_p, \bar{\Sigma}\Sigma^{-1} z - (I - \bar{\Sigma}\Sigma^{-1})\varepsilon_p] \\ &= -\hat{\Sigma}_j \Sigma^{-1} \Sigma \bar{\Sigma} \Sigma^{-1} - \hat{\Sigma}_j \Sigma_p^{-1} \Sigma_p (I - \bar{\Sigma}\Sigma^{-1})] \\ &= -\hat{\Sigma}_j \bar{\Sigma}\Sigma^{-1} - \hat{\Sigma}_j (I - \bar{\Sigma}\Sigma^{-1}) \\ &= -\hat{\Sigma}_j \end{aligned}$$

Substituting the three variance and covariance terms into (40), we find that the variance of excess return is $V_1[E_j[\tilde{f}] - \tilde{p}r] = \hat{\Sigma}_j + V - 2\hat{\Sigma}_j = V - \hat{\Sigma}_j$. Note that this is a diagonal matrix. Substituting the expressions (32) and (24) for the diagonal elements of V and $\hat{\Sigma}_j$ we have

$$(V - \hat{\Sigma}_j)_{ii} = (\bar{\sigma}_i - \hat{\sigma}_i) + (\rho^2 \sigma_x + \bar{K}_i)\bar{\sigma}_i^2$$

A sufficient condition for the expression to hold for informed investors is $K_{ij} > \bar{K}_i$ (implying that $\hat{\sigma}_i^{-1} > \bar{\sigma}_i^{-1}$, this is, they most allocate more capacity to each shock than the average agent. Given the symmetry across investors, they will all learn about the same shocks and thus this condition holds. Therefore, the variance matrix is positive definite. In summary, the excess return is normally distributed as $E_j[\tilde{f}] - \tilde{p}r \sim \mathcal{N}(w, V - \hat{\Sigma}_j)$.

Step 3: Compute ex-ante expected utility Ex-ante expected utility for investor j is $U_{1j} = E_1\left[\rho E_j[W_j] - \frac{\rho^2}{2} V_j[W_j]\right]$. In period 2, the investor has chosen his portfolio and the price is in his information set, therefore the only random variable is z. We substitute the budget constraint, the optimal portfolio choice from (36) and take expectation and variance conditioning on $E_j[\tilde{f}]$ and $\hat{\Sigma}_j$ to obtain $U_{1j} = \rho r W_0 + \frac{1}{2} E_1[(E_j[\tilde{f}] - \tilde{p}r)' \hat{\Sigma}_j (E_j[\tilde{f}] - \tilde{p}r)]$.

Define $m \equiv \hat{\Sigma}_j^{-1/2}(E_j[\tilde{f}] - \tilde{p}r)$ and note that $m \sim \mathcal{N}(\hat{\Sigma}_j^{-1/2} w, \hat{\Sigma}_j^{-1} V - I)$. The second term in the U_{ij} is equal to $E[m'm]$, which is the mean of a non-central Chi-square. Using the formula, if $m \sim \mathcal{N}(E[m], Var[m])$, then $E[m'm] = tr(Var[m]) + E[m]'E[m]$, we get

$$U_{1j} = \rho r W_0 + \frac{1}{2} tr(\hat{\Sigma}_j^{-1} V - I) + \frac{1}{2} w' \hat{\Sigma}_j^{-1} w.$$

Finally we substitute the expressions for $\hat{\Sigma}_j^{-1}$ and w from (24) and (30):

$$\begin{aligned} U_{1j} &= \rho r W_0 - \frac{N}{2} + \frac{1}{2} \sum_{i=1}^{N} \left(\sigma_i^{-1} + K_{ij} + \frac{\bar{K}_i^2}{\rho^2 \sigma_x} \right) V_{ii} + \frac{\rho^2}{2} \sum_{i=1}^{N} \bar{x}_i^2 \bar{\sigma}_i^2 \left(\sigma_i^{-1} + K_{ij} + \frac{\bar{K}_i^2}{\rho^2 \sigma_x} \right) \\ &= \frac{1}{2} \sum_{i=1}^{N} K_{ij} [V_{ii} + \rho^2 \bar{x}_i^2 \bar{\sigma}_i^2] + \rho r W_0 - \frac{N}{2} + \frac{1}{2} \sum_{i=1}^{N} \left(\sigma_i^{-1} + \frac{\bar{K}_i^2}{\rho^2 \sigma_x} \right) [V_{ii} + \rho^2 \bar{x}_i^2 \bar{\sigma}_i^2] \\ &= \frac{1}{2} \sum_{i=1}^{N} K_{ij} \lambda_i + constant \end{aligned} \tag{43}$$

where the weights $\lambda_i = \bar{\sigma}_i [1 + (\rho^2 \sigma_x + \bar{K}_i) \bar{\sigma}_i] + \rho^2 \bar{x}_i^2 \bar{\sigma}_i^2$ are given by the variance of expected excess return V_{ii} from 32 plus a term that depends on the supply of the risk.

Step 4: Information choices The attention allocation problem maximizes ex-ante utility in 43 subject to the information capacity and no-forgetting constraints:

$$\begin{aligned} \max_{\{K_{ij}\}_{i=1}^N} \quad & \frac{1}{2} \sum_{i=1}^{N} K_{ij} \lambda_i + \text{constant} \\ \text{subject to} \quad & \sum_{i=1}^{N} K_{ij} \leq K \\ & K_{ij} \geq 0 \quad \forall i \end{aligned}$$

Observe that λ_i depends only on parameters and on aggregate average precisions. Since the investor has zero mass in a continuum of investors, he takes λ_i as given. Since the constant is irrelevant, the optimal choice maximizes a weighted sum of attention allocations, where the weights are given by λ_i (equation 14), subject to a constraint on an un-weighted sum. This is not a concave objective, so a first-order approach will not deliver a solution. A simple variational argument reveals that allocating all capacity to the risk(s) with the highest λ_i achieves the maximum utility. For a formal proof of this result, see Van Nieuwerburgh and Veldkamp (2010). Thus the solution is given by: $K_{ij} = K$ if $\lambda_i = \max_k \lambda_k$ and $K_{ij} = 0$ otherwise. There may be multiple risks i that achieve the same maximum value of λ_i. In that case, the manager is indifferent about how to allocate attention between those risks.

A.3 Proofs

Proof of Lemma 1

Proof. Following Admati (1985), we know that the equilibrium price takes the following form $\tilde{p}r = A + Bz +$

Cx where

$$\begin{aligned} A &= \Gamma^{-1}\mu - \rho\bar{\Sigma}\bar{x} \\ B &= I - \bar{\Sigma}\Sigma^{-1} \\ C &= -\rho\bar{\Sigma}\left(I + \frac{1}{\rho^2\sigma_x}\bar{\Sigma}_\eta^{-1'}\right) \end{aligned} \tag{44}$$

and therefore the price is given by

$$\tilde{pr} = \bar{\Sigma}\left[(\bar{\Sigma}^{-1} - \Sigma^{-1})z - \rho(\bar{x} + x) - \frac{1}{\rho\sigma_x}\bar{\Sigma}_\eta^{-1'}\right]$$

Furthermore, the precision of the public signal is

$$\Sigma_p^{-1} \equiv \left(\sigma_x B^{-1} C C' B^{-1'}\right)^{-1} = \frac{1}{\rho^2\sigma_x}\bar{\Sigma}_\eta^{-1'}\bar{\Sigma}_\eta^{-1} \tag{45}$$

□

Proof of Proposition 1 *For a given investor j, the optimal choice of attention allocation to risk i is weakly increasing in its variance σ_i.*

Proof. From step 4 of the model solution, we know that the optimal information choice is $K_{ij} = K$ if $\lambda_i = \max_k \lambda_k$ and $K_{ij} = 0$ otherwise. It remains to be shown that the weight λ_i is increasing in σ_i:

$$\frac{\partial\lambda_i}{\partial\sigma_i} = [1 + 2\bar{\sigma}_i(\rho^2(\sigma_x + \bar{x}_i^2) + \bar{K}_i)]\left(\frac{\bar{\sigma}_i}{\sigma_i}\right)^2 > 0$$

and $\frac{\partial\lambda_k}{\partial\sigma_i} = 0$. Therefore, an increase in the variance of risk i only increases the weight associated with K_{ij}, and either shifts the attention towards that risk (if it was not already allocated) or it remains at its previous level.

□

Proof of Proposition 2 *For a given investor j, an increase in risk aversion ρ weakly increases the attention allocated to risk i if its supply $\bar{x}_i$ is sufficiently high.*

Proof. From step 4 of the model solution, we know that all capacity is allocated to the risk i with the highest weight λ_i. The derivative of the weight with respect to risk aversion is given by:

$$\frac{\partial\lambda_i}{\partial\rho} = [1 + 2\bar{\sigma}_i(\rho^2(\sigma_x + \bar{x}_i^2) + \bar{K}_i)]\frac{2\bar{\sigma}_i^2}{\rho\sigma_{lp}} + 2\rho\bar{\sigma}_i^2(\sigma_x + \bar{x}_i^2)$$

In this case, an increase in risk aversion will increase the weight for all risks. However, the magnitude of the increase in the weight will depend positively on the supply of such risk:

$$\frac{\partial^2\lambda_i}{\partial\bar{x}_i\partial\rho} = 4\rho\bar{\sigma}_i^2\bar{x}_i\left(1 + \frac{2\bar{\sigma}_i}{\sigma_{lp}}\right) > 0$$

In conclusion, an increase in risk aversion shifts attention towards the risk with highest supply (if it was not already allocated) or it remains at its previous level.

□

Derivation of excess returns and their dispersion We begin by calculating the portfolio excess return. Note that the return of the portfolio expressed in terms of assets is equal to the return expressed in risk factors:

$$\begin{aligned} (q_j - \bar{q})'(f - pr) &= (q_j - \bar{q})'\Gamma^{-1}(\Gamma f - \Gamma pr) &(46)\\ &= (\tilde{q}_j - \int_l \tilde{q}_l)'(\tilde{f} - \tilde{p}r) &(47) \end{aligned}$$

Substitute (29) and (39) into (47) to get

$$\begin{aligned} E[(\tilde{q}_j - \int_l \tilde{q}_l)'(\tilde{f} - \tilde{p}r)] &= \frac{1}{\rho}E\left[\left(\Sigma_{\eta j}^{-1}\epsilon_j + \Delta(V^{1/2}u + w)\right)'(V^{1/2}u + w)\right] \\ &= \frac{1}{\rho}E\Big[\epsilon_j'\Sigma_{\eta j}^{-1}w + \epsilon_j'\Sigma_{\eta j}^{-1}V^{1/2}u + 2w'\Delta V^{1/2}u + w'\Delta w + u'V^{1/2}\Delta V^{1/2}u\Big] \\ &= \frac{1}{\rho}E\Big[w'\Delta w + u'V^{1/2}\Delta V^{1/2}u\Big] \\ &= \frac{1}{\rho}\Big[\rho^2\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x} + Tr\Big(V^{1/2}\Delta V^{1/2}E(uu')\Big)\Big] \\ &= \frac{1}{\rho}\Big[\rho^2\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x} + Tr(\Delta V)\Big] \\ &= \rho Tr(\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x}) + \frac{1}{\rho}Tr(\Delta V) &(48) \end{aligned}$$

where the third equality comes from the fact that ϵ_j and u are mean zero and uncorrelated.

To get return dispersion, we substitute (29) and (39) into (47), then square the excess return and take the expectation:

$$E[((q_j - \bar{q})'(\tilde{f} - \tilde{p}r))^2] = E\left[\left(\frac{1}{\rho}[\Sigma_\eta^{-1}e_j + \Delta V^{1/2}u + \Delta w]'(w + V^{1/2}u)\right)^2\right]$$

Using the fact that for any random variable x we have that $V(x) = E(x^2) - E^2(x)$, the dispersion of funds' portfolio returns is equal to:

$$\begin{aligned} E[((q_j - \bar{q})'(\tilde{f} - \tilde{p}r))^2] &= \frac{1}{\rho^2}V\Big([\Sigma_{\eta j}^{-1}e_j + \Delta V^{1/2}u + \Delta w]'(V^{1/2}u + w)\Big) \\ &\quad + \frac{1}{\rho^2}\Big(E[\Sigma_{\eta j}^{-1}e_j + \Delta V^{1/2}u + \Delta w]'(V^{1/2}u + w)\Big)^2 \end{aligned}$$

We compute each term separately.

$$\begin{aligned}
V(\cdot) &= V\left[e_j'\Sigma_{\eta j}^{-1}w + e_j'\Sigma_{\eta j}^{-1}V^{1/2}u + 2w'\Delta V^{1/2}u + w'\Delta w + u'V^{1/2}\Delta V^{1/2}u\right] \\
&= w'\Sigma_{\eta j}^{-1}w + 0 + 4w'\Delta V\Delta w + 0 + 2Tr(\Delta V\Delta V) \\
&= \rho^2 Tr(\bar{x}'\bar{\Sigma}\Sigma_{\eta j}^{-1}\bar{\Sigma}\bar{x}) + 4\rho^2 Tr(\bar{x}'\bar{\Sigma}\Delta V\Delta\bar{\Sigma}\bar{x}) + 2Tr(\Delta V\Delta V)
\end{aligned}$$

and

$$\begin{aligned}
E(\cdot)^2 &= (w'\Delta w)^2 + Tr^2(\Delta V) + 2w'\Delta w Tr(\Delta V) \\
&= \rho^4 Tr^2(\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x}) + Tr^2(\Delta V) + 2\rho^2 Tr(\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x})Tr(\Delta V)
\end{aligned}$$

Substituting back we have:

$$\begin{aligned}
E[((q_j-\bar{q})'(\tilde{f}-\tilde{p}r))^2] &= Tr(\bar{x}'\bar{\Sigma}\Sigma_{\eta j}^{-1}\bar{\Sigma}\bar{x}) + 4Tr(\bar{x}'\bar{\Sigma}\Delta V\Delta\bar{\Sigma}\bar{x}) + 2Tr(\bar{x}'\bar{\Sigma}'\Delta\bar{\Sigma}\bar{x})Tr(\Delta V) \\
&+ \rho^2 Tr^2(\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x}) + \frac{1}{\rho^2}Tr^2(\Delta V) + \frac{2}{\rho^2}Tr(\Delta V\Delta V) \\
&= 6\bar{x}_l{}^2\bar{\sigma}_l^2(K_{lj}-\bar{K}_l)^2 V_{ii} + 3(K_{lj}-\bar{K}_l)^2\frac{V^2}{\rho^2} + \bar{x}_l^2\bar{\sigma}_l^2 K_{lj} + \rho^2\bar{x}_l^4\bar{\sigma}_l^4(K_{lj}-\bar{K}_l)^2 \\
&= \sum_{l=1}^{n}(K_{lj}-\bar{K}_l)^2\bar{x}_l{}^2\bar{\sigma}_l^2\left[6V_{ll} + \rho^2\bar{x}_l^2\bar{\sigma}_l^2 + 3\frac{V_{ll}^2}{\rho^2\bar{x}_l{}^2\bar{\sigma}_l^2}\right] + \bar{x}_l{}^2\bar{\sigma}_l^2 K_{lj} \qquad (49)
\end{aligned}$$

Proof of Proposition 3 *For given precisions of an investor j, an increase in variance σ_i: a) increases the dispersion of fund portfolios $E[(q_j-\bar{q})'(q_j-\bar{q})]$ and b) increases the dispersion of portfolio excess returns $E[((q_j-\bar{q})'(\tilde{f}-\tilde{p}r))^2]$.*

Proof. Part a) Our measure of portfolio dispersion is $E[(q_j-\bar{q})'(q_j-\bar{q})]$. Transforming asset quantities into risk factor quantities yield the equivalent expression $E[(\tilde{q}_j-\int_l\tilde{q}_l)'\Gamma^{-1\prime}\Gamma^{-1}(\tilde{q}_j-\int_l\tilde{q}_l)]$.

Using (38) to substitute out $(\tilde{q}_j-\int_l\tilde{q}_l)$, we find that

$$\begin{aligned}
E[(q_j-\bar{q})'(q_j-\bar{q})] &= \frac{1}{\rho^2}\sum_i\left(\Delta_{ii}E[(\tilde{f}_i-\tilde{p}_i r)^2] + \Sigma_{\eta j}^{-2}(i,i)E[\epsilon_{ij}^2]\right)(\Gamma^{-1}(i,i))^2 \qquad (50) \\
&= \frac{1}{\rho^2}\sum_i\left(\Delta_{ii}(w_i^2+V_{ii}) + \Sigma_{\eta j}^{-2}(i,i)K_{ij}^{-1}\right)(\Gamma^{-1}(i,i))^2 \qquad (51)
\end{aligned}$$

Note that the matrices are all diagonal and the cross terms drop out because the market return $(\tilde{f}_i-\tilde{p}_i r)$ is independent of agent j's signal noise ϵ_j.

The only components that are affected by the increase in σ_i are V and w. (33) shows that $\partial V_{ii}/\sigma_i > 0$. From (30), we see that, $w_i = \rho\bar{\Sigma}_{ii}\bar{x}_i + \Gamma^{-1}(i,:)\mu$. Prior variance σ_i enters w_i only through the average posterior variance $\bar{\Sigma}$. (25) establishes that $\partial\bar{\Sigma}_{ii}/\partial\sigma_i > 0$. Since both derivatives are positive and all the other terms in (51) are positive, we conclude that portfolio dispersion increases when the variance of any risk does.

Part b) Recall the dispersion of portfolio excess returns given by (49). The only variables that depend on σ_i are $\bar{\sigma}_i$ and V_{ii}, both of which are increasing in σ_i. Since all other terms in the expression are positive, dispersion of portfolio excess returns is increasing in σ_i. □

Proof of Proposition 4 *If σ_x is sufficiently large, then an increase in risk aversion ρ increases the dispersion of portfolio excess returns $E[((q_j - \bar{q})'(\tilde{f} - \tilde{p}r))^2]$.*

Proof. Dispersion of excess returns given in (49):

$$E[((q_j - \bar{q})'(\tilde{f} - \tilde{p}r))^2] = \sum_{l=1}^{n}(K_{lj} - \bar{K}_l)^2 \left\{ \bar{x_l}^2 \bar{\sigma}_l^2 \left[6V_{ll} + \rho^2 \bar{x}_l^2 \bar{\sigma}_l^2 + \frac{K_l}{(K_{lj} - \bar{K}_l)^2} \right] + \underbrace{3\frac{V_{ll}^2}{\rho^2}}_{sign} \right\}$$

The only terms affected by ρ are $\bar{\sigma}_l$ and V_{ll}, and both are increasing in ρ as shown in (26) and (34). Therefore the only term of the derivative that we need to sign corresponds to the last one, whose derivative is:

$$3\sum_{l=1}^{n}(K_{lj} - \bar{K}_l)^2 \left[\frac{2V_{ll}}{\rho^2}\frac{\partial V_{ll}}{\partial \rho} - \frac{2}{\rho^3}V_{ll}^2 \right] = \frac{6}{\rho^2}\sum_{l=1}^{n}(K_{lj} - \bar{K}_l)^2 V_{ll} \left[\frac{\partial V_{ll}}{\partial \rho} - \frac{V_{ll}}{\rho} \right]$$

This expression is positive if the elasticity of V_{ll} with respect to ρ is larger than one for all l, which is ensured if σ_x is sufficiently large, i.e. satisfies (35). □

Proof of Propositions 5 and 6 *An increase in the variance of any shock σ_i increases the portfolio excess return of an informed fund, $E[(q_j - \bar{q})'(\tilde{f} - \tilde{p}r)]$. Furthermore, if σ_x is sufficiently large, an increase in risk aversion ρ also increases this excess return.*

Proof. From (48) we have that expected excess returns are given by

$$E[(\tilde{q}_j - \int_l \tilde{q}_l)'(\tilde{f} - \tilde{p}r)] = \rho Tr(\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x}) + \frac{1}{\rho}Tr(\Delta V)$$

Taking a derivative with respect to σ_i, recalling that Δ does not depend on it:

$$\frac{\partial E[(\tilde{q}_j - \int_l \tilde{q}_l)'(\tilde{f} - \tilde{p}r)]}{\partial \sigma_i} = 2\rho Tr\Big(\bar{x}'\Delta \left[\frac{\partial \bar{\Sigma}}{\partial \sigma_i}\right] \bar{\Sigma}\bar{x}\Big) + \frac{1}{\rho}Tr\Big(\Delta \left[\frac{\partial V}{\partial \sigma_i}\right]\Big) > 0$$

This expression is positive since both $\bar{\Sigma}$ and V are increasing in σ_i as we know from (25) and (33).

Taking a derivative with respect to ρ, recalling that Δ does not depend on it:

$$\begin{aligned}
\frac{\partial E[(\tilde{q}_j - \int_l \tilde{q}_l)'(\tilde{f} - \tilde{p}r)]}{\partial \rho} &= Tr(\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x}) + 2\rho Tr\Big(\bar{x}'\Delta \left[\frac{\partial \bar{\Sigma}}{\partial \rho}\right] \bar{\Sigma}\bar{x}\Big) - \frac{1}{\rho^2}Tr(\Delta V) + \frac{1}{\rho}Tr\Big(\Delta \left[\frac{\partial V}{\partial \rho}\right]\Big) \\
&= Tr(\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x}) + 2\rho Tr\Big(\bar{x}'\Delta \left[\frac{\partial \bar{\Sigma}}{\partial \rho}\right] \bar{\Sigma}\bar{x}\Big) + \frac{1}{\rho}\left[Tr\Big(\Delta \left[\frac{\partial V}{\partial \rho} - \frac{V}{\rho}\right]\Big)\right]
\end{aligned}$$

A sufficient condition for this expression to be positive is $\frac{\partial V}{\partial \rho} - \frac{V}{\rho} > 0$, which is equivalent to the elasticity of V_{ii} with respect to ρ larger than one for each i. This holds if σ_x is sufficiently large, i.e. satisfies (35). □

Supplementary Appendix for "Rational Attention Allocation over the Business Cycle": Model Simulation
Not for publication

In this supplementary appendix, we do two things. In section 1, we use a numerical example to illustrate the model's predictions for the same measures of attention, portfolio dispersion, and performance as the ones we measure in the data. The goal of this exercise is to confirm that the model makes the same qualitative predictions for these observables as for the slightly different measures of attention allocation, portfolio dispersion, and fund performance for which we formally proved our propositions. Notably, we do not attempt to quantitatively account for all time-series and cross-sectional moments of actively managed fund portfolio holdings and returns. Such a task would be beyond the scope of this paper and indeed beyond the current state of the literature. Our model is too stylized along many dimensions to deliver on such a task. For example, it has only three assets and no heterogeneity in risk aversion, prior beliefs, or initial wealth among funds, and no heterogeneity in information capacity among skilled managers. Adding such features could improve the predictions, but only at the cost of obscuring the main mechanisms operating in the model.

The second section analyzes a model where the agent has expected exponential utility. It re-derives the main results of the paper analytically for that model.

S.1 Numerical Analysis

S.1.1 Parameter Choices

The following explains how we choose the parameters of our model. The simplicity of the model prevents a full calibration. Instead, we pursue a numerical example that matches some salient properties of stock return data. Our benchmark parameter choices are listed in Table S.1. Section S.1.3 below shows that the qualitative results are robust to a wide range of parameter choices.

Our procedure is to simulate 3000 draws of the shocks $(x_1, x_2, x_c, s_1, s_2, a)$ in recessions and 3000 draws of the shocks in expansions. Since our model is static, each simulation is best interpreted as different draws of a random variable, and not as a period (months). The model's recessions differ from expansions in *two respects*.

First, the variance of the aggregate payoff shock σ_a is higher. It is set to replicate the fact the market return volatility is about 25% higher in recessions than in expansions. In the numerical example, the volatility of the market return is 4.0% in expansions and 5.0% in recessions, straddling the observed market return volatility of 4.5%. Setting the variance of the asset supply vector $\sigma_x = .05^2\bar{x}$ allows us to match this level of market return volatility.

Second, recessions are also characterized by lower *realized* stock market returns (despite high expected returns). In order to generate lower realized market returns and higher expected returns in a static model, we have to assume that agents are surprised by unexpectedly low returns in recessions. We accomplish this in the numerical example by replicating the bottom $m = 2.5\%$ of market return realizations among the 3000 simulations of the model in recessions, in effect simulating the economy in recessions for $3000 * (1 + .025) > 3000$ draws. This choice for m is conservative because the 0.03% difference between market returns in expansions (0.87% per month) and recessions (0.84% per month) it generates is lower than the 0.20% per month difference in the data. In the robustness section below, we consider a case that generates a 0.20% return difference. The results are qualitatively and quantitatively similar.

To get the average market return right, we choose the mean of asset payoffs μ (equal for all assets) and the coefficient of absolute risk aversion, ρ, to achieve an average equilibrium market return of about 0.85% per month.

Table S.1: **Numerical Example**

The first column lists the parameter in question, the second column is its symbol, the third column lists its numerical value, and the last column briefly summarizes how we chose that value.

Parameter	*Symbol*	*Value*	*How Chosen?*
CARA	ρ	0.525	*Asset return mean*
mean of payoffs 1,2,c	μ_1, μ_2, μ_c	10, 10, 10	*Asset return mean*
variance aggr. payoff comp. a	σ_a	0.1225 (E), 0.2625 (R)	*Market return vol in expansions vs. recessions*
variance idio. payoff comp. s_i	σ_i	0.25	*Asset return vol vs. market return vol*
a-sensitivity of payoffs	b_1, b_2	0.25,0.50	*Asset beta level + dispersion*
mean asset supply 1,2	$\bar{x}_1 = \bar{x}_2$	1,1	*Normalization*
mean asset supply c	$\bar{x}_c$	7	*Asset return volatility*
variance asset supply	σ_x	$(.05 * \bar{x})^2$	*Asset return idio vol*
risk-free rate	r	0.0022	*Average T-bill return*
initial wealth	W_0	90	*Average cash position*
difficulty learning aggr. info	ψ	1	*Simplicity*
information capacity	K	1	
skilled fraction	χ	0.20	

We consider 3 assets ($n = 3$). We think of assets 1 and 2 as two large industries and the composite asset as summarizing all other industries. Therefore, we normalize the mean asset supply of assets 1 and 2 to 1, and set the supply of the composite asset, $\bar{x}_c$, to 7. The variance of the firm-specific shocks is chosen to match the fact that individual industry returns are about 30% more volatile than the market return over our sample from 1980 to 2005. We use data from the 30 industry portfolios of Fama and French (1997). In the example, the average volatility of assets 1 and 2 is 6.5% in recessions and 5.8% in expansions, 29% and 45% higher than that of the market return. This choice matches the proportion of the average industry's return variance that is idiosyncratic. We choose the asset loadings on the aggregate payoff shock, b_1 and b_2, to be different from each other so as to generate some spread in asset betas. The chosen values generate average market betas of 0.9 and a dispersion in betas of 33%. This is reasonably close to the average beta of 0.95 and the dispersion of 23% for the 30-industry portfolios.

We set the average risk-free rate equal to 0.22% per month, the average of the 1-month yield minus inflation in our sample. We set initial wealth, W_0, to generate average holdings in the risk-free asset around 0%.

For simplicity, we set capacity K for skilled investment managers equal to 1. This implies that learning can increase the precision of one of the idiosyncratic shocks (or the aggregate shock) by 25% (by 18%). We will vary K in our robustness exercise below. Likewise, we have no strong prior on the fraction of skilled funds, χ. In our benchmark, we set it equal to 20%, and we will vary it for robustness. The model is simulated for 800 investors, of which 175 are skilled (20%). We assume that 20% of all investors are non-investment managers ("other investors"). The unskilled managers (60% of the populations) and other investors differ in name only. We note that the parameter conditions in Propositions 2 through 4 are satisfied by these parameter choices.

As in our empirical work on mutual funds in Section 3, we compute all statistics of interest as equally-weighted averages across all investment managers (i.e., without the 20% other investors). We also report results separately for skilled and unskilled investment managers.

S.1.2 Main Simulation Results

Every skilled manager ($K > 0$) solves for the choice of signal precisions $K_{aj} \geq 0$ and $K_{1j} \geq 0$ that maximize time-1 expected utility (11). We assume that these choice variables lie on a 25×25 grid in $\mathbb{R}^2_+$. The signal precision choice $K_{2j} \geq 0$ is implied by the capacity constraint (7).

We simulate a sequence of $T = 3000$ draws (months) for the random variables in each of the recession and expansion states, as explained above. We form a $T \times 1$ time series for the three individual asset returns, for the market return, for each fund's return, and for each fund's (and the market's) portfolio weights in each asset. For each asset i, we then estimate a CAPM regression of the asset's excess return on the market excess return. This delivers the asset's CAPM beta, β_i; one value in expansions and one in recessions. We define the systematic component of returns as $\beta_i R^m_t$, for $t = 1, \cdots, T$ and $i = 1, 2, 3$. Stacking the different is and ts results in a $3T \times 1$ vector of systematic returns. Similarly, we define the idiosyncratic return as $R^i_t - \beta_i R^m_t$.

To compute *Ftiming* in equation (16) for fund j, we stack its portfolio weights in deviation from the market's weights for the three assets and the T draws into a $3T \times 1$ vector. We also create a $3T \times 1$ vector of aggregate shocks by stacking three identical repetitions of each aggregate shock realization a. We calculate *Ftiming* as the covariance between these two variables. Likewise, we form *Timing* as in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2012) as the covariance between the time series of portfolio weights, in deviation from the market's weights, and the systematic component of returns. The procedure delivers one *Ftiming* and one *Timing* measure per fund in recessions and one set of measures in expansions. We multiply *Ftiming* by 1000 and *Timing* by 10,000 because the aggregate shocks are an order of magnitude larger than the systematic returns.

Table S.2 summarizes the predictions of the model for the main statistics of interest. The left panel shows the results for recessions, while the right panel shows the results for expansions. In each panel, we present three columns. Column *skilled* reports the equally weighted average of the statistic in question for the group of skilled investors (20% of investors have $K > 0$ in our benchmark parametrization). Column *unskilled* is the equally weighted average across the unskilled funds (60% of investors are unskilled investment managers). Column *all* is the equally weighted average across all funds (80% of investors). The 20% unskilled other investors are excluded from the table because we do not observe them in the data. However, the model's predictions for this group are identical to those for the unskilled funds. These two groups differ in name only.

Rows 1 and 2 of Table S.2 show that *Ftiming* and *Timing* are higher for skilled investors in recessions (left panel) than in expansions (right panel). Because of market clearing, unskilled investors are the flip side of the skilled ones, their *Ftiming* and *Timing* measures are negative. Since no investors learn about the aggregate shock in expansions, $Ftiming$ and $Timing$ are essentially zero for both skilled and unskilled. The net effect of the skilled and unskilled is listed in Column *all*. This combination of all investment managers, skilled and unskilled, is what we have data on. Hence, the first testable implication of the model is that $Ftiming$ and $Timing$ should be higher for all funds in recessions than in expansions.

In a similar fashion, we construct $Fpicking$ measure, defined in equation (17), and the stock-picking measure $Picking$, defined in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2012). That is, we stack the stock-specific shocks, s_i, the idiosyncratic returns, $R_t^i - \beta_i R_t^m$, and the fund's portfolio weights into $3T \times 1$ vectors and compute the respective covariances. Rows 3 and 4 summarize the predictions of the model for $Fpicking$ (multiplied by 1000) and $Picking$ (multiplied by 10,000). Across all funds (skilled and unskilled), the model predicts lower $Fpicking$ and $Picking$ in recessions. Skilled funds have a high $Fpicking$ and $Picking$ ability in expansions, when they allocate their attention to stock-specific information. Unskilled investors exhibit a negative $Picking$ in expansions for the same reason that they have a negative $Timing$ in recessions: Price fluctuations induce them to buy when returns are low and sell when returns are high. The $Fpicking$ and $Picking$ measures are close to zero for all investors in recessions. Hence, the second testable implication of the model is that $Fpicking$ and $Picking$ should be lower for all funds in recessions than in expansions.

Table S.2: **Benchmark Simulation Results from the Model**

This table provides the main statistics for a simulation of the model under the benchmark parameter values summarized in Table S.1. Panel A reports moments related to attention allocation, Panel B reports the moments related to portfolio dispersion, and Panel C reports moments related to performance. The first column lists the moments in question, as defined in the main text, the next three columns report the predictions for the model simulated in a recession, the last three columns report the results for the model simulated in an expansion. All moments are generated from a simulation of 3,000 draws and 800 investors. For both recessions and expansions, we list the equally-weighted average across *all* investment managers (the 20% skilled and the 60% investment managers), and separately for the *skilled* and the *unskilled* investment managers.

	Recessions			**Expansions**		
	All managers	Skilled	Unskilled	All managers	Skilled	Unskilled
Panel A: Attention Allocation						
1. Ftiming	10.55	162.54	-40.11	0.07	0.94	-0.22
2. Timing	9.91	155.78	-38.72	0.06	3.31	-1.02
3. Fpicking	2.15	33.46	-8.28	15.61	249.72	-62.42
4. Picking	1.66	25.84	-6.40	10.02	160.11	-40.01
Panel B: Dispersion						
5. Concentration	3.75	13.15	0.00	3.12	11.48	0.00
6. Idiosyncratic volatility	5.09	15.21	1.72	4.33	13.50	1.28
7. Dispersion in abnormal return	3.54	10.05	1.36	3.37	9.82	1.22
8. Dispersion in CAPM alpha	2.52	5.05	1.68	2.28	4.55	1.52
9. Dispersion in CAPM beta	6.37	13.20	4.09	1.46	4.30	0.51
Panel C: Performance						
10. Abnormal return	0.346	5.471	-1.363	0.302	4.867	-1.220
11. CAPM Alpha	0.353	5.401	-1.330	0.307	4.861	-1.211

Next, we turn to the measures of portfolio and return dispersion. Row 5 of Table S.2 shows the results for the $Concentration$ measure, defined in equation (18), in our numerical example. We calculate

Concentration^j for fund j by stacking all squared deviations of fund j's portfolio from the market portfolio into a $3T \times 1$ vector, and by summing over its entries, and dividing by T. We obtain one number for recessions and one for expansions. We find that *Concentration* is higher for all funds in recessions than in expansions. This increase is driven entirely by the informed; the uninformed are all holding the exact same portfolio because of common prior beliefs.

More concentrated portfolios are also less diversified. For each fund j, we estimate CAPM regression (19) by regressing the fund's excess return on the market's excess return. This delivers the fund's α^j, β^j, and σ_ε^j. We use the idiosyncratic risk σ_ε^j as our second measure of portfolio dispersion. If all funds held the market portfolio, their idiosyncratic risk would be zero, and there would be zero cross-sectional dispersion. In simulation, the skilled funds take on more idiosyncratic risk than the unskilled ones, and more in recessions than in expansions. As a result, idiosyncratic risk is higher in recessions than in expansions for all funds.

Rows 7 through 9 report the results for the dispersion across funds' abnormal returns, CAPM alphas, and CAPM betas. All three metrics show increasing dispersion in recessions, driven largely by the heterogeneity in the choices of the skilled investors.

Finally, we study performance measures. Rows 10 and 11 of Table S.2 show that skilled investment managers have large excess returns, as measured by abnormal fund returns or fund alphas ($R^j - R^m$ and α^j), at the expense of the uninformed. The average investment manager has a slightly higher alpha in recessions than in expansions. While quantitatively modest (4.6bp per month or 55bp per year), the positive difference in average alphas between recessions and expansions is a robust finding of the model.

The numerical results also reveal that the regression residual variance $(\sigma_\varepsilon^j)^2$ is higher in recessions. This effect arises because a fund that gets different signal draws (information) in each period holds a portfolio with a beta that varies over time. The CAPM equation (19) estimates an unconditional beta instead. The difference between the true, conditional beta and the estimated, constant beta shows up in the regression residual. Since recessions are times when funds learn more new information each period about the aggregate shock, these are times when true fund betas fluctuate more and the regression residuals are more volatile.

S.1.3 Robustness of Simulation Results

This section discusses the robustness of the model to alternative parameter choices. We conduct several experiments in which we vary one key parameter at a time, while holding all other parameters fixed at their benchmark levels. Table S.3 summarizes these robustness checks. For brevity, we only report the results averaged over all investment managers and omit the results broken out for skilled and unskilled managers, separately. We find that none of the comparative statics are sensitive to variation in the key parameters of the model.

Varying the fraction of skilled managers In our benchmark model, we assume that 20% ($\chi = .20$) of investors are skilled mutual funds (60% are unskilled mutual funds and 20% unskilled other investors). We first study two different values for the fraction of skilled investment managers: $\chi = 10\%$ and $\chi = 30\%$. When there are fewer skilled funds, they have a comparatively larger advantage over the unskilled. This results in investment choices that exploit their informational advantage more aggressively. *Timing* for the skilled increases from 156 to 210 in recessions while their *Picking* reading in expansions increases from 160 in the baseline to 181. At the same time, there are fewer skilled investors exploiting more unskilled investors than in the baseline, so that the unskilled investors have less negative average *Timing* values in recessions and less negative *Picking* values in expansions. As a result, the *Timing* value in recessions and *Picking* value in expansions, averaged across across *all* investment managers (80% of the investor population), fall relative to the benchmark (from 9.9 to 5.8). Similarly, *Ftiming* increases in recessions for all funds and *Fpicking* decreases, but the changes are smaller than in the benchmark case. Likewise, our measures of portfolio dispersion continue to be higher in recessions than in expansions, but all dispersion levels are somewhat lower than before. The reason is that there is no dispersion among the unskilled, and there are more of them than in the benchmark. Finally, the performance results remain intact as well. The skilled investors make higher abnormal returns and alphas than in the benchmark, which means the unskilled loose more in total. However, they loose less per unskilled investor. As a result, alphas averaged across all funds are lower than in the benchmark: 18.6bp per month in recessions (versus 35.3bp) and 15.6bp in expansions (versus 30.7bp).

The opposite effects occur when we increase the fraction of skilled investors to 30 percent. The increase in *Ftiming* and *Timing* and the decrease in *Fpicking* and *Picking* in recessions are larger than those in the benchmark model. The same is true for portfolio dispersion and performance. For example, the average alpha is now 50.6bp per month in recessions and 45.4bp in expansions; the difference is slightly higher than in the benchmark. In expansions, all skilled investors continue to learn about the stock-specific information. In recessions, about 70% of attention is allocated to the aggregate shock in recessions and 15% to each of the stock-specific shocks. This 70% is lower than the 87% of skilled managers who learn about the aggregate shock in recessions in our benchmark parametrization. This is a general equilibrium effect, which we label *strategic substitutability.* When many informed investors learn about the aggregate shock, and buy assets that load heavily on that shock, they push up the price of these assets, making it less desirable to learn about for other informed investors ceteris paribus. This leads some to learn about the stock-specific shocks instead. Hence, the higher average Fpicking of the informed in recessions compared to the benchmark. Why is the reverse not happening in expansions? Because the volatility of the aggregate shock is low enough in expansions that it turns out not to be optimal for any of the 30% informed investors to deviate from the full attention allocation to the idiosyncratic shocks.

Varying capacity K The second variational experiment is to decrease and increase the amount of attention allocation capacity K that skilled investors have. In our benchmark, $K = 1$, which amounts to the ability to increase the precision on any one signal by 25% of the prior precision of the stock-specific information through learning. We now consider $K = .5$ and $K = 2$. When the 20% of skilled have twice as much capacity, their *Ftiming* and *Timing* increase substantially in recessions (*Timing* goes up from 157 in the benchmark to 220), and their *Fpicking* and *Picking* increase in expansions (*Picking* goes up from 160 in the benchmark to 311). In contrast to the previous exercise, the *Timing* measure for the unskilled becomes more negative in recessions and their *Picking* more negative in expansions than in the benchmark. The reason is that there are as many unskilled as in the benchmark, but they are now at a larger informational disadvantage. The net effect of the skilled and the unskilled is an increase in *Timing* in recessions from 9.9 in the benchmark to 14.0. Likewise, *Picking* in expansions increases from 10.0 to 19.5. Giving 30% of investors $K = 1$ has similar effects as giving 20% of investors $K = 2$. Portfolio dispersion increases substantially with higher K. The result is driven by the more concentrated portfolios of the skilled, which creates both more dispersion among the skilled and a bigger difference with the unskilled. The skilled investors make abnormal returns and alphas that are about twice as high as those in the benchmark, and the unskilled loose about twice as much. The net effect are average fund alphas that are substantially higher than in the benchmark: 67.2bp per month in recessions (versus 35.3bp) and 59.3bp in expansions (versus 30.7bp). The opposite happens when we lower K to 0.5.

Recessions are times with low returns. We recall that recessions in the model are periods with not only a higher variance of the aggregate shock, but also with lower realized market returns. We implement the latter by first simulating the model in recessions for 3000 periods, then taking the bottom m% of return realizations, and adding them to the 3000 draws when calculating the moments of interest. In our third robustness check we verify how robust our results are to different values for m. We explore $m = 0$ and $m = 0.08$, while our benchmark is $m = 0.025$. When $m = .08$, realized market returns are 22 basis points per month lower in recessions than in expansions (0.54 versus 0.76% per month). This corresponds to the return difference in the data. The results for *Timing*, *Picking*, *Ftiming*, and *Fpicking* are slightly stronger, but the magnitudes are quite close to the benchmark. The same is true for all dispersion measures, except for the beta dispersion. The latter is quite a bit lower in recessions than in the benchmark (3.91 instead of 6.37), driven by a reduction in the beta dispersion of the skilled. Because of the lower returns in recessions, skilled managers have both lower betas and less differences in their betas compared to the unskilled in recessions. Finally, the performance results are similar to the benchmark. Alphas are slightly higher than in the benchmark: 38.5bp per month in recessions (versus 35.3bp) and 31.6bp in expansions (versus 30.7bp). The difference between recessions and expansions grows to 7bp per month.

The case of $m = 0$ corresponds to a world in which assets have realized payoffs that are symmetrically distributed around the same mean in expansions and in recessions. However, because recessions are times in which returns are more volatile, *expected* (and unconditional average) returns must be higher to compensate the investors for bearing higher risk. In particular, the average market return is 1.30% in recessions and 0.95% in expansions. The results on the fund moments are opposite from the case with higher m, but still quantitatively similar to our benchmark case. For example, the difference in average alphas between recessions and expansions is 4.1bp per month compared to 4.6bp in the benchmark.

Table S.3: **Robustness of Predictions of the Model**

Panel A reports moments related to attention allocation, Panel B reports the moments related to dispersion, and Panel C reports moments related to performance. The first column lists the moments in question, as defined in the main text. The other pairs of columns report results for the benchmark parameters and for six robustness exercises. In each pair of columns, the first column reports the predictions for the model simulated in a recession and the second column for the model simulated in an expansion. All moments are generated from a simulation of 3000 draws and 750 investors. For both recessions and expansions, we list the equally weighted average across all investment managers (the 20% skilled and the 60% investment managers). The parameters are the same as in the benchmark model, except for the parameter listed in the first row.

	baseline		$\chi = .10$		$\chi = .30$		$K = 0.5$		$K = 2$		$m = 0$		$m = .08$	
	R	E	R	E	R	E	R	E	R	E	R	E	R	E
Panel A: Attention Allocation														
1. Ftiming	10.55	0.07	6.11	0.04	12.96	-0.08	6.23	0.01	14.53	-0.08	9.30	-0.04	11.01	-0.02
2. Timing	9.91	0.06	5.77	-0.09	11.56	0.12	5.93	-0.05	13.96	0.13	9.34	0.09	10.93	0.06
3. Fpicking	2.15	15.61	0.01	7.76	7.23	23.21	-0.03	7.89	12.66	30.14	2.08	16.16	2.14	15.76
4. Picking	1.66	10.02	0.13	5.04	5.10	14.83	0.18	5.04	8.19	19.45	1.68	9.31	1.65	10.28
Panel B: Dispersion														
5. Concentration	3.75	3.12	1.99	1.59	5.27	4.62	1.78	1.50	7.71	6.74	3.74	3.12	3.77	3.14
6. Idiosyncratic volatility	5.09	4.33	2.90	2.27	6.66	6.07	3.34	2.80	7.38	6.71	4.83	4.09	5.29	4.33
7. Dispersion in abnormal return	3.54	3.37	1.93	1.82	4.94	4.71	2.21	2.13	5.97	5.51	3.42	3.21	3.74	3.44
8. Dispersion in CAPM alpha	2.52	2.28	1.52	1.35	3.06	2.84	1.32	1.15	4.85	4.46	2.22	2.15	2.76	2.35
9. Dispersion in CAPM beta	6.37	1.46	4.70	1.03	5.76	2.45	3.45	0.98	10.84	2.67	21.03	2.08	3.91	1.38
Panel C: Performance														
10. Abnormal return	0.346	0.302	0.178	0.150	0.500	0.450	0.184	0.151	0.664	0.589	0.330	0.283	0.379	0.312
11. CAPM Alpha	0.353	0.307	0.186	0.156	0.506	0.454	0.192	0.156	0.672	0.593	0.329	0.288	0.385	0.316

S.1.4 Endogenous Capacity Model

Finally, we consider an extended model in which skilled managers can freely choose not only how to allocate their information processing capacity, but also how much capacity to acquire. We let the cost of acquiring K units of capacity be $\mathcal{C}(K)$. Each skilled fund solves for the choice of signal precisions $K_{aj} \geq 0$ and $K_{1j} \geq 0$, and capacity K that maximize time-1 expected utility, as in (11) but adjusted for a penalty term $-\mathcal{C}(K)$. In our numerical work, we assume that these choice variables lie on a 25×25 grid in $\mathbb{R}^3_+$. The choice of signal precision $K_{2j} \geq 0$ is implied by the capacity constraint (7).

In our numerical exercise, we consider two different functional forms for $\mathcal{C}(K)$. The first one is $\mathcal{C}_1(K) = c_1 \exp(K)$ and the second one is $\mathcal{C}_2(K) = c_2 K^{\psi}$. For ease of comparison with our exogenous K results, we choose the scalars c_1 and c_2 such that the optimal capacity choice is $K = 1$ on average across expansions and recessions. This is the same capacity choice we assume in our benchmark parametrization. Clearly, increasing (lowering) the scalars c_1 and c_2 will lead to lower (higher) optimal capacity choice. These scalars can be interpreted as (shadow) prices of capacity. All other parameters are the same as in our benchmark model.

More interesting than the level of K that is chosen is how that choice differs between recessions and expansions. We find that for both cost functions, investors acquire more capacity in recessions than in expansions. Nothing in the cost function makes it cheaper to acquire capacity in either expansions or recessions. This result is solely driven by the fact that the higher (aggregate) uncertainty in recessions makes it optimal to acquire more capacity and to allocate it to the aggregate shock. This extensive-margin effect acts as an amplification to our intensive-margin effect. How elastic capacity choice is to changes in prior aggregate uncertainty, and hence how large the amplification effect is, *does* depend on the functional form of the cost function. For cost function 1, we find that capacity choice is 1.02 in recessions and 0.97 in expansions. For cost function 2, the elasticity is much higher, with a capacity choice of 1.15 in recessions and 0.92 in expansions. The reason for the higher elasticity is that the marginal cost function 2 is less steep in capacity. As a result, a given change in the marginal benefit of acquiring information leads to larger equilibrium changes in capacity. Since we have no strong prior over the functional form, we conduct our numerical simulation for both cost functions.

Table S.4 summarizes the main moments of interest for the endogenous K model, alongside the benchmark, exogenous K results. For brevity, we only report the results averaged over all investment managers and omit the results broken out for skilled and unskilled managers, separately. Overall, we find that the results are very similar to those in our exogenous K model, not only qualitatively, but also quantitatively. The moments for cost function 2 (two most right columns) tend to be higher in recessions than do the benchmark numbers, and lower in expansions. Hence, there is amplification of the difference between recessions and expansions. For example, average fund alphas are somewhat higher than in the benchmark in recessions (40.7bp per month versus 35.3bp) and somewhat lower in expansions (27.7bp versus 30.7bp). The resulting difference between recessions and expansions grows substantially from 4.6bp to 13bp per month. For cost function 1 (two middle columns), the moments are slightly higher in recessions since the skilled investment managers choose to acquire slightly more capacity than what they are endowed with in the benchmark ($K = 1.02$ versus 1). The moments are slightly lower in expansions, since they have slightly lower capacity ($K = 0.97$ versus 1). Overall, the difference in our key variables between recessions and expansions is usually very similar to that in our benchmark model.

Table S.4: **Endogenous Capacity Model**

This table provides the results from an extension of the model where skilled funds endogenously choose how much capacity to acquire. It reports on the main predictions of the model. Panel A reports moments related to attention allocation, Panel B reports the moments related to dispersion, and Panel C reports moments related to performance. The first column lists the moments in question, as defined in the main text. The other pairs of columns report results for the benchmark parameters and for two versions of the endogenous K model with different cost functions. The cost function in the first one is $\mathcal{C}_1(K) = c_1 \exp(K)$, while the cost function in the second one is $\mathcal{C}_2(K) = c_2 K^{\psi}$. We set $c_1 = 1.057$, $c_2 = 2.4$, and $\psi = 1.2$. All other parameters are the same as in the benchmark model. In each pair of columns, the first column reports the predictions for the model simulated in a recession (R) and the second column for the model simulated in an expansion (E). All moments are generated from a simulation of 2000 draws and 100 investors. For both recessions and expansions, we list the equally weighted average across all investment managers (the 20% skilled and the 60% investment managers).

	Baseline		$\mathcal{C}_1(K) = c_1 \exp(K)$		$\mathcal{C}_2(K) = c_2 K^{\psi}$	
	R	E	R	E	R	E
Panel A: Attention Allocation						
1. Ftiming	10.55	0.07	10.53	-0.12	11.58	-0.01
2. Timing	9.91	0.06	9.83	0.10	10.65	0.13
3. Fpicking	2.15	15.61	2.35	15.43	3.69	14.37
4. Picking	1.66	10.02	1.76	9.91	2.65	8.87
Panel B: Dispersion						
5. Concentration	3.75	3.12	3.82	3.13	4.37	2.85
6. Idiosyncratic volatility	5.09	4.33	5.20	4.53	5.34	4.15
7. Dispersion in abnormal return	3.54	3.37	3.59	3.34	3.97	3.16
8. Dispersion in CAPM alpha	2.52	2.28	2.57	2.29	2.92	2.06
9. Dispersion in CAPM beta	6.37	1.46	5.18	3.01	5.28	1.94
Panel C: Performance						
10. Abnormal return	0.346	0.302	0.348	0.302	0.399	0.271
11. CAPM Alpha	0.353	0.307	0.355	0.307	0.407	0.277

S.2 An Expected Utility Model

With expected utility, the time-2 utility is the same as in the main text. Utility U_{2j} is a log-transformation of expected exponential utility. Maximizing the log of expected utility is equivalent to maximizing expected utility because log is a monotonic transformation. However, period-1 utility U_{1j} is the time-1 expectation of the log of time-2 expected utility. That is a transformation that induces a preference for early resolution of uncertainty. When thinking about information acquisition, considering agents who have such a preference is helpful. The expected utility model has some very undesirable features and, although versions of the main results still hold, the intuition for why they hold has less useful economic content to it.

The problem is that, at the time when he chooses information, an expected utility investor does not value being less uncertain when he invests. He only cares about the uncertainty he faces initially (exogenous prior uncertainty) and how much uncertainty there is at the end (none, payoffs are observed). Of course, he values information that will help him to increase expected return. But if a piece of information might lead the investor to take an aggressive portfolio position, the investor will be averse to learning this information because given his current information, the portfolio he expects his future self to choose looks too risky. The upshot of all of this is some weird behavior. For example, if we introduce an asset that is very uncertain, but is in near-zero supply. Expected utility investors might all use all of their capacity to study this asset that is an infinitesimal part of their portfolio. Since we want to base our analysis on a plausible description of how financial market participants make decisions, we stuck with mean-variance utility in the main text.

That being said, the purpose of this section is to show that the results are robust to the expected utility formulation of the model. Since the time-2 utility functions are equivalent, the results for optimal portfolio holdings, portfolio dispersion and expected profits are identical. In other words, because lemma 1 and propositions 3, 4, 5 and 6 take arbitrary information choices as given, changes in the model that only affect the information choices do not affect these results. What does change is the proofs of propositions 1 and 2, the results about how attention is allocated.

Utility We begin with a derivation of time-1 expected utility. We compute ex-ante utility for investor j as $U_{1j} = E[-e^{-\rho W}]$ where the expectation is unconditional. First we substitute the budget constraint and obtain $U_{1j} = E[-e^{-\rho \tilde{q}'(\tilde{f}-\tilde{p}r)}]$, where we he omitted the constant term $-e^{-\rho r W_0}$ since it will not change the optimization problem. In period 2, the investor has chosen his portfolio and the price is in his information set, therefore the only random variable is z. Conditioning on $\hat{z}_j$ and $\hat{\Sigma}_j$ and using the formula for the expectation of a log-normal variable we obtain:

$$\begin{aligned} U_{1j} &= E\left[E\left[-e^{-\rho\tilde{q}'(\tilde{f}-\tilde{p}r)}\middle|\hat{z}_j, \hat{\Sigma}_j\right]\right] \\ &= E\left[-e^{-\rho\tilde{q}'(\tilde{f}-\tilde{p}r)+\frac{\rho^2}{2}\tilde{q}'\hat{\Sigma}_j\tilde{q}}\right] \\ &= E\left[-e^{-\frac{1}{2}(E_j[\tilde{f}]-\tilde{p}r)'\hat{\Sigma}_j^{-1}(E_j[\tilde{f}]-\tilde{p}r)}\right] \end{aligned}$$

where the third line substitutes the optimal portfolio choice $\tilde{q} = \rho^{-1}\hat{\Sigma}_j^{-1}(\tilde{f} - \tilde{p}r)$. Now we compute expectations in period 1. Note that both the expected return and the price are random variables and that both are correlated since they contain information about the true payoffs. Recall from the previous section that $E_j[\tilde{f}] - \tilde{p}r \sim \mathcal{N}(w, V - \hat{\Sigma}_j)$, then we have to compute the expectation of the exponential of the square of a normal variable. We will rewrite the expression in terms of the zero mean random variable $y \equiv E_j[\tilde{f}] - \tilde{p}r - w \sim \mathcal{N}(0, V - \hat{\Sigma}_j)$ and use the formula in p.102 of Veldkamp (2011) with $F = -\frac{1}{2}\hat{\Sigma}_j^{-1}$, $G' = -w'\hat{\Sigma}_j^{-1}$ and $H = -\frac{1}{2}w'\hat{\Sigma}_j^{-1}w$:

$$\begin{aligned} U_{1j} &= E\left[-e^{-\frac{1}{2}(E_j[\tilde{f}]-\tilde{p}r)'\hat{\Sigma}_j^{-1}(E_j[\tilde{f}]-\tilde{p}r)}\right] \\ &= E\left[-e^{-\frac{1}{2}y'\hat{\Sigma}_j^{-1}y-w\hat{\Sigma}_j^{-1}y-\frac{1}{2}w'\hat{\Sigma}_j^{-1}w}\right] \\ &= -|I + (V - \hat{\Sigma}_j)\hat{\Sigma}_j^{-1}|^{-\frac{1}{2}} \exp\left\{\frac{1}{2}w'\hat{\Sigma}_j^{-1}\hat{\Sigma}_j V^{-1}(V - \hat{\Sigma}_j)\hat{\Sigma}_j^{-1}w - \frac{1}{2}w'\hat{\Sigma}_j^{-1}w\right\} \\ &= -\left(\frac{|\hat{\Sigma}_j|}{|V|}\right)^{\frac{1}{2}} \exp\left(-\frac{1}{2}w'V^{-1}w\right) \end{aligned}$$

In the proofs below, we will work with a monotonic transformation $\tilde{U} \equiv -2\log(-U_{1j})$ given by

$$\tilde{U} = -\log|\hat{\Sigma}_j| + \log|V| + w'V^{-1}w$$

We now show the computation of each term in utility.

- $|\hat{\Sigma}_j^{-1}| = \prod_{l=1}^n \hat{\sigma}_l^{-1} \implies -\log|\hat{\Sigma}_j| = \sum_{l=1}^n \log \hat{\sigma}_l^{-1}$
- $|V| = \prod_{l=1}^n \bar{\sigma}_l[1+(\rho^2\sigma_x + \bar{K}_l)\bar{\sigma}_l] \implies \log|V| = \sum_{l=1}^n \log\left(\bar{\sigma}_l[1+(\rho^2\sigma_x+\bar{K}_l)\bar{\sigma}_l]\right)$
- $w'V^{-1}w = \sum_{l=1}^n \left(\frac{\rho^2\bar{x}_l^2}{\rho^2\sigma_x+\bar{K}_l+\bar{\sigma}_l^{-1}}\right)$

With all these elements, the transformation of utility reads:

$$\tilde{U} = \sum_{l=1}^{n}\left\{-\log\hat{\sigma}_l + \log\bar{\sigma}_l[1+(\rho^2\sigma_x+\bar{K}_l)\bar{\sigma}_l] + \frac{\rho^2\bar{x}_l^2}{\rho^2\sigma_x+\bar{K}_l+\bar{\sigma}_l^{-1}}\right\} \tag{S.1}$$

Observe that the only utility component affected by the actions of the investor is the first.

S.2.1 Proof of Proposition 1

For a given investor j, the marginal value of allocating an increment of capacity to shock i is increasing in its variance σ_i, this is: $\partial^2 U/\partial K_{ij}\partial\sigma_i > 0$.

Proof. Recall that transformed utility is given by: $\tilde{U} = -\log|\hat{\Sigma}_j| + \log|V| + \sum_{l=1}^n\left\{\frac{\rho^2\bar{x}_l^2}{\rho^2\sigma_x+\bar{K}_l+\bar{\sigma}_l{}^{-1}}\right\}$. We start by taking the derivative of utility with respect to K_{ij}, noting that K_{ij} only affects the investor's posterior variance (it does not affect any average precision inside V because the investor has measure zero):

$$\frac{\partial\tilde{U}}{\partial K_{ij}} = -\frac{\partial\log|\hat{\Sigma}_j|}{\partial K_{ij}} = \hat{\sigma}_i > 0$$

Now we take derivative of the previous expression with respect to σ_i:

$$\frac{\partial^2\tilde{U}}{\partial K_{ij}\partial\sigma_i} = \left(\frac{\hat{\sigma}_i}{\sigma_i}\right)^2 > 0$$

To show the result holds also for the original utility U, first observe that $U = -e^{-\tilde{U}/2}$. Second, we will use Faà di Bruno's formula for the derivative of a composition:

$$\begin{aligned}
\frac{\partial^2 U}{\partial K_{ij}\partial\sigma_i} &= \frac{\partial U}{\partial\tilde{U}}\frac{\partial^2\tilde{U}}{\partial K_{ij}\partial\sigma_i} + \frac{\partial^2 U}{\partial\tilde{U}^2}\frac{\partial\tilde{U}}{\partial K_{ij}}\frac{\partial\tilde{U}}{\partial\sigma_i} \\
&= \frac{1}{2}e^{-\frac{\tilde{U}}{2}}\left(\frac{\hat{\sigma}_i}{\sigma_i}\right)^2 - \frac{1}{4}e^{-\frac{\tilde{U}}{2}}\hat{\sigma}_i\left(-\frac{1}{\hat{\sigma}_i}\left(\frac{\hat{\sigma}_i}{\sigma_i}\right)^2 + \left(\frac{\bar{\sigma}_i}{\sigma_i}\right)^2\left(\frac{1+2(\rho^2\sigma_x+\bar{K}_{i'})\bar{\sigma}_i}{\bar{\sigma}_i(1+(\rho^2\sigma_x+\bar{K}_i)\bar{\sigma}_i)}\right) + \frac{\rho^2\bar{x}_i^2}{\sigma_i^2(\rho^2\sigma_x+\bar{K}_i+\bar{\sigma}_i^{-1})^2}\right) \\
&= \frac{3}{4}e^{-\frac{\tilde{U}}{2}}\left(\frac{\hat{\sigma}_i}{\sigma_i}\right)^2 - \frac{1}{4}e^{-\frac{\tilde{U}}{2}}\hat{\sigma}_i\left[\left(\frac{\bar{\sigma}_i}{\sigma_i}\right)^2\left(\frac{1+2(\rho^2\sigma_x+\bar{K}_i)\bar{\sigma}_i}{\bar{\sigma}_i(1+(\rho^2\sigma_x+\bar{K}_i)\bar{\sigma}_i)}\right) + \frac{\rho^2\bar{x}_i^2}{\sigma_i^2(\rho^2\sigma_x+\bar{K}_i+\bar{\sigma}_i^{-1})^2}\right]
\end{aligned}$$

where we have substituted all the terms. A sufficient condition for this expression to be positive is $\hat{\sigma}_{i'} < \hat{\sigma}_i < 3\hat{\sigma}_{i'}$. Under this condition, the marginal utility of reallocating capacity from shock i to i' is increasing in $\sigma_{i'}$. □

S.2.2 Proof of Proposition 2

An increase in risk aversion ρ increases the marginal utility for investor j of reallocating capacity from shocks with high posterior precision to shocks with low posterior precision: If $K_{i'j} = \tilde{K}$ and $K_{ij} = K - \tilde{K}$, then $\frac{\partial^2 \tilde{U}}{\partial \rho \partial \tilde{K}} > 0$ as long as $\hat{\sigma}_i^{-1} > \hat{\sigma}_{i'}^{-1}$.

Proof. As before, the chain rule implies that $\frac{\partial^2 \tilde{U}}{\partial \rho \partial K_{ij}} = \frac{\partial^2 \tilde{U}}{\partial \rho \partial K_{i'j}} - \frac{\partial^2 \tilde{U}}{\partial \rho \partial K_{ij}}$. For each i, we have that

$$\frac{\partial^2 \tilde{U}}{\partial \rho \partial K_{ij}} = \frac{\partial \left(\frac{\partial \tilde{U}}{\partial K_{ij}} \right)}{\partial \rho} = \frac{\partial \hat{\sigma}_i}{\partial \rho} = \frac{2}{\rho} \frac{\hat{\sigma}_i^2}{\sigma_{ip}} > 0$$

Since each investor has measure zero, his reallocation of capacity does not change the average, which we write as: $\bar{K} \equiv \bar{K}_{ij} = \bar{K}_{i'j}$. Therefore the difference is given by:

$$\frac{\partial^2 \tilde{U}}{\partial \rho \partial \tilde{K}} = \frac{2}{\rho \sigma_{ip}} \left[\hat{\sigma}_{i'}^2 - \hat{\sigma}_i^2 \right]$$

This expression is positive as long as the difference inside the brackets is positive, which is equivalent to $\hat{\sigma}_i^{-1} > \hat{\sigma}_{i'}^{-1}$. To show the result holds also for the original utility U, first observe that $U = -e^{-\tilde{U}/2}$. Second, we will use Faà di Bruno's formula for the derivative of a composition:

$$\begin{aligned} \frac{\partial^2 U}{\partial \rho \partial \tilde{K}} &= \frac{\partial U}{\partial \tilde{U}} \frac{\partial^2 \tilde{U}}{\partial \rho \partial \tilde{K}} + \frac{\partial^2 U}{\partial \tilde{U}^2} \frac{\partial \tilde{U}}{\partial \tilde{K}} \frac{\partial \tilde{U}}{\partial \rho} \\ &= e^{-\frac{\tilde{U}}{2}} \frac{\left[\hat{\sigma}_{i'}^2 - \hat{\sigma}_i^2 \right]}{\rho \sigma_{ip}} - \frac{1}{4} e^{-\frac{\tilde{U}}{2}} (\hat{\sigma}_{i'} - \hat{\sigma}_i) \left\{ \frac{2}{\rho} \sum_{l=1}^{n} \frac{(\bar{\sigma}_l - \hat{\sigma}_l)}{\sigma_{lp}} + \sum_{l=1}^{n} \frac{2\rho \sigma_x \bar{\sigma}_l + \frac{\bar{\sigma}_l^2}{\sigma_{lp}} (\rho^2 \sigma_x + \bar{K}_l)}{1 + (\rho^2 \sigma_x + \bar{K}_l) \bar{\sigma}_l} + 2\rho \sum_{l=1}^{n} \frac{\bar{x}_l^2 (\bar{K}_l + \bar{\sigma}_l^{-1} + \sigma_{lp}^{-1})}{(\rho^2 \sigma_x + \bar{K}_l + \bar{\sigma}_l^{-1})^2} \right. \end{aligned}$$

Thus, if aggregate shocks have lower posterior precision, an increase in risk aversion will make learning about them more valuable. □