Econometrica Supplementary Material

SUPPLEMENT TO "A RATIONAL THEORY OF MUTUAL FUNDS' ATTENTION ALLOCATION"

(*Econometrica*, Vol. 84, No. 2, March 2016, 571–626)

BY MARCIN KACPERCZYK, STIJN VAN NIEUWERBURGH, AND LAURA VELDKAMP

S.1. A NUMERICAL EXAMPLE

IN THIS SECTION, we use a numerical example to illustrate the model's predictions for the objects of the six main propositions: attention allocation, portfolio return dispersion, and abnormal return. The numerical example serves to illustrate that the parameter conditions under which the results are derived are not too restrictive, and that the results hold for plausible variations in these parameters.

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.I. Below, we show that the qualitative results are robust to plausible variations in these parameter choices.

The example features 2500 investors of which a fraction $\chi = 0.2$, or 500, are informed or skilled funds. The remaining 2000 investors are uninformed, and composed of 1500 uninformed funds and 500 uninformed non-fund investors. The example features three assets, two stocks (assets 1 and 2), and one composite asset (asset 3). There also is a risk-free asset whose net return is set to 1%. Our procedure is to simulate 10,000 draws of the shocks $(x_1, x_2, x_3, z_1, z_2, z_3)$, where z_3 is the aggregate shock, in recessions and 10,000 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. Recessions differ from expansions in that (i) the variance of the aggregate payoff shock σ_n is higher, and (ii) the market price of risk is higher, here governed by the coefficient of absolute risk aversion. In expansions, we set risk aversion $\rho = 0.175$ and $\sigma_n = 0.25$. Alongside the other parameters, that delivers an equity risk premium in expansions of 4%. To study the effect of recessions, we conduct three exercises. In the first exercise, we set $\sigma_n = 0.50$ in recessions, double its value in expansion, holding fixed ρ across expansions and recessions. In the second exercise, we set $\rho = 0.35$, double its value in expansions, holding fixed σ_n across expansions and recessions. In the third exercise, we increase both parameters simultaneously.

We normalize the mean asset supply of assets 1 and 2 to 1, and set the supply of the aggregate asset, $\bar{x}_3$, to 15. We set the variance of the asset supply noise

 DOI: 10.3982/ECTA11412

TABLE S.I
NUMERICAL EXAMPLE: PARAMETERS[a]

Parameter	Symbol	Value	Mainly affects
Risk aversion	ρ	0.175 (E), 0.35 (R)	Asset return mean
Variance aggr. payoff comp. z_3	σ_n	0.25 (E), 0.50 (R)	Market return vol in expansions vs. recessions
Mean of payoffs 1, 2, 3	μ_1, μ_2, μ_3	15	Asset return mean
Variance idio. payoff comp. z_1, z_2	σ_i	0.55	Asset return vol. vs. market return vol.
Sensitivity of payoffs to z_3	b_1, b_2	0.7, 1.0	Asset beta level + dispersion
Mean asset supply 1,2	$\bar{x}_1 = \bar{x}_2$	1	Normalization
Mean asset supply 3	$\bar{x}_3$	15	Asset return volatility
Variance asset supply	σ_x	0.5	Asset return idio vol.
Risk-free rate	$r - 1$	0.01	Average T-bill return
Initial wealth	W_0	220	Average cash position
Information capacity	K	1	Information advantage of skilled
Skilled fraction	χ	0.20	Information advantage of skilled

[a]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 what features of the model are predominantly affected by that value.

equal to $\sigma_x = 0.5$. We vary both of these parameters below. The variance of the firm-specific payoff shocks is 0.55 in expansions and recessions, making the volatility of assets 1 and 2 about 80% larger than that on the market portfolio. We normalize mean asset payoffs $\mu_1 = \mu_2 = \mu_3 = 15$. We choose the asset loadings on the aggregate payoff shock, $b_1 = 0.70$ and $b_2 = 1$, to be different from each other so as to generate some spread in asset betas. The chosen values generate average market betas of 1.0 and a standard deviation in betas of 25%. We choose initial wealth $W_0 = 220$ to generate risk-free asset holdings that are close to zero in expansions.

Skilled fund investors ($K > 0$) solve for the choice of signal precisions $K_{ij} \geq 0$ and that maximize time-1 expected utility (11). We assume that these choice variables lie on a 100×100 grid in $\mathbb{R}^2_+$. The signal precision choice $K_{2j} \geq 0$ is implied by the capacity constraint (7). For simplicity, we set capacity K for skilled funds equal to 1. This value implies that learning can increase the precision of one of the idiosyncratic shocks by 55% or the precision of the aggregate shock by 44% (assuming 85% of the periods are expansions and 15% recessions). We will vary K in our robustness exercise below. Likewise, we have no strong prior on the fraction of informed funds, χ, and we will vary it for robustness.

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).

TABLE S.II
BENCHMARK SIMULATION RESULTS[a]

	Change in Aggregate Risk		Change in Risk Aversion		Change in Both	
State	R	E	R	E	R	E
σ_n	0.50	0.25	0.25	0.25	0.50	0.25
ρ	0.175	0.175	0.35	0.175	0.35	0.175
Attention	0.94	0.50	1.00	0.50	1.00	0.50
Dispersion	9.07	6.88	7.56	6.88	38.36	6.88
Performance	0.164	0.148	0.199	0.148	0.541	0.148

[a]The table reports key outcome variables in Propositions 1–6 of the main text, for 10,000 periods of simulation of the model in Recessions (R) and in Expansions (E). The parameters are those reported in Table S.I. The only parameters that change between expansions and recessions are those reported in the rows σ_n and ρ. The model is simulated for 2500 investors of which 500 are skilled fund investors, 1500 are unskilled fund investors, and 500 are unskilled non-fund investors. All moments reported in the table are averages over the 2000 fund investors.

Main Simulation Results

Table S.II summarizes the predictions of the model for the main statistics of interest. We are interested in testing the three main predictions of the model relating to (i) attention allocation (Attention), (ii) portfolio dispersion (Dispersion), and (iii) fund performance (Performance). For each of these outcomes, we investigate the effect of increasing σ_n in recessions (columns 1 and 2), increasing ρ (columns 3 and 4), and increasing both (columns 5 and 6). The first two rows report the values of the variance of the aggregate risk factor z_3, σ_n, and risk aversion, ρ, in each exercise. The attention measure reported in row 3 is the fraction of capacity the average skilled fund devotes to learning about the aggregate risk factor, $\bar{K}_3$. The dispersion measure reported in row 4 is the same as in Propositions 3 and 4, the dispersion of portfolio excess returns $E[((\tilde{q}_j - \bar{\tilde{q}})'(\tilde{f} - \tilde{p}r))^2]$ averaged among all funds. Finally, the performance measure reported in row 5 is the same as in Propositions 5 and 6, the portfolio excess return, averaged among all funds, $E[(\tilde{q}_j - \bar{\tilde{q}})'(\tilde{f} - \tilde{p}r)]$.

Under the chosen parameters, skilled funds choose to allocate half of their capacity to learning about the aggregate risk in expansion, splitting the remaining 50% equally among risk factors 1 and 2. In recessions, whether they be periods with more aggregate risk or higher prices of risk (risk aversion), or both, attention is reallocated towards learning about aggregate risk. In the first experiment, 94% of capacity is allocated to risk factor 3, while in the other two experiments all capacity is allocated to aggregate risk. This attention reallocation confirms Propositions 1 and 2.

The simulation also provides support for the propositions relating to fund return dispersion (Propositions 3 and 4) and fund return performance (Propositions 5 and 6), showing that the portfolio dispersion increases in recessions in all experiments, and fund performance increases in recessions. These effects are driven by the skilled funds who have stronger outperformance in recessions at the expense of unskilled funds (and unskilled non-fund investors).

The last two columns show that the effects of increases in risk aversion and the variance of aggregate risk mutually reinforce one another. Thus there are interaction effects, which we discuss in Section 1.7 of the main text, and associated propositions proven in this Supplemental Material (Section S.7), which are confirmed in our numerical simulation.

Variations on Benchmark Parameters

This section discusses the robustness of the simulation results to alternative parameter choices. The four most natural parameters to vary are the amount of capacity the skilled funds have, K, the fraction of skilled funds, χ, the supply of the aggregate risk, $\bar{x}_3$, and the volatility of the noisy risk factor supply, σ_x. Propositions 2–6 require $\bar{x}_3$ to be sufficiently high, while Propositions 4 and 6 also require σ_x to be sufficiently high. We explore these changes in the four panels of Table S.III.

Panel A explores halving the capacity of skilled investors to $K = 0.5$. This value implies that learning can increase the precision of one of the idiosyncratic shocks (or the aggregate shock) by 28% (by 22%) compared to 55% (44%) in the benchmark. One might worry that the benchmark example gives too much capacity to skilled investors. As we see in the first two columns, there is more attention paid to the aggregate shock in expansions than before. With less capacity overall, learning about the most abundant risk becomes more valuable. Yet, there still is reallocation towards the aggregate shock in recessions. Average fund outperformance is weaker in expansions than in the benchmark since the skilled funds have a smaller advantage over unskilled fund and non-fund investors. The same is true for portfolio dispersion. However, dispersion and performance continue to increase sharply going from expansions to recessions. Columns 3 and 4 show that increasing risk aversion in recessions continues to drive attention reallocation towards the aggregate shock, increases dispersion, and increases performance.

In the second variational exercise, we lower the fraction of skilled investors to 10%. The amount of capacity of these skilled investors is set back to its benchmark value of 1. Compared to the previous variational exercise, the overall capacity of the skilled investor base is the same ($0.5 \times 0.20 = 1 \times 0.10$), but capacity is now more concentrated in the hands of fewer investors. As panel B shows, this parameter choice leads to the same capacity allocation outcome and outperformance in expansions and recessions, but it leads to more portfolio dispersion, and a larger increase therein in recessions.

In the benchmark model, assets 1 and 2 represented each 6.5% of the overall market capitalization. In panel C, we explore a lower share of 5% for the individual assets (and a larger 90% share for the composite asset) by increasing $\bar{x}_3$ from 15 to 20. We simultaneously increase W_0 to keep the risk-free asset

TABLE S.III

ROBUSTNESS SIMULATION RESULTS[a]

	Change in Aggregate Risk		Change in Risk Aversion		Change in Both	
State	R	E	R	E	R	E
σ_n	0.50	0.25	0.25	0.25	0.50	0.25
ρ	0.175	0.175	0.35	0.175	0.35	0.175
			Panel A: Change in capacity to $K = 0.5$			
Attention	1.00	0.82	1.00	0.82	1.00	0.82
Dispersion	10.43	3.22	3.62	3.22	23.46	3.22
Performance	0.185	0.080	0.126	0.080	0.420	0.080
			Panel B: Change in fraction informed to $\chi = 0.1$			
Attention	1.00	0.82	1.00	0.82	1.00	0.82
Dispersion	17.20	4.44	6.02	4.44	46.43	4.44
Performance	0.185	0.080	0.126	0.080	0.420	0.080
			Panel C: Change in size of aggregate asset to $\bar{x}_3 = 20$			
Attention	1.00	0.77	1.00	0.77	1.00	0.77
Dispersion	13.04	7.81	15.88	7.81	91.76	7.80
Performance	0.214	0.159	0.330	0.159	0.918	0.159
			Panel D: Change in variance of supply noise $\sigma_x = 1.0$			
Attention	1.00	0.60	1.00	0.60	1.00	0.60
Dispersion	17.11	7.58	9.28	7.58	57.71	7.58
Performance	0.248	0.158	0.227	0.158	0.684	0.157

[a]The table reports key outcome variables in Propositions 1–6 of the main text, for 10,000 periods of simulation of the model in Recessions (R) and in Expansions (E). The parameters are those reported in Table S.I. The only parameters that change between expansions and recessions are those reported in the rows σ_n and ρ, as well as the parameter listed in the first row of each panel. The model is simulated for 2,500 investors of which 500 are skilled fund investors, 1,500 are unskilled fund investors, and 500 are unskilled non-fund investors. All moments reported in the table are averages over the 2,000 fund investors.

allocation in expansions close to zero, but this does not affect any of the entries in the table. The larger size of the aggregate asset makes it more valuable to learn about, so that the fraction of capacity devoted to learning about this asset rises in expansions (from 50% in the benchmark to 77% here). Yet, there is still attention reallocation towards the aggregate asset going from expansions to recession. Similarly, dispersion and outperformance continue to increase going from expansion to recession.

Finally, in panel D we explore sensitivity to the volatility of noisy risk factor supply. We increase σ_x from 0.5 to 1.0, which makes prices less informative than in the benchmark. Prices convey about half as much information as the private signals informed investors receive, compared to them being slightly more informative than private information in the benchmark. On the margin, the increase in supply noise makes allocating attention to aggregate information more valuable in equilibrium. Dispersion and performance continue to increase in recessions.

We conclude that our main results survive across a range of parameters.

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 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 issue 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. This feature generates some undesirable behavior. For example, if an asset is introduced that is very uncertain but that 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 use mean-variance utility in the main text.

Putting this issue aside, 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, and 5 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 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

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

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 on page 102 of Veldkamp (2011) with $F = -\frac{1}{2}\hat{\Sigma}_j^{-1}$, $G' = -w'\hat{\Sigma}_j^{-1}$, and $H = -\frac{1}{2}w'\hat{\Sigma}_j^{-1}w$:

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

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:

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

With all these elements, the transformation of utility reads

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

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; that is, $\partial^2 U / \partial K_{ij}\, \partial \sigma_i > 0$.

PROOF: Recall that transformed utility is given by $\tilde{U} = -\log |\hat{\Sigma}_j| + \log |V| + \sum_{l=1}^{n} \{ \frac{\rho^2 \bar{x}_l^2}{\rho^2 \sigma_x + \bar{K}_l + \bar{\sigma}_l{}^{-1}} \}$. 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:

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

$$= \frac{3}{4} e^{-\tilde{U}/2} \left(\frac{\hat{\sigma}_i}{\sigma_i} \right)^2 - \frac{1}{4} e^{-\tilde{U}/2} \hat{\sigma}_i \left[\left(\frac{\bar{\sigma}_i}{\sigma_i} \right)^2 \left(\frac{1 + 2(\rho^2 \sigma_x + \bar{K}_i) \bar{\sigma}_i}{\bar{\sigma}_i (1 + (\rho^2 \sigma_x + \bar{K}_i) \bar{\sigma}_i)} \right) + \frac{\rho^2 \bar{x}_i^2}{\sigma_i^2 (\rho^2 \sigma_x + \bar{K}_i + \bar{\sigma}_i^{-1})^2} \right],$$

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 $\sigma_{i'}$. *Q.E.D.*

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:

$$\frac{\partial^2 U}{\partial \rho \, \partial \tilde{K}} = \frac{\partial U}{\partial \tilde{U}} \frac{\partial^2 \tilde{U}}{\partial \rho \, \partial \tilde{K}} + \frac{\partial^2 U}{\partial \tilde{U}^2} \frac{\partial \tilde{U}}{\partial \tilde{K}} \frac{\partial \tilde{U}}{\partial \rho}$$

$$= e^{-\tilde{U}/2}\frac{\left[\hat{\sigma}_{i'}^2 - \hat{\sigma}_i^2\right]}{\rho\sigma_{ip}} - \frac{1}{4}e^{-\tilde{U}/2}(\hat{\sigma}_{i'} - \hat{\sigma}_i)\left\{\frac{2}{\rho}\sum_{l=1}^{n}\frac{(\bar{\sigma}_l - \hat{\sigma}_l)}{\sigma_{lp}}\right.$$

$$\left. + \sum_{l=1}^{n}\frac{2\rho\sigma_x\bar{\sigma}_l + \dfrac{\bar{\sigma}_l^2}{\sigma_{lp}}\left(\rho^2\sigma_x + \bar{K}_l\right)}{1 + \left(\rho^2\sigma_x + \bar{K}_l\right)\bar{\sigma}_l} + 2\rho\sum_{l=1}^{n}\frac{\bar{x}_l^2\left(\bar{K}_l + \bar{\sigma}_l^{-1} + \sigma_{lp}^{-1}\right)}{\left(\rho^2\sigma_x + \bar{K}_l + \bar{\sigma}_l^{-1}\right)^2}\right\}.$$

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

S.3. SIGNALS ABOUT ASSET PAYOFFS

Suppose asset payoffs f have the structure given by equations (1) and (2) in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2014). But instead of having signals about the independent risk factors, each investor j gets a vector of signals η, where each entry η_i is an unbiased signal about the payoff of asset i:

$$\eta_j = f + \varepsilon_j, \tag{S.2}$$

where $f \sim N(\mu, \Sigma)$ and $\varepsilon_j \sim$ iid $N(0, K^{-1})$. Note that K is a diagonal matrix, implying that each signal has noise that is uncorrelated with other signals, but Σ is not diagonal, meaning that asset payoffs are correlated with each other.

The optimal portfolio choice first-order condition still takes the standard form

$$q_j = \frac{1}{\rho}\hat{\Sigma}_j^{-1}\left(E[f] - pr\right). \tag{S.3}$$

Substituting this optimal portfolio and equilibrium price into the budget constraint (4) and substituting that into time-2 expected utility (3):

$$\frac{1}{2}\left(E_j[f] - pr\right)\hat{\Sigma}_j^{-1}\left(E_j[f] - pr\right). \tag{S.4}$$

The expectation (posterior belief at time 2) $E_j[f]$ and posterior variance $\hat{\Sigma}_j^{-1}$ are still computed using Bayes's law. But unlike before, $\hat{\Sigma}_j$ will no longer be a diagonal matrix.

From Admati (1985), we know that for an arbitrary asset covariance and posterior belief covariance structure, prices are a linear function of asset payoffs and noisy asset supply shocks:

$$p = A + Bf + Cx. \tag{S.5}$$

Therefore, at time 1, the variable $E_j[f] - pr$ is a multivariate normal. Thus, time-1 expected utility is the expectation of a non-central chi-square:

$$U_{1j} = \frac{1}{2}\operatorname{trace}\big(\hat{\Sigma}_j^{-1} V_1\big[E_j[f] - pr\big]\big) + \frac{1}{2} E_1\big[E_j[f] - pr\big]' \hat{\Sigma}_j^{-1} E_1\big[E_j[f] - pr\big]. \tag{S.6}$$

Note that beliefs are a martingale. Thus, $E_1[E_j[f]] = E_1[f] = \mu$. Similarly, using (S.5), we can write $E_1[p] = A + B\mu + Cx$. Combining these expressions and using the fact that x is a mean-zero shock with variance Σ_x, we get $E_1[E_j[f] - pr] = (I - rB)\mu - A$ and $V_1[E_j[f] - pr] = (I - rB)\Sigma(I - rB)' + C\Sigma_x C'$. Notice that neither of these expressions contains choice variables. They depend on equilibrium pricing coefficients A, B, and C, which depend on aggregate information choices, but not on agent j's choice and on prior variance Σ, which is exogenous. The choice variable K_j shows up only in $\hat{\Sigma}_j^{-1}$, which is the posterior precision of beliefs. According to Bayes's law, this is

$$\hat{\Sigma}_j^{-1} = \Sigma^{-1} + \Sigma_p^{-1} + K_j. \tag{S.7}$$

Writing Expected Utility as a Separable Sum

Thus, we can write time-1 expected utility as

$$U_{1j} = c + \frac{1}{2}\operatorname{trace}\big(K V_1\big[E_j[f] - pr\big]\big) + \frac{1}{2} E_1\big[E_j[f] - pr\big]' K E_1\big[E_j[f] - pr\big], \tag{S.8}$$

where c is a constant that depends on parameters and equilibrium price coefficients. Note that K_j is a diagonal matrix. Therefore, this matrix expression can be written in the form of a simple sum:

$$U_{1j} = c + \frac{1}{2}\sum_{i=1}^{N} K_i\big[V_1\big[E_j[f_i] - p_i r\big] + E_1\big[E_j[f_i] - p_i r\big]^2\big]. \tag{S.9}$$

This illustrates that the marginal value of an additional unit of signal precision is the prior variance of the return on that asset, plus the squared expected return.

Effect of Higher Aggregate Risk Variance on Attention Allocation

Now that we have the model in this simple sum form, the question becomes, how do changes in aggregate risk and risk aversion change this marginal value of signal precision? Assets whose payoffs are most sensitive to aggregate risk, meaning that $(\partial V_1/\partial\sigma_a^2)$ and $(\partial E_1^2/\partial\sigma_a^2)$ are high, will become more valuable to learn about and will thus have weakly higher attention allocated to them.

Effect of Risk Aversion on Attention Allocation

The effect of risk aversion on attention allocation is more subtle to see. Risk aversion ρ enters expected utility through the pricing coefficients A and C. Thus,

$$\frac{\partial E_1[f_i - p_i r]}{\partial\rho} = \sum_{l=1}^{N} \bar{\bar{\Sigma}}(i,l)\bar{x}_l.$$

From (31) in the Appendix, we know that $V_1[f - pr] = \bar{\bar{\Sigma}}[\rho^2\sigma_x I + \bar{\bar{\Sigma}}_\eta^{-1'} + \bar{\Sigma}^{-1}]\bar{\bar{\Sigma}}$. Thus,

$$\frac{\partial V_1[f_i - p_i r]}{\partial\rho} = 2\rho\sigma_x\bar{\bar{\Sigma}}\bar{\bar{\Sigma}}.$$

Combining these two results allows us to describe how risk aversion changes the marginal value of signal precision:

$$\frac{\partial^2 U_{1j}}{\partial K_i\,\partial\rho} = 2\sigma_x(\bar{\bar{\Sigma}}\bar{\bar{\Sigma}})(i,i) + 2E_1[f_i - p_i r]\sum_{l=1}^{N}\bar{\bar{\Sigma}}(i,l)\bar{x}_l. \tag{S.10}$$

Assets for which (S.10) is high will become more valuable to learn about when risk aversion rises in recessions. Note that this change in the marginal value of information is greater for assets in abundant supply $\bar{x}_l$.

These results demonstrate that if signals are about asset payoffs (or any linear combination of asset payoffs), then attention will be reallocated in recessions. Although agents cannot learn more about aggregate risk directly (by assumption), the nature of the predictions in the same: In recessions, fund managers will learn more about assets whose payoffs are sensitive to aggregate risk and assets that are in abundant supply.

S.4. ENTROPY-BASED INFORMATION CONSTRAINT

The model in the main text features an attention limit that is a constraint on the sum of signal precisions. The Lagrangian problem therefore takes the form of the objective, plus the attention constraint, plus the nonnegativity constraints on all the signal precisions:

$$\mathcal{L} = c + \frac{1}{2}\sum_{i=1}^{N}\lambda_i K_i + \theta\left(\kappa - \sum_{i=1}^{N} K_i\right) + \sum_i \phi_i K_i \tag{S.11}$$

for a constant c and weights λ_i that depend on parameters and equilibrium aggregate attention choices. All the reallocation of attention arises when changes in parameters change the λ_i's.

We could instead constrain the entropy of signals. Entropy-based constraints are used by Mondria (2010), Sims (2003), or Maćkowiak and Wiederholt (2009, 2015). For an n-dimensional normal variable $\eta \sim N(f, K^{-1})$, entropy is a simple function of the determinant of variance–covariance matrix: $H(\eta) = 1/2 \ln[(2\pi e)^n |K^{-1}|]$, where $|\cdot|$ denotes the matrix determinant. Recognizing that $-\ln[|\hat{\Sigma}|] = \ln[|\hat{\Sigma}^{-1}|]$ and exponentiating both sides, we can write the entropy constraint $H(\eta) \geq \tilde{\kappa}$ as

$$|K| \leq \kappa,$$

where $\kappa = (2\pi e)^n \exp(-2\tilde{\kappa})$. Note that K is a diagonal matrix. Therefore, its determinant is the product of its elements.

The new Lagrangian problem takes the form

$$\mathcal{L} = c + \frac{1}{2}\sum_{i=1}^{N} \lambda_i K_i + \theta\left(\kappa - \prod_{i=1}^{N} K_i\right) + \sum_i \phi_i K_i. \tag{S.12}$$

The choice of each K_i must respect the no-forgetting constraint: $K_i \geq 1$, which is captured in the last term of the Lagrangian.

Solution to the Entropy Model

Notice that in all the preliminaries in Appendix A.1, there is no reference to the information constraint. The preliminary results are manipulations of expected utility, given some information allocation. As such, they apply verbatim to the entropy-constraint model. Similarly, in Appendix A.2, steps 1, 2, and 3 do not depend on the information constraint and therefore, apply equally to both models. In step 4, the information constraint enters in the information choice problem. The equivalent optimization for the entropy problem uses Lagrangian (S.12) instead of Lagrangian (S.11).

This problem maximizes a weighted sum of K_i's, subject to a product and an inequality constraint. The second-order condition for this problem is positive, meaning the optimum is a corner solution. A simple variational argument shows that the maximum is attained by maximizing the K_i with the highest λ_i weight in the sum (marginal value of precision). 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. We focus on symmetric equilibria.

Proposition 1 With the Entropy Constraint

Proposition 1 states that K_{ij} is weakly increasing in σ_i for a skilled investor j. In the linear-precision constraint model, the marginal value of signal precision is λ_i. Note that marginal utility is the same in this entropy-constrained problem. Therefore, the proof of Proposition 1 (in Appendix A.3) shows that $\partial\lambda_i/\partial\sigma_i > 0$ applies here as well. The rest of the proof is devoted to explaining why, if this marginal value increases, then K_{ij} is weakly increasing.

Thus, after establishing that the marginal value of a unit of precision is λ_i, the rest of the proof of Proposition 1 follows verbatim.

Proposition 2 With the Entropy Constraint

The proof of Proposition 2 in Appendix A.2 shows that if $\bar{x}_i$ is sufficiently large, $\partial\lambda_i/\partial\rho > 0$. Now that we have established that the marginal value of information λ_i is identical in this problem, the rest of the proof, arguing that K_i is weakly increasing, follows as before.

Propositions 3–7 With the Entropy Constraint

These propositions do not depend on the information constraint at all. They examine how a change in σ_i or ρ affects portfolio choice, something that is entirely separate from the form of the information choice. They rely on information choice only through the result that if σ_i or ρ increases, k_i weakly increases as well. Since that effect has been established in Propositions 1 and 2 above, the rest of the results follow, as in Appendix A.3.

S.5. COSTLY LEARNING FROM PRICES

This section shows that if we change the information constraint so that it requires capacity to process information from prices, then investors would choose not to process that information and to obtain independent signals instead. The idea behind this result is that an investor who learns from price information will infer that the asset is valuable when its price is high and infer that the asset is less valuable when its price is low. Buying high and selling low is generally not a way to earn high profits. This effect shows up as a positive correlation between $\hat{\mu}$ and pr, which reduces the variance $V_1[\hat{\mu}_j - pr]$.

Mathematical Preliminaries: Note that $B^{-1}(pr - A) = f + B^{-1}Cx$. Since x is a mean-zero shock, this is an unbiased signal about the true asset payoff f. The precision of this signal is $\Sigma_p^{-1} \equiv \sigma_x^{-1}B'(CC')^{-1}B$.

LEMMA 3: *A manager who could choose either learning from prices and observing a signal $\tilde{\eta}|f \sim N(f, \tilde{\Sigma}_\eta)$ or not learning from prices and instead getting a* *higher-precision signal* $\eta | f \sim N(f, \Sigma_\eta)$, *where the signals are conditionally independent across agents*, *and where* $\Sigma_\eta^{-1} = \Sigma_p^{-1} + \tilde{\Sigma}_\eta^{-1}$, *would prefer not to learn from prices*.

PROOF: From (11) in the main text, we know that expected utility is

$$U_{1j} = \frac{1}{2}\operatorname{trace}\big(\hat{\Sigma}_j^{-1} V_1[\hat{\mu}_j - pr]\big) + \frac{1}{2} E_1[\hat{\mu}_j - pr]' \hat{\Sigma}_j^{-1} E_1[\hat{\mu}_j - pr].$$

Since the two options yield equally informative signals, by Bayes's rule, they yield equally informative posterior beliefs: $\hat{\Sigma}_j^{-1} = \Sigma^{-1} + \Sigma_\eta^{-1}$, which is also equal to $\Sigma^{-1} + \Sigma_p^{-1} + \tilde{\Sigma}_\eta^{-1}$. Likewise, since both possibilities give the manager unbiased signals, beliefs are a martingale, meaning that $E_1[\hat{\mu}_j - pr]$ is identical under the two options.

Thus, the only term in expected utility that is affected by the decision to learn information from prices is $V_1[\hat{\mu}_j - pr]$. Let $\hat{\mu}_j = E[f|\eta]$ be the posterior expected value of payoffs for the manager who learns from the conditionally independent signal. By Bayes's law,

$$\hat{\mu}_j = \hat{\Sigma}_j\big(\Sigma^{-1}\mu + \Sigma_\eta^{-1}\eta\big).$$

The signal η can be broken down into the true payoff, plus noise: $\eta = f + \varepsilon$, where $\varepsilon \sim N(0, \Sigma_\eta)$. Using the expression for $\hat{\mu}_j$ and the pricing equation $pr = A + Bf + Cx$, we write

$$\hat{\mu}_j - pr = \hat{\Sigma}_j \Sigma^{-1}\mu - A + \hat{\Sigma}_j \Sigma_\eta^{-1}\varepsilon + \big(\hat{\Sigma}_j \Sigma_\eta^{-1} - B\big) f - Cx.$$

Since μ and A are constants, and ε, f, and x are mutually independent, the variance of this expression is

$$V_1[\hat{\mu}_j - pr] = \hat{\Sigma}_j \Sigma_\eta^{-1} \hat{\Sigma}_j + \big(\hat{\Sigma}_j \Sigma_\eta^{-1} - B\big)\Sigma\big(\hat{\Sigma}_j \Sigma_\eta^{-1} - B\big)' - \sigma_x CC'.$$

Next, consider the manager who chooses to learn information in prices. This person will have different posterior belief about f. Let $E[f|p, \tilde{\eta}] = \tilde{\mu}$. Using Bayes's law, he will combine information from his prior, prices, and the signal $\tilde{\eta}$ his posterior belief:

$$\tilde{\mu} = \hat{\Sigma}_j\big(\Sigma^{-1}\mu + \Sigma_p^{-1} B^{-1}(pr - A) + \tilde{\Sigma}_\eta^{-1}\tilde{\eta}\big).$$

Again, breaking up the signal into truth and noise ($\tilde{\eta} = f + \tilde{\varepsilon}$), and using the price equation, we can write

$$\begin{aligned}\hat{\mu}_j - pr = {} & \hat{\Sigma}_j \Sigma^{-1}\mu + \hat{\Sigma}_j \tilde{\Sigma}_\eta^{-1}(f + \varepsilon) \\ & + \big(\hat{\Sigma}_j \Sigma_p^{-1} B^{-1} - I\big)(A + Bf + Cx) - \hat{\Sigma}_j \Sigma_p^{-1} B^{-1} A\end{aligned}$$

$$= \hat{\Sigma}_j \Sigma^{-1}\mu + \big(\hat{\Sigma}_j \Sigma_p^{-1}(I - B^{-1}) - I\big)A$$
$$+ \big(\hat{\Sigma}_j \tilde{\Sigma}_\eta^{-1} + \hat{\Sigma}_j \Sigma_p^{-1} - B\big)f + \hat{\Sigma}_j \tilde{\Sigma}_\eta^{-1}\tilde{\varepsilon} + \big(\hat{\Sigma}_j \Sigma_p^{-1} - I\big)Cx.$$

Since μ and A are constants, and ε, f, and x are mutually independent, the variance of this expression is

$$V_1[\tilde{\mu} - pr] = \big(\hat{\Sigma}_j \tilde{\Sigma}_\eta^{-1} + \hat{\Sigma}_j \Sigma_p^{-1} - B\big)\Sigma\big(\hat{\Sigma}_j \tilde{\Sigma}_\eta^{-1} + \hat{\Sigma}_j \Sigma_p^{-1}B - B\big)'$$
$$+ \hat{\Sigma}_j \tilde{\Sigma}_\eta^{-1}\hat{\Sigma}_j + \sigma_x\big(\hat{\Sigma}_j \Sigma_p^{-1} - I\big)CC'\big(\hat{\Sigma}_j \Sigma_p^{-1} - I\big)'.$$

We have assumed that $\tilde{\Sigma}_\eta^{-1} + \Sigma_p^{-1} = \Sigma_\eta^{-1}$. Therefore, the first term $\hat{\Sigma}_j \tilde{\Sigma}_\eta^{-1} + \hat{\Sigma}_j \Sigma_p^{-1} - B = \hat{\Sigma}_j \Sigma_\eta^{-1} - B$, which is the same quantity as in the first term of $V_1[\hat{\mu}_j - pr]$.

Thus, when we subtract one expression from the other,

$$V_1[\hat{\mu}_j - pr] - V_1[\tilde{\mu} - pr]$$
$$= \hat{\Sigma}_j\big(\Sigma_\eta^{-1} - \tilde{\Sigma}_\eta^{-1}\big)\hat{\Sigma}_j - \sigma_x\big(\hat{\Sigma}_j \Sigma_p^{-1}CC'\Sigma_p^{-1}\hat{\Sigma}_j - 2\hat{\Sigma}_j \Sigma_p^{-1}CC'\big).$$

Since $\Sigma_\eta^{-1} = \tilde{\Sigma}_\eta^{-1} + \Sigma_p^{-1}$ and Σ_p^{-1} is positive semi-definite (an inverse variance matrix always is), $\hat{\Sigma}_j(\Sigma_\eta^{-1} - \tilde{\Sigma}_\eta^{-1})\hat{\Sigma}_j$ is positive semi-definite. Thus, the difference is positive semi-definite if $2I - \Sigma_p^{-1}\hat{\Sigma}_j$ is. Since for the investor that learns about prices, Bayes's rule tells us that $\hat{\Sigma}_j^{-1} = \Sigma^{-1} + \Sigma_p^{-1} + \tilde{\Sigma}_\eta^{-1}$, this means that $I - \Sigma_p^{-1}\hat{\Sigma}_j = (\Sigma^{-1} + \tilde{\Sigma}_\eta^{-1})\hat{\Sigma}_j$, which is positive semi-definite. Therefore, $2I - \Sigma_p^{-1}\hat{\Sigma}_j$ is also positive semi-definite.

Thus, the difference in utility from learning conditionally independent information and learning price information is $1/2\,\mathrm{trace}(\hat{\Sigma}_j^{-1}(V_1[\hat{\mu}_j - pr] - [V_1[\tilde{\mu} - pr]))$. Since the expression inside the trace is a product of positive semi-definite matrices, the trace and therefore the difference in expected utilities is positive. *Q.E.D.*

S.6. DISPERSION AND PERFORMANCE RESULTS

In this section, we reprove Propositions 4 and 6 from the main text for dispersion and outperformance in certainty equivalent units, for less restrictive parameter assumptions.

Prove: If $\bar{x}_a$ is sufficiently large, a marginal increase in risk aversion, ρ, increases the difference in expected certainty equivalent returns between informed and uninformed investors, $U_{1I} - U_{1U}$.

Recall that U_{1j} is the expected utility of investor j at the end of period 1 (i.e., after he has chosen his attention allocation but before he receives his signals). Fix j to be for an informed investor and let U_{1U} be the expected utility of an uninformed investor at the end of period 1. Using equation (40) from the Appendix, we have

$$U_{1j} - U_{1U} = \frac{1}{2}\sum_{i=1}^{N} K_{ij}\lambda_i = \frac{1}{2}\sum_{i=1}^{N} K_{ij}\big(\bar{\sigma}_i\big[1 + \big(\rho^2\sigma_x + \bar{K}_i\big)\bar{\sigma}_i\big] + \rho^2\bar{x}_i^2\bar{\sigma}_i^2\big). \tag{S.13}$$

If there is no change in attention allocation when ρ increases marginally, $U_{1j} - U_{1U}$ will increase because of the direct effect of ρ in equation (S.13) and because $\bar{\sigma}_i$ increases in ρ for all i.

It remains to consider the case in which the attention allocation changes after a marginal increase in ρ. This can happen when attention is allocated to multiple risks before the change in ρ. From the proof of Proposition 2, we know that after a marginal increase in ρ, λ_i will be higher for all risks that receive attention. Since $\sum_{i=1}^{N} K_{ij} = K$ both before and after the increase in ρ, $U_{1j} - U_{1U}$ increases.

Prove: If $\bar{x}_a$ is sufficiently large, a marginal increase in risk aversion, ρ, increases the dispersion in expected certainty equivalent returns $E[(U_{1j} - \bar{U}_1)^2]$, where $\bar{U}_1 \equiv \int U_{1j}\,dj$.

The average certainty equivalent return is $\bar{U}_1 = \lambda U_{1I} + (1-\lambda)U_{1U}$, since all informed agents I have the same expected utility and all uninformed agents U have the same expected utility. For an informed agent j, $(U_{1j} - \bar{U}_1) = (1-\lambda)(U_{1I} - U_{1U})$. From the previous result, we know that an increase in ρ increases $U_{1I} - U_{1U}$. Thus, dispersion in certainty equivalent returns increases as well.

S.7. INTERACTION EFFECTS: RISK AVERSION AND AGGREGATE RISK

This section shows that the effect of aggregate risk on attention, dispersion, and returns is greater when risk aversion is high. The testable prediction that follows from these results is that the effect of aggregate volatility should be greater in recessions, because these are times when the price of risk (governed by risk aversion in the model) is high. Those empirical findings are presented in the next section.

RESULT S.7.1: Aggregate volatility σ_n has a larger effect on the marginal utility of an additional unit of precision in the signal about aggregate risk λ_n, when risk aversion is high: $\frac{\partial^2\lambda_n}{\partial\rho\,\partial\sigma_n} > 0$.

Recall that

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

$$\bar{\sigma}_i = \frac{1}{\sigma_i^{-1} + \bar{K}_i + \dfrac{\bar{K}_i^2}{\rho^2\sigma_x}}. \tag{S.15}$$

From equations (S.14) and (S.15), it is immediate that $\partial^2\lambda_i/\partial\rho\,\partial\sigma_n = 0$ for all $i \neq n$. For λ_n,

$$\frac{\partial^2\lambda_n}{\partial\rho\,\partial\sigma_n} = \frac{\partial^2\bar{\sigma}_n}{\partial\rho\,\partial\sigma_n} + 4\rho\bar{\sigma}_n\big(\sigma_x + \bar{x}_n^2\big)\frac{\partial\bar{\sigma}_n}{\partial\sigma_n} + 2\big(\rho^2\big(\sigma_x + \bar{x}_n^2\big) + \bar{K}_n\big)\left(\frac{\partial\bar{\sigma}_n}{\partial\sigma_n}\frac{\partial\bar{\sigma}_n}{\partial\rho} + \bar{\sigma}_n\frac{\partial^2\bar{\sigma}_n}{\partial\rho\,\partial\sigma_n}\right).$$

All of the derivatives on the right-hand side of this expression can be evaluated using equation (S.15):

$$\frac{\partial\bar{\sigma}_n}{\partial\sigma_n} = \left(\frac{\bar{\sigma}_n}{\sigma_n}\right)^2 > 0, \tag{S.16}$$

$$\frac{\partial\bar{\sigma}_n}{\partial\rho} = \frac{2\bar{K}_n^2\bar{\sigma}_n^2}{\rho^3\sigma_x} > 0, \tag{S.17}$$

$$\frac{\partial^2\bar{\sigma}_n}{\partial\rho\,\partial\sigma_n} = \left(\frac{4\bar{K}_n^2\bar{\sigma}_n}{\rho^3\sigma_x}\right)\frac{\partial\bar{\sigma}_n}{\partial\sigma_n} > 0. \tag{S.18}$$

Rearranging these expressions delivers the result.

RESULT S.7.2: *If risk aversion ρ is sufficiently high, then volatility σ_n has a larger effect on portfolio return dispersion V_{nn} when risk aversion is higher: $\partial^2 V_{nn}/\partial\rho\,\partial\sigma_n > 0$.*

Portfolio excess return dispersion is described in equation (48). We can write this equation out as

$$E\big[\big((q_j - \bar{q})'(\tilde{f} - \tilde{p}r)\big)^2\big] = \sum_{l=1}^{n} 6(K_{lj} - \bar{K}_l)^2 V_{ll}\bar{x}_l^2\bar{\sigma}_l^2 + (K_{lj} - \bar{K}_l)^2\rho^2\bar{x}_l^4\bar{\sigma}_l^4 + 3(K_{lj} - \bar{K}_l)^2\frac{V_{ll}^2}{\rho^2} + \bar{x}_l^2\bar{\sigma}_l^2 K_{lj}.$$

The expression in the summation only depends on σ_n when $l = a$, so we can restrict attention to

$$F_1 \equiv 6(K_{nj} - \bar{K}_n)^2 V_{nn} \bar{x}_n^2 \bar{\sigma}_n^2 + (K_{nj} - \bar{K}_n)^2 \rho^2 \bar{x}_n^4 \bar{\sigma}_n^4 + 3(K_{nj} - \bar{K}_n)^2 \frac{V_{nn}^2}{\rho^2} + \bar{x}_n^2 \bar{\sigma}_n^2 K_{nj}, \tag{S.19}$$

knowing that

$$\frac{\partial^2 E\big[\big((q_j - \bar{q})'(\tilde{f} - \tilde{p}r)\big)^2\big]}{\partial\rho\,\partial\sigma_n} = \frac{\partial^2 F_1}{\partial\rho\,\partial\sigma_n}.$$

To evaluate $\partial^2 F_1/\partial\rho\,\partial\sigma_n$ six derivatives are needed. Three of them are given in equations (S.16), (S.17), and (S.18). The other three are

$$\frac{\partial V_{nn}}{\partial\sigma_n} = \left(\frac{\bar{\sigma}_n}{\sigma_n}\right)^2 \big[1 + 2\big(\rho^2\sigma_x + \bar{K}_n\big)\bar{\sigma}_n\big] > 0,$$

$$\frac{\partial V_{nn}}{\partial\rho} = 2\rho\sigma_x\bar{\sigma}_n^2\left[1 + \frac{\bar{K}_n^2}{\rho^4\sigma_x^2}\big(1 + 2\big(\rho^2\sigma_x + \bar{K}_n\big)\bar{\sigma}_n\big)\right] > 0,$$

$$\frac{\partial^2 V_{nn}}{\partial\rho\,\partial\sigma_n} = 4\bar{\sigma}_n\left[\rho\sigma_x\left(1 + \frac{\bar{K}_n^2}{\rho^4\sigma_x^2}\right) + \frac{3\bar{\sigma}_n\bar{K}_n^2\big(\rho^2\sigma_x + \bar{K}_n\big)}{\rho^3\sigma_x}\right]\frac{\partial\bar{\sigma}_n}{\partial\sigma_n} > 0.$$

We will now evaluate the cross-partial derivative of each of the four terms on the right-hand side of equation (S.19) with respect to ρ and σ_n. The cross-partial deriviative of the first, second, and fourth terms are all strictly positive because they are all products of strictly positive constants and positive powers of V_{nn} and $\bar{\sigma}_n$. The sign of the third term is not clear on inspection because ρ has a negative power. The cross-partial derivative for the third term is

$$\frac{\partial^2}{\partial\rho\,\partial\sigma_n}\left[\frac{V_{nn}^2}{\rho^2}\right] = \frac{2}{\rho^2}\frac{\partial V_{nn}}{\partial\sigma_n}\frac{\partial V_{nn}}{\partial\rho} + \frac{2}{\rho^2}V_{nn}\frac{\partial^2 V_{nn}}{\partial\rho\,\partial\sigma_n} - \frac{4}{\rho^3}V_{nn}\frac{\partial V_{nn}}{\partial\sigma_n}.$$

In sum, this cross-partial has seven terms, six of which are unambiguously positive. The seventh term may be positive or negative. But for ρ sufficiently high, the negative term, which is multiplied by $1/\rho^3$, vanishes to zero and other positive terms become more positive. Thus, there must exist some finite $\bar{\rho}$ such that $\forall\rho > r\bar{h}o$, $\partial^2 V_{nn}/\partial\rho\,\partial\sigma_n > 0$.

RESULT S.7.3: If the supply of the aggregate risk σ_n is sufficiently large, then volatility has a larger effect on expected returns when risk aversion is high: $\partial^2 E[(q_j - \bar{q})'(f - pr)]/\partial\rho\,\partial\sigma_n > 0$.

Equations (47) and (48) from the paper provide that

$$E\big[(q_j - \bar{q})'(f - pr)\big] = \rho\,\mathrm{Tr}\big(\bar{x}'\bar{\Sigma}\Delta\bar{\Sigma}\bar{x}\big) + \frac{1}{\rho}\mathrm{Tr}(\Delta V). \tag{S.20}$$

This expression depends on σ_i through $\bar{\Sigma}$, Δ, and V. All of these objects are diagonal matrices. Using equations (25), (32), and (28) from the paper, their ith diagonal elements are, respectively,

$$\bar{\sigma}_i = \frac{1}{\sigma_i^{-1} + \bar{K}_i + \dfrac{\bar{K}_i^2}{\rho^2\sigma_x}},$$

$$(V)_{ii} = \bar{\sigma}_i\big[1 + \big(\rho^2\sigma_x + \bar{K}_i\big)\bar{\sigma}_i\big],$$

$$(\Delta)_{ii} = K_{ij} - \int_j K_{ij}\,dj.$$

Since K_{ij} is the same for all skilled investors, $K_{ij} = 0$ for all unskilled investors and the fraction of skilled investors is χ, the third equation can be written as

$$(\Delta)_{ii} = (1 - \chi)K_{ij}.$$

Equation (S.20) provides an expression for portfolio excess return. For current purposes, we can restrict attention to the terms in this expression that depend on σ_n. This leaves the following:

$$F_2 \equiv \rho\bar{x}_n^2\bar{\sigma}_n^2(\Delta)_{nn} + \frac{1}{\rho}(\Delta)_{nn}V_{nn}.$$

Since the omitted terms do not depend on σ_n, we know that

$$\frac{\partial^2 E\big[(q_j - \bar{q})'(f - pr)\big]}{\partial\rho\,\partial\sigma_n} = \frac{\partial^2 F_2}{\partial\rho\,\partial\sigma_n}.$$

The cross-partial derivative of F_2 with respect to ρ and σ_n is

$$\begin{aligned}\frac{\partial^2 F_2}{\partial\rho\,\partial\sigma_n} &= 2\bar{x}_n^2\bar{\sigma}_n + (\Delta)_{nn}\frac{\partial\bar{\sigma}_n}{\partial\sigma_n} + 2\rho\bar{x}_n^2(\Delta)_{nn}\frac{\partial\bar{\sigma}_n}{\partial\rho}\frac{\partial\bar{\sigma}_n}{\partial\sigma_n} \\ &\quad + 2\bar{x}_n^2\bar{\sigma}_n(\Delta)_{nn}\frac{\partial^2\bar{\sigma}_n}{\partial\rho\,\partial\sigma_n} - \frac{1}{\rho^2}(\Delta)_{nn}\frac{\partial V_{nn}}{\partial\sigma_n} + \frac{1}{\rho}(\Delta)_{nn}\frac{\partial^2 V_{nn}}{\partial\rho\,\partial\sigma_n}.\end{aligned}$$

There are five positive terms and one negative. The only negative term on the right-hand side of this expression is $-\frac{1}{\rho^2}(\Delta)_{nn}\frac{\partial V_{nn}}{\partial\sigma_n}$. So this cross-partial derivative is positive, as long as that one term is sufficiently small relative to the other terms.

A sufficient conditions for this to be positive is that the supply of aggregate risk $\bar{x}_n$ is sufficiently large. Notice that V_{nn} and Δ_{nn} do not depend on $\bar{x}_n$. Since this is a partial derivative, we are holding the learning choices K_i and $\bar{K}$ fixed. However, some of the positive terms are increasing in $\bar{x}_n$. Thus, for some level of $\bar{x}_n$, the positive terms must be larger and the cross-partial derivative must be positive.

S.8. CYCLICALITY OF IDIOSYNCRATIC RISK

There are asset pricing papers that adopt the same notion of idiosyncratic risk in returns that we do and come to different conclusions about its cyclicality. One of the most well-known papers on idiosyncratic risk is Campbell, Lettau, Malkiel, and Xu (2001, CLMX). CLMX found that aggregate market volatility, industry volatility, and firm-specific volatility all rise in recessions but they provided no standard errors on these increases. As such, while our results on aggregate market volatility are consistent with theirs, we find less support for the countercyclicality of idiosyncratic volatility since we find no statistically significant difference across the cycle for stock-specific variance. This appendix explores the differences between our paper and CLMX in detail. There are several. The first four are minor measurement issues. The last point is that the sample periods differ, and that seems to account for most of the difference in our answers.

First, our methodology is slightly different. We use the more standard CAPM, whereas they set all firms' market beta equal to 1 in order to avoid problems with measuring betas. As a result of this procedure, their measures of firm-specific and industry risk are upward biased (see their equations (15) and (16)). Moreover, because market variance is higher in recessions and the cross-sectional dispersion of betas is higher in recessions, the bias increases in recessions.

Second, to calculate our volatility measures, we use twelve-month rolling-window regressions of monthly data; CLMX, in turn, used daily data.

Third, because of their methodology, CLMX were forced to compute value-weighted volatility measures, while we focus on equally weighted volatility measures. Nevertheless, we show that our results are robust to using value weighting in our sample.

Fourth, we control for important factor exposures that may potentially vary with the business cycle.

Fifth, and most importantly, the sample periods are different. The CLMX results are for 1962–1997, while ours are for 1980–2005. It turns out that the countercyclicality of the idiosyncratic volatility of stock returns is a fragile result. In Table S.IV, we provide the time-series results of Table I in the main text across different sample periods. The top panel uses equal weight-

TABLE S.IV

RISK IN INDIVIDUAL STOCKS: COMPARISON WITH CLMX (2001)[a]

	Aggregate Risk				Idiosyncratic Risk			
	1980–2005	1927–2008	1962–1997	1962–2005	1980–2005	1927–2008	1962–1997	1962–2005
				Equal-weighted Results				
Recession	1.253	1.766	1.523	1.665	0.128	0.153	0.564	0.741
	(0.705)	(0.688)	(0.303)	(0.427)	(0.976)	(0.598)	(0.444)	(0.691)
MKT	−3.879	−3.528	2.458	−0.230	−1.695	−3.164	4.804	1.237
	(3.125)	(3.282)	(2.174)	(2.116)	(2.941)	(2.673)	(2.747)	(2.891)
SMB	10.028	16.143	2.875	6.730	11.538	14.238	0.959	6.673
	(4.116)	(6.234)	(2.734)	(2.670)	(4.771)	(5.133)	(4.092)	(4.498)
HML	3.448	−1.834	−1.956	1.660	8.873	4.649	2.789	5.100
	(6.364)	(6.869)	(3.175)	(4.289)	(7.878)	(5.497)	(5.240)	(7.293)
MOM	−4.897	−10.755	−8.479	−6.023	−0.967	−2.927	−4.806	−1.944
	(2.500)	(4.424)	(2.542)	(2.090)	(3.686)	(3.323)	(3.043)	(3.528)
Constant	7.030	6.897	6.351	6.667	13.050	10.173	10.785	11.444
	(0.224)	(0.257)	(0.182)	(0.178)	(0.269)	(0.268)	(0.256)	(0.281)
Observations	309	979	432	528	309	979	432	528

(*Continues*)

TABLE IV—*Continued*

	Aggregate Risk				Idiosyncratic Risk			
	1980–2005	1927–2008	1962–1997	1962–2005	1980–2005	1927–2008	1962–1997	1962–2005
	Value-weighted Results							
Recession	0.958	1.742	1.436	1.395	0.417	0.320	0.754	0.650
	(0.440)	(0.556)	(0.288)	(0.324)	(0.557)	(0.285)	(0.209)	(0.372)
MKT	−1.685	−0.694	3.450	1.427	−1.510	−1.496	1.416	−0.267
	(2.659)	(2.763)	(2.022)	(1.855)	(2.026)	(1.411)	(1.161)	(1.485)
SMB	8.797	11.412	1.205	4.665	9.205	6.002	1.660	5.132
	(3.437)	(5.473)	(2.451)	(2.400)	(4.062)	(2.527)	(1.811)	(2.989)
HML	2.043	−1.003	−2.620	0.096	6.960	3.190	1.164	3.409
	(6.013)	(5.785)	(2.920)	(4.105)	(7.172)	(2.949)	(2.266)	(5.397)
MOM	−4.994	−8.733	−8.110	−5.341	−0.514	−1.500	−3.227	−1.120
	(2.242)	(3.828)	(2.306)	(1.899)	(3.308)	(1.784)	(1.508)	(2.614)
Constant	5.365	5.236	4.778	5.039	6.959	5.764	5.821	6.259
	(0.231)	(0.195)	(0.173)	(0.171)	(0.215)	(0.127)	(0.112)	(0.166)
Observations	309	979	432	528	309	979	432	528

[a]The variable definitions and the empirical specification are as defined in the caption of Table I, panel A in the main text.

ing, the bottom panel value weighting. The left four columns show that aggregate risk is significantly higher in recessions in all eight specifications. The same is not true for idiosyncratic risk in the four rightmost columns. Using equal weighting, we find that idiosyncratic volatility is *not* statistically higher in recessions for the 1962–1997 period, or in our sample, or for the full sample 1927–2008. Using value weighting, our methodology recovers the CLMX findings of higher idiosyncratic risk in recessions for the sample 1962–1997. This result suggests that the differences between our and CLMX results are not (primarily) driven by methodological differences. What seems more important is time period. As in our 1980–2005 sample, we find no statistical difference in idiosyncratic volatility between expansions and recessions for the full 1927–2008 sample using value weighting. The inclusion of the late 1990s in our sample, a period of high stock-specific volatility and economic expansion, weakens the evidence for countercyclical idiosyncratic risk. Indeed, the last column shows that even the value-weighted idiosyncratic risk measure becomes insignificant once the CLMX sample is extended from 1997 to the end of our sample, 2005. In sum, there seems to be no consistent evidence for countercyclical variation in idiosyncratic stock risk, a message that also comes out of our 1980–2005 results (Table I of the main text).

S.9. NONLINEAR VOLATILITY EFFECTS

In this section, we expand on the volatility results discussed in Section 3 in the main text. Specifically, we estimate a nonlinear volatility specification in which we include top-5%, 5–10%, 10–30% , and 30–70% volatility indicator variables. The omitted volatility category is the bottom 30%. We first present results without the recession indicator variable in column 1 and then results with the recession variable added in column 2. We also consider an additional specification in which we interact the continuous volatility measure with the recession indicator and interact volatility with 1 minus the recession indicator. This specification asks whether the effect of volatility on our outcome variables is different in recessions and expansions, and its results are reported in column 3. All regressions have our usual set of control variables. The results for *Ftiming* and *Fpicking* are in Table S.V, the results for return dispersion are in Table S.VI, and the results for the four-factor alphas are in Table S.VII.

With the nonlinear specification, we find nicely monotonic results. *Fpicking* is lower when volatility is high, and more so at the top of the temporal volatility distribution than at the bottom (column 4). The effect is still negative for the 5–10%, but not for low-volatility periods. For *Ftiming*, we find a positive (albeit insignificant) volatility effect at the top of the volatility distribution,

TABLE S.V

NONLINEAR VOLATILITY EFFECTS: *Ftiming* AND *Fpicking*

	Ftiming			*Fpicking*		
	(1)	(2)	(3)	(4)	(5)	(6)
Recession		0.011			−0.607	
		(0.003)			(0.129)	
Vol. Top-5	0.001	−0.002		−0.414	−0.190	
	(0.003)	(0.004)		(0.140)	(0.140)	
Vol. 5–10	−0.000	−0.001		−0.120	−0.101	
	(0.005)	(0.005)		(0.201)	(0.198)	
Vol. 10–30	−0.006	−0.006		0.258	0.282	
	(0.003)	(0.003)		(0.181)	(0.179)	
Vol. 30–70	0.000	0.000		0.528	0.530	
	(0.003)	(0.003)		(0.165)	(0.165)	
Vol. ∗ Rec.			0.014			−4.854
			(0.014)			(0.804)
Vol. ∗ Exp.			−0.011			−1.448
			(0.015)			(0.723)
log(Age)	−0.001	−0.001	−0.001	0.437	0.444	0.437
	(0.001)	(0.001)	(0.001)	(0.062)	(0.061)	(0.062)
log(TNA)	−0.001	−0.001	−0.001	−0.121	−0.126	−0.124
	(0.000)	(0.000)	(0.000)	(0.030)	(0.029)	(0.030)
Expenses	−0.223	−0.203	−0.224	99.900	98.547	97.605
	(0.220)	(0.220)	(0.218)	(11.076)	(11.041)	(11.221)
Turnover	−0.004	−0.004	−0.004	−0.259	−0.259	−0.258
	(0.001)	(0.001)	(0.001)	(0.063)	(0.063)	(0.063)
Flow	−0.011	−0.010	−0.011	0.686	0.638	0.710
	(0.011)	(0.011)	(0.011)	(0.645)	(0.647)	(0.650)
Load	0.012	0.008	0.010	−10.244	−9.990	−10.091
	(0.022)	(0.022)	(0.022)	(1.931)	(1.933)	(1.949)
Flow Vol.	−0.001	−0.005	−0.002	6.205	6.425	6.465
	(0.017)	(0.017)	(0.017)	(1.034)	(1.037)	(1.037)
Constant	0.001	0.000	0.000	2.858	2.881	3.204
	(0.002)	(0.002)	(0.002)	(0.113)	(0.114)	(0.107)
Observations	221,488	221,488	221,488	165,029	165,029	165,029
R-squared	0.001	0.001	0.001	0.003	0.003	0.002

TABLE S.VI

NONLINEAR VOLATILITY EFFECTS: DISPERSION

	Return Dispersion		
	(1)	(2)	(3)
Recession		0.278	
		(0.141)	
Vol. Top-5	0.646	0.597	
	(0.257)	(0.274)	
Vol. 5–10	0.432	0.427	
	(0.337)	(0.335)	
Vol. 10–30	−0.136	−0.138	
	(0.191)	(0.189)	
Vol. 30–70	0.039	0.041	
	(0.207)	(0.207)	
Vol. ∗ Rec.			3.533
			(1.168)
Vol. ∗ Exp.			2.292
			(1.113)
log(Age)	−0.104	−0.106	−0.107
	(0.018)	(0.018)	(0.018)
log(TNA)	0.034	0.036	0.034
	(0.011)	(0.010)	(0.010)
Expenses	24.856	25.158	24.701
	(2.751)	(2.715)	(2.647)
Turnover	0.074	0.074	0.074
	(0.015)	(0.014)	(0.014)
Flow	−0.280	−0.268	−0.289
	(0.224)	(0.224)	(0.216)
Load	−3.429	−3.496	−3.490
	(0.549)	(0.542)	(0.541)
Flow Vol.	2.013	1.939	1.953
	(0.272)	(0.287)	(0.260)
Constant	1.837	1.828	1.699
	(0.168)	(0.169)	(0.124)
Observations	227,141	227,141	227,141
R-squared	0.082	0.083	0.076

TABLE S.VII

NONLINEAR VOLATILITY EFFECTS: PERFORMANCE

	Four-Factor Alpha		
	(1)	(2)	(3)
Recession		0.092	
		(0.035)	
Vol. Top-5	0.145	0.120	
	(0.058)	(0.061)	
Vol. 5–10	0.065	0.062	
	(0.105)	(0.105)	
Vol. 10–30	0.052	0.051	
	(0.053)	(0.053)	
Vol. 30–70	−0.008	−0.007	
	(0.038)	(0.038)	
Vol. ∗ Rec.			1.092
			(0.205)
Vol. ∗ Exp.			0.388
			(0.296)
log(Age)	−0.031	−0.032	−0.031
	(0.006)	(0.006)	(0.006)
log(TNA)	0.017	0.018	0.017
	(0.003)	(0.003)	(0.003)
Expenses	−6.979	−6.819	−6.705
	(0.665)	(0.663)	(0.651)
Turnover	−0.080	−0.080	−0.080
	(0.008)	(0.008)	(0.008)
Flow	1.369	1.376	1.370
	(0.097)	(0.095)	(0.096)
Load	−0.215	−0.250	−0.239
	(0.127)	(0.123)	(0.127)
Flow Vol.	1.326	1.287	1.306
	(0.109)	(0.107)	(0.104)
Constant	−0.073	−0.077	−0.087
	(0.019)	(0.019)	(0.026)
Observations	224,130	224,130	224,130
R-squared	0.057	0.058	0.057

as predicted by the theory, but not in the rest of the volatility distribution (column 1). These results highlight the nonlinearity: we find evidence of our volatility channel, but it is concentrated in high-volatility periods. Once the recession indicator is added in columns 2 and 5, we lose statistical significance. As we argued in the paper, recessions are often (but not always) periods of high volatility and the recession effect takes explanatory power and statistical significance away from the volatility effect. Yet, the monotonicity of the volatility effect remains. Columns 3 and 6 explore the interaction between volatility and recession further. Column 3 shows that volatility has the predicted positive effect on *Ftiming* in recessions, but not in expansions. The point estimate remains insignificant, but the effect of volatility in recessions on *Ftiming* has a higher t-statistic than in any of the other specifications. Volatility has a negative effect on *Fpicking*, and the effect is four times larger in absolute value in recessions than in expansions. While both volatility coefficients in column 6 are significant, the t-statistic is three times larger in recessions.

Table S.VI shows that there is a positive effect of high volatility on return dispersion (column 1). Again, the effect is concentrated in the top-10% volatility periods. The monotonicity of the effect as well as its statistical strength are preserved once we add a recession indicator in column 2. Column 3 shows stronger effect of volatility on dispersion in expansions than in recessions: the point estimate is 50% larger, but both effects are significant. Again, this suggests that there is a separate role for volatility outside recessions, but that the volatility effect is strongest in recessions.

Finally, Table S.VII shows similar results for the four-factor alpha measure. Unreported results for CAPM and three-factor alphas are along the same lines. High-volatility periods are associated with statistically and economically significant outperformance (column 1). The nonlinear volatility effect survives inclusion of a recession indicator (column 2). The effect of volatility is more than twice as strong in recessions as in expansions. It is highly significant in recessions, but loses significance in expansions.

S.10. RESULTS WITH MANAGERS AS THE UNIT OF OBSERVATION

In our final set of results (Table S.VIII), we measure fundamentals-based market-timing ability (*Ftiming*), fundamentals-based stock-picking ability (*Fpicking*), portfolio dispersion (*Dispersion*), and the alpha (4-*Factor Alpha*), all at the manager level. We find that for these results, the distinction between measuring the behavior of a manager or the behavior of a fund makes little difference quantitatively.

TABLE S.VIII

ROBUSTNESS: MANAGERS AS THE UNIT OF OBSERVATION[a]

	Ftiming		*Fpicking*		Return Dispersion		4-Factor Alpha	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Recession	0.008	0.007	−0.701	−0.824	0.105	0.142	0.167	0.141
	(0.003)	(0.002)	(0.130)	(0.132)	(0.031)	(0.024)	(0.038)	(0.035)
log(Age)	−0.002	−0.004	0.460	0.268	0.154	0.017	−0.032	−0.069
	(0.001)	(0.001)	(0.066)	(0.065)	(0.032)	(0.021)	(0.006)	(0.008)
log(TNA)	−0.000	−0.000	−0.126	−0.139	−0.131	−0.093	0.007	0.003
	(0.000)	(0.000)	(0.032)	(0.033)	(0.018)	(0.011)	(0.004)	(0.005)
Expenses	0.130	0.060	127.222	45.894	43.920	13.241	−8.225	−10.590
	(0.105)	(0.146)	(13.972)	(13.867)	(6.012)	(4.504)	(0.794)	(1.138)
Turnover	−0.005	−0.004	−0.287	0.210	−0.127	−0.014	−0.081	-0.035
	(0.002)	(0.002)	(0.077)	(0.090)	(0.030)	(0.020)	(0.010)	(0.009)
Flow	−0.009	−0.016	1.037	1.186	0.154	−0.374	1.832	1.483
	(0.009)	(0.009)	(0.613)	(0.554)	(0.121)	(0.101)	(0.097)	(0.089)
Load	−0.009	−0.069	−16.064	−5.832	−4.674	−0.231	−0.426	0.097
	(0.017)	(0.023)	(2.393)	(2.307)	(1.284)	(0.809)	(0.151)	(0.172)
Fixed Effect	N	Y	N	Y	N	Y	N	Y
Constant	−0.002	−0.002	2.966	2.977	1.438	1.447	−0.045	−0.043
	(0.001)	(0.001)	(0.072)	(0.068)	(0.026)	(0.008)	(0.024)	(0.021)
Observations	332,676	332,676	249,942	249,942	332,776	332,776	332,776	332,776

[a]The dependent variables are fundamentals-based market-timing ability (*Ftiming*), fundamentals-based stock-picking ability (*Fpicking*), return dispersion (*Portfolio Dispersion*), and the four-factor alpha (4-*Factor Alpha*), all of which are tracked at the manager level. Columns with a "Y" include manager-fixed effects. The independent variables, the sample period, and the standard error calculations are the same as in Table II.

REFERENCES

ADMATI, A. (1985): "A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets," *Econometrica*, 53 (3), 629–657. [10]

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. [21]

KACPERCZYK, M., S. VAN NIEUWERBURGH, AND L. VELDKAMP (2014): "Time-Varying Fund Manager Skill," *Journal of Finance*, 69, 1455–1484. [10]

MAĆKOWIAK, B., AND M. WIEDERHOLT (2009): "Optimal Sticky Prices Under Rational Inattention," *American Economic Review*, 99 (3), 769–803. [13]

——— (2015): "Business Cycle Dynamics Under Rational Inattention," *Review of Economic Studies*, 82 (4), 1502–1532. [13]

MONDRIA, J. (2010): "Portfolio Choice, Attention Allocation, and Price Comovement," *Journal of Economic Theory*, 145, 1837–1864. [13]

SIMS, C. (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50 (3), 665–690. [13]

VAN NIEUWERBURGH, S., AND L. VELDKAMP (2010): "Information Acquisition and Under-Diversification," *Review of Economic Studies*, 77 (2), 779–805. [13]

VELDKAMP, L. (2011): *Information Choice in Macroeconomics and Finance*. Princeton, NJ: Princeton University Press. [7]

Imperial College London, South Kensington Campus, London, SW7 2AZ, U.K. and NBER; m.kacperczyk@imperial.ac.uk; http://www3.imperial.ac.uk/people/m.kacperczyk,

Dept. of Finance, Stern School of Business, New York University, 44 W. 4th Street, New York, NY 10012, U.S.A., NBER, and CEPR; svnieuwe@stern.nyu.edu; http://www.stern.nyu.edu/~svnieuwe,

and

Dept. of Economics, Stern School of Business, New York University, 44 W. 4th Street, New York, NY 10012, U.S.A., NBER, and CEPR; lveldkam@stern.nyu.edu; http://www.stern.nyu.edu/~lveldkam.

Co-editor Lars Peter Hansen handled this manuscript.

Manuscript received February, 2013; final revision received October, 2015.