The Review of Asset Pricing Studies

Investor Information Choice with Macro and Micro Information

Paul Glasserman
Columbia Business School, USA

Harry Mamaysky
Columbia Business School, USA

We develop a model of information and portfolio choice in which ex ante identical investors choose to specialize because of fixed attention costs required in learning about securities. Without this friction, investors would invest in all securities and would be indifferent across a wide range of information choices. When securities' dividends depend on an aggregate (macro) risk factor and idiosyncratic (micro) shocks, fixed attention costs lead investors to specialize in either macro or micro information. Our results favor Samuelson's dictum that markets are more micro than macro efficient. We derive testable predictions from our model and find empirical support for our predictions in specialization by U.S. equity mutual funds. (*JEL* G12, G14, G23)

Received December 27, 2019; editorial decision February 1, 2022 by Editor Thierry Foucault. Authors have furnished an Internet Appendix, which is available on the Oxford University Press Web site next to the link to the final published paper online.

Skilled investors specialize. Fund managers typically invest in stocks or bonds, but not both. Some look for value stocks and others look for growth stocks. Equity analysts focus on a sector and follow a limited number of companies. Some investors focus on marketwide information, whereas others gather company-specific information. Among hedge funds, for example, global macro funds look for broad trends, whereas equity long-short funds seek to hedge out changes in the overall market.

We thank Jiashu Sun for excellent research assistance. We thank Thierry Foucault (editor) and two anonymous referees for their valuable suggestions. We thank Patrick Bolton, Charles Calomiris, Larry Glosten, Bob Hodrick, Gur Huberman, Tomasz Piskorski, Dimitri Vayanos, and Laura Veldkamp, as well as seminar participants at the University of Amsterdam, Johns Hopkins University, the Federal Reserve Bank of New York, Columbia University, and the Copenhagen Business School FRIC conference for valuable comments. This is a revised version of a paper formerly titled "Market Efficiency with Micro and Macro Information." Send correspondence to Harry Mamaysky, hm2646@columbia.edu.

The Review of Asset Pricing Studies 13 (2023) 1–52

https://doi:10.1093/rapstu/raac009 Advance Access publication 14 March 2022

An evident reason for specialization is the time and effort required to learn about each investment opportunity. Yet a common prediction of many models of investor decisions, including all representative agent models, is that all investors hold the same portfolio, or at least that differences in holdings may be ignored in determining asset prices.

We develop a model of investor information acquisition and portfolio selection in which ex ante identical investors choose to specialize in equilibrium. Their decision to specialize is driven by fixed attention costs associated with investing in each security; these fixed costs are the key friction that drives our analysis. Moreover, the resultant specialization takes a specific form: some investors acquire only macro information, whereas other learn only about a single stock, which alternatively could be interpreted as a single sector. The impact of investor specialization is reflected in equilibrium asset prices.

We contrast our specialization results with a benchmark framework that lacks our key friction. In the benchmark framework, investors first decide whether to pay to become informed or to remain uninformed. Becoming informed grants an investor a limited capacity to acquire information about dividends beyond what is available in security prices. An informed investor chooses a consideration set of securities and a set of signals, subject to the capacity constraint, and then sets demands for securities in the consideration set based on the signals. We assume normally distributed payoffs and signals and CARA preferences.

Within this benchmark framework, we show that informed investors always prefer larger consideration sets. This conclusion is not obvious because with a smaller consideration set an informed investor can potentially acquire more precise signals about fewer securities. Nevertheless, we show that no investor has reason to specialize. We show this first in a general model of security payoffs and then apply the result when the payoffs have a factor structure, in which each stock's payoff depends on a common (macro) factor and a stock-specific (micro) shock. Investors can trade the macro factor through an index fund.

We then introduce our key friction, the fixed attention costs for information acquisition. We introduce an important asymmetry between macro and micro information. With individual stocks, investors incur a fixed attention cost (consuming part of their information capacity) for each stock they consider. With the index fund, investors can costlessly make inferences from the price of the fund; they incur a fixed attention cost only in acquiring information *beyond* the price.

Our asymmetric treatment of micro and macro information is motivated by the observation that information about the overall level of the stock market is widely available and widely discussed in the media. In contrast, most people know little about most public companies; acquiring even basic information about individual stocks consumes some attention.

With these frictions, we show that specialization is a necessary condition for an informative equilibrium, meaning an equilibrium in which informed investors collectively learn something about every security's payoff. Informed investors concentrate in two types: the macro-informed who learn about the macro factor and trade only in the index fund; and the micro-informed, who learn about a single stock and trade in that stock and the index fund. An informative equilibrium may also contain *oblivious* investors, who trade stocks without learning about fundamentals. The main alternative to an informative equilibrium is a scenario in which *all* investors are oblivious, resulting in prices uncorrelated with dividends.

Having established that specialization is a necessary property for an informative equilibrium, we then show that such an equilibrium indeed exists through an explicit construction of market-clearing prices and optimal information and portfolio choices by investors. Security prices are not fully revealing due to the presence of exogenously specified supply shocks. We identify restrictions on the cost of becoming informed and the attention cost parameters under which all investors endogenously choose to be macro-informed, micro-informed, or uninformed investors who invest solely in the index fund. In particular, no investor will choose to be oblivious.

We calibrate the model's parameters to plausible values using market data. With this calibration, we compare measures of our market's macro-efficiency and micro-efficiency, based on the informativeness of prices. Interestingly, over a wide parameter range, the model supports what Jung and Shiller (2005) call Samuelson's dictum, the idea that the stock market is more micro-efficient than macro-efficient. We analytically characterize the boundary of the parameter region between micro and macro efficiency and show that the tendency toward micro efficiency is a general result independent of most of the model's parameters.

This outcome is perhaps surprising in our setting because we attach a *lower* attention cost to macro information than micro information. But this very asymmetry drives the result by keeping the uninformed out of the market for individual stocks, thus leading to specialization. Specialization allows micro informed to be the sole liquidity providers for micro supply shocks, whereas macro informed share in liquidity provision for macro supply shocks with other agents who learn about macro fundamentals from the index price. This creates an incentive to become micro-informed, despite the higher attention cost associated with micro information. At the same time, the higher fixed cost of micro information forces micro prices to be more informative than the index price to allow the micro informed to fully utilize the information content of their signals.

We test predictions from our model through an analysis of U.S. equity mutual fund holdings, focusing on the question of investor specialization. We think of fund managers as informed investors who incur fixed attention costs in adding stocks to their portfolios. To control for the effect of ordinary transaction costs, we introduce a novel measure, *percent bought new* (PBN), which looks at the number of stocks bought by a fund in a quarter and measures the proportion that are new to the fund, relative to the proportion of all stocks in which the fund is not invested. In the absence of specialization, we would expect a PBN close to one.

Our model makes four predictions that are testable in the fund data: (1) If there were no attention costs, mutual funds would prefer larger consideration sets; (2) because of attention costs, mutual funds typically have low PBNs; (3) because of a declining cost in becoming informed, PBNs have decreased over time; and (4) funds with lower PBNs have higher returns. We find that most funds invest in a relatively small number of stocks, which argues in favor of attention costs in (1). We find strong evidence supporting (2)–(4). In the case of (4), we find that funds in the lowest-PBN quintile earn 1.5%–2% higher annualized returns than funds in the highest-PBN quintile. This result is not driven by fund characteristics, but rather reflects the underlying informational disadvantage of large consideration sets.

Our work bears on several lines of research. Our model of limited information capacity draws on the rational inattention literature, as in Sims (2011) and Maćkowiak and Wiederholt (2009). Peng and Xiong (2006) combine an investor attention allocation problem with a distinction between macro and micro information; their model features a representative investor and does not incorporate heterogeneous information choices. Our theoretical conclusions contrast with the work of Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016), which also allows informed (or skilled) investors to choose signals and portfolios. Their equilibrium does not require specialization, and they focus on the case of a symmetric equilibrium in which all skilled investors make identical choices. Some of our indifference results in the benchmark framework generalize a partial equilibrium comparison with fixed (uninformative) prices in Van Nieuwerburgh and Veldkamp (2010), but their work does not consider comparisons with different consideration sets, nor does it consider fixed attention costs.

In Peng and Xiong (2006), investors allocate more attention to sector or marketwide information and less attention to firm-specific information. Their conclusion contrasts with ours (and with the Jung-Shiller discussion of Samuelson's dictum) because in their setting a representative investor diversifies away all micro uncertainty. Gârleanu and Pedersen (2018, 2021) extend the Grossman and Stiglitz (1980) model to link market efficiency and asset management through search costs incurred by investors in selecting fund managers, and they find that micro portfolios are more price-efficient than macro portfolios. The investors in Gârleanu and Pedersen (2018, 2021) "make an all-or-nothing" information choice to acquire an exogenously specified signal about all asset payouts: all their informed investors have the same information by assumption, while our focus is on how informed investors endogenously specialize in their information choices. Schneemeier's

(2015) model predicts greater micro- than macro-efficiency when managers use market prices in their investment decisions. Bhattacharya and O'Hara (2016) and Glosten, Nallareddy, and Zou (2020) study a model with an exchange-traded fund (ETF), as well as macro- and micro-informed agents. The ETF has higher liquidity than its constituent stocks, which prevents its price from equaling that of the underlying security basket. In this setting, ETF prices can be informative about individual stock prospects, a dynamic that is absent in our model in which agents are atomic and trade with no price impact.

Our empirical results are consistent with the finding of Kacperczyk, Sialm, and Zheng (2005) that mutual funds with higher industry concentrations perform better. But our PBN measure is agnostic to a stock's industry. Within a large fund complex, a fund manager may draw on the expertise of the complex's industry analysts, forming a portfolio that is spread across multiple sectors but concentrated within sectors, reflecting the stocks followed by each analyst. The analysis of Kacperczyk et al. (2005) is based on data ending in 1999; ours uses more recent data and examines trends over time. Although Kacperczyk et al. (2005) do not propose a model predicting concentration, they note the likely role of informational advantages, which would be consistent with our work. Our 1.5%–2% return differential between low and high PBN funds is of the same order of magnitude as past findings of mutual fund return predictability.[1] Kacperczyk et al. (2005) document that funds with high industry concentrations earn roughly 1.1-1.2% higher returns than funds with low industry concentrations. Carhart (1997) documents that the risk-adjusted return differential between past winning and past losing funds is around 3.4% per year, though this is caused largely by the poor performance of past losing funds. Daniel et al. (1997) find that the top-performing fund category by their selectivity measure—aggressive-growth funds—owns stocks that outperform a characteristic-based benchmark by 1.5% per year. Ferson and Mo (2016) document a before-fee annual return differential between stocks held by funds with selection ability versus stocks held by funds without selection ability of 0.3% for value weighted and 2.6% for equal weighted risk-adjusted returns.

Our model sheds light on why mutual funds hold concentrated portfolios, and on why they face diminishing returns to scale. Koijen and Yogo (2019) note that "institutions hold a small set of stocks and that the set of stocks that they have held in the recent past (e.g., over the past 3 years) hardly changes over time." In our setting, such behavior is optimal in the presence of fixed attention costs. Berk and Green (2004) and Berk and van Binsbergen (2015) assume that mutual fund managers face decreasing returns to scale, with

[1] The results hold with either gross or net of fee returns, as well as with either raw returns or risk-adjusted returns, where the risk adjustment uses either the market factor or the Fama and French (2015) five-factor model augmented with momentum.

returns diminishing as fund size grows. Berk and Green (2004) offer one potential explanation: "with a sufficiently large fund, a manager will spread his information gathering activities too thin." In our framework, if a large fund approaches liquidity limits for investing in a concentrated portfolio and begins to expand its consideration set in order to grow, the incurred attention costs will use up information capacity and lead to worse performance.

1. Benchmark Model of Portfolio and Information Choice

This section presents a benchmark model against which to compare the main results we will introduce later. The setting can be viewed as a multivariate version of Grossman and Stiglitz (1980), similar to that in Admati and Pfleiderer (1987) and many others, combined with an information capacity constraint similar to constraints used in the rational inattention literature and related work (e.g., Mondria 2010, Peng and Xiong 2006, Sims 2011, and Van Nieuwerburgh and Veldkamp 2009). We show that within the benchmark model, informed investors are indifferent to their choice of information, and they prefer to consider investments in more assets rather than focus their attention on fewer assets. In short, this framework provides no incentive for specialization in information or investments. We establish these results in a general setting and then specialize to a factor structure for assets with a common (macro) factor and idiosyncratic (micro) shocks. We present these results as a contrast to Section 2, where we introduce an information friction that leads to the opposite conclusion, with investors specializing in macro or micro information and investments.

1.1 Indifference to information choice

We assume throughout that all investors have constant absolute risk aversion (CARA) preferences with the risk-aversion parameter $\gamma > 0$: the expected utility of terminal wealth W is

$$J = -\mathrm{E}[e^{-\gamma W}]. \tag{1}$$

A financial market trades N risky securities with with payoffs $Y_1, \ldots, Y_N$ at the end of the period and prices $P_{Y_1}, \ldots, P_{Y_N}$ at the beginning of the period. A single riskless asset pays a gross return of $R > 1$ over the period. At this point, we do not specify how prices are determined. We have in mind a setting in which prices are informative about payoffs, but not fully revealing. We make the following normality assumption, but no assumptions about independence:

(a0) Normality: $(Y_1, P_{Y_1}, \ldots, Y_N, P_{Y_N})$ are jointly normal.

Investors observe market prices costlessly. They may choose to acquire additional information about the security payoffs at a cost $c > 0$. Additional information takes the form of a set $\mathcal{I}$ of random variables, which we call signals. At this point, we can keep the types of signals available general and simply strengthen (a0) as follows:

(a1) Normality with signals: $(Y_1, P_{Y_1}, \ldots, Y_N, P_{Y_N}, \mathcal{I})$ are jointly normal.

We will measure the informativeness of a signal set $\mathcal{I}$ through the resultant reduction in uncertainty about security payoffs. To make this explicit, let

$$\Sigma = \operatorname{cov}[Y_1, \ldots, Y_N | P_{Y_1}, \ldots, P_{Y_N}]$$

and

$$\widehat{\Sigma} = \operatorname{cov}[Y_1, \ldots, Y_N | P_{Y_1}, \ldots, P_{Y_N}, \mathcal{I}].$$

The covariance matrix Σ reflects the payoff uncertainty faced by an investor who conditions on prices. Its determinant $|\Sigma|$ provides a scalar summary of this uncertainty. Similarly, the determinant $|\widehat{\Sigma}|$ measures the posterior uncertainty after observing the signals $\mathcal{I}$ as well as prices. Assuming $|\Sigma| > 0$, we measure the informativeness of $\mathcal{I}$ through the variance reduction ratio $|\widehat{\Sigma}|/|\Sigma|$. The smaller this ratio, the more informative the signal set.[2] In the case of a single security, this ratio becomes $\mathrm{V}[Y|P_Y, \mathcal{I}]/\mathrm{V}[Y|P_Y]$, where $\mathrm{V}[\cdot]$ denotes variance; if Σ and $\widehat{\Sigma}$ are both diagonal matrices, the ratio of their determinants becomes a product of similar terms.

Investors may choose to acquire information at a cost $c > 0$. Paying this cost endows an investor with *capacity* $\kappa \in (0, 1)$. This capacity allows the investor to select any signal set for which

$$\frac{|\widehat{\Sigma}|}{|\Sigma|} \geq \kappa. \tag{2}$$

In other words, the capacity κ limits the informativeness of the signals available to the investor, with smaller κ allowing greater informativeness (more variance reduction).[3] A smaller κ should therefore be interpreted as greater capacity. We will say that a signal set *uses the investor's full capacity* if equality holds in (2).

After choosing a signal set, an informed investor chooses an optimal portfolio. See Figure 1 for a summary of the timeline. We consider an atomistic investor whose choice does not affect prices. To evaluate the resultant

[2] The ratio $|\widehat{\Sigma}|/|\Sigma|$ can be interpreted as a multivariate analog of $1 - R^2$ in a multivariate regression of $(Y_1, \ldots, Y_N)$ on $(P_{Y_1}, \ldots, P_{Y_N}, \mathcal{I})$. A lower bound on the ratio is thus an upper bound on the informativeness of the signals. The determinant ratio also arises naturally when attention capacity is measured through entropy, as in Section 3.1 of Sims (2011).

[3] An alternative formulation would constrain ratios of the form $\mathrm{V}[Y|P_Y, \mathcal{I}]/\mathrm{V}[Y]$, with ex ante uncertainty in the denominator. The alternative formulation implies that learning from prices consumes capacity, whereas in (2) only learning beyond prices consumes capacity. We return to this point at the end of Section 2.1.

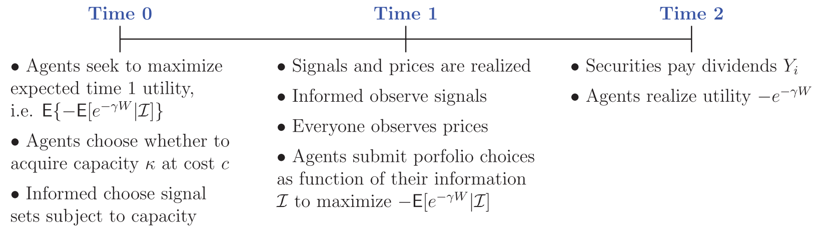

Figure 1
Model timing

expected utility, we need the unconditional mean and covariance of the net payoffs,

$$\mu = \mathrm{E}(Y_1 - RP_{Y_1}, \ldots, Y_N - RP_{Y_N})^\top,$$

and

$$\Psi = \mathrm{cov}[Y_1 - RP_{Y_1}, \ldots, Y_N - RP_{Y_N}].$$

We require the following condition:

(a2) $\widehat{\Sigma}$ and $\Psi - \widehat{\Sigma}$ are positive definite, for all signal sets $\mathcal{I}$ available to investors.

Having $\widehat{\Sigma}$ positive definite (and therefore also Σ) ensures that prices and signals are not so precise that they render some portfolio of risky assets riskless. The last part of (a2) ensures that the conditioning information in $\widehat{\Sigma}$ is not vacuous. Under these minimal conditions, investors are indifferent between focusing their information acquisition or spreading their capacity widely:

Proposition 1.1 (Indifference to information choice). Under (a1), an informed investor is indifferent among all signal sets that satisfy (a2) and use the investor's full capacity.

Proof. A direct application of Proposition 3.1 of Admati and Pfleiderer (1987) shows that under (a2) the squared expected utility for an informed investor is given by

$$J^2 = e^{2\gamma c} \frac{|\widehat{\Sigma}|}{|\Psi|} e^{-\mu^\top \Psi^{-1} \mu}, \tag{3}$$

where $e^{2\gamma c}$ reflects the utility loss from paying the cost c to become informed. For any signal set that uses the full capacity κ, we have equality in (2), so we may replace $|\widehat{\Sigma}|$ with $\kappa|\Sigma|$ to get

$$J^2 = e^{2\gamma c} \kappa \frac{|\Sigma|}{|\Psi|} e^{-\mu^\top \Psi^{-1} \mu}, \tag{4}$$

which does not depend on the choice of signals. □

Since maximizing investor utility in (1) is equivalent to minimizing (3), using the full information capacity is optimal. Van Nieuwerburgh and Veldkamp (2010) prove a similar indifference result in the special case that prices are constants and thus not informative about payoffs.

We can gain further insight by specializing Proposition 1.1 to the case of independent payoffs and prices and considering an increasing number of assets N. Suppose $(Y_1, P_{Y_1}), \ldots, (Y_N, P_{S_N})$ are i.i.d. The matrices Σ and Ψ are then diagonal, and (4) simplifies to

$$J^2 = e^{2\gamma c}\kappa\left(\frac{\mathrm{V}[Y|P_Y]}{\mathrm{V}[Y - RP_Y]}e^{-(\mathrm{E}[Y-RP_Y])^2/\mathrm{V}[Y-RP_Y]}\right)^N, \tag{5}$$

where (Y, P) has the common distribution of the (Y_i, P_{Y_i}). The variance ratio in (5) never exceeds 1. Also, $(\mathrm{E}[Y - RP_Y])^2/\mathrm{V}[Y - RP_Y] \geq 0$. It follows that J^2 in (5) is nonincreasing in N. A smaller value of J^2 implies a greater expected utility because J is negative. Thus, given the choice to participate in a market with N securities, the investor prefers[4] to take N as large as possible and is indifferent to the information gained about these securities, as long as the information uses the investor's full capacity κ. The investor has no incentive to specialize in the choice of securities or information.

1.2 Utility of uninformed investors

Our main focus at this point is the allocation of information processing capacity for an informed investor. In anticipation of a later discussion regarding the decision to acquire capacity in the first place, we digress to consider the expected utility of the uninformed.

An uninformed investor avoids the cost c of becoming informed but has no capacity to acquire signals. For the uninformed, the posterior covariance $\widehat{\Sigma}$ always equals Σ. (Equivalently, we can think of an uninformed investor as having $\kappa = 1$.) As Σ is conditioned on prices, the uninformed learn from prices, but they do not learn more than the information in prices. Thus, for an uninformed investor, the squared expected utility (4) becomes

$$J_U^2 = \frac{|\Sigma|}{|\Psi|}e^{-\mu^\top \Psi^{-1}\mu}. \tag{6}$$

Comparing this expression with (4), we see that the value of becoming informed depends on the cost of capacity: remaining uninformed is preferable to becoming informed if capacity is expensive, in the sense that $e^{2\gamma c}\kappa > 1$.

1.3 Preference for larger consideration sets

We noted that in the i.i.d. case (5), investors have no incentive to focus on a subset of securities. We now show that this holds more generally to

4 This preference is strict if either $\mathrm{E}[Y - RP_Y] = 0$ or $\mathrm{V}[Y|P_Y] < \mathrm{V}[Y - RP_Y]$.

contrast this setting with the information frictions we introduce in Section 2.

By a consideration set $\mathcal{C}$ we mean a subset of the securities. Restricting attention to a subset of securities may be beneficial if it allows an investor to acquire more precise information about a fewer payoffs. Suppose $\mathcal{C} = \{Y_{i_1}, \ldots, Y_{i_K}\}$ for $1 \leq K \leq N-1$ and $i_j \in \{1, \ldots, N\}$. Let

$$\Sigma_{\mathcal{C}} = \mathrm{cov}[Y_{i_1}, \ldots, Y_{i_K} | P_{Y_{i_1}}, \ldots, P_{Y_{i_K}}], \tag{7}$$

and, for any signal set $\mathcal{I}$, let

$$\widehat{\Sigma}_{\mathcal{C}} = \mathrm{cov}[Y_{i_1}, \ldots, Y_{i_K} | P_{Y_{i_1}}, \ldots, P_{Y_{i_K}}, \mathcal{I}]. \tag{8}$$

An informed investor with consideration set $\mathcal{C}$ faces the capacity constraint

$$\frac{|\widehat{\Sigma}_{\mathcal{C}}|}{|\Sigma_{\mathcal{C}}|} \geq \kappa. \tag{9}$$

We do not require that $\mathcal{I}$ contain information about every security in the consideration set. For example, an investor could apply the full capacity κ to learning about a single security and invest in the other securities in the consideration set as an uninformed investor who conditions only on prices.

It is not immediately obvious how the constraint in (9) compares with (2). The denominator in (9), based on a subset of securities, could be larger or smaller than the denominator in (2) for the full set of securities. This opens the possibility that a signal set $\mathcal{I}$ could satisfy (9), but not (2), meaning that an investor could acquire more precise information by focusing on a subset of securities. An example helps illustrate this point. We compare consideration sets $\{Y_1\}$ and $\{Y_1, Y_2\}$, with independent payoffs Y_i. Suppose for simplicity that the prices P_1, P_2 are independent of the payoffs and may be ignored. Let $\mathcal{I}_\rho = \{X_\rho, Y_2\}$, where X_ρ has correlation ρ with Y_1 and is independent of Y_2. Then the determinant ratios for the two consideration sets become

$$\frac{\mathrm{V}[Y_1|X_\rho, Y_2]}{\mathrm{V}[Y_1]} = 1 - \rho^2, \qquad \frac{\mathrm{V}[Y_1|X_\rho, Y_2]\mathrm{V}[Y_2|X_\rho, Y_2]}{\mathrm{V}[Y_1]\mathrm{V}[Y_2]} = 0.$$

The signal set $\mathcal{I}_\rho$ is thus infeasible for the larger consideration set (assuming $\kappa > 0$) but, depending on the value of ρ, could be feasible and very informative for the smaller consideration set.

Enlarging the consideration set thus has the benefit of expanding an investor's portfolio choices, but the potential downside of changing the set of feasible signals. The net effect is ambiguous, so we compare expected utilities. For a consideration set $\mathcal{C} = \{Y_{i_1}, \ldots, Y_{i_K}\}$, let

$$\mu_{\mathcal{C}} = \mathrm{E}(Y_{i_1} - RP_{Y_{i_1}}, \ldots, Y_{i_K} - RP_{Y_{i_K}})^\top,$$

and

$$\Psi_{\mathcal{C}} = \mathrm{cov}[Y_{i_1} - RP_{Y_{i_1}}, \dots, Y_{i_K} - RP_{Y_{i_K}}].$$

We restrict attention to information sets satisfying

(a3) $\widehat{\Sigma}_{\mathcal{C}}$ and $\Psi_{\mathcal{C}} - \widehat{\Sigma}_{\mathcal{C}}$ are positive definite.

In several places, we state conclusions that rely on informed investors being able to fully utilize their information capacity. To avoid repetition, we will assume that a sufficiently rich set of signals is always available to meet this condition, so investors are merely constrained by their capacity κ: investors never learn everything that can be known.

(a4) Signal availability: for any consideration set $\mathcal{C}$ and any $\epsilon \in (0, 1)$, there is a signal set $\mathcal{I}$ satisfying (a3) for which the posterior covariance satisfies $|\widehat{\Sigma}_{\mathcal{C}}|/|\Sigma_{\mathcal{C}}| = \epsilon$.

Proposition 1.2 (Preference for larger consideration sets). Suppose (a1), (a3), and (a4) hold. The squared expected utility for an informed investor with consideration set $\mathcal{C}$ *is given by*

$$J_{\mathcal{C}}^2 = e^{2\gamma c}\kappa \frac{|\Sigma_{\mathcal{C}}|}{|\Psi_{\mathcal{C}}|} e^{-\mu_{\mathcal{C}}^\top \Psi_{\mathcal{C}}^{-1} \mu_{\mathcal{C}}}. \tag{10}$$

If $\mathcal{C} \subseteq \mathcal{C}'$*, then* $J_{\mathcal{C}}^2 \geq J_{\mathcal{C}'}^2$*, so the investor prefers the larger consideration set* $\mathcal{C}'$*.*

This result generalizes the conclusion we reached in the i.i.d. case using the expression (5) for squared utility under the i.i.d. assumption. Proposition 1.2 shows that regardless of the covariance structure and regardless of the types of signals, investors prefer larger consideration sets and are otherwise indifferent to information choices: there is no incentive to specialize. We view this as a negative result, in the sense that it shows that this approach to information choice does not lead to interesting conclusions. These observations lead us to introduce information frictions in Section 2. To put these frictions in context, we first introduce a factor structure to the securities.

1.4 Factor structure: Macro and micro shocks

The market trades N stocks with payoffs

$$u_i = M' + S_i', \quad i = 1, \dots, N,$$

with $M', S_1', \dots, S_N'$ mutually independent and normally distributed, and S_1' $, \dots, S_N'$ identically distributed. We think of M' as a common macro factor and the S_i' as stock-specific risks, though we will modify this formulation shortly. We let $\mathrm{E}[M'] = \bar{m}$ and assume $\mathrm{E}[S_i'] = 0$, $i = 1, \dots, N$, consistent with the interpretation of S_i' as idiosyncratic.

The market also trades an index fund. One share of the index fund holds $1/N$ shares of each of the N stocks. The fund's payoff is therefore

$$u_F = \frac{1}{N}\sum_{i=1}^{N} u_i = M' + \bar{S}', \qquad \text{with} \quad \bar{S}' = \frac{1}{N}\sum_{i=1}^{N} S_i'. \tag{11}$$

This expression makes clear that the sample mean $\bar{S}'$ is an undiversifiable common factor of the stock returns and is therefore more appropriately treated as part of the macro factor. We therefore move this component into the macro factor by setting

$$M = M' + \bar{S}', \quad S_i = S_i' - \bar{S}', \; i = 1, \dots, N. \tag{12}$$

This reformulation allows us to write

$$u_i = M + S_i, \quad i = 1, \dots, N, \tag{13}$$

without changing the stock payoffs — we have only adjusted the decomposition into macro and micro components, M and S_i.

With this formulation, the payoff of the index fund is

$$u_F = \frac{1}{N}\sum_{i=1}^{N} u_i = M, \quad \text{because} \quad \bar{S} = \frac{1}{N}\sum_{i=1}^{N} S_i \equiv 0. \tag{14}$$

Thus, we have gained two important benefits: the payoff of the index fund is now exactly the macro factor, and the micro shocks S_i are fully diversifiable, in the sense that $\bar{S} = 0$. The micro shocks are still uncorrelated with the macro shock because

$$\begin{aligned} \mathrm{E}[MS_i] &= \mathrm{E}[(M' + \bar{S}')(S_i' - \bar{S}')] = \mathrm{E}[\bar{S}' S_i'] - \mathrm{V}[\bar{S}'] = \mathrm{V}[S_i']/N - \mathrm{V}[S_i']/N \\ &= 0. \end{aligned}$$

The micro shocks $S_1, \dots, S_N$ are not independent of each other but they are *exchangeable*, meaning that permuting their order does not change their joint distribution.[5] In particular, all pairs of micro shocks have the same correlation, which evaluates to

$$\mathrm{corr}[S_i, S_j] = -1/(N-1), \quad i = j. \tag{15}$$

We write $\sigma_S^2 = \mathrm{V}[S_i]$ and $\sigma_M^2 = \mathrm{V}[M]$.[6]

[5] The same idea has been used to formulate the capital asset pricing model (CAPM) with a finite number of securities, as in Ross (1978), Chen and Ingersoll (1983), and Kwon (1985), ensuring that idiosyncratic risks are fully diversifiable.

[6] Equation (15) follows directly from (12). But it is also a general property of exchangeable random variables that add to a constant: the correlation between any two such variables is $-1/(N-1)$. Furthermore, exchangeable random variables must have a common mean.

If we worked instead with the initial formulation (11), the payoff of the index fund $M' + \bar{S}'$ would be correlated with the stock-specific shocks S_i', which would undermine the distinction between macro and micro payoffs. We prefer to make the index fund a pure position in the macro factor M, which is uncorrelated with the S_i, and to have the S_i fully diversifiable, in the sense that $\bar{S} = 0$.

This reformulation avoids a further difficulty that we would otherwise have to deal with once we introduce prices. Write P_F for the price of the index fund, P_i for the price of the ith stock, and P_{S_i} for the price of the ith *hedged security*, which is long stock i, short the index fund, and has payoff $u_i - u_F$. The price of this hedged security is $P_{S_i} = P_i - P_F$. Absence of arbitrage requires

$$P_F = \frac{1}{N}\sum_{i=1}^{N} P_i = P_F + \frac{1}{N}\sum_{i=1}^{N} P_{S_i}, \tag{16}$$

and thus $\sum_i P_{S_i}/N \equiv 0$. In other words, the average price of the hedged securities must be identically zero. This would be a difficult condition to impose if the hedged payoffs were independent, but it becomes natural in our setting because the average of the hedged payoffs S_i is identically zero. The absence of arbitrage makes exchangeability, rather than independence, the natural condition for the hedged securities and their prices. Next, we will show that utility calculations are about as straightforward in the exchangeable setting as in the independent case.

1.5 Expected utility with macro and micro shocks

Because the index fund holds a portfolio of the N stocks, the market of $N+1$ securities is spanned by any N of the securities, and we may assume that no investor invests in more than N securities. To simplify the analysis, we will assume from now on that investors directly trade in the index fund and the hedged (or micro) securities, so they have access to securities with payoffs $M, S_1, \ldots, S_N$. No investor invests in more than $N-1$ of the micro securities S_i because the Nth would be spanned by the others.

More generally, we distinguish two types of consideration sets, based on whether they include M. A consideration set consisting of K hedged securities, $1 \leq K \leq N-1$, has the form

$$\mathcal{C}_{0,K} = \{S_{i_1}, \ldots, S_{i_K}\}; \tag{17}$$

with the index fund included and $0 \leq K \leq N-1$, it takes the form

$$\mathcal{C}_{1,K} = \{M, S_{i_1}, \ldots, S_{i_K}\}. \tag{18}$$

We will refer to these sets as having *type* $(0, K)$ or $(1, K)$.

Once we restrict ourselves to any of these consideration sets, the covariance matrix of the payoffs has full rank, and we are in the general setting of

Section 1.1. However, the factor structure provides additional simplification. We will use the following condition:

(a5i) Macro-micro independence: $(S_1, P_{S_1}), \ldots, (S_N, P_{S_N})$ are independent of (M, P_F);

(a5ii) Exchangeability: $(S_1, P_{S_1}), \ldots, (S_N, P_{S_N})$ are exchangeable.

These properties hold for $M, S_1, \ldots, S_N$ because of the factor structure we introduced in Section 1.4. The content of (a5) is the extension of the same structure to include prices. When we have exchangeability, we often use (S, P_S) to refer to the common distribution of any of the (S_i, P_{S_i}). We construct an equilibrium in which (a5) holds in Section 3.

A general implication of (a5) is that hedged securities have zero expected risk premiums, or

$$\mathrm{E}[S_i - RP_{S_i}] = 0. \tag{19}$$

The no-arbitrage condition in (16) implies that $\sum_i P_{S_i} = 0$, and then $\mathrm{E}\sum_i P_{S_i} = \sum_i \mathrm{E}P_{S_i} = 0$. Exchangeability implies that all P_{S_i} have the same mean, which must then be zero. Finally, $\mathrm{E}S_i = 0$ by construction.

Given our factor structure, the next proposition shows that expected utilities have the same form as in the i.i.d. case in (5). This greatly simplifies our subsequent analysis of the impact of information frictions. However, in the absence of such frictions, the imposition of a factor structure leaves the information indifference result unchanged. Consistent with the discussion following (2), we say that an investor with consideration set $\mathcal{C}$ uses the full capacity κ if equality holds in (9). The signal availability condition (a4) ensures that using the full capacity is always feasible.

Proposition 1.3 (Expected utilities in factor model). Suppose conditions (a1) and (a3)–(a5) hold and consider an informed agent who uses the full capacity κ. The agent's expected utility with consideration set $\mathcal{C}$ depends on $\mathcal{C}$ only through its type, $(0, K)$ *or* $(1, K)$. For a consideration set of type $(0, K)$,

$$J_{0,K}^2 = e^{2\gamma c}\kappa\left(\frac{\mathrm{V}[S|P_S]}{\mathrm{V}[S - RP_S]}\right)^K. \tag{20}$$

And for a consideration set of type $(1, K)$,

$$\begin{aligned} J_{1,K}^2 &= e^{2\gamma c}\kappa e^{-Q_F}\frac{\mathrm{V}[M|P_F]}{\mathrm{V}[M - RP_F]}\left(\frac{\mathrm{V}[S|P_S]}{\mathrm{V}[S - RP_S]}\right)^K, \\ Q_F &= (\mathrm{E}[M - RP_F])^2/\mathrm{V}[M - RP_F]. \end{aligned} \tag{21}$$

In (20) and (21), squared expected utility factors *as if* we had independence. These expressions are analogous to the i.i.d. case in (5) because, under (a5), $\Sigma_{\mathcal{C}}$ and $\Psi_{\mathcal{C}}$ have the same correlation matrix, so the correlations effectively cancel when we take the ratio of their determinants in (4). One difference between (20) and (21) and the i.i.d. result in (5) is that (20)–(21) reflect the zero risk premium condition in (19). The risk premium from the macro factor appears in Q_F in (21).

Moreover, it is evident from (20) and (21) that, given a consideration set, an informed investor is indifferent to the choice of information set (among signals that use the investor's full capacity), and investors prefer larger consideration sets over smaller ones.[7] These conclusions are special cases of our earlier results, but the factor structure leads to more explicit expressions.

To summarize, in this section we have shown that under the basic framework of normally distributed risks, CARA preferences, and variance-ratio information processing constraints, investors are indifferent to information choices and prefer larger consideration sets. They have no reason to specialize in information or securities. This holds under essentially arbitrary covariance structures, including the factor structure of (13). We lay out these results as a benchmark against which to compare the consequences of introducing information frictions.

2. Information Frictions

Thus far, we have seen that, under general conditions, a standard limited information capacity framework leads investors to be indifferent about information choices and to prefer to consider the largest set of securities possible. As we argue in Section 4, these conclusions are unrealistic. Investors do not make rational inferences from countless prices, nor do they invest in all available securities. The framework we have considered thus far lacks the fixed attention costs entailed in learning even basic information about a stock and the market.

To address these points, we introduce frictions in the form of fixed attention costs to information acquisition. Importantly, we differentiate the frictions for micro and macro information. We posit that information about the overall level of the stock market is widely available — it is regularly discussed in the media and in public commentary. The same is not true for most stocks — learning even basic information about most companies requires an investor to make a special effort. To address this distinction, we continue to use the factor structure of Section 1.4, and the consideration sets $\mathcal{C}_{0,K}$ and $\mathcal{C}_{1,K}$ from (17) and (18), respectively.

[7] This follows from Proposition 1.2, but also it follows directly from the fact that $\mathrm{V}[S|P_S] \leq \mathrm{V}[S - RP_S]$.

2.1 Fixed attention costs

We begin with micro information. We assume that there is a fixed attention cost to following a stock. Investors are free to invest in as many stocks as they wish, but to follow a stock and make rational inferences from its price, an informed investor must dedicate at least a minimum amount of attention. For example, an investor must learn what the company's line of business is, who serves on its senior management team, who are the company's main competitors, and so on. We capture this idea by introducing a factor $\delta \in (\kappa, 1)$ and changing the capacity constraint (9) to

$$\frac{|\widehat{\Sigma}_{\mathcal{C}_{0,K}}|}{|\Sigma_{\mathcal{C}_{0,K}}|} \geq \kappa/\delta^K. \tag{22}$$

Recall that a smaller bound on the right allows for more informative signals by allowing a smaller posterior $|\widehat{\Sigma}_{\mathcal{C}_{0,K}}|$. Dividing by δ^K thus consumes capacity, and the loss of capacity is greater with larger K. Each factor of δ reflects the fixed cost of following a stock.

Whereas we assume that each stock requires at least a minimum amount of attention, we allow investors to follow the index fund—and to make inferences from its price—costlessly. This assumption reflects the widespread availability of information about the overall stock market. Whereas investors may need to spend some time learning about a new company, they already know about and follow the overall stock market. However, investors who seek to acquire additional information about the macro factor M beyond its market price face a fixed attention cost $\delta_F \in (\kappa, 1)$. We will say that signal set $\mathcal{I}$ is informative about the macro factor if $\mathrm{V}[M|P_F, \mathcal{I}] < \mathrm{V}[M|P_F]$.[8] The capacity constraint becomes

$$\frac{|\widehat{\Sigma}_{\mathcal{C}_{1,K}}|}{|\Sigma_{\mathcal{C}_{1,K}}|} \geq \begin{cases} \kappa/(\delta^K \delta_F), & \text{if } \mathrm{V}[M|P_F, \mathcal{I}] < \mathrm{V}[M|P_F]; \\ \kappa/\delta^K, & \text{otherwise.} \end{cases} \tag{23}$$

Condition (23) reflects two asymmetries in how we model macro and micro information. In the first case of (23), acquiring macro information (beyond the price P_F of the index fund) requires a fixed cost δ_F rather than δ. But the more important asymmetry is the second case: an investor may invest "passively" in the index fund, learning about M only through its price, without incurring a fixed cost. In contrast, following an individual stock always entails a fixed cost of δ.

We emphasize that once an investor decides to follow a stock or to acquire additional information about the macro factor, conditioning on the price of the stock or of the index fund is costless. The investor is only charged capacity for acquiring a signal with additional information content beyond the price.

[8] When we make the macro-micro independence assumption (a5i), this definition is unchanged if we include other prices in the conditioning information and write $\mathrm{V}[M|P_F, P_{S_1}, \ldots, P_{S_K}, \mathcal{I}] < \mathrm{V}[M|P_F, P_{S_1}, \ldots, P_{S_K}]$.

Furthermore, it is important to keep in mind the distinction between the cost c of becoming informed, and the capacity charges δ and δ_F that are incurred from following a stock or acquiring additional macro information. The quantity c is the standard cost of becoming an informed investor, and thus acquiring information capacity κ. The capacity charges δ and δ_F consume part of κ when investors increase their consideration sets, or when they acquire macro information.

We will say that an informed investor uses the full capacity κ if equality holds in (22) or (23) for the investor's choice of consideration set and signals. For a given consideration set, this condition constrains K through the requirement that

$$\delta^K \geq \kappa \quad \text{or} \quad \delta_F \delta^K \geq \kappa. \tag{24}$$

We illustrate the constraint in (23) with some examples. Investing passively in the index fund ($\mathcal{C}_{1,0}$ with $\mathcal{I}$ empty) consumes no capacity because it corresponds to the second case in (23) with $K = 0$. A *macro-informed* investor chooses consideration set $\mathcal{C}_{1,0}$ and acquires information $\mathcal{I}$ about M, facing the constraint $\mathrm{V}[M|P_F, \mathcal{I}]/\mathrm{V}[M|P_F] \geq \kappa/\delta_F$, which is the first case in (23) with $K = 0$. Next, consider a *micro-informed* investor who follows a single stock ($K = 1$), while costlessly investing in the index fund by not seeking macro information. Suppose (M, P_F) and (S_i, P_{S_i}) are independent and remain so given the micro information $\mathcal{I}$. The micro-informed investor faces the constraint

$$\frac{\mathrm{V}[S_i|P_{S_i}, \mathcal{I}]}{\mathrm{V}[S_i|P_{S_i}]}\frac{\mathrm{V}[M|P_F, \mathcal{I}]}{\mathrm{V}[M|P_F]} \equiv \frac{\mathrm{V}[S_i|P_{S_i}, \mathcal{I}]}{\mathrm{V}[S_i|P_{S_i}]} \geq \kappa/\delta.$$

The following proposition characterizes the utility that investors gain from different information sets:

Proposition 2.1 (Utility of the informed with fixed attention costs). Assume conditions (a1) and (a3)-(a5) hold with the information constraints (22)-(23). Then the expected utility of an informed agent who uses the full capacity κ depends on the agent's consideration set and information choice as follows. For a consideration set of type $(0, K)$, $1 \leq K \leq N-1$, *with* $\kappa \leq \delta^K$,

$$J^2_{0,K} = e^{2\gamma c}\kappa\left(\frac{\mathrm{V}[S|P_S]}{\mathrm{V}[S - RP_S]}\right)^K/\delta^K; \tag{25}$$

for a consideration set of type $(1, K)$, $1 \leq K \leq N-1$, with $\kappa \leq \delta^K$, *and macro-independent information*,

$$J^2_{1,K} = e^{2\gamma c}\kappa e^{-Q_F}\frac{\mathrm{V}[M|P_F]}{\mathrm{V}[M - RP_F]}\left(\frac{\mathrm{V}[S|P_S]}{\mathrm{V}[S - RP_S]}\right)^K/\delta^K; \tag{26}$$

for a consideration set of type $(1, K)$, $0 \leq K \leq N-1$, $\kappa \leq \delta_F \delta^K$, *with macro information,*

$$J^2_{M,K} = e^{2\gamma c} \kappa e^{-Q_F} \frac{\mathrm{V}[M|P_F]}{\mathrm{V}[M - RP_F]} \left(\frac{\mathrm{V}[S|P_S]}{\mathrm{V}[S - RP_S]} \right)^K / (\delta_F \delta^K). \tag{27}$$

In particular, $J^2_{0,K} \geq J^2_{1,K}$, *so an investor in micro shocks can always do better by including a position in the index fund without acquiring macro information.*

These expressions follow directly from combining Proposition 1.3 with the costs in (22)-(23). The result that $J^2_{0,K} \geq J^2_{1,K}$ holds because $Q_F \geq 0$, as both its numerator and denominator are positive. Furthermore $\mathrm{V}[M|P_F] = \mathrm{V}[M - RP_F|P_F] \leq \mathrm{V}[M - RP_F]$. The inequality $J^2_{0,K} \geq J^2_{1,K}$ is strict if $\mathrm{E}[M - RP_F] = 0$ (so $Q_F > 0$) or if P_F has nonzero correlation with M.

Equations (26) and (27) both correspond to consideration sets $\{M, S_1, \ldots, S_K\}$, but the expected utilities differ because they correspond to the two cases in the capacity constraint (23), determined by whether the investor selects macro-informative signals. Our discussion will therefore refine investor types to distinguish between these two cases: consideration set $\mathcal{C}_{1,K}$ with or without macro-informative signals.

We close this section with a comment on interpreting the fixed costs in (22)-(23). For simplicity, consider the case of a single security with payoff Y and price P_Y. Given a signal set $\mathcal{I}$, the total amount an investor learns about the payoff Y is measured by

$$\frac{\mathrm{V}[Y|P_Y, \mathcal{I}]}{\mathrm{V}[Y]} = \frac{\mathrm{V}[Y|P_Y, \mathcal{I}]}{\mathrm{V}[Y|P_Y]} \frac{\mathrm{V}[Y|P_Y]}{\mathrm{V}[Y]}. \tag{28}$$

The original capacity constraint (2) limits the first factor on the right, that is, the variance reduction attributable to $\mathcal{I}$. We can think of δ in (22) as capturing the second factor on the right, that is, the attention required to make inferences from prices, even before any other information is collected. For example, knowing how to interpret a stock drop of 5% requires knowing whether the company had earnings, or whether a drug trial result came out, or whether the CEO resigned, or whether the fall happened on no news at all. Not all 5% price drops mean the same thing; δ is the cost of knowing enough about the company to make these distinctions.[9] In this interpretation, (22) puts a lower bound on the total variance reduction on the left side of (28). The correspondence is not exact because the variance ratios are ultimately equilibrium outcomes, whereas δ is fixed. Nevertheless, the decomposition in (28)

[9] Chan (2003) and Tetlock (2010) document that stock moves on no news are quickly reversed, whereas stock news accompanied by news are not. Context matters.

should be helpful in thinking about δ as the minimal attention required to follow a stock.[10]

2.2 Uninformed and oblivious investors

The information constraints (23) apply only to investors who have paid the cost c to acquire capacity κ. Investors who do not buy capacity cannot acquire signals. They can still invest in the index fund without acquiring macro information because there are no fixed capacity costs to doing so. This option corresponds to $K = 0$ in (26), with $c = 0$ and $\kappa = 1$, which yields squared expected utility of

$$J_U^2 = e^{-Q_F} \frac{\mathrm{V}[M|P_F]}{\mathrm{V}[M - RP_F]}. \tag{29}$$

This is a special case of our expression (6) for the utility of the uninformed, specialized to the consideration set $\{M\}$. Here, as in (6), the uninformed investor makes inferences from prices — from the index fund price P_F in the case of (29).

We will also allow investors without information capacity to buy stocks. However, they cannot cover the fixed attention cost δ of following a stock, so they do not have the capacity to make inferences from stock prices. Instead, they simply take prices as given. We call these investors *oblivious*.

To make the distinction precise, consider a vector of payoffs Y with a vector of prices P_Y. An investor who learns from prices chooses a portfolio vector q to maximize $-\mathrm{E}[\exp(-q^\top (Y - RP_Y)/\gamma | P_Y]$, leading to $q = \mathrm{cov}[Y|P_Y]^{-1}(\mathrm{E}[Y|P_Y] - RP_Y)/\gamma$. An oblivious investor solves the problem unconditionally and arrives at demands

$$q = \mathrm{cov}[Y]^{-1}(\mathrm{E}[Y] - RP_Y)/\gamma. \tag{30}$$

The oblivious investor thus responds to prices and sets demands rationally, but fails to process all the information in prices.

These oblivious demands apply to any S_i in the oblivious investor's consideration set. An oblivious investor may also invest in the index fund and, in doing so, make inferences from P_F. This formulation is consistent with the constraints in (23) and with our treatment of the uninformed in (29): an uninformed investor in the sense of (29) is simply an oblivious investor with consideration set $\{M\}$.

Proposition 2.2 (Utility of the oblivious). Suppose (a0) and (a5) hold. For an oblivious investor with consideration set $\mathcal{C}_{1,K} = \{M, S_{i_1}, \ldots, S_{i_K}\}$,

[10] An earlier version of this paper placed capacity constraints on the left side of (28). That formulation implies that learning from prices (the last factor in Equation (28)) consumes capacity. But the two versions of the model lead to similar conclusions.

$$J^2_{U,K} = e^{-Q_F} \frac{\mathrm{V}[M|P_F]}{\mathrm{V}[M - RP_F]} \left(\frac{\mathrm{V}[S]}{\mathrm{V}[S - RP_S] + \mathrm{cov}[S, RP_S]^2/\mathrm{V}[S]} \right)^K \tag{31}$$

The utility in (31) is evaluated by taking into account the true joint distribution of prices, security payoffs, and oblivious investor demands. If we excluded M from the consideration set, the exponential factor and the first variance ratio on the right side of (31) would be absent. It follows that omitting M is never preferred (see the discussion following Proposition 2.1), so we may assume that oblivious investors always include the index fund in their consideration sets. We will sometimes refer to the special case $K = 0$ in (29) as the ordinary uninformed investors. In Theorem 3.1, we will give conditions under which being informed or (ordinarily) uninformed dominates being oblivious.

2.3 Price informativeness requires specialization

We investigate the choices of consideration sets and information that are compatible with equilibrium, meaning that they are not strictly dominated by any other choices. It suffices to consider which *types* of consideration sets are undominated because the expected utility for a consideration set depends only on its type. Here, the type refers to the combination of consideration set and information choice: an informed investor with set $\mathcal{C}$ and an uninformed investor with set $\mathcal{C}$ are different types.

We will use λs to indicate the fractions of investors who select each type, as follows:

$$\lambda_{0,K} : \text{informed choosing } \{S_{i_1}, \ldots, S_{i_K}\};$$

$$\lambda_{1,K} : \text{informed choosing } \{M, S_{i_1}, \ldots, S_{i_K}\}, \text{ as in (26)};$$

$$\lambda_{M,K} : \text{informed choosing } \{M, S_{i_1}, \ldots, S_{i_K}\}, \text{ as in (27)};$$

$$\lambda_U : \text{uninformed choosing } \{M\};$$

$$\lambda_{U,K} : \text{oblivious choosing } \{M, S_{i_1}, \ldots, S_{i_K}\}.$$

In particular, investors of type $\lambda_{M,K}$ learn about the macro factor (and up to K stocks), whereas investors of type $\lambda_{1,K}$ invest in the index fund (and up to K stocks) without learning more about M than the information in the price of the index fund.

We assume that, for each type, λ is evenly distributed among all consideration sets consistent with that type; e.g., for $\lambda_{1,2}$ we will have an equal share of investors with consideration sets $\{M, S_i, S_j\}$, for all pairs (S_i, S_j) with $i \neq j$.

We call a collection of nonnegative λs summing to one an *allocation*. We have not yet introduced a notion of equilibrium, but we can define basic requirements that investor choices (as reflected in an allocation of λs) should satisfy to be compatible with equilibrium. The first requirement is that investors do not make suboptimal choices:

Definition 2.1 (Choices Compatible with Equilibrium). An allocation of λs is **compatible with equilibrium** if no type with a positive λ is dominated by any other type: $\lambda_x > 0 \Rightarrow J_x^2 \leq J_{x'}^2$, for all types x, x'.

We established through Propositions 2.1 and 2.2 that omitting the index fund is never advantageous, and it is strictly dominated, except in the degenerate case that the fund price P_F and its payoff M are uncorrelated and satisfy $\mathrm{E}[M - RP_F] = 0$. Outside of this degenerate case, compatibility with equilibrium would require

$$\lambda_{0,K} = 0, \text{ for all } K \geq 1. \tag{32}$$

Definition 2.1 allows the possibility that all investors remain uninformed, in which case any associated equilibrium would presumably result in uninformative prices. We would like to further characterize allocations under which informed investors collectively learn about all securities. This consideration is motivated by the idea that if no informed investor learns about a security's payoff, than the price of that security cannot be informative about the payoff. Condition (32) must hold if P_F carries any information about M. Moreover, only investors of type $\lambda_{M,K}$, for some $K \geq 0$, learn about the macro factor, whereas all investors of types $\lambda_{1,K}$ and $\lambda_{M,K}$, $K \geq 1$, can learn about the S_i. The condition we require is therefore

$$\lambda_{M,K} > 0, \text{ for some } K \geq 1; \text{ or } \lambda_{M,0} > 0 \text{ and } \lambda_{1,K} > 0 \text{ for some } K \geq 1. \tag{33}$$

Definition 2.2 (Choices Compatible with an Informative Equilibrium). An allocation of λs is **compatible with an informative equilibrium** if it is compatible with equilibrium and satisfies (32) and (33).

The condition in (33) allows a wide range of possibilities in the allocation of informed investors to types, including the possibilities that investors choose maximal consideration sets, as we saw in previous sections, or that some investors choose to be oblivious. For the following result, we highlight two special types: $\lambda_M = \lambda_{M,0}$ are the purely macro-informed investors with consideration sets $\{M\}$; and $\lambda_S = \lambda_{1,1}$ are the purely micro-informed investors with consideration sets $\{M, S_i\}$. These λ_M investors only trade the index fund, and spend their information capacity to learn about M, in the sense of the first case in (23) with K

$= 0$. These λ_S investors select macro-uninformative signals, in the sense of the second case in (23). Each follows a single stock.

Proposition 2.3 (Compatibility with informative equilibrium requires specialization). Suppose conditions (a1) and (a3)–(a5) hold with capacity constraints (22) and (23). In any allocation compatible with an informative equilibrium, $\lambda_M > 0$ *and* $\lambda_S > 0$, and there are no other informed investors.

Proof. From (26) and (27), we find that $J^2_{M,K} = J^2_{1,K}/\delta_F > J^2_{1,K}$, for all $K \geq 1$, so all such $J^2_{M,K}$ are dominated. Compatibility with equilibrium then implies $\lambda_{M,K} = 0$, for all $K \geq 1$. Condition (33) therefore requires $\lambda_{M,0} \equiv \lambda_M > 0$ and $\lambda_{1,K} > 0$, for some $K \geq 1$.
To have $\lambda_{1,K} > 0$, we need $J^2_{1,K}$ to be undominated. Let $r_S = \mathrm{V}[S|P_S]/(\mathrm{V}[S - RP_S]\delta)$. Using first $0 < \delta_F < 1$, then the expressions for $J^2_{1,K}$ and $J^2_{M,K}$ in (26) and (27), we get

$$J^2_{1,K=1} < J^2_{1,K=1}/\delta_F = J^2_{M,K=1} = J^2_{M,K=0} r_S.$$

If $r_S \leq 1$, we would conclude that $J^2_{1,K=1} < J^2_{M,K=0}$, contradicting $\lambda_M > 0$. For compatibility with an informative equilibrium, we must therefore have $r_S > 1$. But then the only undominated case of $J^2_{1,K}$ in (26) is $K = 1$; in other words $\lambda_S > 0$. Moreover, we have shown that all choices of consideration sets by informed investors other than sets $\{M\}$ and $\{M, S_i\}$ are dominated. □

This result presents a stark contrast with the results of Section 1. In the previous results, we consistently found that informed investors are indifferent to their information choices and prefer larger consideration sets. Here we find that, once we introduce fixed attention costs, an informative equilibrium requires just the opposite: informed investors must specialize in macro or micro information, and the micro-informed focus on a single stock.[11] The two optimal information sets are therefore extreme specialization achieved by investing in a single stock, along with the index fund, or complete diversification achieved by investing only in the index fund. Note that investors have the option to use all their capacity on a subset of stocks and invest in other securities as uninformed investors who condition only on prices, but they reject this option, and the micro-informed focus only on a single stock about which they obtain a private signal. According to Proposition 2.3, the only alternative to this macro/micro specialization is to drop the "informative" requirement and assume that there are some securities that no investors learn about.

[11] Recall that the λ_S micro-informed investors are evenly divided among the N consideration sets $\{M, S_i\}$, $i = 1, \ldots, N$, all of which yield the same expected utility.

We briefly contrast Proposition 2.3 with prior work. In Goldstein and Yang (2015), a security's payoff depends on two types of fundamentals; investors acquire information about both—and do not specialize—unless the cost of acquiring two types of information is greater than the sum of the costs of acquiring each separately. Mondria (2010) finds cases of asymmetric equilibria numerically, but these are outside the scope of his theoretical analysis, which focuses on identical signal choices by investors. Investors in Van Nieuwerburgh and Veldkamp (2009) specialize, but their specialization, unlike ours, depends on differences in prior information.

It is important to clarify what we are not claiming in Proposition 2.3. The proposition does not demonstrate that an informative equilibrium exists — we will do that in the next section — so the result should be read as a necessary property of any such equilibrium. Nor are we claiming that all prices must carry information about payoffs in equilibrium. Indeed, we will see that, depending on parameter choices in (26)-(27), it is possible to have all informed investors driven out of the market and an equilibrium with entirely uninformative prices. We view the most realistic and most interesting outcome as an informative equilibrium that includes some ordinary uninformed investors — those who invest solely in the index fund:

Corollary 2.1 An allocation compatible with an informative equilibrium requires $e^{2\gamma c}\kappa \leq \delta_F$, and it admits ordinary uninformed investors (meaning that $\lambda_U > 0$) *if*

$$e^{2\gamma c}\kappa = \delta_F. \tag{34}$$

Proof. We know from Proposition 2.3 that compatibility with an informative equilibrium requires that $J^2_{M,0}$ be undominated. Comparison with (29) shows that $J^2_{M,0} = J^2_U e^{2\gamma c}\kappa/\delta_F$. If $e^{2\gamma c}\kappa/\delta_F > 1$, remaining uninformed would dominate acquiring information, so compatibility with an informative equilibrium requires $e^{2\gamma c}\kappa/\delta_F \leq 1$. If the ratio equals 1, then $J^2_U = J^2_{M,0}$, so $0 < \lambda_U < 1$ is compatible with an informative equilibrium. □

The condition in (34) equalizes the utility loss from spending c to become informed ($e^{2\gamma c}$) with the utility gain from informational capacity applied to the macro signal (κ/δ_F).

3. Equilibrium

In this section we construct a rational-expectations equilibrium in which agents' information choices are subject to the constraint in (23),[12] and in which assumptions (a1) and (a3)–(a5) are satisfied. The equilibrium is *informative*, in the sense that investors collectively acquire information about the

[12] Proposition 2.1 showed that investors are no worse off by including the index fund in their consideration sets, and, therefore, the information constraint in (22) is not binding.

payoff of every security. As in Section 1.4, we assume the existence of N stocks and an index fund holding $1/N$ shares of each stock. The payoffs M of the index fund and S_i of each stock are defined in (12).

Our equilibrium construction proceeds in two steps. We first suspend investors' flexibility to make information choices and simply assign all investors to three types, the macro-informed $\lambda_M > 0$, the micro-informed $\lambda_S > 0$, and the uninformed $\lambda_U \geq 0$, with $\lambda_M + \lambda_S + \lambda_U = 1$. This allocation is motivated by the necessary specialization proved in Proposition 2.3. We continue to assume that the λ_S micro-informed are evenly divided among the N individual stocks. We derive market clearing prices and demands in this constrained setting. We will then reinstate the flexibility for investors to make information choices and provide conditions under which our market clearing prices and demands yield an equilibrium in this unconstrained setting.

3.1 Signals

Macro-informed investors observe the signal m, with

$$M = m + \epsilon_M, \tag{35}$$

where m and ϵ_M are uncorrelated and normally distributed. We write $\mathrm{V}[m] = f_M \mathrm{V}[M]$ to use $f_M \in (0, 1)$ as a measure of the precision of the macro signal.

Similarly, we will use the decomposition

$$S_i = s_i + \epsilon_i, \quad i = 1, \ldots, N, \tag{36}$$

where s_i and ϵ_i are uncorrelated with each other and normally distributed. We write $\mathrm{V}[s_i] = f_S \mathrm{V}[S_i]$, taking f_S to be common to all i. The λ_S/N investors that invest in the ith stock observe the signal s_i.

Starting with a representation $u_i = M' + s_i' + \epsilon_i'$, we can extend the argument in (13) to the s_i and ϵ_i, so that

$$\sum_{i=1}^{N} s_i = \sum_{i=1}^{N} \epsilon_i = 0, \tag{37}$$

and $\mathrm{corr}(s_i, s_j) = \mathrm{corr}(\epsilon_i, \epsilon_j) = -1/(N-1)$ for $i \neq j$. As in our discussion in Section 1.4, (37) should be viewed as the result of removing a common component from a finite number of idiosyncratic terms. Because the S_i are exchangeable, but not independent, it is not possible for all the s_i and ϵ_i to be independent.

The ratios f_M and f_S determine how much information is learnable about payoffs from signals. For now, we treat f_M and f_S as fixed parameters. In Section 3.4, we discuss how f_M and f_S are determined in equilibrium to be consistent with the information frictions setting of Section 2.

3.2 Supply shocks

We now detail the supply of the securities. We suppose that the supply has a factor structure similar to that of the dividends in (12), with the supply of the ith stock given by

$$\frac{1}{N}(X_F + X_i), \tag{38}$$

where X_F is the common supply shock, normally distributed with mean $\bar{X}_F$ and variance $\sigma^2_{X_F}$, and the X_i are normally distributed idiosyncratic shocks, each with mean 0 and variance σ^2_X. Supply shocks are independent of cash flows, and X_i is independent of X_F for all i. Following the procedure in Section 1.4, we define the common factor X_F so that the idiosyncratic shocks are exchangeable and fully diversifiable, in the sense that

$$\sum_{i=1}^{N} X_i = 0, \tag{39}$$

and $\mathrm{corr}(X_i, X_j) = -1/(N-1)$ for $i \neq j$. We make the standard assumption that supply shocks are unobservable by the agents.

The aggregate portion of supply shocks, X_F, is standard in the literature—as will become clear, it is analogous to the single security supply shock in Grossman and Stiglitz (1980). The idiosyncratic portion of the supply shock, X_i, proxies for price-insensitive noise trading in individual stocks. Some of this noise trading may be liquidity driven (for example, individuals needing to sell their employer's stock to pay for unforeseen expenditures), and some may originate from incorrect expectations or from other value-irrelevant triggers, such as an affinity for trading or fads. Empirical studies (Brandt et al. 2010; Foucault, Sraer, and Thesmar 2011) document a link from retail trading (proxied in our setting by X_i) to idiosyncratic volatility of stock returns. Our equilibrium stock price in Equation (43) will have this feature as well.

3.3 Market clearing and prices

Given prices, investors set their demands by maximizing the expected utility of terminal wealth, as indicated at time 1 in Figure 1. The uninformed choose their demand q_F^U for the index fund conditional on $\mathcal{I}_U = \{P_F, P_{S_1}, \ldots, P_{S_N}\}$; the macro-informed choose their demand q_F^M conditional on $\mathcal{I}_M = \{m, P_F, P_{S_1}, \ldots, P_{S_N}\}$; and micro-informed investors informed about stock i choose their demands q_F^i (for the index fund) and q_i^i (for stock i) conditional on $\mathcal{I}_i = \{s_i, P_F, P_{S_1}, \ldots, P_{S_N}\}$. Aggregate holdings of the index fund are given by

$$q_F \equiv \lambda_U q_F^U + \lambda_M q_F^M + \frac{\lambda_S}{N} \sum_{i=1}^{N} q_F^i. \tag{40}$$

Although the market trades the index fund along with individual stocks, only stocks experience exogenous supply shocks. The market should therefore clear for each stock. Each share held in the index fund requires $1/N$ shares of each of the N stocks. The market clearing condition for stock i is therefore

$$\frac{\lambda_S}{N} q_i^i + \frac{q_F}{N} = \frac{1}{N}(X_F + X_i), \quad i = 1, \ldots, N. \tag{41}$$

The first term on the left is the direct demand for stock i from investors informed about that stock; these are the only investors who invest directly in the stock. The second term is the amount of stock i held in the index fund. The right side is the supply shock from (38). The direct and indirect demand for stock i must equal its supply.

We now exhibit a set of prices for which the resultant optimal demands clear the market:

$$P_F = a_F + b_F(m - \bar{m}) + c_F(X_F - \bar{X}_F), \quad \text{with} \quad \frac{c_F}{b_F} = -\frac{\gamma(1-f_M)\sigma_M^2}{\lambda_M}, \tag{42}$$

and with a_F, b_F, and c_F to be specified. For $i = 1, \ldots, N$, set

$$P_{S_i} = \frac{s_i}{R} - \frac{\gamma(1-f_S)\sigma_S^2}{\lambda_S R} X_i. \tag{43}$$

Proposition 3.1 With the prices in (42) and (43) and appropriate coefficients a_F, $b_F > 0$, and c_F, investors' optimal demands clear the market.

These prices and the associated investor demands thus constitute a market equilibrium for fixed λ_M, λ_S, and λ_U. The form of the index fund price P_F (and its coefficients) follows from Grossman and Stiglitz (1980). The price P_{S_i} has the same general form but without a risk premium and with $1/R$ in place of the b coefficient. Both features are consequences of the fact that only the micro-informed trade individual stocks. The last term in (43) reflects the price discount they receive for absorbing the idiosyncratic supply shocks.

Equation (43) gives the price of the hedged security paying S_i. The unhedged stock pays $M + S_i$ and has price $P_i = P_F + P_{S_i}$. The no-arbitrage condition (16) between the index fund and its constituent stocks holds because the average of the P_{S_i}, $i = 1, \ldots, N$, equals zero. Had we not imposed the exchangeability conditions (37) and (39), the price of the index fund would deviate from the average of the stock prices.

It is straightforward to verify that the signals in (35)–(36) and prices in (42)–(43) satisfy assumptions (a1), (a3), and (a5) (see the Internet Appendix for details). We use (a4) to ensure that informed investors can use their full capacity. In the present context, that condition becomes condition (47) (see below), which is satisfied through the choice of f_S and f_M.

3.4 Sufficient conditions for equilibrium

We now extend the preceding analysis to construct an equilibrium when investors choose whether to become informed and how to use their capacity if they do become informed. Note that in the constrained economy of Sections 3.1–3.3, if information frictions were absent, then Proposition 1.3 implies that the marginal agent would be no worse off by deviating from being either micro- or macro-informed to a less concentrated information set. The information frictions of Section 2 are necessary in order for the information choices in the constrained setting to be optimal.

Definition 3.1. An equilibrium consists of prices for all securities and an allocation of investors to types such that (1) the market clears when investors set demands to maximize their expected utilities and (2) Definition 2.1 holds: no investors are allocated to suboptimal types.

We established (1) in Proposition 3.1. We will extend our analysis to satisfy (2). The key feature of the resultant equilibrium will be specialization: all informed investors will be either macro-informed or micro-informed; some investors will remain uninformed and invest in the index fund only. In short, we will have

$$J_M^2 = J_S^2 = J_U^2 < J_x^2, \quad \text{for all other types } x, \tag{44}$$

where we have written

$$J_M^2 \equiv J_{M,0}^2 = e^{2\gamma c}(\kappa/\delta_F)e^{-Q_F}\frac{\mathrm{V}[M|P_F]}{\mathrm{V}[M-RP_F]} \tag{45}$$

for the squared expected utility (27) of the macro-informed, and

$$J_S^2 \equiv J_{1,1}^2 = e^{2\gamma c}(\kappa/\delta)e^{-Q_F}\frac{\mathrm{V}[M|P_F]}{\mathrm{V}[M-RP_F]}\frac{\mathrm{V}[S|P_S]}{\mathrm{V}[S-RP_S]} \tag{46}$$

for the squared expected utility (26) of the micro-informed, and J_U for the expected utility (29) of the ordinary uninformed. In particular, (44) rules out the presence of uninformed investors who invest in individual stocks, that is, the oblivious investors.

In Section 3.1, the informativeness ratios f_M and f_S were specified exogenously. When we embed the market of Sections 3.1–3.3 in the information choice framework of Section 2, the achievable informativeness depends on the capacity κ and the cost parameters c, δ, and δ_F. In particular, when micro-informed and macro-informed investors use their full capacity κ, the constraints (22) and (23) become

$$\frac{\mathrm{V}[S_i|s_i]}{\mathrm{V}[S_i|P_{S_i}]} = \kappa/\delta \quad \text{and} \quad \frac{\mathrm{V}[M|m]}{\mathrm{V}[M|P_F]} = \kappa/\delta_F, \tag{47}$$

respectively. Because the prices P_F and P_{S_i} in (42) and (43) depend on f_M and f_S, the conditions in (47) constrain the values of these parameters. We will see that f_S is pinned down to a specific value, whereas f_M may vary within a limited range.

Define

$$f_S(c) = \frac{\left(e^{2\gamma c} - 1\right)\left(1 - \delta_F e^{-2\gamma c}/\delta\right)}{\left(e^{2\gamma c} - 1\right) - \left(1 - \delta_F e^{-2\gamma c}/\delta\right)}, \tag{48}$$

wherever the denominator is nonzero. For c such that $f_S(c) \in (0, 1)$, define

$$L_S(c) = \left[\frac{1 - f_S(c)}{e^{2\gamma c} - 1}\gamma^2 \mathrm{V}[S]\mathrm{V}[X]\right]^{1/2}; \tag{49}$$

and, for any $f_M \in (1 - e^{-2\gamma c}, 1)$, define

$$L_M(f_M, c) = \left[\frac{\left(f_M + e^{-2\gamma c} - 1\right)(1 - f_M)}{(1 - e^{-2\gamma c})f_M}\gamma^2 \mathrm{V}[M]\mathrm{V}[X_F]\right]^{1/2}. \tag{50}$$

The variables L_S and L_M correspond to λ_S and λ_M once we account for the constraints that these proportions fall between 0 and 1 and that their sum not exceed 1.

Theorem 3.1 (Existence of oblivious-free equilibrium with specialization). *Suppose* $e^{2\gamma c}\kappa = \delta_F < 1$, *and suppose* $8/9 < \delta/\delta_F < 1$. *Let*

$$\beta_\pm = \frac{3 \pm \sqrt{9 - 8\delta_F/\delta}}{2},$$

and suppose $\beta_- < e^{2\gamma c} < \beta_+$. *Then* $\kappa < \delta$. *Furthermore, within this valid cost interval, an equilibrium is given as follows:*

- $f_S(c)$ is as given in (48) and is contained in $(e^{2\gamma c} - 1, 1)$; $\lambda_S(c) = \min(1, L_S(c))$;
- f_M *is any value in* $(1 - e^{-2\gamma c}, 1)$ *for which* $L_M(f_M, c) \leq 1 - \lambda_S(c)$, *and* $\lambda_M = L_M(f_M, c)$;

$$\lambda_U = 1 - \lambda_M - \lambda_S.$$

This result states that for every δ/δ_F ratio in $(8/9, 1)$ there is a range of costs c consistent with an informative equilibrium (meaning that informed investors collectively learn something about every security) with no oblivious investors. In all equilibria described in the proposition, informed investors are micro-informed or macro-informed — the equilibria are characterized by specialization. The value of λ_S is determined by c and δ/δ_F. In contrast, we may have a range of values of λ_M consistent with equilibrium at fixed pair $(c, \delta/\delta_F)$. At each c, the largest value of $L_M(f_M, c)$ (and thus the largest potential value of

λ_M) is attained at $f_M^* = \sqrt{1 - e^{-2\gamma c}}$; and the smallest value is $L_M(f_M, c) = 0$ for $f_M = 1 - e^{-2\gamma c}$. The range of allowable f_M and λ_M represent a range of equilibria, but all with the essentially the same structure. We derive (48)–(50) using the equalities in (44) and (47).

We met the condition $e^{2\gamma c}\kappa = \delta_F$ in Corollary 2.1 in balancing the expected utility of the uninformed and the informed. Similarly, setting $J_M^2 = J_S^2$ in (45)–(46) yields

$$\delta/\delta_F = \frac{\mathrm{V}[S|P_S]}{\mathrm{V}[S - RP_S]} = \frac{\mathrm{V}[S - RP_S|P_S]}{\mathrm{V}[S - RP_S]} \leq 1,$$

with strict inequality if S and P_S have nonzero correlation; hence the condition $\delta/\delta_F < 1$ in the proposition. This, together with the Theorem 3.1 result that $\kappa < \delta$, ensures that (24) is satisfied, which guarantees a net information gain from becoming informed. We show in Lemma A.2 that the key to keeping out oblivious investors is having $f_S > e^{2\gamma c} - 1$; in other words, the signal-informativeness measure f_S must be sufficiently large, relative to the cost of becoming informed. The valid cost interval defined by $\beta_\pm$ enforces this condition for $f_S(c)$ in (48). Outside this interval oblivious investors drive out all other investors leading to a degenerate scenario in which no investor learns anything about any security's dividend.[13]

3.5 Analysis of equilibrium

To gain intuition into Theorem 3.1, we now construct a numerical example of equilibrium, using the parameter values in Table 1 of the Internet Appendix. We may take as a model primitive the ratio δ/δ_F and note that for each value of this ratio, the limits $\beta_\pm$ in Theorem 3.1 define an admissible interval of costs c. The allowable set of pairs $(c, \delta/\delta_F)$ are the points above the U-shaped curve in the left panel of Figure 2. We refer to this as the *admissible cost region*.

At each admissible pair $(c, \delta/\delta_F)$, recalling that $f_M^*(c) = \sqrt{1 - e^{-2\gamma c}}$, the largest possible share of macro-informed λ_M^* is given by

$$\lambda_M^*(c) = \begin{cases} L_M(f_M^*(c), c), & \text{if } L_M(f_M^*(c), c) < 1 - \lambda_S(c); \\ \max(0, 1 - \lambda_S(c)), & \text{otherwise (constrained case).} \end{cases} \tag{51}$$

We consider f_M^* the largest plausible value for f_M in the unconstrained region: at any larger value of f_M, macro prices would be so informative

[13] For example, with $f_S = 0$, prices are uninformative about dividends and the term inside parentheses in (31) becomes $\mathrm{V}[S]/(\mathrm{V}[S] + \mathrm{V}[RP_S]) < 1$, so the oblivious prefer to invest in as many stocks as possible. Moreover, comparing the utility of the micro-informed in (46) with the utility of the oblivious in (31), we find that $J_S^2 = \delta_F J_{U,1}^2/\delta > J_{U,1}^2 > J_{U,K}^2, K \geq 2$, so being oblivious is strictly preferred to being informed about individual stocks. The result is an equilibrium in which no investors learn about stocks and prices are uninformative about dividends. In this equilibrium, prices change solely in response to supply shocks, which makes trading in these stocks attractive to the oblivious.

that investors would be discouraged from acquiring macro information and λ_M would decrease. The constrained case occurs at low values of c where λ_S becomes large, thus squeezing out the macro-informed. There are two values of $f_M(c)$ consistent with each $\lambda_M(c)$ in the constrained case. (Setting $L_M(f_M, c)$ equal to a constant in Equation 50 yields a quadratic equation in f_M.) The larger root exceeds $f^*_M(c)$, so we consider the smaller root more meaningful. The equilibrium share of uninformed is given by $\lambda_U = 1 - \lambda_S - \lambda^*_M$, so $\lambda_U = 0$ in the constrained region.

The right panel of Figure 2 shows the equilibrium shares of uninformed, macro- and micro-informed as functions of c when $\delta/\delta_F = 0.98$. The constrained region of λ^*_M occurs for c to the left of the peak in the λ^*_M curve (slightly to the left of $c = 0.02$). At this point, the share of micro-informed is so large that, even with the uninformed share at zero, it begins to squeeze out the macro-informed. The two vertical bars represent the maximum and minimum allowable c (the valid cost interval) for the delta ratio of 0.98. In the constrained region, the maximum λ^*_M is increasing in c, which may be surprising. But it is important to remember that the share λ^*_M is based on the macro-informativeness $f_M = f^*_M(c)$, which is increasing in c. We consider the unconstrained region to the right of the peak in λ^*_M to be the more interesting region because it includes uninformed investors, $\lambda_U > 0$, even at the maximum level of macro-informed investors.

3.6 Micro versus macro efficiency

A natural measure of price efficiency is the fraction of price variation that comes from the signal of the informed, which also equals the square of the correlation of the price and the signal. In the case of m and the index fund price P_F from (42) the squared correlation is

$$\rho_F^2 = \frac{f_M}{f_M + \gamma^2(1-f_M)^2\sigma_M^2\sigma_{X_F}^2/\lambda_M^2}; \tag{52}$$

this follows from dividing the last term in (A19), which equals $\mathrm{cov}[M, P_F]^2/\mathrm{V}[P_F]$, by $\mathrm{V}[m] = f_m\sigma_M^2$ and noting that $\mathrm{cov}[M, P_F] = \mathrm{cov}[m, P_F]$. Similarly, for individual stocks we define ρ_S^2 to be the squared correlation between s_i and P_{S_i} in (43). The last term in (A14) yields

$$\rho_S^2 = \frac{f_S}{f_S + \gamma^2(1-f_S)^2\sigma_S^2\sigma_X^2/\lambda_S^2}. \tag{53}$$

In both cases, as $\rho^2 \to 1$ prices become fully revealing. We say that the market is *macro-efficient* if $\rho_F^2 > \rho_S^2$; it is *micro-efficient* if the inequality is reversed.

In the left panel of Figure 2, B marks the constrained region (with no uninformed); we focus on regions A and C. Throughout C, the market is

Table 1
Mutual fund performance and PBN

Gross returns: high- minus low-PBN performance

	TR full idx (1)	TR full nonidx (2)	TR 2011 idx (3)	TR 2011 nonidx (4)	CRSP 2011 idx (5)	CRSP 2011 nonidx (6)
Excess return	−0.0042**	−0.0041**	−0.0059***	−0.0059***	−0.0014	−0.0012
	(0.0018)	(0.0018)	(0.0017)	(0.0016)	(0.0027)	(0.0027)
α_{ff}	−0.0034**	−0.0037**	−0.0026*	−0.0037***	−0.0026*	−0.0042***
	(0.0017)	(0.0018)	(0.0015)	(0.0014)	(0.0016)	(0.0015)
α_{mkt}	−0.0050***	−0.0051***	−0.0051***	−0.0056***	−0.0057**	−0.0061***
	(0.0016)	(0.0016)	(0.0017)	(0.0016)	(0.0024)	(0.0023)

Net returns: high- minus low-PBN performance

	TR full idx (1)	TR full nonidx (2)	TR 2011 idx (3)	TR 2011 nonidx (4)	CRSP 2011 idx (5)	CRSP 2011 nonidx (6)
Excess return	−0.0050***	−0.0046***	−0.0067***	−0.0063***	−0.0024	−0.0017
	(0.0018)	(0.0018)	(0.0017)	(0.0016)	(0.0027)	(0.0027)
α_{ff}	−0.0042**	−0.0041**	−0.0034**	−0.0041***	−0.0036**	−0.0047***
	(0.0017)	(0.0018)	(0.0015)	(0.0015)	(0.0016)	(0.0015)
α_{mkt}	−0.0058***	−0.0056***	−0.0058***	−0.0060***	−0.0067***	−0.0066***
	(0.0015)	(0.0016)	(0.0017)	(0.0016)	(0.0025)	(0.0023)

The table shows the quarter t average excess returns and alphas of a portfolio that is long funds in the top quintile based on their quarter $t-1$ PBNs and short funds that are in the bottom quintile based on their quarter $t-1$ PBNs. Returns are weighted by each bucket's funds' total net assets at the end of quarter $t-1$. Alphas are calculated using the five-factor Fama-French model augmented with momentum (α_{ff}) or using the CAPM (α_{mkt}). The columns correspond to the full-sample Thomson-Reuters data set, the Thomson-Reuters data set starting in 2011 and forward, or the CRSP data set starting in 2011. All three variants are shown including (idx) or excluding (nonidx) index funds. The top panel uses gross (before fees are subtracted) mutual fund returns, and the bottom panel uses mutual fund returns net of fees (i.e., after fees have been subtracted from returns). The returns data are quarterly. Standard errors, shown in parentheses, use Newey-West with two lags.
$*p < 1$, $**p < 0.05$, $***p < 0.01$

micro-efficient, and in A it is macro-efficient.[14] In the Internet Appendix, we show that the boundary separating C from A is defined by the inequality

$$\delta/\delta_F \geq \frac{1+e^{-2\gamma c}\sqrt{1-e^{-2\gamma c}}}{1+\sqrt{1-e^{-2\gamma c}}}. \tag{54}$$

We emphasize three observations from Figure 2: (1) The micro-efficient region C is much larger than the macro-efficient region A. Moreover, (54) shows that the only model parameter that affects the A–C boundary is the risk-aversion parameter, which enters only through the product γc; thus, the comparison in the figure is quite robust to parameter choices.[15] (2) At any

[14] In region B, ρ_F^2 depends on which of the two roots mentioned following (51) is chosen for f_M. We have observed that in most cases the larger root results in macro-efficiency and the smaller root in micro-efficiency. As noted previously, we consider the smaller root more relevant.

[15] The boundary defining B, which we explicitly derive in the Internet Appendix, depends on other model parameters.

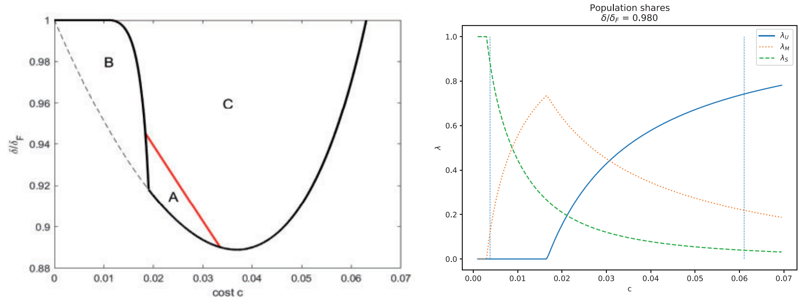

Figure 2
Admissible parameter region and population shares
In the left panel, the U-shaped curve shows the boundary of admissible pairs, $(c, \delta/\delta_F)$. Region B corresponds to the constrained case in (51). Micro-efficiency holds in C and macro-efficiency in A. The right panel shows the population shares as a function of c. The vertical bars in the right panel show the allowable range of cs for the given δ/δ_F ratio.

cost level c, increasing δ/δ_F moves the market toward micro-efficiency; this is evident in the figure and in (54). (3) At any level of δ/δ_F, increasing c moves the market toward micro-efficiency; this is evident in the figure and confirmed by (54) because the expression on the right side of (54) is decreasing in c.

To gain further insight into observation (2), it is helpful to consider the effect of varying the macro informativeness parameter f_M, recalling that Theorem 3.1 allows a range of values for f_M. In (52) and (53), we can replace λ_M and λ_S with their equilibrium values to get

$$\rho_F^2 = \frac{1 - e^{2\gamma c}(1 - f_M)}{f_M}, \quad \rho_S^2 = \frac{1 - (\delta/\delta_F)e^{2\gamma c}(1 - f_S)}{f_S}. \tag{55}$$

Each ρ^2 is increasing in the corresponding f. If we have $f_M = f_S \equiv f$, then

$$\frac{\rho_S^2}{\rho_F^2} = \frac{1 - (\delta/\delta_F)e^{2\gamma c}(1 - f)}{1 - e^{2\gamma c}(1 - f)} > 1, \tag{56}$$

for $\delta/\delta_F < 1$. In other words, the greater attention required to learn about a stock compared with learning about the overall market biases the market toward micro-efficiency in equilibrium. Equivalently, to have $\rho_F^2 = \rho_S^2$ requires $f_M > f_S$: achieving equal price efficiency requires that the macro signal be more informative than the micro signal.

This effect can be seen in condition (47), which reflects the requirement that informed investors use their full capacity. We can rewrite (47) as

$$\frac{\mathrm{V}[S_i|P_{S_i}]}{\mathrm{V}[S_i]} = \frac{\delta}{\kappa}(1 - f_S), \quad \frac{\mathrm{V}[M|P_F]}{\mathrm{V}[M]} = \frac{\delta_F}{\kappa}(1 - f_M).$$

With $\delta < \delta_F$ and $f_M = f_S$, the micro prices P_{S_i} need to be relatively more informative about dividends than the macro price. Equally informative signals *require* micro prices to be more informative because of the steeper micro fixed attention cost δ. As δ increases, the fixed attention cost consumes less of the micro-informed's capacity, allowing them to acquire more informative signals. In other words, as δ increases, f_S increases, increasing ρ_S^2 and moving the market toward micro-efficiency.[16]

Observation (2) holds for any choice of f_M. For our discussion of observation (3), we focus on the λ_M-maximizing choice f_M^*, which we argued following (51) is the largest plausible value for f_M and thus produces the largest plausible level of macro-efficiency ρ_M^2. Both $f_S(c)$ and $f_M^*(c)$ are increasing in c: investors need to be induced to bear the higher cost of becoming informed through more informative signals. In Theorem 3.1 we have the first inequality in

$$f_S(c) > e^{2\gamma c} - 1 > \sqrt{1 - e^{-2\gamma c}} = f_M^*(c), \tag{57}$$

and the second inequality holds for $e^{2\gamma c}$ greater than (the golden mean) $(1 + \sqrt{5})/2$.[17] With $\gamma = 5.5$, this threshold becomes approximately $c \geq 0.044$. Thus, for sufficiently large c, we are guaranteed to have $f_S(c) > f_M^*(c)$ and thus micro-efficiency $\rho_S^2 > \rho_F^2$. (The precise threshold is provided by equation 54.)

In understanding the bias toward micro-efficiency, it is helpful to contrast the equilibrium of Theorem 3.1 with an equilibrium in which oblivious investors drive out the micro-informed and stock prices carry no information about dividends; see the discussion at the end of Section 3.4. The first inequality in (57) (which holds in the valid cost interval) ensures that being micro-informed is preferred to being oblivious; see Lemma A.2. In studying the equilibrium of Theorem 3.1, we are conditioning on the micro signal being sufficiently informative to attract informed investors, despite the cost and attention required to be micro-informed. Moreover, the micro-informed are not deterred by prices becoming overly revealing because no other investors learn from micro prices: oblivious investors ignore the information in prices, ordinary uninformed investors do not have the capacity to follow individual stocks, and other informed investors specialize in other securities. In contrast, macro information is partly revealed to the uninformed and the micro-informed, who are able to trade in the index fund because doing so does not entail an attention penalty. The micro informed therefore monopolize liquidity provision for micro supply shocks, whereas the macro informed must share in liquidity provision for macro supply shocks. This induces some

16 As δ/δ_F approaches 1, we have $f_s(c) \to 1$ in (48), and then $\rho_S^2 \to 1$, indicating perfect micro-efficiency. However, right at $\delta/\delta_F = 1$, Theorem 3.1 fails to describe an equilibrium. In particular, (49) would imply $\lambda_S = 0$, so the market for individual stocks could not clear and (43) would not be a valid price.

17 Letting $\beta = e^{2\gamma c} > 1$, the middle inequality in (57) requires $\beta^3 - 2\beta^2 + 1 > 0$. This factors to $(\beta - 1)(\beta^2 - \beta - 1)$.

agents to become micro informed despite the higher fixed cost. In turn, the higher fixed cost forces micro prices to be more informative to allow the micro informed to fully utilize the information content of their signals.

4. Tests of Model Predictions about Specialization

According to our model, investors should either specialize or completely diversify by owning an index fund.[18] Intermediate consideration sets are not optimal. In this section, we use mutual fund holdings data as a proxy for the portfolios of informed investors and explore several implications of our framework. We focus on specialization; our empirical investigation does not address micro- versus macro-efficiency. The core friction in our model is a fixed attention cost of following a stock: adding a new security to an investor's consideration set consumes a portion of that investor's information capacity. Absence of information frictions, as highlighted in Proposition 1.2, suggests that, in a general setting, investors prefer larger consideration sets.

Prediction 1 *(of frictionless model)*. Mutual funds prefer larger consideration sets.

On the other hand, according to Proposition 2.3, investors who face the information frictions of Section 2 should specialize. To operationalize this idea in a dynamic setting, we assume that investors' attention costs need to be paid only once, or perhaps only periodically. Once a security enters an investor's consideration set, it stays there for some time. Our prediction that investors specialize suggests that investors should own a small set of stocks, and that this set of stocks should not change much over time. A problem with this observation as an empirical strategy is that investor portfolios may be persistent due to the costs of trading, and not because of informational frictions. To isolate the effect of information frictions rather than transaction costs, we propose a novel measure that looks at the percentage of all firms in which a mutual fund buys shares in a given quarter that are new to the fund's portfolio. We refer to this percentage of firms bought that are new as a fund's raw *PBN*. For fund i in quarter t, this is defined as

$$PBN_{i,t}^{raw} = \frac{\text{number of firms bought in } t \text{ that are new to the portfolio}}{\text{number of firms in which the fund bought stocks in quarter } t}.$$

For example, if a fund purchased stocks in eight firms in a quarter, and five of these firms were new to the fund's portfolio, the fund's PBN would

[18] In practice, we do not expect the prediction that it is optimal to hold a single stock to be literally true. If we were to allow a fixed attention cost to cover multiple, related stocks, we would find that it is optimal to hold such related stocks, as opposed to a single stock. In the model, we sacrificed this degree of realism for the benefit of simplicity.

equal 62.5%. We measure the PBN using numbers of firms (rather than dollar amounts invested, or number of shares bought) because in our model investing in a new firm requires the payment of a fixed attention cost, regardless of the size of the position. If a fund did not purchase any shares of stock in a quarter, its PBN for that quarter is missing.
We normalize the raw PBN by dividing it by the percentage of firms the fund does not own,

$$PFNO_{i,t} = 1 - \frac{\text{number of firms owned by the fund}}{\text{total number of firms owned by all funds in a given quarter}},$$

to get the adjusted PBN, which we use in our analysis:

$$PBN_{i,t} = \frac{PBN^{raw}_{i,t}}{PFNO_{i,t}}. \tag{58}$$

Under the null hypothesis that a given fund is as likely to buy stock in a firm that it does not own as in a firm that is does own, the normalized PBN should be equal to 100%. To the extent that, conditional on the decision to buy, funds are more likely to invest in firms that they already own, the PBN should be less than 100%. Thus, PBN naturally controls for portfolio persistence due to transaction costs because it conditions on a fund's decision to buy stocks in the first place.[19] PBN is a noisy measure of the information acquisition activity of funds because it only identifies research efforts that result in share purchases, and not those that result in no purchases. Nevertheless, funds with low PBNs focus on stocks they already own, and are therefore specialized; and funds with high PBNs frequently buy stocks outside of their existing area of expertise, and are therefore less specialized.

Prediction 2 *(of model with attention costs)*. Mutual funds typically have low PBNs.

Corollary 2.1 shows that in order to have ordinary uninformed investors in equilibrium the cost of becoming informed c should be proportional to the log of the attention cost of macro information δ_F. Theorem 3.1 shows that to construct an equilibrium that is not overrun by oblivious investors who crowd out all informed investors—which we believe to be the empirically relevant case—the ratio δ/δ_F should be in the interval $(8/9, 1)$. Together these two results suggest that the fixed

[19] We do need to assume that both types of stocks have similar liquidity profiles. But based on how concentrated fund holdings are, as shown in Section 4.2, it is fair to assume that many non-owned stocks have very similar liquidity characteristics to those of owned stocks.

attention cost parameters δ and δ_F are, over long periods of time, monotonically increasing functions of c.[20]

According to figure 1 of Philippon (2012), the proportion of the U.S. economy represented by the financial sector[21] increased from 4% to 5% in the 1920s and 1930s to 8%–9% by the 1990s, with the finance industry share of the economy showing a clear increasing trend throughout the twentieth century. Greenwood and Scharfstein (2013) and Philippon and Reshef (2013) provide similar U.S. and international evidence, respectively, for growth in the finance sector share of the economy. We interpret these results as indicating that society has devoted an increasing share of its productive resources, including human capital, to the finance sector over the prior century. Further evidence of an increase in informed investors comes from Lakonishok, Shleifer, and Vishny (1992). They document that from 1955 to 1990 institutional ownership of equities increased from 23% to 53%, while equity ownership by individuals fell from 77% to 47%. The trend from 1990 to today goes in the same direction: the Investment Company Institute (2018, p. 36) documents that the share of household financial assets held in investment companies increased from 3% in 1980 to 24% in 2017. Together, these results suggest that λ_U has fallen over time. According to Figure 2, this implies that c fell.

As c falls, δ_F and δ should also fall, which, according to the information constraint in (23), *increases* the cost of nonspecialization. We therefore expect to see specialization increasing in our mutual fund sample over time.

Prediction 3 *(of model with attention costs)*. Because of a falling c, mutual funds' PBNs have decreased over time.

Finally, a lack of specialization reduces the amount of investors' information capacity that can be spent on acquiring payoff-relevant information. Having worse information decreases investors' expected profits. This is exactly true when $\gamma \to 0$ because investors then become risk-neutral, and approximately true when $\gamma > 0$.[22] On the other hand, specialization should increase investors' profits.

[20] With κ fixed, δ_F is monotonic in c, and $8/9\delta_F < \delta < \delta_F$. Therefore, for large movements in δ_F caused by changing c, δ must move move as well to stay within the prescribed range.

[21] The Philippon (2012) series splices together the financial sector value added over GDP ratio and the labor compensation share of the finance industry. When both series are available (in the latter part of the sample) they track each other very closely.

[22] In a related finding, Kacperczyk, Sialm, and Zheng (2005) find that funds with higher industry concentrations tend to outperform less concentrated funds.

Prediction 4 *(of model with attention costs)*. Mutual funds with lower PBNs earn higher returns.

Our empirical tests focus on Predictions 1–4, which are grounded in the attention costs of micro information. Mutual fund holdings are less well suited to finding evidence of specialization in macro information. In practice, we expect that macro-informed investors would vary their overall level of investment in the stock market based on macro signals, possibly trading through an index fund, but this variation would not be reflected in the composition of stock holdings of index funds.

4.1 Mutual fund data

Our mutual fund holdings come from the Thomson Reuters Mutual Fund Ownership Data (S12), and from the CRSP Survivor-Bias-Free U.S. Mutual Fund database. The Thomson Reuters (TR) holdings data run from 1980 to 2018 and the CRSP holdings data from 2011 to 2018.[23] The TR database provides a longer time series of holdings data, but in the period of overlap, the CRSP database offers better coverage. We perform all subsequent analysis using both databases, and the results are consistent between the two.

We obtain mutual fund returns from the CRSP mutual fund returns database, which starts in 1983. We map the TR funds to the CRSP returns data using a mapping table provided by WRDS. We perform all of our tests separately for the entire universe of mutual funds, as well as for all mutual funds but with index funds excluded. We exclude all ETFs and exchange-traded notes (ETNs) from the analysis. Fund performance is analyzed in two ways, by using either gross returns (before fees have been charged) or net returns (after fees have been charged). Net returns and fees (expense ratios) are obtained from the CRSP mutual fund returns database, and gross returns are calculated as the sum of net returns and expense ratios. All fund returns are aggregated to the quarterly level.

The Internet Appendix contains further details, including our fund selection and data-cleaning methodology, index fund and ETF classification scheme, and fund mapping process.

4.2 Results

Prediction 1 says that in the frictionless setting of Section 1 investors should prefer larger consideration sets. Figure 3 shows the 25th, 50th, and 75th percentiles of mutual fund security holdings over time, using data aggregated at the quarterly level. The number of securities (left panel) held by the median mutual fund hovers between 60 and 80 for most of the sample. The number of

[23] The CRSP holdings data start in 2008, but there is a large jump in the number of funds that are covered only in late 2010.

distinct stocks held by the median mutual fund increases from roughly 40 to the mid-60s by the end of the sample. The 25th and 75th percentiles of security and stock holdings by mutual funds range from 20 to around 150. Given that throughout this time period, there are thousands of stocks (as well as thousands of additional securities, like bonds) that funds can hold, Figure 3 provides strong evidence against a setting without information frictions. As we show in the Internet Appendix, the average fund's PFNO (percentage of firms not owned) hovers around 98%. Funds focus on a very small subset of all possible securities and clearly do not prefer ever larger consideration sets. This is in stark contrast to the prediction of the no-information-friction benchmark model that every fund should have positions in thousands of individual stocks.

The essence of Prediction 2 is that with fixed attention costs, mutual funds should prefer to specialize. We test this by checking whether mutual fund PBNs are "low." Consider a fund that owns stocks of 50 firms and receives investor inflows. The fund looks to invest this money. We assume that the fund faces transaction costs to rebalance its portfolio, which are currently high enough that it does not sell stock in any of the 50 firms currently in its portfolio. If the fund's consideration set consists of all traded stocks, as would be the case in the no-frictions benchmark, the fund then observes a signal about each of several thousand stocks, say 3,000, and chooses to invest its new capital into whichever one of these stocks has the highest expected return. Under an assumption of independence, the probability that the new investment is made into a currently held stock is 50/3,000, and the probability that the investment is made in the stock of a new stock, not currently held, is 2,950/3,000. Repeating this experiment many times, we would find this fund's PBN to be close to 100%, under the null hypothesis of no information frictions.

Figure 4 shows the median, and the 25th and 75th percentile, of fund PBNs over time. The left panel shows the PBNs in the sample including index funds, and the right panel shows the PBNs in the sample excluding index funds. We note that the median fund's PBN is typically between 30%–60%. Assuming the percentage of firms the fund currently does not own is 98% (the historical average), a PBN of 0.5 means for every 100 newly initiated buy transactions, 49 (0.5×98) come from the fund's existing set of owned firms. Under the null hypothesis that the fund is equally likely to buy any stock, this number should be two. Thus, funds with PBNs in the vicinity of the 25th, 50th, or 75th percentile of the PBN distribution initiate buy transactions in firms they already own much more frequently than they would if they had very large consideration sets. This is strong evidence in favor of the hypothesis that adding a new position to an existing set of firms is costly, not because of the transaction costs of trading, but because of the informational resources that are required.

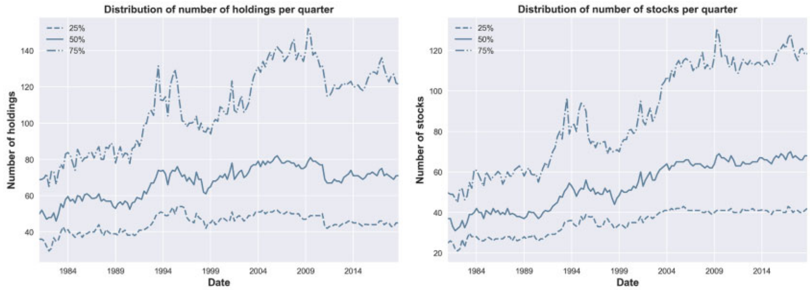

Figure 3
Mutual fund holdings
The left (right) panel shows the distribution of the number of securities (unique stocks) held by mutual funds over time. These data are from the Thomson Reuters database.

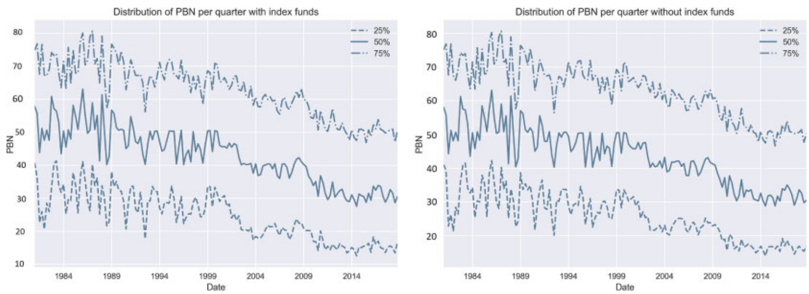

Figure 4
PBN distribution
Distribution of PBN (percentage of stock bought that is new to the portfolio in each quarter) of mutual funds over time. The mutual fund data are from the Thomson Reuters database. The left panel includes index funds, and the right panel excludes index funds.

Another notable trend in Figure 4 is that PBNs have been falling over time. The median fund's PBN started the sample at close to 60% in the early 1980s and has steadily declined over time to a 2020 level of around 30%. This is consistent with Prediction 3 that a fall in the cost of becoming informed will decrease δ, making nonspecialization more costly, and pushing fund managers toward more concentrated portfolios. The drop in mutual fund PBNs over the last several decades is as pronounced a pattern as the well-documented movements of assets under management toward passive strategies (note that the trend in PBNs is nearly identical in the right panel of Figure 4, which excludes index funds). Thus, the mutual fund industry has been transitioning over the last several decades toward index investing and toward more concentrated portfolios among the nonindexers. These are the two optimal information sets from Proposition 2.3.

According to Proposition 2.3, nonindex funds find it optimal to specialize, and, therefore, nonspecialized funds are utilizing their information capacity suboptimally and should earn lower returns. To test this last (4th) prediction, we classify funds into quintile buckets based on their PBNs in quarter $t-1$, and check whether the low-PBN (specialized) funds have higher risk-adjusted returns than the high-PBN (nonspecialized) funds in quarter t. The quarter $t-1$ classification is based on pooled, full-sample PBN quintiles. In quarter t we look at the difference in returns between bucket 5 (nonspecialized) funds and bucket 1 (specialized) funds. Each bucket's return is calculated as an average quarter t return where the weights are the total net assets at the end of quarter $t-1$ of each of the bucket's funds.

We look at six different fund subsets for this analysis: the full sample of TR fund returns (TR full idx); the full sample of TR fund returns excluding index funds (TR full nonidx); the same two fund groupings using TR holdings data starting in 2011 (TR 2011 idx and TR 2011 nonidx); and the same two fund groupings using CRSP holdings data starting in 2011 (CRSP 2011 idx and CRSP 2011 nonidx). We restrict the TR data to the same window as the CRSP data as a robustness check. We perform the analysis using either gross returns (before fees) or net returns (after fees), and using different risk adjustment methods. Gross returns are a cleaner test of the effect of information choice on fund performance because funds have discretion in the fees that they charge investors. However, net returns are the more relevant measure for investors. For these reasons, we report both sets of results.

The first row of each panel of Table 1 shows the average excess returns of PBN bucket 5 over PBN bucket 1. This excess return cannot be achieved via a trading strategy because mutual funds cannot be sold short, but it highlights the difference in returns between less and more concentrated mutual funds. The second row of each table reports the quarterly alpha of 5-1 excess returns after controlling for the Fama and French (2015) five-factor model augmented with momentum. The third row shows the alpha of the 5-1 excess returns after controlling only for the market exposure. Returns are quarterly, and shown as decimals, e.g., -0.05 means a negative 5% return. Each column corresponds to a database (TR or CRSP) and time period (full sample or 2011 and after) combination. The top panel in the table shows the results using gross returns, and the bottom panel uses net returns.[24]

Let us focus first on the "TR full nonidx" column in the gross returns panel. High PBN funds underperform low PBN funds by 41 basis points per quarter. If we adjust for the five Fama-French factors and momentum, the underperformance of less concentrated funds declines to 37 basis points, and controlling for only the market, the underperformance increases to 51 basis points. In all cases, the results are highly statistically significant. Multiplying these quarterly results by four shows that *more* concentrated funds (low PBNs) earn 1.5%–2% higher annualized returns than *less* concentrated

[24] Factor data are obtained from Kenneth French's website. The full regression results are in the Internet Appendix.

funds (higher PBNs). The "TR full nonidx" column in the bottom panel shows that this return differential is higher by about 20 basis points per year when looking at net returns, the more relevant measure for investors. The other columns in Table 1 show all other data set and time period combinations. The results are very consistent. Using either the TR or CRSP database, over the full or partial sample, with or without index funds, low PBN mutual funds outperform high PBN mutual funds by at least 1.5% per year. The evidence is highly supportive of Prediction 4: larger consideration sets are wasteful of scarce information capacity, and funds that trade too many stocks underperform those that are more focused.

4.3 PBNs and fund characteristics

To better understand the types of funds that tend to be low or high PBN, we classify all fund-quarters into PBN buckets based on full-sample pooled PBN quintiles and the fund's prior quarter PBN. We then analyze the characteristics of funds that fall in each quintile (with 1 the lowest, and 5 the highest). For this analysis, we restrict attention to the CRSP data because of its more complete coverage of fund characteristics. Compared to low-PBN funds, the highest PBN funds are smaller based on their total net assets, are managed by smaller fund groups, charge higher fees, and have considerably higher turnover. Surprisingly, high-PBN funds also have longer-serving portfolio managers. Using CRSP mutual fund objective codes, we find that growth funds have high PBNs, while income funds have low PBNs. Large cap funds rarely have high PBNs, while mid- and small-cap funds have hump-shaped PBN distributions, with more middle PBN funds than either high or low ones. Finally, funds that are classified as having a sector specialization have low PBNs, which is consistent with low PBN funds being more concentrated. This rich heterogeneity in fund characteristics across PBN quintiles is analyzed further in Section I.3.8 of the Internet Appendix.

To gauge whether the performance results in Table 1 are driven by funds' PBNs or characteristics, we double sort our fund-quarter observations into PBN-characteristics groups. The three characteristics we consider are the size of the fund's adviser group, the fund's total net assets, and the fund's CRSP objective code. It is possible, for example, that high PBN funds underperform because they are managed by smaller advisors who have fewer resources to devote to research. A similar argument can be made about high PBN funds that have low total net assets. Alternatively, the fund style, for example, growth versus income, may determine performance, and style may happen to be correlated with PBN. Double sorting allows us to measure the performance of high- versus low-PBN funds, while controlling for fund characteristics by focusing on a particular characteristic group. In the results detailed in Section I.3.8 and Table I11 of the Internet Appendix, we show that high-PBN funds underperform low-PBN funds across virtually all fund characteristics that we consider.[25] Combined with the results in Table 1, this is strong evidence in support of Prediction 4.

5. Conclusion

Our model predicts that when information capacity is limited, but in the absence of other information frictions, investors are indifferent to their choice of information, as long as their full information capacity is used. Our information capacity constraint limits the variance reduction investors can obtain from their signals, but does not penalize investors for conditioning on prices. Furthermore, investors prefer larger consideration sets to smaller consideration sets. These results continue to hold after we impose a factor structure on returns, where idiosyncratic shocks are exchangeable and fully diversifiable in the sense that they sum to zero by construction.

We argue that ever-increasing consideration sets are implausible because, even when making inference from prices is costless, as in our setting, investors must still learn some context about each stock that they trade, and this process is costly. We model this via a fixed attention cost for each new stock that an investor chooses to trade. Investors can trade the index fund without any information penalty, but learning macro information relevant to the index payoff also carries a fixed attention cost. In this setting, we obtain a specialization result in an informative equilibrium: investors strictly prefer to learn macro information and invest in the index fund, or to learn micro information and trade in a single stock. We do not take the "single" stock aspect of this result literally, but interpret this as proxying for a collection of related stocks.

We then construct an economy that satisfies the conditions of an informative equilibrium, and show that the equilibrium is generally characterized by micro- rather than macro-efficiency. This result is consistent with Samuelson's dictum that markets are better at figuring out the "correct" prices for individual stocks than for aggregate market indexes.

Finally, we use U.S. mutual fund data to test several predictions of our model, and find that mutual funds: do not favor large consideration sets; trade as if adding a new security to an existing portfolio is costly; have tended to become more specialized over time; outperform when they are more specialized. The latter result is not driven by fund characteristics. All of the empirical findings are consistent with our assumption that adding a new stock to an investment portfolio entails a fixed attention cost.

[25] Of our 18 characteristic sorts, only "Hedge & short" objective funds, of which there are very few, have a positive and significant high- versus low-PBN return, and then only in the sample without index funds.

A Appendix

A.1 Expected Utility Comparison

We begin with a preliminary lemma concerning an arbitrary set of random variables $\mathcal{S} = \{X_1, \ldots, X_K, Y_1, \ldots, Y_N\}$ whose covariance matrix has full rank.

Lemma A.1. *The determinant ratios*

$$\frac{|\mathrm{cov}[Y_1, \ldots, Y_n | X_1, \ldots, X_K]|}{|\mathrm{cov}[Y_1, \ldots, Y_n]|} \quad \text{and} \quad \frac{|\mathrm{cov}[Y_1, \ldots, Y_n | X_1, \ldots, X_n]|}{|\mathrm{cov}[Y_1, \ldots, Y_n]|}. \tag{A1}$$

are decreasing in $n = 1, \ldots, N$.

Proof. The proof uses two properties. The first is the identity

$$|\mathrm{cov}[Y_1, \ldots, Y_n | X_1, \ldots, X_k]| = \frac{|\mathrm{cov}[Y_1, \ldots, Y_n, X_1, \ldots, X_k]|}{|\mathrm{cov}[X_1, \ldots, X_k]|}, \tag{A2}$$

which holds, in particular, with $k = n$ or $k = K$ (see, e.g., Anderson 2003, theorem A.3.2). It follows that the two ratios in (A1) can be written as

$$\frac{|\mathrm{cov}[Y_1, \ldots, Y_n, X_1, \ldots, X_k]|}{|\mathrm{cov}[Y_1, \ldots, Y_n]| \cdot |\mathrm{cov}[X_1, \ldots, X_k]|}, \tag{A3}$$

with $k = K$ and $k = n$, respectively.

To state the second property we need, for any subset $A \subseteq \mathcal{S}$, let $D(A)$ be the determinant of the covariance matrix of the random variables in A. Then for any $A, B \subseteq \mathcal{S}$,

$$D(A \cap B) D(A \cup B) \leq D(A) D(B); \tag{A4}$$

(see, e.g., Fan 1968). Sometimes called "Koteljanskii's inequality," this is equivalent to the statement that the differential entropy of a set of Gaussian random variables is submodular, because the entropy is proportional to the log determinant of the covariance matrix. Inequality (A4) captures the idea reflected in (A1) of diminishing value of conditioning information.

Write X for $X_1, \ldots, X_K$. It follows from (A4) that

$$D(Y_1, \ldots, Y_{n-1}) D(Y_1, \ldots, Y_n, X) \leq D(Y_1, \ldots, Y_{n-1}, X) D(Y_1, \ldots, Y_n).$$

So,

$$\frac{D(Y_1, \ldots, Y_n, X)}{D(X) D(Y_1, \ldots, Y_n)} \leq \frac{D(Y_1, \ldots, Y_{n-1}, X)}{D(X) D(Y_1, \ldots, Y_{n-1})}.$$

The two sides of this inequality have the form in (A3) with $k = K$. We have therefore shown that the first ratio in (A1) is decreasing in n. Similarly, two applications of (A4) yield

$$\begin{aligned} & D(Y_1, \ldots, Y_n, X_1, \ldots, X_n) D(X_1, \ldots, X_{n-1}) D(Y_1, \ldots, Y_{n-1}) \\ \leq\ & D(Y_1, \ldots, Y_n, X_1, \ldots, X_{n-1}) D(X_1, \ldots, X_n) D(Y_1, \ldots, Y_{n-1}) \\ \leq\ & D(Y_1, \ldots, Y_{n-1}, X_1, \ldots, X_{n-1}) D(X_1, \ldots, X_n) D(Y_1, \ldots, Y_n). \end{aligned}$$

Rearranging the first and last expressions, we get

$$\frac{D(Y_1,\ldots,Y_n,X_1,\ldots,X_n)}{D(X_1,\ldots,X_n)D(Y_1,\ldots,Y_n)} \le \frac{D(Y_1,\ldots,Y_{n-1},X_1,\ldots,X_{n-1})}{D(X_1,\ldots,X_{n-1})D(Y_1,\ldots,Y_{n-1})}.$$

In other words, the second ratio in (A1) is decreasing in n. □

Proof of Proposition 1.2. It suffices to treat the case of $\mathcal{C}=\{Y_1,\ldots,Y_n\}$ and $\mathcal{C}'=\{Y_1,\ldots,Y_{n+1}\}$. Recall that $\Psi_{\mathcal{C}}$ is the unconditional covariance matrix of the $Y_i - RP_{Y_i}$. The matrix $\Sigma_{\mathcal{C}}$ is the conditional covariance matrix of the Y_i given prices, but then it is also the conditional covariance matrix of the net payoffs $Y_i - RP_{Y_i}$ given prices. It follows that $|\Sigma_{\mathcal{C}}|/|\Psi_{\mathcal{C}}|$ has exactly the form considered in Lemma A.1, a determinant ratio for a conditional and unconditional covariance matrix. The ratio is therefore decreasing in n.

Next we compare values of $\mu_{\mathcal{C}}^\top \Psi_{\mathcal{C}}^{-1}\mu_{\mathcal{C}}$. We will use the subscripts n and $n+1$ to distinguish the two consideration sets, and we will show that $\mu_{n+1}^\top \Psi_{n+1}^{-1}\mu_{n+1} \ge \mu_n^\top \Psi_n^{-1}\mu_n$. Write Ψ_{n+1} in block form as

$$\Psi_{n+1} = \begin{pmatrix} \Psi_n & v \\ v^\top & a \end{pmatrix},$$

where a $(= \mathrm{V}[Y_{n+1} - RP_{Y_{n+1}}])$ is a scalar and v is a vector (of covariances) of length n. The inverse can be derived from theorem A.3.3 of Anderson (2003) to get

$$\Psi_{n+1}^{-1} = \begin{pmatrix} \Psi_n^{-1} & 0_n \\ 0_n^\top & 0 \end{pmatrix} + yy^\top, \quad y = \frac{1}{\sqrt{a - v^\top \Psi_n^{-1} v}} \begin{pmatrix} \Psi_n^{-1} v \\ -1 \end{pmatrix},$$

where 0_n is an n-vector of zeros. The expression inside the square root is

$$a - v^\top \Psi_n^{-1} v = \mathrm{V}[Y_{n+1} - RP_{n+1} | Y_1 - RP_1, \ldots, Y_n - RP_n] \ge 0.$$

To verify that Ψ_{n+1}^{-1} has the claimed form, we evaluate the products

$$\Psi_{n+1} \begin{pmatrix} \Psi_n^{-1} & 0_n \\ 0_n^\top & 0 \end{pmatrix} = \begin{pmatrix} I & 0_n \\ v^\top \Psi_n^{-1} & 0 \end{pmatrix}$$

and

$$\Psi_{n+1} yy^\top = \frac{1}{a - v^\top \Psi_n^{-1} v} \begin{pmatrix} 0_n \\ v^\top \Psi_n^{-1} v - a \end{pmatrix} \begin{pmatrix} \Psi_n^{-1} v \\ -1 \end{pmatrix}^\top = \begin{pmatrix} 0_{n\times n} & 0_n \\ -v^\top \Psi_n^{-1} & 1 \end{pmatrix}.$$

Adding the two term yields the identity matrix. We now have

$$\mu_{n+1}^\top \Psi_{n+1}^{-1} \mu_{n+1} = \mu_{n+1}^\top \begin{pmatrix} \Psi_n^{-1} & 0_n \\ 0_n^\top & 0 \end{pmatrix} \mu_{n+1} + \mu_{n+1}^\top yy^\top \mu_{n+1} = \mu_n^\top \Psi_n^{-1} \mu_n + (\mu_{n+1}^\top y)^2 \ge \mu_n^\top \Psi_n^{-1} \mu_n,$$

as claimed. We have thus shown that adding a security to the consideration set decreases $J_{\mathcal{C}}^2$ in (10). □

A.2 Expected Utility in the Factor Model

Proof of Proposition 1.3. Under the exchangeability and independence condition, $\Sigma_{\mathcal{C}}$ and $\Psi_{\mathcal{C}}$ depend on $\mathcal{C}$ only through its type, and the same is true for the vector μ of expected net payoffs of the securities in $\mathcal{C}$. It then follows from the general expression for J^2 in (4) that the investor's expected utility is determined by the consideration set's type.

Evaluation of $\mu^\top \Psi^{-1} \mu$. To see why the risk premium is missing from (20) and has the indicated form in (21), recall from (19) that $\mathrm{E}[S_i - RP_{S_i}] = 0$. It follows that for a consideration set of type $(0, K)$, μ is identically zero; for type $(1, K)$, only the first component, $\mathrm{E}[M - RP_F]$, is nonzero. For type $(0, K)$, the term $\mu^\top \Psi^{-1} \mu$ in (4) becomes zero. For type $(1, K)$, it becomes Q_F because the independence of (M, P_F) from the micro shocks and prices implies that $\Psi^{-1}(1,1) = 1/\Psi(1,1) = 1/\mathrm{V}[M - RP_F]$.

Evaluation of $|\Sigma|$. We begin with consideration sets that do not include M. Let Σ_K^S be the covariance matrix of $S_1, \ldots, S_K$ and thus (by exchangeability) the covariance of any K micro shocks. Let G_K denote the $K \times K$ matrix with 1s on the diagonal and all other entries equal to $-1/(N-1)$, and notice that G_K has full rank for $1 \le K \le N-1$. It follows from (15) that G_K is the correlation matrix of $S_1, \ldots, S_K$, so $\Sigma_K^S = \sigma_S^2 G_K$, where σ_S^2 is the common variance of the S_i. As $P_{S_1}, \ldots, P_{S_N}$ are exchangeable and sum to zero, they have the same correlations as $S_1, \ldots, S_N$. It follows that the covariance matrix of $P_{S_1}, \ldots, P_{S_K}$ is given by $\Sigma_K^P = \sigma_P^2 G_K$, where σ_P^2 is the common variance of the prices.

For the cross-covariance between prices and payoffs, observe that exchangeability implies

$$0 = \mathrm{cov}[0, S_j] = \mathrm{cov}\left[\sum_{i=1}^N P_{S_i}, S_j\right] = \sum_{i=1}^N \mathrm{cov}[P_{S_i}, S_j] = \mathrm{cov}[P_{S_j}, S_j] + (N-1)\mathrm{cov}[P_{S_k}, S_j],$$

with $k \neq j$. This says that the off-diagonal entries $\mathrm{cov}[P_{S_k}, S_j]$ are all equal to $-1/(N-1)$ times the common value $\sigma_{PS} = \mathrm{cov}[P_{S_i}, S_i]$ of the diagonal entries. Thus, the cross-covariance of $P_{S_1}, \ldots, P_{S_K}$ with $S_1, \ldots, S_K$ is given by $\Sigma_K^{PS} = \sigma_{PS} G_K$. This matrix is symmetric. The full covariance matrix of $S_1, \ldots, S_K, P_{S_1}, \ldots, P_{S_K}$ is given by

$$\begin{pmatrix} \Sigma_K^S & \Sigma_K^{PS} \\ \Sigma_K^{PS} & \Sigma_K^P \end{pmatrix} = \begin{pmatrix} \sigma_S^2 G_K & \sigma_{PS} G_K \\ \sigma_{PS} G_K & \sigma_S^2 G_K \end{pmatrix} = \begin{pmatrix} \sigma_S^2 & \sigma_{PS} \\ \sigma_{PS} & \sigma_S^2 \end{pmatrix} \otimes G_K, \tag{A5}$$

where $\otimes$ denotes the Kronecker product of matrices. By the formula in (A2) for the determinant of a conditional covariance matrix, $|\Sigma|$ is the ratio of the determinant of (A5) to the determinant of Σ_K^P, the covariance of the conditioning variables. The determinant of the Kronecker product (A5) is given by (Anderson 2003, Theorem A.4.5),

$$\begin{vmatrix} \sigma_S^2 & \sigma_{PS} \\ \sigma_{PS} & \sigma_S^2 \end{vmatrix}^K \times \left| G_K \right|^2 = (\sigma_S^2 \sigma_P^2 - \sigma_{PS}^2)^K \left| G_K \right|^2,$$

and

$$|\Sigma_K^P| = \sigma_P^{2K} |G_K|,$$

so

$$\left|\Sigma\right| = \frac{(\sigma_S^2 \sigma_P^2 - \sigma_{PS}^2)^K |G_K|^2}{\sigma_P^{2K} |G_K|} = (\sigma_S^2 - \sigma_{PS}^2/\sigma_P^2)^K \left| G_K \right| = (\mathrm{V}[S|P_S])^K \left| G_K \right|.$$

Evaluation of $|\Psi|$. Recalling that Ψ is the covariance of the $S_i - RP_{S_i}$ and using the covariance calculations above,

$$\Psi = \Sigma_K^S + R^2 \Sigma_K^P - 2R\Sigma_{PS}^K = (\sigma_S^2 + R^2 \sigma_P^2 - 2R\sigma_{PS}) G_K, \tag{A6}$$

$$\text{so } |\Psi| = (\mathrm{V}[S - RP_S])^K |G_K|, \text{ and}$$

$$\frac{|\Sigma|}{|\Psi|} = \left(\frac{\mathrm{V}[S|P_S]}{\mathrm{V}[S - RP_S]} \right)^K,$$

as required for (20).

Including M. As (M, P_F) are independent of the micro shocks and prices, when we add M to the consideration set $S_1, \ldots, S_K$, we multiply $|\Sigma|$ by $\mathrm{V}[M|P_F]$ and we multiply $|\Psi|$ by $\mathrm{V}[M - RP_F]$. The determinant ratio is thus multiplied by $\mathrm{V}[M|P_F]/\mathrm{V}[M - RP_F]$, as in (21). □

A.3 Utility of the Oblivious

Proof of Proposition 2.2. Under the macro-micro independence assumption (a5i), the expected utility for a CARA investor factors into the contribution from investing in M and a contribution from investing in the micro securities. The contribution from M is given by (29), so it suffices to consider the second factor.

Under exchangeability, we may assume the consideration set is $\{S_1, \ldots, S_K\}$, $1 \leq K \leq N-1$. Write S for the vector of payoffs and P for the vector of corresponding prices. Write Σ_S, Σ_P, and $\Sigma_{PS} = \Sigma_{SP}$ for the indicated covariance and cross-covariance matrices, which we derived in the proof of Proposition 1.3,

Because $\mathrm{E}[S_i] = 0$, the oblivious investor's demands (30) become

$$q = -\frac{1}{\gamma}\Sigma_S^{-1} RP,$$

which earns expected utility

$$-\mathrm{E}[e^{-\gamma q^\top (S-RP)}] = -\mathrm{E}[e^{RP^\top \Sigma_S^{-1}(S-RP)}]. \tag{A7}$$

We will evaluate this expectation by first conditioning on P and then taking the expectation over P.

Conditional expectation given prices P. Recall that if $X \sim N(\mu, \Sigma)$ and v is a vector of the same length as X, then

$$\mathrm{E}[e^{v^\top X}] = e^{v^\top \mu + v^\top \Sigma v/2}.$$

We will apply this identity to evaluate the conditional expectation $\mathrm{E}[e^{RP^\top \Sigma_S^{-1} S}|P]$ by setting $v = RP^\top \Sigma_S^{-1}$ and letting X have the conditional distribution of S given P, which is

$$S|P \sim N(\mathrm{E}[S|P], \Sigma_{S|P}).$$

Thus,

$$\mathrm{E}[e^{RP^\top \Sigma_S^{-1} S}|P] = e^{v^\top \mathrm{E}[S|P]} e^{v^\top \Sigma_{S|P} v/2} = e^{RP^\top \Sigma_S^{-1} \mathrm{E}[S|P]} e^{RP^\top \Sigma_S^{-1} \Sigma_{S|P} \Sigma_S^{-1} RP/2}.$$

Recalling our target in (A7),

$$\mathrm{E}[e^{RP^\top \Sigma_S^{-1}(S-RP)}|P] = e^{-RP^\top \Sigma_S^{-1} RP} e^{RP^\top \Sigma_S^{-1} \mathrm{E}[S|P]} e^{RP^\top \Sigma_S^{-1} \Sigma_{S|P} \Sigma_S^{-1} RP/2}. \tag{A8}$$

Now $\mathrm{E}[S|P] = \Sigma_{SP} \Sigma_P^{-1} P$, so making this substitution we can write (A8)

$$\mathrm{E}[e^{RP^\top \Sigma_S^{-1}(S-RP)}|P] = e^{RP^\top ARP} \tag{A9}$$

with

$$A = \frac{1}{2}\Sigma_S^{-1}\Sigma_{S|P}\Sigma_S^{-1} - \Sigma_S^{-1} + \frac{1}{R}\Sigma_S^{-1}\Sigma_{SP}\Sigma_P^{-1}. \tag{A10}$$

Expectation over prices P. It remains to evaluate the expectation over P in (A9). We use the following property: if $X \sim N(0, \Sigma)$ is n-dimensional and A is $n \times n$ and symmetric, then

$$\mathrm{E}[e^{X^\top AX}] = |I - 2\Sigma A|^{-1/2},$$

provided $I - 2\Sigma A$ is positive definite (Anderson 2003, Theorem 7.3.1). We will apply this result with $X = RP$ and A as in (A10). We therefore need to evaluate $I - 2R^2\Sigma_P A$.

We showed in the proof of Proposition 1.3 that $\Sigma_S = \sigma_S^2 G_K$, $\Sigma_P = \sigma_P^2 G_K$, and $\Sigma_{SP} = \Sigma_{PS} = \sigma_{PS} G_K$ are all proportional to the matrix G_K. The same then holds for the conditional covariance matrix $\Sigma_{S|P}$, as

$$\Sigma_{S|P} = \Sigma_S - \Sigma_{SP}\Sigma_P^{-1}\Sigma_{SP} = (\sigma_S^2 - \sigma_{PS}^2/\sigma_P^2)G_K.$$

Making this substitution in (A10), we find that A is proportional to G_K^{-1}, with

$$A = \left(\frac{1}{2}\left(\frac{1}{\sigma_S^2} - \frac{\sigma_{PS}^2}{\sigma_S^4\sigma_P^2}\right) - \frac{1}{\sigma_S^2} + \frac{1}{R}\frac{\sigma_{PS}}{\sigma_S^2\sigma_P^2}\right)G_K^{-1}.$$

Since $\Sigma_P = \sigma_P^2 G_K$, Σ_{PA} is a multiple of the identity, with

$$\begin{aligned} I - 2R^2\Sigma_P A &= \left(1 + R^2\frac{\sigma_P^2}{\sigma_S^2} - 2R\frac{\sigma_{PS}}{\sigma_S^2} + R^2\frac{\sigma_{PS}^2}{\sigma_S^4}\right)I \\ &= \frac{\sigma_S^2 + R^2\sigma_P^2 - 2R\sigma_{PS} + R^2\sigma_{PS}^2/\sigma_S^2}{\sigma_S^2}I \\ &= \frac{\mathrm{V}[S_i - RP_{S_i}] + R^2\mathrm{cov}[S_i, P_{S_i}]^2/\mathrm{V}[S_i]}{\mathrm{V}[S_i]}I. \end{aligned}$$

This $K \times K$ matrix is diagonal. The diagonal entries are strictly positive because $\mathrm{V}[S_i - RP_{S_i}] = 0$ implies $\mathrm{cov}[S_i, P_{S_i}]^2 > 0$. The matrix is therefore positive definite. Its determinant is the Kth power of any of the identical diagonal entries. Similarly, for its inverse we have

$$|I - 2R^2\Sigma_P A|^{-1} = \left(\frac{\mathrm{V}[S_i]}{\mathrm{V}[S_i - RP_{S_i}] + R^2\mathrm{cov}[S_i, P_{S_i}]^2/\mathrm{V}[S_i]}\right)^K,$$

as claimed in (31). Recall that (31) shows the squared expected utility, so we are evaluating the square of (A7), which is why we do not need to take the square root of the determinant. □

A.4 Market Clearing Prices in the Constrained Model

Proof of Proposition 3.1. The prices $P_{S_1}, \ldots, P_{S_N}$ given by (43) are independent of the index fund price P_F given by (42). It follows that when a micro-informed investor sets demands for the index fund and a hedged security S_i, the investor may set these two demands independently of each other. In setting their index fund demands, the micro-informed choose the same quantity as the uninformed because the signal s_i is independent of M. The market for the index fund thus has the form of the single-security market of Grossman and Stiglitz (1980), with λ_M informed investors and $\lambda_S + \lambda_U$ uninformed. The Grossman-Stiglitz price takes the form in (42), with coefficients a_F, b_F, and c_F given in Section I.1.1 of the Internet Appendix. The supply of the index fund cleared at this price is X_F. We can think of this supply as X_F shares of M; more precisely in our setting, this supply is X_F/N shares of every stock.

The λ_S/N investors in stock i face the price P_{S_i} in (43) for the hedged security S_i. A standard calculation shows that each such investor's demand is given by

$$q_{S_i} = \frac{\mathrm{E}[S_i|s_i] - RP_{S_i}}{\gamma\mathrm{V}[S_i|s_i]} = X_i/\lambda_S.$$

The total demand for S_i is therefore $q_{S_i}\lambda_S/N = X_i/N$.

Each stock i pays $M + S_i$ and has supply $(X_F + X_i)/N$, $i = 1, \ldots, N$, so the value of the total supply is

$$\frac{1}{N}\sum_{i=1}^{N}(X_F+X_i)(M+S_i) = X_F M + \frac{1}{N}\sum_{i=1}^{N}X_iS_i + X_F\frac{1}{N}\sum_{i=1}^{N}S_i + M\frac{1}{N}\sum_{i=1}^{N}X_i$$
$$= X_F M + \frac{1}{N}\sum_{i=1}^{N}X_iS_i,$$

using the exchangeability properties (14) and (39). In other words, clearing the stocks is equivalent to clearing X_F shares of M and X_i/N shares of each S_i, and we have shown that these are indeed the demands that follow from the prices (42) and (43). □

A.5 Existence of Equilibrium

We begin with a condition that keeps oblivious investors out of the market.

Lemma A.2. *Suppose prices are as in (42) and (43), and suppose $J_S^2 = J_U^2$ when the micro-informed use their full capacity, in the sense that the first equality in (47) holds. If*

$$e^{2\gamma c} - 1 < f_S < 1, \tag{A11}$$

then $J_U^2 < J_{U,K}^2$, for all $K \geq 1$, so being an ordinary uninformed investor is strictly preferred over being an oblivious investor.

Proof. In light of the expression for J_U^2 in (29) and for $J_{U,K}^2$ in (31), we need to show that

$$\mathrm{V}[S] > \mathrm{V}[S - RP_S] + \mathrm{cov}[S, RP_S]^2/\mathrm{V}[S]. \tag{A12}$$

Equating the expressions for J_U^2 in (29) and J_S^2 in (46), we get

$$1 = J_U^2/J_S^2 = \frac{\mathrm{V}[S - RP_S]}{\mathrm{V}[S|P_S]}e^{-2\gamma c}\frac{\delta}{\kappa}.$$

Making the substitution $\mathrm{V}[S|P_S] = \mathrm{V}[S|s]\delta/\kappa$ from the capacity constraint (47), we get

$$V[S - RP_S] = e^{2\gamma c}\mathrm{V}[S|s] = e^{2\gamma c}(1 - f_S)\mathrm{V}[S].$$

Also, from (43) we get $\mathrm{cov}[S, P_S] = \mathrm{cov}[S, s] = f_S\mathrm{V}[S]$. Thus, (A12) holds if

$$\mathrm{V}[S] > e^{2\gamma c}(1 - f_S)\mathrm{V}[S] + f_S^2\mathrm{V}[S];$$

that is, if $1 - f_S^2 > e^{2\gamma c}(1 - f_S)$, which holds if $1 + f_S > e^{2\gamma c}$ and $f_S < 1$. □

Proof of Theorem 3.1. We need to show that under the conditions of the theorem, (44) and (47) hold. Equation (44) specifies the conditions we need on the expected utilities of the different types of investors, and Equation (47) confirms that the expected utilities of the informed investors are achievable under the information constraints. We divide the proof into three parts: showing that $J_S^2 = J_U^2$ and the first equality in (47) hold; showing that $J_U^2 < J_{U,K}^2$, and $J_S^2 < J_{0,K}^2$, for all $K \geq 1$ (so all types other micro-informed, macro-informed, and ordinary uninformed are strictly dominated); showing that $J_M^2 = J_U^2$ and the second equality in (47) holds.

Part 1: proving that $J_S^2 = J_U^2$ and the first equality in (47) hold. To address (47), we begin by showing that the condition $\kappa < \delta$ (which we imposed when we introduced δ in Equation 22) holds within the valid cost interval. This condition ensures that informed investors can learn about at least one stock. As $\kappa e^{2\gamma c} = \delta_F$, it suffices to show that $\delta_F/\delta < e^{2\gamma c}$, which holds throughout the interval provided $\delta_F/\delta \leq \beta_-$. Equivalently, we need

$$\sqrt{9 - 8\delta_F/\delta} \leq 3 - 2\delta_F/\delta,$$

and squaring both sides shows that this follows from $\delta_F/\delta \in (1, 9/8)$.

We could prove the first equality in (47) and the condition $J_S^2 = J_U^2$ by substituting the claimed expressions for f_S and λ_S and verifying that the needed equations hold. The argument is clearer if we proceed in the opposite direction, deriving the claimed expressions for λ_S and f_S from these equations and noting that each step could be reversed.

Comparing (29) and (46), we find that $J_S^2 = J_U^2$ is equivalent to

$$\frac{\mathrm{V}[S|P_S]}{\mathrm{V}[S - RP_S]} = \frac{\delta}{\delta_F}, \tag{A13}$$

because $e^{2\gamma c}\kappa = \delta_F$. We use (A13) to solve for λ_S^2. In the numerator of (A13), we make the substitution $\mathrm{V}[S|P_S] = \mathrm{V}[S|s]\delta/\kappa = (1-f_S)\mathrm{V}[S]\delta/\kappa$ from the capacity constraint (47); in the denominator, we use the expression for P_S in (43). Equation (A13) then becomes

$$\frac{(1-f_S)\mathrm{V}[S]\delta/\kappa}{(1-f_S)\mathrm{V}[S] + \gamma^2(1-f_S)^2\mathrm{V}[S]^2\mathrm{V}[X]/\lambda_S^2} = \frac{\delta}{\delta_F}.$$

Noting that $\kappa/\delta_F = e^{-2\gamma c}$ and taking the reciprocal of both sides, we get

$$1 + \frac{\gamma^2(1-f_S)\mathrm{V}[S]\mathrm{V}[X]}{\lambda_S^2} = e^{2\gamma c}.$$

Solving for λ_S^2 yields $\lambda_S^2 = L_S^2$, with L_S as defined in (49) and evaluated at f_S.

We derive a second expression for λ_S^2 starting from (47). First observe that

$$\mathrm{V}[S|P_S] = \mathrm{V}[S] - \frac{\mathrm{cov}[S, P_S]^2}{\mathrm{V}[P_S]} = \mathrm{V}[S] - \frac{f_S^2\mathrm{V}[S]^2}{f_S\mathrm{V}[S] + \gamma^2\mathrm{V}[S]^2\mathrm{V}[X](1-f_S)^2/\lambda_S^2}. \tag{A14}$$

Making this substitution in (47) together with $\mathrm{V}[S|s] = (1-f_S)\mathrm{V}[S]$, yields

$$1 - f_S = (\kappa/\delta)\left(1 - \frac{f_S^2}{f_S + \gamma^2\mathrm{V}[S]\mathrm{V}[X](1-f_S)^2/\lambda_S^2}\right).$$

Replacing κ with $e^{-2\gamma c}\delta_F$ and solving for λ_S^2, we get

$$\lambda_S^2 = \frac{(f_S - 1 + \delta_F e^{-2\gamma c}/\delta)}{(1 - \delta_F e^{-2\gamma c}/\delta)}\frac{(1-f_S)}{f_S}\gamma^2 V[S]V[X], \tag{A15}$$

whenever

$$1 - \delta_F e^{-2\gamma c}/\delta < f_S < 1; \tag{A16}$$

we will verify in Part 2 of the proof that (A16) holds in the valid cost interval when f_S is given by (48).

Equating the two expressions we have derived for λ_S^2 yields

$$\frac{1-f_S}{e^{2\gamma c}-1} = \frac{(f_S - 1 + \delta_F e^{-2\gamma c}/\delta)}{(1 - \delta_F e^{-2\gamma c}/\delta)}\frac{(1-f_S)}{f_S},$$

and thus

$$f_S(1 - \delta_F e^{-2\gamma c}/\delta) = (\delta_F e^{-2\gamma c}/\delta - 1)(e^{2\gamma c} - 1) + f_S(e^{2\gamma c} - 1).$$

By solving for f_S we get (48). As the equations derived for λ_S rely on (A16), (48) is valid only for f_S in this range. Substituting $f_S(c)$ in (49) yields λ_S as a function of c.

Part 2: proving that $J_S^2 < J_{0,K}^2$ and $J_U^2 < J_{U,K}^2$, for all $K \geq 1$. We know from Grossman and Stiglitz (1980) that the coefficient b_F in the price P_F in (42) is nonzero. It follows that P_F has nonzero correlation with M and, therefore, as explained following Proposition 2.1, that $J_S^2 < J_{0,K}^2$.

Next we show $J_U^2 < J_{U,K}^2$. In light of Lemma A.2, it suffices to show that our choice of $f_S(c)$ in (48) satisfies (A11) on the valid cost interval.

With $f_S(c)$ in (48), the lower bound in (A11) becomes

$$\frac{\left(1-\delta_F e^{-2\gamma c}/\delta\right)}{(e^{2\gamma c}-1)-(1-\delta_F e^{-2\gamma c}/\delta)} > 1. \tag{A17}$$

We showed at the beginning of Part 1 that $\kappa < \delta$ in the valid cost interval, so the numerator of (A17) is positive in the valid cost interval. If we set $\beta = e^{2\gamma c}$, the condition that the denominator in (A17) be positive becomes $\beta^2 - 2\beta + \delta_F/\delta > 0$, which holds for all β because $\delta_F/\delta > 1$. The condition in (A17) can therefore be written as

$$2(1-\delta_F e^{-2\gamma c}/\delta) > e^{2\gamma c} - 1. \tag{A18}$$

Again setting $\beta = e^{2\gamma c}$, we can write (A18) as $\beta^2 - 3\beta + 2\delta_F/\delta < 0$. This condition holds for $\beta_- < \beta < \beta_+$, the valid cost interval (and only in this interval). We have thus shown that $f_S(c)$ in (48) satisfies the lower bound in (A11) on the valid cost interval. Also, because the denominator of (A17) is positive, we have $f_S(c) > e^{2\gamma c} - 1 > 1 - \delta_F e^{-2\gamma c}/\delta$, which proves the lower bound in (A16).

Next we show that the upper bound $f_S(c) < 1$ required for (A11) also holds in this interval. Write $f_S(c)$ in (48) as

$$\frac{(\beta-1)(\beta-\delta_F/\delta)}{\beta^2-2\beta+\delta_F/\delta} = \frac{\beta^2-(1+\delta_F/\delta)\beta+\delta_F/\delta}{\beta^2-2\beta+\delta_F/\delta} = 1 - \frac{(\delta_F/\delta-1)\beta}{\beta^2-2\beta+\delta_F/\delta}.$$

The last expression is strictly bounded by 1 because $\beta > 0$, $\delta_F > \delta$, and we established previously that the denominator is positive. We have thus shown that $f_S(c)$ satisfies (A11) for c in the valid cost interval.

Part 3: proving conditions on f_M and λ_M. For any $f_M \in (1-e^{-2\gamma c}, 1)$, we derive the values of λ_M compatible with the second equality in (47). From the price P_F in (42), we get

$$\begin{aligned} \mathrm{V}[M|P_F] &= \mathrm{V}[M] - \frac{\mathrm{cov}[M,P_F]^2}{\mathrm{V}[P_F]} \\ &= \mathrm{V}[M] - \frac{b_F^2 f_M^2 \mathrm{V}[M]^2}{b_F^2 f_M \mathrm{V}[M] + c_F^2 \mathrm{V}[X_F]} \\ &= \mathrm{V}[M] - \frac{f_M^2 \mathrm{V}[M]^2}{f_M \mathrm{V}[M] + \gamma^2 \mathrm{V}[M]^2 \mathrm{V}[X_F](1-f_M)^2/\lambda_M^2}. \end{aligned} \tag{A19}$$

Also, $\mathrm{V}[M|m] = (1-f_M)\mathrm{V}[M]$. Making these substitutions in the condition $\mathrm{V}[M|m]/\mathrm{V}[M|P_F] = \kappa/\delta_F$ in (47) and solving for λ_M, we find that λ_M is given by (50), using the condition $f_M \geq 1 - e^{-2\gamma c}$ to ensure that the expression inside the square-root in (50) is nonnegative. The resultant value for λ_M is valid if it does not exceed $1 - \lambda_S(c)$. Thus, for any f_M and $\lambda_M = L_M(f_M, c)$ satisfying

$$1 - e^{-2\gamma c} < f_M < 1, \quad \lambda_M < 1 - \lambda_S(c), \tag{A20}$$

the second equality in (47) holds (macro-informed are able to use their full capacity) so J_M^2 is given by (45). Comparing the expressions for J_U^2 and J_M^2 in (29) and (45), we see that $e^{2\gamma c}\kappa = \delta_F$ implies $J_U^2 = J_M^2$. Thus, for f_M in the range where (A20) holds, we have an equilibrium with $J_M^2 = J_S^2 = J_U^2$ and $\lambda_M + \lambda_S + \lambda_U = 1$. □

In the discussion following the statement of Theorem 3.1, we commented that the largest potential value of λ_M is attained at $f_M = f_M^*(c) \equiv \sqrt{1-e^{-2\gamma c}}$. We now show why this holds. With $\lambda_M = L_M$ as given by (50), we can write

$$\lambda_M^2 = \frac{2 - e^{-2\gamma c} - 2\sqrt{1 - e^{-2\gamma c}}}{1 - e^{-2\gamma c}} \gamma^2 \mathrm{V}[M]\mathrm{V}[X_F] - \frac{(f_M - \sqrt{1 - e^{-2\gamma c}})^2}{(1 - e^{-2\gamma c}) f_M} \gamma^2 \mathrm{V}[M]\mathrm{V}[X_F]. \quad \text{(A21)}$$

The term subtracted on the right is positive, so λ_M is indeed maximized at $f_M = f_M^*(c)$, where that term vanishes. The maximum is given by

$$\lambda_M^{*2}(c) = \frac{(1 - \sqrt{1 - e^{-2\gamma c}})^2}{1 - e^{-2\gamma c}} \gamma^2 \mathrm{V}[M]\mathrm{V}[X_F] = \left(\frac{1 - f_M^*(c)}{f_M^*(c)} \right)^2 \gamma^2 \mathrm{V}[M]\mathrm{V}[X_F]. \quad \text{(A22)}$$

Every value of λ_M between 0 and $\min(\lambda_M^*(c), 1 - \lambda_S(c))$ is consistent with equilibrium for some value f_M in $(1 - e^{-2\gamma c}, f_M^*(c)]$ and some value of f_M in $[f_M^*(c), 1)$.

References

Admati, A., and P. Pfleiderer. 1987. Viable allocations of information in financial markets. *Journal of Economic Theory* 43:76–115.

Anderson, T. W. 2003. *An introduction to multivariate statistical analysis*, 3rd ed. New York: Wiley Interscience.

Berk, J., and R. Green. 2004. Mutual fund flows and performance in rational markets. *Journal of Political Economy* 112:1269–95.

Berk, J., and J. van Binsbergen. 2015. Measuring skill in the mutual fund industry. *Journal of Financial Economics* 118:1–20.

Bhattacharya, A., and M. O'Hara. 2016. Can ETFs increase market fragility? Effect of information linkages on ETF markets. Working Paper, Baruch College.

Brandt, M., A. Brav, J. Graham, and A. Kumar. 2010. The idiosyncratic volatility puzzle: Time trend or speculative episodes? *Review of Financial Studies* 23:863–99.

Carhart, M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52:57–82.

Chan, W. 2003. Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics* 70:223–60.

Chen, N.-F., and J. Ingersoll. 1983. Exact pricing in linear factor models with finitely many assets: A note. *Journal of Finance* 38:985–8.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers. 1997. Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52:1035–58.

Fama, E., and K. French, 2015. A five-factor asset pricing model. *Journal of Financial Econonometrics* 16:1–22.

Fan, K. 1968. An inequality for subadditive functions on a distributive lattice, with application to determinantal in equalities. *Linear Algebra and its Applications* 1:33–38.

Ferson, W., and H. Mo, 2016. Performance measurement with selectivity, market and volatility timing. *Journal of Financial Economics* 121:93–110.

Foucault, T., D. Sraer, and D. Thesmar. 2011. Individual investors and volatility. *Journal of Finance* 66:1369–406.

Gârleanu, N., and L. H. Pedersen. 2018. Efficiently inefficient markets for assets and asset management. *The Journal of Finance* 73:1663–712.

Gârleanu, N., and L. H. Pedersen. 2021. Active and passive investing: understanding Samuelson's dictum. *Review of Asset Pricing Studies.*

Glosten, L., S. Nallareddy, and Y. Zou. 2020. ETF trading and informational efficiency of underlying securities. *Management Science* 67:22–47.

Goldstein, I., and L. Yang. 2015. Information diversity and complementarities in trading and information acquisition. *Journal of Finance* 70:1723–65.

Greenwood, R., and D. Scharfstein. 2013. The growth of finance. *Journal of Economic Perspectives* 27:3–28.

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

Investment Company Institute. 2018. *Investment Company Fact Book*, 58th ed. Washington, DC.

Jung, J. and R. Shiller. 2005. Samuelson's dictum and the stock market. *Economic Inquiry* 43:221–8.

Kacperczyk, M., C. Sialm, and L. Zheng. 2005. On the industry concentration of actively managed equity mutual funds. *Journal of Finance* 60:1983–2011.

Kacperczyk, M. T., S. Van Nieuwerburgh, and L. Veldkamp. 2016. A rational theory of mutual funds' attention allocation. *Econometrica* 84:571–626.

Koijen, R., and M. Yogo. 2019. A demand system approach to asset pricing. *Journal of Political Economy* 127:1475–515.

Kwon, Y. 1985. Derivation of the capital asset pricing model without normality or quadratic preference: A note. *Journal of Finance* 40:1505–9.

Lakonishok, J., A. Shleifer, and R. Vishny. 1992. The structure and performance of the money management industry. *Brookings Papers: Microeconomics* 1992:339–91.

Maćkowiak, B., and M. Wiederholt. 2009. Optimal sticky prices under rational inattention. *American Economic Review*. 99:769–803.

Mondria, J. 2010. Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory* 145:1837–64.

Peng, L., and W. Xiong. 2006. Investor attention, overconfidence, and category learning. *Journal of Financial Economics* 80:563–602.

Philippon, T. 2012. Finance vs. Wal-Mart: Why are financial services so expensive?. In *Rethinking the financial crisis*, eds. A. Blinder, A. Lo, and R. Solow. Albany, NY: Russell Sage.

Philippon, T., and A. Reshef. 2013. An international look at the growth of modern finance. *Journal of Economic Perspectives* 27:73–96.

Ross, S. A. 1978. Mutual fund separation in financial theory – The separating distributions. *Journal of Economic Theory* 17:254–86.

Schneemeier, J. 2015. Firm investment and price informativeness. Working Paper.

Sims, C. 2011. Rational inattention and monetary economics. In *Handbook of Monetary Economics*, chap. 4, vol. 3A. Amsterdam, the Netherlands: Elsevier.

Tetlock, P. 2010. Does public financial news resolve asymmetric information? *Review of Financial Studies* 23:3520–57.

Van Nieuwerburgh, S., and L. Veldkamp. 2009. Information immobility and the home bias puzzle. *Journal of Finance* 64:1187–215.

———. 2010. Information acquisition and portfolio under-diversification. *Review of Economic Studies* 77:779–805.