Investor Information Choice
with Macro and Micro Information

Paul Glasserman        Harry Mamaysky*

Abstract

We study information and portfolio choices in a market of securities whose dividends depend on an aggregate (macro) risk factor and idiosyncratic (micro) shocks. Investors can acquire information about dividends at a cost. We first establish a general result showing that investors endogenously choose to specialize in either macro or micro information. We then develop a specific model with this specialization and study the equilibrium mix of macro-informed and micro-informed investors and the informativeness of macro and micro prices. Our results favor Samuelson’s dictum, that markets are more micro efficient than macro efficient. We calibrate the model to market data and show that it can reproduce important features of the widely studied decomposition of variability in stock returns into cash flow variance and discount rate variance.

Keywords: Information choice; asset pricing; price efficiency; attention

*Glasserman: Columbia Business School, pg20@columbia.edu. Mamaysky: Columbia Business School, hm2646@columbia.edu. 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.”
1 Introduction

Samuelson’s dictum, as discussed in Shiller (2000), is the hypothesis that the stock market is “micro efficient” but “macro inefficient.” More precisely, the dictum holds that the efficient markets hypothesis describes the pricing of individual stocks better than it describes the aggregate stock market. Jung and Shiller (2005) review and add to empirical evidence that supports the dictum, including evidence of macro inefficiency in Campbell and Shiller (1988) and evidence for somewhat greater micro efficiency in Vuolteenaho (2002) and Cohen et al. (2003).

We develop a model of investor information choice to study a potential wedge between micro and macro price efficiency. Our setting may be viewed as a multi-security generalization of the classical model of Grossman and Stiglitz (1980). Our market consists of a large number of individual stocks, each of which is exposed to a macro risk factor and an idiosyncratic risk. The macro risk factor is tradeable through an index fund that holds all the individual stocks and diversifies away their idiosyncratic risks.

We begin with a very general formulation in which investors may choose to acquire information processing capacity at a cost. This capacity allows an investor to observe and make inferences from signals about fundamentals. Subject to their capacity constraint, informed investors may choose to learn about the macro risk factor, about the micro (idiosyncratic) risks of individual stocks, or any combination of the two. The capacity constraint limits the fraction of uncertainty about dividends an informed investor can remove from a collection of securities.

In formulating this capacity constraint, we differentiate the index fund from individual stocks. We posit that the capacity consumed in making inferences from the price of the index fund is fixed, irrespective of the informativeness of the price. This assumption is based on the view that the implications of the overall level of the stock market are widely discussed and accessible in a way that does not apply to individual stocks. In conditioning demand for the index fund on its price, an investor allocates a fixed capacity to paying attention to this information.

Our first main result shows that investors endogenously choose to specialize in either macro or micro information. Our investors are ex ante identical, and once they incur the cost of becoming informed they are free to choose general combinations of signals, yet in equilibrium they concentrate in two groups, macro-informed and micro-informed investors. The macro-informed use all their capacity to learn about the macro factor and invest only in the index fund; a micro-informed investor acquires a signal about a single stock and invests in that stock and the index fund; some investors choose to remain
This outcome — heterogeneous information choices among ex ante identical investors — contrasts with the related literature, as we explain later. In particular, it contrasts with the model of Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016), in which all informed investors choose the same information, and all investors will focus on the macro factor if it carries significantly more risk than individual stocks.

Having demonstrated that specialization in macro and micro information is a general phenomenon in our framework, we construct a specific model by imposing this specialization as a constraint. In other words, under the conditions of our general result, specialization is a necessary property in equilibrium, and the constrained model demonstrates that such an equilibrium is in fact feasible.

The constrained model has three types of investors: uninformed, macro-informed, and micro-informed, as required by our general result. To solve the model, we first take the fractions of each type as given and solve for an explicit market equilibrium, assuming all agents have CARA preferences. Shares of individual stocks and the index fund are subject to exogenous supply shocks. The exogenous supply shocks themselves exhibit a factor structure. A common component, reflecting the aggregate level of supply, affects the supply of shares for all firms. In addition, noise trading in individual stocks contributes an idiosyncratic component to the supply of each stock. These micro supply shocks have smaller variance than macro supply shocks; investor specialization provides an incentive to nevertheless acquire information about stock-specific risks. Supply shocks are not observable to investors; as a consequence, equilibrium prices are informative about, but not fully revealing of, the micro or macro information acquired by informed agents.

We then allow informed investors to choose between being micro-informed and macro-informed, and we characterize the equilibrium in which a marginal agent is indifferent between the two types of information. In practice, developing the skills needed to acquire and apply investment information takes time — years of education and experience. In the near term, these requirements leave the total fraction of informed investors relatively fixed. By contrast, we suppose that informed investors can move comparatively quickly and costlessly between being macro-informed or micro-informed by shifting their focus of attention. Endogenizing this focus gives rise to an attention equilibrium centered on the choice between macro and micro information.

Over a longer horizon, agents choose whether to gain the skills to become informed, as well as the type of information to acquire. We therefore study an information equilibrium that endogenizes both decisions to determine equilibrium proportions of macro-informed, micro-informed, and uninformed investors. An information equilibrium in the constrained...
model delivers an explicit case of the necessary specialization established in our general formulation: the investors in this model would not prefer to deviate from their specialization and select other combinations of signals that consume the same capacity.

Working with the constrained model, we find a recurring asymmetry between micro and macro information. For example, we show that the information equilibrium sometimes has no macro-informed agents, but some fraction of agents will always choose to be micro-informed. We show that increasing the precision of micro information makes micro-informed investors worse off — we say that the micro-informed overtrade their information, driving down their compensation for liquidity provision. In contrast, macro-informed investors may be better or worse off as a result of more precise macro information: they are better off when the fraction of macro-informed agents or, equivalently, the informativeness of the price of the index fund is sufficiently low. Similarly, the equilibrium fraction of macro-informed agents always increases with the precision of micro information, but it can move in either direction with an increase in the precision of macro information. A simple condition on the relative precision of micro and macro information determines whether the market is more micro efficient or more macro efficient.

We use our model to analyze variance ratios that have been widely studied empirically. Campbell (1991) and Vuolteenaho (2002) decompose the variance of index-level and idiosyncratic stock returns into variance from cash flow news and variance from news about discount rates. Their estimates show that the ratio of cash flow variance to discount rate variance is larger for individual stocks than for the aggregate market, a pattern predicted by our model. We also argue that the trends in variance ratios are consistent with a declining cost of becoming informed over the twentieth century combined with increasing indirect index trading, meaning trading in the index that does not involve trading in individual stocks. Jung and Shiller (2005) call the cash flow and discount rate components of returns the efficient market and inefficient market components, respectively. They therefore interpret the larger variance ratio for individual stocks — found empirically and in our model — as evidence for Samuelson’s dictum.

Our work is related to several strands of literature. Our model effectively nests Grossman and Stiglitz (1980) if we take the index fund as the single asset in their model. We also draw on the analysis of Hellwig (1980), Admati (1985) and Admati and Pfleiderer (1987) but address different questions; see the books by Brunnermeier (2001), Vives (2008) and Veldkamp (2011) for a survey of related literature. Admati and Pfleiderer (1987) and Goldstein and Yang (2015) focus on understanding when signals are complements or substitutes; our specialization result — that an informed investor will choose
either macro or micro information, but not both – makes macro and micro information strategic substitutes. Schneemeier’s (2015) model predicts greater micro than macro efficiency when managers use market prices in their investment decisions. As in Kyle (1985), our noise traders are price insensitive, and gains from trade against them accrue to the informed, which provides an incentive to collect information. We shed light on the discussion in Black (1986) of the crucial role that “noise” plays in price formation by proposing a model in which the factor structure of noise trading plays an important role in determining the relative micro versus macro efficiency of markets.

Van Nieuwerburgh and Veldkamp (2009) analyze how investors’ choices to learn about the domestic or foreign market in the presence of asymmetric prior knowledge may explain the home bias puzzle, and Van Nieuwerburgh and Veldkamp (2010) use related ideas to explain investor under-diversification. Kacperczyk et al. (2016) develop a model of rational attention allocation in which fund managers choose whether to acquire macro or stock specific information before making investment decisions. Their model, like ours, has multiple assets subject to a common cash flow factor; but, in marked contrast to our setting, their agents ultimately all acquire the same information. We also compare variance ratios (cash flow variance to discount rate variance) in the two models and find that they show qualitatively different patterns as a result of the differences in investor information choices.

Peng and Xiong (2006) also use a model of rational attention allocation to study portfolio choice. In their framework, 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 and the Maćkowiak and Wiederholt 2009 model of sticky prices under rational intattention) primarily because in their setting a representative investor makes the information allocation decision; since macro uncertainty is common to all securities, while micro uncertainty is diversified away, the representative investor allocates more attention to macro and sector level information. Gărleanu and Pedersen (2016) extend the Grossman-Stiglitz model to link market efficiency and asset management through search costs incurred by investors in selecting fund managers, in a model with a single risky asset.

Bhattacharya and O’Hara (2016) and Glosten, Nallareddy, and Zou (2016) study a Kyle-type model with an ETF and multiple underlying securities. Their models contain macro-informed and micro-informed agents, as well as supply shocks in the ETF and the underlying securities. In their setting the liquidity of the ETF is higher than that of the constituent stocks and there is price impact from trade. Therefore the price of the ETF
is not always equal to the price of the underlying basket of securities, and ETF prices can be informative about individual stock prospects. In our model agents are atomic and trade with no price impact; therefore no arbitrage requires that the index price is equal to the price of the underlying securities. We consequently do not address these important microstructure effects.

Section 2 describes our securities and the information choices available to investors, and it then presents our general result showing that investors endogenously specialize in macro or micro information. Section 3 introduces the constrained model and adds additional features (supply shocks and market clearing) that lead to explicit expressions for prices and price efficiency in the market equilibrium of Section 4. Sections 5 and 6 investigate the attention equilibrium and information equilibrium, respectively, in the constrained model. Section 7 discusses model implications for variance ratios. Proofs are deferred to an appendix.

2 The economy

Securities

We assume the existence $N$ risky securities — called stocks — indexed by $i$. There is also an index fund, $F$, one share of which holds $1/N$ shares of each of the $N$ stocks. There is a riskless security with a gross return of $R$.

The time 2 dividend payouts of the stocks are given by

$$u_i = M + S_i, \quad i = 1, \ldots, N. \quad (1)$$

We interpret $M$ as a macro factor and the $S_i$ as idiosyncratic contributions to the dividends. The random variables $M, S_1, \ldots, S_N$ are jointly normal, with $E[M] = \bar{m}$, $E[S_i] = 0$, $\text{var}[M] = \sigma_M^2$, $\text{var}[S_i] = \sigma_S^2$, and $E[MS_i] = 0$, $i = 1, \ldots, N$. To make the idiosyncratic shocks fully diversifiable with a finite number of stocks, we assume that

$$\text{corr}(S_i, S_j) = -\frac{1}{N-1}, \quad i \neq j, \quad (2)$$

Our results hold with minor modifications if $M$ is replaced with $\beta_i M$, provided the $\beta_i$s average to 1.

This condition makes $(S_1, \ldots, S_N)$ exchangeable random variables, meaning that their joint distribution is invariant under permutations of the variables. The correlation matrix specified by (2) is diagonally dominant and therefore positive semidefinite.
which implies that $\sum_{i=1}^{N} S_i = 0$. As a consequence, the index fund $F$ pays

$$u_F = \frac{1}{N} \sum_{i=1}^{N} u_i = M + \frac{1}{N} \sum_{i=1}^{N} S_i = M; \quad (3)$$

the last equality is the benefit of imposing (2).

Prices of individual stocks and of the index fund are realized at time 1; in Section 3 we detail price formation through market clearing, but at this point we keep the setting general. Individual stock prices are given by $P_i$. The index fund price is $P_F$, and precluding arbitrage requires that

$$P_F = \frac{1}{N} \sum_{i=1}^{N} P_i. \quad (4)$$

We also define the price $P_{S_i} = P_i - P_F$, $i = 1, \ldots, N$, of a security paying $u_i - M = S_i$, the idiosyncratic portion of the dividend of stock $i$.

**Agents and information sets**

At time 0, a unit mass of agents maximize expected utility, $-E[\exp(-\gamma \tilde{W}_2)]$, over time time 2 wealth

$$\tilde{W}_2 = W_1 R + q_F(u_F - RP_F) + \sum_{i=1}^{N} q_i (u_i - RP_i),$$

where $q_F$ and $q_i$ are the shares invested in the index fund and stock $i$, and which are chosen given the information $\mathcal{I}$ available to investors at time 1 to maximize $-E[\exp(-\gamma \tilde{W}_2)|\mathcal{I}]$. The initial wealth $W_1$ does not affect an investor’s decisions. The risk aversion parameter $\gamma > 0$ is common to all investors.

At time 0, agents can choose to acquire information capacity $\kappa$, $0 < \kappa < 1$, by incurring a cost $c$. This capacity allows an agent to select signals $m'$ about $M$ and signals $s'_i$ about the $S_i$. We measure the informativeness of signals $m'$ and $s'_i$ through the variance reduction ratios $(\text{var}[M] - \text{var}[M|m'])/\text{var}[M]$ and $(\text{var}[S_i] - \text{var}[S_i|s'_i])/\text{var}[S_i]$. Informativeness will be constrained by $\kappa$, and the available signals will allow full use of $\kappa$.

In more detail, for any level of informativeness $f \in [0,1]$, there is a signal $s_i(f)$, with $s_i(0) = 0$ and $s_i(1) = S_i$. Each $s_i(f)$ has mean zero and variance $f \sigma^2_i$, with

---

3This condition ensures that idiosyncratic terms are fully diversifiable with $N$ finite; otherwise, the index fund would include some indiosyncratic risk. This assumption is common in the asset pricing literature when considering multi-security economies with a finite number of assets; see, for example, Ross (1978), Chen and Ingersoll (1983), and Kwon (1985). We think of $N$ as large, so the correlation required by (2) is small. Correlations of this form, which we will encounter in several places, may be interpreted as “independence for large $N$.”
\[ \text{var}[S_i|s_i(f)] = (1 - f)\sigma_S^2. \]

Similarly, the signal \( m(f) \) has \( \text{E}[m(f)] = \bar{m}, \text{var}[m(f)] = f\sigma_M^2, \) and \( \text{var}[M|m(f)] = (1 - f)\sigma_M^2. \) All macro signals \( m(f) \) are independent of signals \( s_i(f') \) about idiosyncratic payouts, and all signals and payouts are jointly normal. We henceforth omit the argument \( f \) from the signals unless needed for clarity.

An informed investor selects a set of securities about which to acquire signals and in which to invest. The consideration set of securities contains \( K \) stocks, \( i_1, \ldots, i_K \), for some \( 0 \leq K \leq N - 1 \), and may contain the index fund, in which case \( K \leq N - 2 \). We assume that prices are freely available, so once an investor chooses to become informed about a security, the investor knows at least the price of the security.

Together with a set of securities, an investor chooses a corresponding information set

\[
\mathcal{T}^{(0)}_K = \{(s_{i_1}', P_{S_{i_1}}), \ldots, (s_{i_K}', P_{S_{i_K}})\},
\]

\[
\mathcal{T}^{(1)}_K = \{P_F, (s_{i_1}', P_{S_{i_1}}), \ldots, (s_{i_K}', P_{S_{i_K}})\},
\]

or

\[
\mathcal{T}^{(2)}_K = \{(m', P_F), (s_{i_1}', P_{S_{i_1}}), \ldots, (s_{i_K}', P_{S_{i_K}})\},
\]

depending on whether the index fund is in the consideration set and, if it is, whether the investor learns more than the fund’s price.

With additional structure (which we introduce later), prices will reflect investors’ information choices. For now, we keep the discussion general and just assume that prices and signals are jointly normal. We also assume that the fund price \( P_F \) is uncorrelated with the prices \( P_{S_i} \) of the idiosyncratic payouts.

Write \( \Sigma^{(\iota,K)} \) for the unconditional covariance matrix of the payouts \( M, S_{i_1}, \ldots, S_{i_K} \) or \( S_{i_1}, \ldots, S_{i_K} \) in the consideration set, with \( \iota \in \{1, 2\} \) if the index fund is in the set, and \( \iota = 0 \) if it is not. The off-diagonal elements of \( \Sigma^{(0,K)} \) are determined by (2), and we have

\[
\Sigma^{(2,K)} = \Sigma^{(1,K)} = \begin{pmatrix}
\text{var}[M] & 0 \\
0 & \Sigma^{(0,K)}
\end{pmatrix}.
\]

After observing signals, the investor evaluates the posterior distribution of the security payoffs and evaluates the conditional covariance matrix for the payoffs in the consideration set, which we denote by \( \hat{\Sigma}^{(\iota,K)} \). Because every macro signal \( m' \) is independent of every micro signal \( s_i' \), we assume that \( P_F \) is independent of \( s_i' \) and \( P_{S_i} \) is independent of \( m' \). The
conditional covariance matrices therefore have the form:

\[
\hat{\Sigma}^{(1,K)} = \begin{pmatrix}
\text{var}[M|P_F] & 0 \\
0 & \hat{\Sigma}^{(0,K)}
\end{pmatrix}, \quad \hat{\Sigma}^{(2,K)} = \begin{pmatrix}
\text{var}[M|m', P_F] & 0 \\
0 & \hat{\Sigma}^{(0,K)}
\end{pmatrix}.
\]

Investors are constrained in how much information they can acquire, and we model this constraint through a bound on signal precision. Using \(|\cdot|\) to indicate the determinant of a matrix, for \(\iota = 0\) or \(\iota = 2\), we impose the constraint

\[
|\hat{\Sigma}^{(\iota,K)}|/|\Sigma^{(\iota,K)}| \geq \kappa,
\]

where \(0 < \kappa < 1\) measures the information capacity an investor attains at the cost \(c\). Smaller \(\kappa\) corresponds to greater variance reduction and thus greater capacity. For the case of \(\iota = 1\), meaning signal set \(I^{(1)}_{K}\), we impose the constraint

\[
\delta_F |\hat{\Sigma}^{(0,K)}|/|\Sigma^{(0,K)}| \geq \kappa,
\]

with \(\delta_F < 1\); in other words, we have replaced \(\text{var}[M|P_F]/\text{var}[M]\) in the determinant ratio with a fixed quantity \(\delta_F\).

This modification treats the index fund price differently from other types of information. We are particularly interested in the case

\[
\frac{\text{var}[M|P_F]}{\text{var}[M]} < \delta_F.
\]

When this inequality holds, making inferences from the price of the index fund consumes less capacity than would be expected from the variance reduction achieved. The idea is that the implications of the overall state of the market, as measured by the index fund, are widely discussed and publicly disseminated; \(\delta_F\) is the capacity consumed by paying attention to this ambient information. If (7) holds, then making inferences from the price of the index fund is at least slightly easier than making inferences from other information,

---

4 The posterior distribution preserves the independence of macro and idiosyncratic sources of risk, but we do not assume that \(\hat{\Sigma}^{(0,K)}\) has the same dependence structure as \(\Sigma^{(0,K)}\), a point emphasized in Sims (2011), p.167.

5 The determinant ratio is a multivariate generalization of a variance ratio, and it generalizes one minus a regression \(R^2\). The constraint in (5) is very similar to the entropy constraint used by Sims (2003), Mondria (2010), Van Nieuwerburgh and Veldkamp (2010), Hellwig et al. (2012) and others. With no information, \(\Sigma = \Sigma\) and the determinant ratio is 1, indicating that no capacity is consumed, whereas the entropy measure includes a term depending on the number of assets \(K\). When the number of assets is fixed, the measures are equivalent. Take the determinant of empty matrices to be one, so \(|\Sigma^{(0,K)}| = |\hat{\Sigma}^{(0,K)}| = 1\, \text{if} \, K = 0.\)
Agents seek to maximize expected time 1 utility, i.e. $E\{-E[e^{-\gamma \tilde{W}_2}|I]\}$

- Agents choose whether to acquire capacity $\kappa$ at cost $c$
- Those who become informed choose an information set from $I_K(0)$, $I_K(1)$, or $I_K(2)$ subject to capacity $\kappa$

Signals and prices are realized

- Informed observe their signals from the set $I^{(i)}_K$
- Uninformed observe either $P_F$ or some combination of prices from $\{P_{S_1}, \ldots, P_{S_N}\}$
- Agents submit portfolio choices as a function of their information $I$ to maximize $-E[e^{-\gamma \tilde{W}_2}|I]$

Stocks pay dividends $M + S_i$
- Index pays dividend $M$
- Agents realize utility $-e^{-\gamma \tilde{W}_2}$

Figure 1: Model timing.

holding fixed the level of variance reduction. We do not assume (7); we will show that it follows from more basic assumptions.

Because we condition on prices as well as nonpublic signals in (5) and (6), our formulation implies that making inferences from prices consumes some of the capacity $\kappa$. This point merits emphasis. The capacity $\kappa$ accounts for two types of effort: the effort required to acquire nonpublic signals $m'$ or $s'_i$, and also the effort required to make inferences from these signals and from publicly available prices. Price information is freely available, but regularly following the prices of hundreds of stocks and extracting investing implications from these prices consumes attention and effort.

Uninformed investors — those who do not incur the cost $c$ to acquire the capacity $\kappa$ — observe market prices, but they cannot observe signals $m(f)$ or $s_i(f)$, $f > 0$. Because making inferences from prices requires some information processing capacity, we endow uninformed investors with capacity $\delta_F$. This allows the uninformed to invest in the index fund and condition their demand on the price of the index. They may also reallocate this capacity to make inferences from the prices of individual stocks. Figure 1 summarizes the sequence of events in our model.

---

6Van Nieuwerburgh and Veldkamp (2010), p.796, propose a capacity constraint in which different variance ratios are raised to different powers. Our constraint can be viewed as raising $\text{var}[M/P_F]/\text{var}[M]$ to a power of zero and scaling it by a constant.

7In Section 4.1, $\delta_F$ has an alternative interpretation as the capacity consumed when an investor in a stock trades the index fund to hedge part of the macro risk in the stock. Trading to hedge rather than invest is less dependent on the information in the price.

8Kacperczyk et al. (2016) also analyze a variant of their main model with costly learning from prices.
Equilibrium

Once investors choose their information sets, their optimal portfolios (chosen to maximize expected utility) are determined by the price system. An investor’s strategy thus reduces to a choice of information set. We use the following definition of equilibrium: An equilibrium consists of a collection of investor information choices and a joint distribution (assumed normal) for prices, dividends, and signals, under which no investor can increase expected utility through a different information choice. We consider equilibria with the following features:

(e1) The joint distribution of the pairs \((S_i, P_{S_i})\), \(i = 1, \ldots, N\), is invariant under permutation of the indices, and every \((m(f), P_F)\) is independent of every \((s_i(f'), P_{S_i})\).

(e2) If no investors choose a macro signal \(m(f), f > 0\), then \(\text{var}[M|P_F] = \text{var}[M]\), and if no investors choose any micro signal \(s_i(f), f > 0, i = 1, \ldots, N\), then \(\text{var}[S_j|P_{S_j}] = \text{var}[S_j]\), for all \(j = 1, \ldots, N\).

(e3) The information cost and capacity parameters satisfy \(e^{2\gamma_c\kappa} < \delta_F\).

(e4) A positive fraction of investors choose to remain uninformed, and a positive fraction of these invest in the index fund.

Condition (e1) restricts attention to equilibria that are symmetric in the individual stocks, which is a reasonable restriction given that their dividends are ex ante identically distributed. This restriction works against finding equilibria in which investors make heterogeneous information choices. The second part of (e1) is consistent with the interpretation of the \(S_i\) as idiosyncratic components. Condition (e2) ensures that prices do not contain exogenous information about dividends — only information acquired by investors; the condition leaves open the possibility of a spillover of information from some \(s_i(f)\) to \(P_{S_j}\), \(j \neq i\).

We have not yet specified how investor information choices affect prices; (e1) and (e2) are minimal consistency requirements between prices and information choices. Theorem 2.1 will establish a necessary condition for equilibrium. If this condition holds under minimal assumptions on prices, it holds when we impose additional structure on price formation, as we do in Section 3.

The last two conditions limit us to “interior” equilibria: we will see that (e3) ensures that there is a benefit to becoming informed, whereas (e4) restricts our focus to equilibria

More precisely, no information set that is selected by a positive fraction of agents is strictly dominated by another information set.
in which not all investors become informed. The second half of (e4) further ensures that at least some of the uninformed participate in the market. We do not know if there are equilibria that violate (e4); but we consider that the most interesting equilibria satisfy (e4), and Theorem 2.1 should be understood as describing these equilibria.

To state the main result of this section, we highlight two types of information choices. Call informed investors who choose the information set $I_0^{(2)} = \{m, P_F\}$ macro-informed, and call informed investors who choose any information set $I_1^{(1)} = \{P_F, (s_i, P_{S_i})\}$, $i = 1, \ldots, N$, micro-informed. Here, $m$ and $s_i$ are the maximally informative macro and micro signals that can be achieved in these information sets with capacity $\kappa$.

**Theorem 2.1.** In any equilibrium satisfying (e1)–(e4), all informed investors choose to be either macro-informed or micro-informed, and both types of investors are present in positive proportions.

Under the conditions in the theorem, all informed investors choose one of two types of information. In particular, we obtain heterogeneous information choices by ex ante identical investors. This result stands in marked contrast to most of the related literature. In a partial equilibrium setting with exogenous prices, Van Nieuwerburgh and Veldkamp (2010) show that investors with exponential utility and a variance-ratio information constraint are indifferent across all feasible information choices: their investors would indeed be indifferent between $I_0^{(2)}$ and $I_1^{(1)}$, but they would also be indifferent between these and any other information set that consumed all available capacity. 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. In Kacperczyk et al. (2016) all informed investors choose the same information; more precisely, only aggregate information choices matter, so a market in which half of investors choose one information set and half choose another is equivalent to one in which all investors split their attention evenly between the two sets. In our setting, specialized micro- and macro-informed investors cannot be replaced with identical investors who divide their attention between micro and macro information.

In Goldstein and Yang (2015), the dividend of a single stock depends on two types of fundamentals. Their interpretation is different, but one could think of the two fundamentals as macro and micro sources of uncertainty. In their equilibrium, investors choose to learn about both fundamentals unless a cost penalty is introduced that makes the cost of acquiring both types of information greater than the sum of the costs of acquiring each type of information separately. Their outcome therefore differs from ours, in which investors choose to focus on one source of uncertainty. Investors in Goldstein and Yang
(2015) have just one security through which to trade on two types of dividend information, so information about one signal can be inferred from the other; the two types of information can be substitutes or complements, depending on the strength of the interaction effect. Our setting has as many securities as sources of dividend information, which removes the interaction effect; because investors specialize, macro and micro signals are strategic substitutes. Investors in Van Nieuwerburgh and Veldkamp (2009) also specialize, but their specialization depends on differences in prior information.

The proof of Theorem 2.1 requires several steps, as detailed in the appendix. Here we provide some brief intuition. We show that conditions (e1)–(e4) imply (7), which means that, in equilibrium, making inferences from the price of the index fund consumes at least slightly less capacity than would be expected from the variance reduction achieved. More surprisingly, in equilibrium, learning about individual stocks effectively introduces a fixed cost (in expected utility) to following each additional a stock, in addition to the variable cost associated with increased signal precision. This effect discourages informed investors from spreading their capacity across multiple stocks.

Going forward, we will denote by \( m = m(f_M) \) the maximally informative macro signal chosen by a macro-informed investor, \( \text{var}[m] = f_M \text{var}[M] \), and we will represent \( M \) as

\[ M = m + \epsilon_M, \tag{8} \]

where \( m \) and \( \epsilon_M \) are uncorrelated. Similarly, we will write

\[ S_i = s_i + \epsilon_i, \quad i = 1, \ldots, N, \tag{9} \]

where \( s_i \) and \( \epsilon_i \) are uncorrelated with each other, and where \( s_i = s_i(f_S) \), with \( \text{var}[s_i] = f_S \text{var}[S_i] \), is the maximally informative micro signal chosen by a micro-informed investor, recalling that the micro-informed also observe the index fund price \( P_F \). The information choices \( \{m, P_F\} \) and \( \{P_F, (s_i, P_{S_i})\} \), consume the investor’s full capacity, so we have

\[ \kappa = \frac{\text{var}[M|m, P_F]}{\text{var}[M]} = \delta_F \frac{\text{var}[S_i|s_i, P_{S_i}]}{\text{var}[S_i]}. \tag{10} \]

### 3 The constrained model

Theorem 2.1 shows that a necessary condition for an equilibrium in our setting is that all informed investors are either macro-informed or micro-informed. We will now show that such an equilibrium does in fact exist. We do so by imposing the necessary condition as a
constraint from the outset. In other words, we now consider a market with just three types of investors: uninformed, macro-informed, and micro-informed, with respective fractions $\lambda_U$, $\lambda_M$, and $\lambda_S = 1 - \lambda_U - \lambda_M$. The macro-informed select the signal $m$ in (8), and a micro-informed investor selects $P_F$ and a signal $s_i$ from (9); no other signals are chosen by any investors. We assume that the mass $\lambda_S$ of micro-informed investors is evenly divided among the $N$ stocks, so $\lambda_S/N$ investors observe each signal $s_i$, $i = 1, \ldots, N$, and only these investors invest directly in stock $i$. We limit the uninformed to investing in the index fund (the alternative of allocating their $\delta_F$ capacity to condition on individual stock prices is suboptimal under the conditions of Theorem 2.1). We extend the correlation condition in (2) to the $s_i$ and $\epsilon_i$, so that

$$\sum_{i=1}^N s_i = \sum_{i=1}^N \epsilon_i = 0.$$  \hfill (11)

Supply shocks

Investor demands for the securities will follow from their utility maximizing decisions. We now detail the supply of the securities. We suppose that the supply has a factor structure similar to that of the dividends in (1), with the supply of the $i^{th}$ stock given by

$$\frac{1}{N} (X_F + X_i).$$  \hfill (12)

Here, $X_F$ is the common supply shock, normally distributed with mean $\bar{X}_F$ and variance $\sigma^2_{X_F}$. The $X_i$ are normally distributed idiosyncratic shocks, each with mean 0 and variance $\sigma^2_X$. Supply shocks are independent of cash flows, and $X_i$ is independent of $X_F$ for all $i$. The $X_i$ have the same correlation structure as the $S_i$ in (2), so the idiosyncratic shocks diversify, in the sense that

$$\sum_{i=1}^N X_i = 0.$$  \hfill (13)

We make the standard assumption that supply shocks are unobservable by the agents.

The idiosyncratic supply shock $X_i/N$ has standard deviation $\sigma_X/N$, which is much smaller than the standard deviation $\sigma_{X_F}$ of the aggregate supply shock if $\sigma_{X_F}$ and $\sigma_X$ have similar magnitudes, as they will in our calibration to market data. With $N$ large and identical investors, the contribution to portfolio risk from idiosyncratic supply shocks would go to zero even in the absence of learning, but investor specialization provides an incentive to acquire information: only $1/N$ investors are informed about each stock.\footnote{In the model of Kacperczyk et al. (2016), individual risk factors are assumed to have the same...}
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), but the majority is likely to come from incorrect expectations or from other value-irrelevant triggers, such as an affinity for trading or fads. Empirical studies (Brandt, Brav, Graham, and Kumar 2010 and Foucault, Sraer, and Thesmar 2011) document a link from retail trading to idiosyncratic volatility of stock returns. We show in Section 7.2 that the volatility of $X_i$ enters directly into the idiosyncratic volatility of stock returns.

**Market clearing**

Let us write $q_j^U$, $q_j^M$, and $q_j^i$ for the demands of each investor group for security $j$, which can be one of the $N$ stocks or the index fund $F$. For each stock $i$, $q_i^i$ denotes the *direct demand* for stock $i$ by investors informed about stock $i$. Each group’s $F$ demand, $q_F$, leads to an *indirect demand* of $q_F/N$ for every stock $i$.

Aggregate holdings of the index fund are given by

$$q_F \equiv \lambda_U q_F^U + \lambda_M q_F^M + \frac{\lambda_S}{N} \sum_{i=1}^{N} q_F^i. \quad (14)$$

The market clearing condition for each stock $i$ is given by

$$\frac{\lambda_S}{N} q_i^i + \frac{q_F}{N} = \frac{1}{N} (X_F + X_i), \quad i = 1, \ldots, N. \quad (15)$$

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 (12). The direct and indirect demand for stock $i$ must equal its supply.

As (15) must hold for all $i$, the quantity $\xi \equiv \lambda_S q_i^i - X_i$ cannot depend on $i$. We can therefore write the direct demand for stock $i$ and the total demand for the index fund as

$$\lambda_S q_i^i = X_i + \xi, \quad q_F = X_F - \xi, \quad (16)$$

standard deviation as the aggregate risk factor. We explore the consequences of these contrasting modeling assumptions in Section 7 and Appendix A.5.
for some $\xi$ that does not depend on $i$.

We will show in Section 4.1 that in equilibrium $\xi$ must be zero, leading to two important implications. It will follow from (16) that micro-informed investors fully absorb the idiosyncratic supply shock $X_i$, and that the index fund holds the aggregate supply shock. We will interpret the first equation in (16) as liquidity provision by the micro-informed investors in the securities in which they specialize.

4 Market equilibrium in the constrained model

We construct an equilibrium in which the index fund price takes the form

$$P_F = a_F + b_F(m - \bar{m}) + c_F(X_F - \bar{X}_F), \quad (17)$$

and individual stock prices are given by

$$P_i = P_F + b_Ss_i + c_S(X_i + \xi), \quad i = 1, \ldots, N. \quad (18)$$

Here, $m$ and $s_i$ are the macro and micro signals in (8) and (9). Equation (17) makes the index fund price linear in the macro shock $m$ and the aggregate supply shock $X_F$. Equation (18) makes the idiosyncratic part of the price of stock $i$, $P_i - P_F$, linear in the micro shock $s_i$ and the idiosyncratic supply shock $X_i + \xi$. These prices satisfy (e1) and, consistent with (e2), the only information they contain about dividends comes from the selected signals $m$ and $s_i$.

4.1 Model solution

Macro-informed and uninformed investors trade only in the index fund, and a micro-informed investor trades the index fund and one security $i$. An investor sets his demand at time 1 by maximizing expected utility conditional on his information set $I$, as illustrated in Figure 1. Here $I = \{P_F\}$ for the uninformed, $I = \{m, P_F\}$ for the macro-informed, and $I = \{P_F, P_{S_i}, s_i\}$ for the micro-informed.

By standard arguments, the macro-informed demand for the index fund is given by

$$q^M_F = \frac{1}{\gamma(1 - f_M)\sigma^2_M} (m - RP_F), \quad (19)$$

as in equation (8) of Grossman and Stiglitz (1980), where $R$ is the risk-free gross return,
and uninformed demand for the index fund is given by

$$q^U_F = \frac{1}{\gamma \text{var}[M|P_F]}(E[M|P_F] - RP_F).$$

If $P_F$ takes the form in (17), then $E[M|P_F] = K_F(P_F - a_F) + \bar{m},$

$$\text{var}[M|P_F] = \text{var}[m|P_F] + \text{var}[\epsilon_M] = f_M \sigma^2_M (1 - K_F b_F) + (1 - f_M) \sigma^2_M,$$

$$K_F = \frac{b_F f_M \sigma^2_M}{b^2_F f_M \sigma^2_M + \sigma^2_X}.\qquad(21)$$

Demands of the micro-informed agents are given by the following proposition.

**Proposition 4.1.** If the prices $P_F$ and $P_i$ take the form in (17) and (18), then the demands of $i$ informed agents are given by

$$q_i = \frac{R}{\gamma(1 - f_S) \sigma^2_S}(P_F + s_i/R - P_i),$$

$$q_i^F = q^U_F - q_i.$$

**Equation (23)** shows that a micro-informed agent’s demand for the index fund consists of two components. The first component is the demand $q^U_F$ of the uninformed agents: neither the micro-informed nor the uninformed have any information about $M$ beyond that contained in $P_F$. The second term $-q_i^F$ offsets the exposure to $M$ that the micro-informed agent takes on by holding stock $i$. We interpret the second term as the micro-informed’s hedging demand: the micro-informed use the index fund to hedge out excess exposure to $M$ that they get from speculating on their signal $s_i$. The net result is that micro-informed and uninformed agents have the same exposure to $M$.

Substituting (23) in (14) — which gives the aggregate index fund demand — and combining this with the index fund market clearing condition in (16) yields

$$(\lambda_U + \lambda_S)q^U_F + \lambda_M q^M_F = X_F.\quad(24)$$

With the demands (19)–(20) for the index fund and demands (22)–(23) for individual securities, market-clearing prices are given by the following proposition:

\[\text{Using (23) we see that } N^{-1}\lambda_S \sum_i q_i \text{ in (14) equals } \lambda_S q^U_F - N^{-1} \sum_i \lambda_S q_i. \text{ Using the first equation in (16) this becomes } \lambda_S q^U_F - N^{-1} \sum_i (X_i + \xi) = \lambda_S q^U_F - \xi, \text{ and the second equation in (16) then yields (24).}\]

16
Proposition 4.2. The market clears at an index fund price of the form (17),

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

(25)

and prices for individual stocks $i$ of the form (18), given by

$$P_i = P_F + s_i \frac{R}{\lambda_M} - \frac{\gamma(1 - f_S)\sigma^2_S}{\lambda_S R}(X_i + \xi).$$

(26)

The no-arbitrage condition (4) is satisfied if and only if $\xi = 0$.

The form of the index fund price $P_F$ follows from Grossman and Stiglitz (1980); explicit expressions for the coefficients $a_F$, $b_F$, and $c_F$, are derived in the appendix. Comparison of (18) and (26) shows that the ratio $c_S/b_S$ in the price of stock $i$ has exactly the same form as $c_F/b_F$ in the price of the index fund in (25). In fact, if $\lambda_M = 1$, then $b_F = 1/R$ and $c_F$ has exactly the same form as $c_S$. The stock $i$ equilibrium is the direct analog of the index fund equilibrium with only macro-informed agents.

When the proportions $\lambda_U$, $\lambda_M$, and $\lambda_S$ are all endogenously positive and $f_M > f_S$, the constrained model solved by Proposition 4.2 realizes the equilibrium conditions of Theorem 2.1. The constrained model is more general in the sense that it does not impose a relationship between the information ratios $f_M$ and $f_S$. With prices as in Proposition 4.2, we can drop $P_F$ and $P_S_i$ from the conditioning in (10) and write (10) as

$$\kappa = 1 - f_M = \delta_F(1 - f_S).$$

(27)

Here we need $f_M > f_S$: the informativeness of the macro signal $m$ is greater than that of the micro signal $s_i$. We will see in Section 5.2 that (e3) leads to an interior equilibrium in the constrained model through (27).

The equality $\kappa = \delta_F(1 - f_S)$ suggests an alternative interpretation of $\delta_F$. To trade on their signal $s_i$, micro-informed investors trade stock $i$, which changes their exposure to macro risk, compared with an uninformed investor. We can interpret $\delta_F$ as the capacity consumed by hedging this extra macro risk, leaving informativeness $f_S$ for $s_i$. A fixed $\delta_F$ then means that hedging capacity does not depend on the informativeness of prices.

4.2 Price efficiency

We will investigate the extent to which prices reflect available information, and to do so we need a measure of price efficiency. For the case of the index fund, we define price
efficiency, $\rho_F^2$, as the proportion of price variability that is due to variability in $m$, the knowable portion of the aggregate dividend. This is the $R^2$ from regressing $m$ on $P_F$.

The squared correlation between $P_F$ in (17) and $m$ is given by

$$\rho_F^2 = \frac{b_F^2 f_M \sigma_M^2}{b_F^2 f_M \sigma_M^2 + c_F \sigma_X^2}.$$  

(28)

This equals $b_F K_F$ in (21), so we can use (21) to write $\text{var}[m|P_F] = f_M \sigma_M^2 (1 - \rho_F^2)$. As the price efficiency goes to 1, $P_F$ becomes fully revealing about $m$. Dividing both sides of (28) by $b_F^2 \sigma_M^2$ and using the expression for $c_F/b_F$ in (25), we get

$$\rho_F^2 = \frac{f_M}{f_M + \gamma^2 (1 - f_M)^2 \sigma_M^2 \sigma_X^2 / \lambda_M^2}.$$  

(29)

For stock $i$ we define price efficiency as the proportion of the variability of the price that is driven by variability in $s_i$, the knowable part of the idiosyncratic dividend shock, once $P_F$ is known. From the functional form of $P_i$ in (18) and the fact that $\xi = 0$, this is given by

$$\rho_S^2 = \frac{b_S^2 f_S \sigma_S^2}{b_S^2 f_S \sigma_S^2 + c_S \sigma_X^2} = \frac{f_S}{f_S + \gamma^2 (1 - f_S)^2 \sigma_S^2 \sigma_X^2 / \lambda_S^2},$$  

(30)

using the expression for $c_S/b_S$ in (26). As in the case of the index fund, as $\rho_S^2$ goes to 1, $P_i$ becomes fully revealing about $s_i$.

Differentiating (29) and (30) and straightforward algebra, yields the following result:

**Proposition 4.3** (When are prices more informative?).

(i) Micro (macro) prices are more efficient as either (a) the fraction of micro (macro) informed increases, or (b) as the micro (macro) signal informativeness improves:

$$d\rho_S^2/d\lambda_S > 0 \quad \text{and} \quad d\rho_F^2/d\lambda_M > 0,$$

and

$$d\rho_S^2/df_S > 0 \quad \text{and} \quad d\rho_F^2/df_M > 0.$$

(ii) Furthermore, when the fraction of micro (macro) informed is zero, or when the signals are non-informative, price efficiency is zero. In other words, $\rho_F^2 \to 0$ as either $\lambda_M \to 0$ or $f_M \to 0$, and $\rho_S^2 \to 0$ as either $\lambda_S \to 0$ or $f_S \to 0$.

(iii) When the signals are perfectly informative, prices become fully revealing. In other words, $\rho_F^2 \to 1$ as $f_M \to 1$ if $\lambda_M > 0$, and $\rho_S^2 \to 1$ as $f_S \to 1$ if $\lambda_S > 0$.  

18
5 Attention equilibrium in the constrained model

In Section 2, we allowed investors to acquire information processing capacity at a cost and then to allocate this capacity. In this section, we focus on the allocation decision, taking the decision to acquire capacity or remain uninformed as given. In other words, we hold $\lambda_U$ fixed and investigate the equilibrium mix of $\lambda_M$ and $\lambda_S$. In the next section, we endogenize $\lambda_U$ as well.

Part of the cost of becoming informed lies in developing the skills needed to acquire and apply investment information, and this process takes time. In the near term, these requirements leave the total fraction of informed investors $\lambda_M + \lambda_S$ fixed. Once investors have the skills needed to become informed, we suppose that they can move relatively quickly and costlessly between macro and micro information by shifting the focus of their attention. We therefore distinguish a near-term attention equilibrium, in which $\lambda_U$ is fixed and the split between $\lambda_M$ and $\lambda_S$ is endogeneous, from a longer-term information equilibrium, in which the decision to become informed is endogenized along with the choice of information on which to focus. We analyze the attention equilibrium in this section and address the information equilibrium in Section 6.

5.1 Relative utility

Recall that an investor’s ex ante expected utility is given by $J \equiv E[-\exp(-\gamma \tilde{W}_2)]$, where the expectation is taken unconditionally over time 2 wealth. Write $J_M$, $J_S$, and $J_U$ for expected utility of macro-informed, micro-informed, and uninformed investors, respectively.

Fixing the fraction of uninformed, the following proposition establishes the relative benefit of being macro- or micro-informed relative to being uninformed.

**Proposition 5.1.** If the cost of becoming informed is given by $c$, then the benefit of being macro-informed relative to being uninformed is given by

$$J_M/J_U = \exp(\gamma c) \left( 1 + \frac{f_M}{1 - f_M} \right)^{-\frac{1}{2}}. \quad (31)$$

The benefit of being micro-informed relative to being uninformed is given by

$$J_S/J_U = \exp(\gamma c) \left( 1 + \frac{f_S}{1 - f_S} \left( \frac{1}{\rho_S^2} - 1 \right) \right)^{-\frac{1}{2}}. \quad (32)$$

Note that because utilities in our model are negative, a decrease in these ratios repre-
\[\gamma = 5.5, \sigma^2_M = 0.01323, \sigma^2_S = 0.004408, f_M = 0.47, f_S = 0.2, \sigma^2_X = 0.648186, \sigma^2_Y = 0.259081\]

Figure 2: The information equilibrium for a fixed number of uninformed investors. Relative utilities are shown assuming cost of becoming informed is \(c = 0\).

...ents a gain in informed relative to uninformed utility. Each of the ex ante utility ratios in the proposition is increasing in the corresponding measure of price efficiency — that is, informed investors become progressively worse off relative to uninformed as micro or macro prices become more efficient. But the dependence on \(\rho^2_S\) in (32) differs from the dependence on \(\rho^2_F\) in (31).

Recalling from Proposition 4.3 that macro and micro price efficiency increase in \(\lambda_M\) and \(\lambda_S\), respectively, we immediately get the following:

**Proposition 5.2** (Benefit of information decreases with number of informed). \(J_S/J_U\) strictly increases (making micro-informed worse off) in \(\lambda_S\). \(J_M/J_U\) strictly increases (making macro-informed worse off) in \(\lambda_M\).

Figure 2 illustrates the results of Propositions 5.1 and 5.2. The figure holds \(\lambda_U\) fixed, and the x-axis is indexed by \(\lambda_M\). As \(\lambda_M\) increases, \(J_M/J_U\) increases, indicating that the macro-informed are becoming worse off. Similarly, at the rightmost point of the graph, \(\lambda_S = 0\), and as we move to the left, \(J_S/J_U\) increases, indicating that the micro-informed are becoming worse off as more of their type enter the economy.\(^{12}\)

\(^{12}\)Our numerical examples use the parameters calibrated to market data in Section 7.1.

\(^{13}\)Many of our comparisons could be recast as statements about trading intensities, in the sense of Goldstein and Yang (2015). Macro and micro trading intensities are given by \(-b_F/e_F\) and \(-b_S/c_S\) in Proposition 4.2.
5.2 Choice between macro and micro information

At an interior equilibrium, the marginal investor must be indifferent between macro and micro information, in which case equilibrium will be characterized by a $\lambda^*_M$ such that with that many macro-informed investors and with $1 - \lambda_U - \lambda^*_M$ micro-informed investors we will have $J_M = J_S$, which just sets (31) equal to (32). To cover the possibility of a corner solution, we define an attention equilibrium by a pair of proportions $\lambda_M \geq 0$ and $\lambda_S = 1 - \lambda_U - \lambda_M \geq 0$ satisfying

$$J_M < J_S \Rightarrow \lambda_M = 0 \quad \text{and} \quad J_S < J_M \Rightarrow \lambda_S = 0. \quad (33)$$

The inequalities in this condition are equivalent to $J_M/J_U > J_S/J_U$ and $J_S/J_U > J_M/J_U$, respectively, because $J_U < 0$.

Recall from Proposition 4.3 that when the fraction of macro- or micro-informed is zero, price efficiency is also zero. From (31) and (32), we see that

$$J_M/J_U(\lambda_M = 0) = e^{\gamma c} \sqrt{1 - f_M} \quad \text{and} \quad J_S/J_U(\lambda_S = 0) = 0. \quad (34)$$

From Proposition 5.2 we know that $J_M/J_U$ and $J_S/J_U$ both increase monotonically (i.e., make the informed worse off) with their respective $\lambda$’s. When $\lambda_M$ is zero, the macro-informed achieve their maximal utility; when $\lambda_M = 1 - \lambda_U$, the micro-informed achieve their maximal utility. As $\lambda_M$ increases from zero to $1 - \lambda_U$, $\lambda_S$ decreases, so the macro-informed become progressively worse off and the micro-informed become progressively better off. If at some $\lambda_M$ the two curves $J_M/J_U$ and $J_S/J_U$ intersect, we will have an interior equilibrium, and it must be unique because of the strict monotonicity in Proposition 5.2. This case is illustrated in Figure 2. If there is no interior equilibrium, then either macro or micro information is always preferred, and no investor will choose the other. Such a scenario is possible in the constrained model of Section 3, though not under the more general information choices in Theorem 2.1.

To make these observations precise, let us define

$$\tilde{\lambda}_M \equiv (1 - \lambda_U) \frac{1 - \sqrt{\varphi + (1 - \varphi) \frac{\gamma^2}{(1-\lambda_U)^2}}}{1 - \varphi}, \quad (35)$$

14 The first equality in (34) sheds additional light on (e3). Combining (34) with (27) yields $J_M/J_U < 1$ at $\lambda_M = 0$. But if $J_M/J_U < 1$ then some uninformed investors will prefer to become macro-informed, resulting in $\lambda_M > 0$. In this sense, (e3) leads to an interior attention equilibrium.

15 This expression has a finite limit as $\varphi \to 1$, and we take that limit as the value of $\tilde{\lambda}_M$ at $\varphi = 1$. 

21
where

$$\varphi = \frac{(1 - f_S)\sigma_S^2\sigma_X^2}{(1 - f_M)\sigma_M^2\sigma_{XF}^2}$$

and

$$\alpha = \frac{1 - f_M}{f_M}(1 - f_S)\sigma_S^2\sigma_X^2.$$  \hfill (36)

Note that $\varphi$ is the ratio of the total risk arising from the unknowable portion of idiosyncratic supply shocks (the variance of $\epsilon_i$ times the variance of $X_i$) to the total risk arising from macro supply shocks (the variance of $\epsilon_M$ times the variance of $X_F$). The larger $\varphi$ the more total unknowable risk comes from idiosyncratic rather than systematic sources.

The following proposition characterizes the equilibrium allocation of attention in the economy between macro information and micro information when the total fraction of informed investors $1 - \lambda_U$ is fixed.

**Proposition 5.3** (Attention equilibrium). Suppose $0 \leq \lambda_U < 1$, so some agents are informed.

(i) Interior equilibrium\footnote{We refer to (i) as the case of an interior equilibrium, even though it includes the possibility of a solution at the boundary. If $\lambda_M = \lambda_M^* = 0$, then $J_M = J_S$, and the marginal investor is indifferent between micro and macro information, which is what we mean by an interior equilibrium. If case (ii) holds, then $\lambda_M^* = 0$ because micro information is strictly preferred over macro information at all $\lambda_M$.}. If $\tilde{\lambda}_M \in [0, 1 - \lambda_U)$, then this point defines the unique equilibrium: at $\lambda_M^* = \tilde{\lambda}_M$, the marginal informed investor will be indifferent between becoming macro- or micro-informed.

(ii) If $\tilde{\lambda}_M \notin [0, 1 - \lambda_U)$, the unique equilibrium is at the boundary $\lambda_M^* = 0$, where all informed agents are micro-informed.

(iii) In equilibrium, we always have $\lambda_M^* < 1 - \lambda_U$. In other words, some informed agents will choose to be micro-informed.

It bears emphasizing that an attention equilibrium — regardless of parameter values — precludes all informed agents from being macro-informed. In contrast, it is possible for all informed agents to be micro-informed. We therefore have, as a fundamental feature of the economy, a bias for micro over macro information. This property holds in the constrained model, where informed investors are limited to macro or micro information, and the micro-informed receive all the benefits of providing liquidity to noise traders in individual stocks. We know from Theorem 2.1 that in a setting with greater information choices, an equilibrium will contain both micro- and macro-informed investors.

An increase in micro (macro) volatility, as measured by $\sigma_S\sigma_X$ ($\sigma_M\sigma_{XF}$), will increase the benefit of information to the micro (macro) informed, and will therefore decrease (increase) $\lambda_M^*$ when the economy is at an interior equilibrium:
**Proposition 5.4** (Effects of risk aversion and risk on the attention equilibrium). We consider the case of an interior equilibrium with $\lambda^*_M > 0$.

(i) Risk aversion pushes investors towards micro information: $d\lambda^*_M/d\gamma < 0$.

(ii) Increase in micro (macro) risk pushes investors towards micro (macro) information:

$$\frac{d\lambda^*_M}{d(\sigma_S\sigma_X)} < 0 \quad \text{and} \quad \frac{d\lambda^*_M}{d(\sigma_M\sigma_{X_F})} > 0.$$  

### 5.3 Relative price efficiency

Define

$$\tau_M \equiv \frac{f_M/(1-f_M)}{f_S/(1-f_S)}$$

where $f_M, f_S$ are the fraction of total information in macro and micro signals, respectively. Under condition \([27]\), $f_M > f_S$ so $\tau_M > 1$; however, in the general setting of the constrained model of Section 3, any $\tau_M \geq 0$ is feasible. We therefore explore the full range of possible $\tau_M$ values but put particular emphasis on the case $\tau_M > 1$.

In an interior attention equilibrium, the marginal informed investor is indifferent between macro and micro information because $J_M/J_U = J_S/J_U$. From Proposition 5.1, we see this condition then implies that micro price efficiency is related to macro price efficiency via

$$\frac{1 - \rho^2_S}{\rho^2} = \tau_M(1 - \rho^2_F),$$

which yields

$$\rho^2 = \frac{1}{1 + \tau_M(1 - \rho^2_F)}. \quad (37)$$

As markets become fully macro efficient ($\rho^2_F \to 1$), they must also become fully micro efficient ($\rho^2_S \to 1$), and vice versa. However, as macro price efficiency tends towards zero, micro price efficiency tends towards $1/(1+\tau_M)$. Since both sides of (37) are decreasing as their respective $\rho^2$ falls, this also represents the lower bound for $\rho^2_S$ in an interior attention equilibrium. When less information is revealed in equilibrium, the economy tends towards micro efficiency – suggesting micro information is in a sense more valuable than macro information.

It follows from (37) that $\rho^2_S > \rho^2_F$, i.e., markets are more micro efficient, whenever $\rho^2_F < 1/\tau_M$, which certainly holds if $\tau_M < 1$. For the calibrated parameter values in Section 7.1, $\tau_M = 3.4$, and the difference in price efficiencies $\rho^2_S - \rho^2_F$ as a function of $\rho^2_F$ is illustrated in Figure 3. The solid (dashed) portion of the curve represents the region in which micro efficiency exceeds (is less than) macro efficiency.

The red dots in the
$\gamma = 5.5$, $\sigma^2_M = 0.01323$, $\sigma^2_S = 0.004408$, $f_M = 0.47$, $f_S = 0.2$, $\omega^2_M = 2.593$, $\omega^2_S = 0.2591$

Figure 3: The figure shows the difference between micro and macro price efficiency, $\rho^2_S - \rho^2_F$, as a function of macro price efficiency $\rho^2_F$. The solid (dashed) portion of the curve represents the region of micro (macro) efficiency. The red points are labeled with the values of $\lambda_U$ corresponding to that particular $\rho^2_F$. The vertical portion of the curve corresponds to corner equilibria with $\lambda_M = 0$.

The figure represents attention equilibria at a given level of $\lambda_U$, with $\lambda_M$ given by equation 35 and $\rho^2_F$ and $\rho^2_S$ determined by (29) and (30), respectively. The vertical portion of the curve represents corner equilibria with no macro informed investors. Unless the number of uninformed investors is implausibly low, the economy is in a region of micro efficiency.

5.4 Impacts of information precision

Recall from (8) and (9) that $f_M$ and $f_S$ measure the fraction of variation in $M$ and $S_i$ that is known to informed investors. We refer to this as information precision. Surprisingly, more precise micro information makes the micro-informed worse off:

**Proposition 5.5** (The micro-informed overtrade on their information). *More precise information is worse for the micro-informed in the sense that*

$$\frac{d(J_S/J_U)}{df_S} > 0 \quad \text{(micro informed are worse off)}.$$  

When investors become micro-informed, the more they know about the ultimate idiosyncratic portion of the payout $S_i$, the less uncertainty they face from owning the stock. From (26) we see that the discount in the stock price due to idiosyncratic supply shocks...
$X_i$ will be zero when the micro information is perfect, i.e., when $f_S = 1$. With no discount in the price, the compensation for liquidity provision goes to zero. Because atomic informed agents cannot act strategically and coordinate to limit their liquidity provision in an optimal (for them) way, uncertainty about the dividend helps them by decreasing the sensitivity of their demand to price shocks, which in turn leads to a higher risk premium in prices. In contrast to the micro-informed, the macro-informed may be better or worse off as their precision, $f_M$, improves (see Figure 4):

**Proposition 5.6** (The macro-informed can be better or worse off with more information).

*More precise information is better for the macro-informed if and only if*

$$\rho_F^2 < \frac{1}{1 + f_M}, \quad \text{or equivalently} \quad \lambda_M < \gamma \sigma_M \sigma_{XF} \frac{1 - f_M}{f_M}. \quad (38)$$

*In this case,*

$$\frac{d(J_M/J_U)}{df_M} < 0 \quad \text{(macro-informed are better off).}$$

To gain intuition into this result recall that at $f_M = 0$ we would have $\rho_F^2 = 0$ (price reveals nothing when nothing about $M$ is knowable), and at $f_M = 1$ we would have $\rho_F^2 = 1$ (prices are fully revealing when $M$ is fully known). Furthermore, from Proposition 4.3 we know $\rho_F^2$ increases monotonically in $f_M$. So (38) implies that the macro-informed benefit from an increase in the precision $f_M$ only when $f_M$ (hence also the price informativeness $\rho_F^2$) is low. Equivalently the condition can be reinterpreted as placing a limit on how many macro-informed investors the economy can support before better macro precision begins to make the macro-informed worse off.

The contrast between micro and macro information in Propositions 5.5 and 5.6 can be understood as follows. In the market for the index fund, informed investors trade against uninformed investors as well as taking the other side of price insensitive liquidity shocks, introducing an effect that is absent in the market for individual stocks. With a low signal precision, prices are not very informative, so a small improvement in precision gives the macro-informed an informational edge over the uninformed, allowing the informed to extract rents in trading. However, as the signal precision improves and price efficiency grows, the incremental ability to extract rents from trading against the uninformed diminishes, while the tendency to overtrade on information (as in the case of the market for individual stocks) grows.

As illustrated in Figure 4, these propositions imply that $\lambda^*_M$ increases in $f_S$, and it
Figure 4: The effect of increasing micro precision $f_S$ (left) or macro precision $f_M$ (right) on the attention equilibrium with fixed $\lambda_U$.

also increases in $f_M$ as long as the condition in \((38)\) holds\(^{17}\)

6 Information equilibrium in the constrained model

We now examine a longer-term equilibrium in which the uninformed can become informed by incurring a cost $c$. In other words, while continuing to work within the constrained model of Section 3, we now endogenize not only the choice between micro and macro information, but also the decision to become informed. An equilibrium in this setting — which we refer to as an information equilibrium — is defined by nonnegative proportions $(\lambda_M, \lambda_S, \lambda_U = 1 - \lambda_M - \lambda_S)$ such that no agent of a type in positive proportion prefers switching to a different type. Extending (33), we require that, for any $i, i' \in \{M, S, U\}$,

\[ J_i/J_{i'} > 1 \Rightarrow \lambda_i = 0. \quad (39) \]

Recall that our utilities are negative, so the inequality on the left implies that type $i'$ is preferred to type $i$.\(^{18}\)

Figure 5 helps illustrate the general results that follow. The figure plots the equilibrium proportion of each type of investor in the constrained model as a function of the cost $c$ of

\(^{17}\)This is discussed in more detail in the Internet Appendix. See Bond and Garcia (2017) for a similar result where increased signal precision can make the informed worse off.

\(^{18}\)The ratios $J_M/J_U$, $J_S/J_U$, and $J_S/J_M$ all have well-defined limits as some or all of $\lambda_M$, $\lambda_S$, and $\lambda_U$ approach zero. (This follows from the expressions for these ratios in (31) and (32) and the dependence of $\rho_F^2$ and $\rho_S^2$ on $\lambda_M$ and $\lambda_S$ in (29) and (30), respectively.) We may therefore evaluate and compare these ratios even in cases where one or more of the proportions $\lambda_i$ is zero.
information acquisition. The figure divides into three regions. At sufficiently low costs, all agents prefer to become informed, so $\lambda_U = 0$. At sufficiently high costs, no investors choose to be macro-informed, so $\lambda_M = 0$. At intermediate costs, we find agents of all three types, and this is the region of overlap with Theorem 2.1. At all cost levels, some fraction of agents choose to be micro-informed.

To justify these assertions and to give an explicit characterization of the information equilibrium at each cost level $c > 0$, we first consider the possibility that all three types of agents are present in positive proportions. To be consistent with equilibrium, this outcome requires $J_M/J_U = J_S/J_U = 1$. Using the expressions for these ratios in (31) and (32), these equalities imply

$$\rho_F^2 = 1 - \frac{1 - f_M}{f_M} [e^{2\gamma c} - 1].$$

and

$$\rho_S^2 = \left(1 + \frac{1 - f_S}{f_S} [e^{2\gamma c} - 1]\right)^{-1}.\quad (41)$$

Setting these expressions equal to (29) and (30), respectively, we can solve for $\lambda_M$ and $\lambda_S$ to get

$$\lambda_M(c) = \gamma(1 - f_M)\sigma_M\sigma_X \frac{1}{(1 - f_M)(e^{2\gamma c} - 1) - \frac{1}{f_M}}^{1/2}.$$

(42)
and
\[ \lambda_S(c) = \gamma(1 - f_S)\sigma_S\sigma_X \left( \frac{1}{(1 - f_S)(e^{2\gamma c} - 1)} \right)^{1/2}. \] (43)

The expression for \( \lambda_M(c) \) is valid for \( c \leq \bar{c} \), with
\[ \bar{c} = -\frac{1}{2\gamma} \log(1 - f_M); \] (44)
set \( \lambda_M(c) = 0 \) for \( c > \bar{c} \). If \( \lambda_M(c) + \lambda_S(c) \leq 1 \) with \( c \leq \bar{c} \), then \((\lambda_M(c), \lambda_S(c), 1 - \lambda_M(c) - \lambda_S(c))\) defines an information equilibrium with \( J_M = J_S = J_U \).

Both \( \lambda_M(c) \) and \( \lambda_S(c) \) increase continuously and without bound as \( c \) decreases toward zero, so the equation
\[ \lambda_M(c) + \lambda_S(c) = 1, \]
defines the lowest cost at which we can meaningfully set \( \lambda_U = 1 - \lambda_M(c) - \lambda_S(c) \). At lower cost levels, we need to consider the possibility of an equilibrium with \( \lambda_U = 0 \).

Once we fix a value for \( \lambda_U \), the split between macro- and micro-informed agents is characterized by Proposition 5.3. Write \( \lambda^*_M(0) \) for the value of \( \lambda^*_M \) in Proposition 5.3 at \( \lambda_U = 0 \); this value is given either by the root \( \tilde{\lambda}_M \) in (35) or zero. Set
\[ \lambda_M, \lambda_S, \lambda_U = \begin{cases} \lambda^*_M(0), 1 - \lambda^*_M(0), 0, & 0 < c < \zeta; \\ \lambda_M(c), \lambda_S(c), 1 - \lambda_M(c) - \lambda_S(c), & \zeta \leq c < \bar{c}; \\ 0, \lambda_S(c), 1 - \lambda_S(c), & c \geq \max\{\zeta, \bar{c}\}. \end{cases} \] (45)

Then (45) makes explicit the equilibrium proportions illustrated in Figure 5 for the case \( \zeta < \bar{c} \): at large cost levels, \( \lambda_M = 0 \); at low cost levels, \( \lambda_U = 0 \) and \( \lambda_M \) and \( \lambda_S \) are constant; at intermediate cost levels, all three proportions are positive; at all cost levels, \( \lambda_S > 0 \). We always have \( \zeta > 0 \) and \( \bar{c} < \infty \), so the low cost and high cost ranges are always present; but it is possible to have \( \zeta \geq \bar{c} \), in which case the intermediate cost range is absent. This occurs when \( \lambda_S(\bar{c}) \geq 1 \). By evaluating (43) at (44), we find that \( \lambda^*_S(\bar{c}) = \gamma^2\alpha \), with \( \alpha \) as in (36). If \( \gamma^2\alpha \geq 1 \), then the root \( \tilde{\lambda}_M \) in (35) evaluated at \( \lambda_U = 0 \) is less than or equal to zero if it is real, so \( \lambda^*_M(0) = 0 \). We summarize these observations in the following result.

Proposition 6.1 (Information equilibrium in the constrained model). At each \( c > 0 \), the proportions in (45) define the unique information equilibrium. If \( \gamma^2\alpha < 1 \), then \( \zeta < \bar{c} \) and all three cases in (45) are present. If \( \gamma^2\alpha \geq 1 \), then \( \zeta \geq \bar{c} \), the second range in (45) is empty, and no investors choose to be macro-informed at any cost level.

With this result, we can revisit some of the conditions in Theorem 2.1 when applied
to the constrained model. Recalling from (27) that $\kappa = 1 - f_M$, condition (e3) implies $c < \bar{c}$, the condition for $\lambda_M(c) > 0$. Combining (e3) with (e4) ensures that we are in the range $c < c < \bar{c}$ and thus that all three investor proportions are positive.

From Proposition 6.1, we can deduce several further properties of the information equilibrium. Let us define $\Pi_M$ as the fraction of informed who are macro-informed, or

$$\Pi_M \equiv \frac{\lambda_M}{\lambda_M + \lambda_S} = \frac{\lambda_M}{1 - \lambda_U}. \quad (46)$$

At an interior attention equilibrium ($\lambda_M^* > 0$), $\Pi_M$ is the coefficient of $(1 - \lambda_U)$ in (35). Differentiating with respect to $\lambda_U$ yields $d\Pi_M/d\lambda_U < 0$, when $\lambda_M > 0$: the more uninformed investors there are in the economy, the greater the fraction of informed investors who choose to be micro-informed. The next result describes the dependence of $\Pi_M$ on $c$.

**Corollary 6.1** (Effect of information cost $c$ on information equilibrium). In equilibrium, with a cost of becoming informed given by $c$, the following will hold:

(i) As $c$ increases, the fraction $\Pi_M$ of informed investors who choose macro information falls; moreover, $\Pi_M$ is strictly decreasing in $c$ if $\lambda_M > 0$ and $\lambda_U > 0$.

(ii) As $c$ increases the fraction of investors who are uninformed increases; moreover $\lambda_U$ is strictly increasing in $c$ wherever $\lambda_U > 0$.

(iii) Micro and macro price efficiency are decreasing in $c$.

As $c$ increases and the number of uninformed grows, the benefit to being informed increases, in the sense that $J_M/J_U$ and $J_S/J_U$ decrease, as shown Proposition 5.2. However, the micro-informed gain more than the macro-informed. In order to maintain the attention equilibrium at a higher $c$, we need more micro-informed to equilibrate the relative benefits of micro versus macro information. Therefore, $\Pi_M$ must fall when $c$ increases.

7 Model implications: Variance decompositions

To study changes in expected stock returns, Campbell (1991) decomposes the variance of aggregate market returns into variance from cash flow news and variance from news about discount rates. (We will use “VR” to abbreviate “variance ratio” in discussing the ratio of cash flow variance to discount rate variance.) Vuolteenaho (2002) estimates a similar decomposition for individual firms and finds a much larger VR for individual firms than for the aggregate market. Jung and Shiller (2005) call the cash flow component the efficient
market component of returns and they call the discount rate component the inefficient market component. They interpret the larger VR for individual stocks as evidence for Samuelson’s dictum: greater micro efficiency than macro efficiency.

In this section, we show that our model produces results consistent with empirical patterns when calibrated to market data. We also compare our model’s results with historical trends. For this comparison, we argue that the past century has seen a reduction in the cost of becoming informed and an increase in the “indirect” supply of the macro factor, by which we mean trading in the index (through ETFs and derivatives) that does not involve trading in the individual stocks. We then examine how VRs in our model respond to these changes and compare the results with historical trends.

We also show that the micro VR and macro VR can respond differently to an increase in the fraction of informed investors. In particular, our model predicts an increase in the micro VR and a U-shaped change in the macro VR as functions of the number of informed. These effects follow from the endogenous specialization in investor information choice in our model. They contrast with the model of Kacperczyk et al. (2016), where all informed investors have the same information, and micro and macro VR always decrease or remain unchanged as the fraction of informed investors increases.

7.1 Calibration

For our calibration we normalize the aggregate mean dividend level to equal one by setting $\bar{m} = \bar{X}_F = 1$. We think of the one period in our model as representing a year.

Supply shocks and turnover

We calibrate the share volatilities $\sigma_{X_F}$ and $\sigma_X$ to annual turnover. Lo and Wang (2000, Table 3) find that value-weighted stock turnover – shares traded divided by shares outstanding – in the US over the period 1987–1996 averaged 1.25% per week. This implies an annual turnover of $52 \times 1.25% = 65\%$. We update their results by calculating the equal-weighted average turnover for the Dow Jones Industrial average from 1980 to 2018. The Dow Industrials turnover has averaged 76% over this time period. In our model, we measure equal-weighted index turnover as $1/N \sum_{i=1}^{N} |X_F - \bar{X} + X_i|$. We therefore require

$$E|X_F - \bar{X} + X_i| = \sqrt{\frac{2}{\pi}} \times \sqrt{\sigma_{X_F}^2 + \sigma_{X}^2} = 0.76,$$
using a standard result for the normal distribution. We then regress firm-level on index
turnover (annualized, in rolling windows) and find that the average $R^2$ in these regression
is 47%\textsuperscript{19}. If we define stock-level turnover in our model as $|X_F - \bar{X} + X_i|/N$, then we
would like the $R^2$ of the regression of stock turnover on index turnover,

$$
|X_F - \bar{X} + X_i| = a + b \sum_{j=1}^{N} |X_F - \bar{X} + X_j| + Noise,
$$

to be 47%. Using results from Kamat (1958) (see the Internet Appendix), this gives us a
second equation in $\sigma_{X_F}$ and $\sigma_X$, which we solve to get $\sigma_{X_F} = 0.805$ and $\sigma_X = 0.509$.

We assume that all idiosyncratic trading demand in an individual stock takes places via
trading in the stock itself, and therefore is captured by our firm-level turnover measure. However, we believe that our bottom-up turnover index meaningfully understates the
actual index level liquidity demand. On June 13, 2018 we looked at the top 15 most
heavily traded ETFs (according to ETF.com), as well as the first two S&P500, S&P500
e-mini and Dow futures. The average daily trading volume over the prior 30 days for these
21 instruments represented an annualized dollar turnover of $18.25$ trillion\textsuperscript{20}. Compared
to the market capitalization of the Russell 3000 index of $30.49$ trillion, this represents an
annual turnover of 60%. It is important to note that while some of this trading results in
creation or redemption of ETFs, which leads to some trading by the ETFs in individual
stocks, the majority of trading in ETFs does not result in any underlying stock trades.
And clearly trading in futures does not directly lead to any individual stock trading.
Compared to our 76% bottom-up turnover estimate for the Dow Industrials, the actual
liquidity demand for index trading – which is the quantity that $X_F$ proxies for in our
model – is conservatively twice as high in the current market\textsuperscript{21}. To proxy for an increase
in this type of indirect index turnover (which does not involve turnover in individual
stocks), we use $X_F$ volatility levels of $\ell \times \sigma_{X_F}$, with $\ell = 1$ or 2.

\textsuperscript{19}Lo and Wang (2000, Table 7) show that the first two principal components of turnover-beta sorted
portfolios account for close to 90% of portfolio weekly turnover. Our number is lower because we are
interested in stock-level turnover.

\textsuperscript{20}These results are available in the Internet Appendix.

\textsuperscript{21}Including index options trading or trade in over-the-counter derivatives would further increase
turnover.
Dividend volatility

To map our single-period model to a multi-period environment, we interpret the aggregate dividend $M$ paid at the end of the period as the discounted value of $N$ months of dividends,

$$M \equiv D_{t+12} + \beta D_{t+13} + \cdots + \beta^{N-1}D_{t+N},$$  \hspace{1cm} (47)$$

where the $D_t$ are monthly dividends, and $\beta = 0.998$ is a discount factor. Glasserman, Mamaysky, and Shen (2018) find that monthly S&P500 dividends are (reasonably) well approximated by an AR(1) process where $D_{t+1} = \mu_D + 0.967 \times D_t + 0.047 \times \epsilon_{t+1}$, where $\epsilon_t$ are iid unit variance shocks.\footnote{Farboodi and Veldkamp (2018) estimate the one month autocorrelation of dividends to be 0.937 with unit dividends having innovations with 7.5% volatility.} With these parameters, we find numerically that the per unit volatility $\sigma_M / E[M]$ peaks in month 54 at a value of 0.1146. Given our normalization $E[M] = \bar{m} = 1$, we calibrate to this normalized volatility measure and set $\sigma_M^2 = 0.115^2$. Details are in the Internet Appendix.

Ball, Sadka, and Sadka (2009) and Bonsall, Bozanic, and Fischer (2013) show that in the US, between 60-80% of firm-level earnings variation (at quarterly or annual frequency) can be explained by contemporaneous macro factors. In light of this, we set

$$\frac{\sigma_M^2}{\sigma_M^2 + \sigma_S^2} = 0.75,$$

which implies $\sigma_S^2 = \sigma_M^2 / 3$.

Dividend forecastability

Fama and French (2000, Table 2) regress year-ahead changes in earnings of US firms on a set of lagged market- and accounting-based explanatory variables (we assume earnings are paid out as dividends). They find that the $R^2$’s of these regressions range from 0.05 to 0.20. We use the high end of this range, $f_S = 0.2$, under the assumption that informed investors have information superior to a simple regression model.\footnote{Nissim and Ziv (2001, Table III) show that lagged accounting variables and dividend changes are able to explain 14.6% of variation in year-ahead earnings growth. Lev and Nissim (2004) study the effects of changes in accounting standards on earnings predictability. Using lagged accounting and market based variables, they find the $R^2$ of forecasting regressions for year-ahead earnings changes to be between 14-18\% (Table 3). Our choice of $f_S = 0.20$ therefore respresents the high-end of year-ahead earnings predictability documented in the literature.}

To come up with a one-year earnings forecastability benchmark for the S&P500 index we run a forecasting model similar to that in Nissim and Ziv (2001). We use market,
analyst and accounting information to forecast one-year ahead S&P500 earnings and dividend growth. In both sets of regressions\textsuperscript{24} we forecast

\[
\frac{CF[t, t + x] - CF[t]}{Book[t]}
\]

where \(CF[t, t + x]\) is the average cashflow-like (either earnings or dividends) variable over the next \(x\) years, \(CF[t]\) is the value of this variable over the prior 12 months, and \(Book[t]\) is the current per share book value for the S&P500. For earnings we find that the \(R^2\) increases from 0.303 to 0.689 as \(x\) increases from 1 to 5 years, whereas for dividends we find that the \(R^2\) falls from 0.639 to 0.296\textsuperscript{25}. The average \(R^2\) between the two regressions is very stable over time, ranging from 0.47 to 0.49. We use \(f_M = 0.47\) in our calibrations.

**Summary of calibration**

Finally, we use an annual (one-period) interest rate of 2\% (which is consistent with our choice of \(\beta\) in \[47\]), and we set \(\gamma = 5.5\). Our choice of risk aversion, interest rate, and aggregate dividend level leads to an annual excess return, \(\bar{m}/\alpha_F - R\), in the range of 6\% (the risk premium depends on the level of \(c\)). To summarize, our parameters choices are

<table>
<thead>
<tr>
<th>(\gamma)</th>
<th>(R)</th>
<th>(\bar{m})</th>
<th>(\bar{X})</th>
<th>(f_M)</th>
<th>(f_S)</th>
<th>(\sigma_M^2)</th>
<th>(\sigma_S^2)</th>
<th>(\sigma_{XF}^2)</th>
<th>(\sigma_X^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5</td>
<td>1.02</td>
<td>1</td>
<td>1</td>
<td>0.47</td>
<td>0.20</td>
<td>0.115(^2)</td>
<td>0.115(^2/3)</td>
<td>((\ell \times 0.805)^2)</td>
<td>0.509(^2)</td>
</tr>
</tbody>
</table>

Table 1: Parameter values from model calibration.

### 7.2 Variance decomposition

The variance decompositions in Campbell (1991) and Vuolteenaho (2002) are based on returns. Since our normally distributed prices have a small probability of being negative, returns in our model are not always well defined. As is customary in this literature (for example, Peng and Xiong 2006 and Veldkamp 2006), we instead analyze profits \((M - RP_F)\) in the case of the index fund) though we refer to them as returns. For the index fund,

\[
M - RP_F = m + \epsilon_M - RP_F = constant + \epsilon_m + (1 - Rb_F)m - Rc_FX_F. \tag{48}
\]

\textsuperscript{24}Detailed results are available in the Internet Appendix.

\textsuperscript{25}The latter finding is consistent with Beeler and Campbell’s (2012) finding that the \(R^2\)’s of dividend and consumption growth rate forecasting regressions fall with the forecasting horizon. Our \(R^2\)’s are higher than their Table 6 partly because we are using average dividends paid over the forecast horizon whereas they are using dividend growth – the latter being less predictable.
For time 0 investors, prior to the realization of prices, the variance of $M - RP_F$ is

$$Vol_{syst}^2 = \sigma_{\epsilon M}^2 + (1 - Rb_F)^2 \sigma_m^2 + R^2 c_F^2 \sigma_{XF}^2. \tag{49}$$

The ratio of the variance of cash flow news to that of discount rate news is

$$VR_M \equiv \frac{\sigma_{\epsilon M}^2 + (1 - Rb_F)^2 \sigma_m^2}{R^2 c_F^2 \sigma_{XF}^2}. \tag{50}$$

Our market-adjusted individual stock return is

$$u_i - RP_i - (M - RP_F) = \epsilon_i + \gamma(1 - f_S)\sigma_{XF}^2 X_i/\lambda_S, \tag{51}$$

which has the same form as the index return, except $1 - Rb_S = 0$. Note that, as in Vuolteenaho (2002), we are measuring “market-adjusted” (and not raw) stock returns. Since $\sigma_{\epsilon S}^2 = (1 - f_S)\sigma_S^2$, the variance of idiosyncratic stock returns is given by

$$Vol_{idio}^2 = \sigma_{\epsilon S}^2 \left(1 + \frac{\gamma^2 \sigma_{\epsilon S}^2 \sigma_{XF}^2}{\lambda_S^2}\right). \tag{52}$$

The ratio of cash flow to discount rate news variance for market-adjusted returns is

$$VR_S \equiv \frac{\lambda_S^2}{\gamma^2 \sigma_{\epsilon S}^2 \sigma_{XF}^2}. \tag{53}$$

The figures on the left side of Figure 6 show the index ($VR_M$) and firm-level ($VR_S$) variance ratios as functions of the equilibrium fraction of informed investors $\lambda_M + \lambda_S$. Each point is calculated by first choosing a $\lambda_U$ in the interior equilibrium region\(^{26}\) and then calculating the equilibrium $\lambda_M^*$ and $\lambda_S^*$. Since the number of informed investors is falling in the cost $c$ of becoming informed, moving right along the x-axis takes us from higher to lower $c$’s. Panel A uses the baseline level of index turnover $\ell = 1$, and Panel B uses the higher level $\ell = 2$ discussed in Section 7.1. The figures on the left thus allow us to compare micro and macro VR (solid versus dashed), the effect of the cost of becoming informed (decreasing from left to right), and the effect of indirect index turnover (increasing from the top panel to the lower panels).

The figures make two main predictions. First, they predict that the single-stock (mi-

\(^{26}\)When $\lambda_u \in [0, 1 - \lambda_S(\max\{c, \bar{c}\})]$, as discussed in Proposition 6.1, we have in equilibrium $\lambda_M^* > 0$ ($\lambda_S^*$ is always positive).
VR is higher than the index-level (macro) VR, unless the fraction of informed investors and $\ell$ are both low. For the second prediction, we will argue that the cost of becoming informed has declined over time (moving us from left to right), and indirect index turnover has increased over time (moving us from the top down). The combined effect can be seen as a move from Point A to Point B in the figure: a large decrease in the macro VR along with a small increase in the micro VR.

We compare these model features to empirical estimates. Campbell (1991) finds that between 1927 and 1951 the VR for the aggregate market was approximately $0.437/0.185 = 2.362$, and that this ratio fell to $0.127/0.772 = 0.165$ in the 1952–1988 time period. Using data from 1954 to 1996 and a similar methodology, Vuolteenaho (2002) finds an aggregate VR of $0.0232/0.0296 = 0.784$, larger than Campbell’s 0.165 but still quite low relative to the 2.362 ratio from the 20’s to the 50’s. These findings support a large decrease over time in the macro VR, as suggested by Points A and B in the figure.

Using 1954–1996 firm-level market-adjusted returns, Vuolteenaho (2002) finds a micro VR of $0.0801/0.0161=4.975$, much higher than the index-level measures. We interpret Panel B as reflective of more recent history, and thus consistent with a micro VR larger than the macro VR, though the model value is not as large as the historical estimate.

We do not have evidence on the single-stock VR in the early twentieth century. We know that since the 1950’s this ratio has been close to 5, and a higher value earlier seems implausible. Indeed, our model predicts that the micro VR would have been lower in the earlier period: with fewer informed investors to trade individual stocks (moving to the left in the figure), liquidity provision would fall, and discount factor variance would rise.

The model of Kacperczyk et al. (2016) makes different predictions. In their setting, the VR decreases (or remains constant) as the fraction of informed investors increases. As all informed investors in their model have the same information and hold the same portfolios, single-stock and index VRs all decrease as the number of informed investors grows. However, unless the number informed (or the total information capacity) is sufficiently large, investors will choose to learn only about the macro factor and not about individual stocks, in which case the micro VR remains constant as the macro VR declines.

\[27\] The variance units in Campbell (1991) and Vuolteenaho (2002) are not comparable, but the ratios are.

\[28\] We only analyze variation in $VR_M$ and $VR_S$ due to investor information choices. Changing variance and correlation of discount rates and cash flows also contribute to changing variance ratios. Therefore our analysis does not generate the empirically observed magnitude of change in $VR_M$.

\[29\] We are not aware of empirical analysis of individual stock variance ratios in an earlier sample than the 1954–1996 time period in Vuolteenaho (2002). One difficulty with extending the Vuolteenaho (2002) study to an earlier period is that the merged CRSP-Compustat data are not available prior to 1954.
are interesting contrasts between our model and that of Kacperczyk et al. (2016). They
are discussed in more detail in Appendix A.5.

The increasing micro VR implied by our model is a consequence of the specialization
in investor information choice and can be understood as follows. As long as some investors
are micro informed ($\lambda_S > 0$), the price of stock $i$ fully incorporates the information $s_i$, and $s_i$ does not enter returns — it is absent from equation (50). However, with few micro informed, there are few investors to absorb the idiosyncratic supply shocks $X_i$, so these shocks lead to large price concessions in order to clear the market. As a result, the innovations in $X_i$ are the key determinant of return variance, making $VR_S$ low when $\lambda_S$ is low. As $\lambda_U$ falls, the number of micro informed grows, meaning $X_i$ shocks can be absorbed by a larger fraction of investors. The requisite price concession therefore decreases, and the proportion of return variance due to cash flow shocks increases.

In contrast, the macro cash flow ratio is a non-monotonic function of $\lambda_M + \lambda_S$. When there are many uninformed investors, little information about $M$ is incorporated into prices, which makes the cash flow contribution to returns in (49) relatively large. At the same time, supply shocks $X_F$ can be absorbed by all investors — since everyone trades in the index fund — therefore preventing the price concession due to $X_F$ from dominating the variance of returns. As $\lambda_U$ decreases and the number of macro informed grows, a larger portion of $m$ is incorporated into the index price, $P_F$, and the ratio of cash flow variation to discount variation falls. At lower $\lambda_U$, with higher $\lambda_M$, the risk effect begins to dominate — as more about $m$ is known (either through prices or through more informed investors), the cash flow risk from owning the index falls, and the supply shock $X_F$ requires less of a price concession to clear the market. As the variation in the price due to $X_F$ falls, the proportion of return variance due to cash flows begins to increase.

The non-monotonicity of the variance ratio for index returns thus reflects an important tradeoff. A decline in uninformed investors makes the index price more informative and makes index returns less sensitive to cash flow news; but with a more informative prices, the index also becomes less sensitive to supply shocks. Without uninformed investors — as is the case for individual stocks — the relationship is monotonic.

If the cost of becoming informed continues to drop while index-related trading stays at the $\ell = 2$ level, we would expect the macro VR to eventually increase. However, if $\ell$ continues to increase, then we expect that macro VR continues to fall, while micro VR remains unchanged, as illustrated by Point C in Figure 6.

$^{30}$When $\lambda_M \to 1$, $(1 - Rb_F)^2$ in (49) goes to zero since $b_F \to 1/R$. For low $\lambda_M$, this term is large.
7.3 Time trends in variance and the cost of becoming informed

Moving from left to right in Figure 6 means increasing the fraction of informed investors and thus decreasing the cost of becoming informed, which we have suggested coincides with experience from the early to later part of the twentieth century. According to Figure 1 of Philippon (2012), the proportion of the US economy represented by the financial sector increased from 4-5% in the 20’s and 30’s to 8-9% by the 1990’s, with the finance industry share of the economy showing a clear increasing trend throughout the entire twentieth century. Greenwood and Scharfstein (2013) and Philippon and Reshef (2013) provide similar US 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. We take this as evidence that $\lambda_U$ fell, and $\lambda_M + \lambda_S$ rose, from the 1920’s and 30’s to the 1990’s.

The charts on the right side of Figure 6 show trends in the total variance of index returns, $\text{Vol}^2_{\text{syst}}$, and market-adjusted stock returns, $\text{Vol}^2_{\text{idio}}$. At our calibrated values of $f_M$ and $f_S$, most of the variation in returns comes from the unknowable components, $\epsilon_M$ and $\epsilon_i$, of the dividend. In fact, our return variance measures are of the same order of magnitude as cash-flow variation $\sigma^2_M$ and $\sigma^2_S$. According to our model, a gradual rise in the informed fraction $\lambda_M + \lambda_S$ combined with an increase in indirect index turnover (a move from point A to point B) should have resulted in a modest increase in market volatility from around $\sqrt{0.016} = 12.6\%$ per year to $\sqrt{0.019} = 13.8\%$ per year (if we assume $\lambda_M + \lambda_S$ increased to around 0.5). At the same time idiosyncratic return volatility would have remained almost constant.

The evidence on whether there has been a long-term trend in idiosyncratic volatility is mixed. Campbell et al. (2001) document an upward trend in US firm level volatility in the time period from 1962 to 1997. However, Brandt et al. (2010) and Bekaert, Hodrick

---

31 The Philippon 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 later part of the sample) they track each other very closely.

32 Further evidence in support of a rising $\lambda_M + \lambda_S$ is in Bai, Philippon, and Savov (2016). They show that the $R^2$ of regressing firm-level earnings in years $t + 1, \ldots, t + 5$ on year $t$ price to book ratios has been increasing since 1960. They interpret this finding as indicating that markets have become more informative about fundamentals. Our $\rho^2_F$ and $\rho^2_S$ measure the $R^2$ of regressing the knowable part of earnings $m$ or $s_i$ on today’s price. These quantities are increasing in $\lambda_M$ and $\lambda_S$ respectively, as is shown in Proposition 4.3.

33 Given the parameter values in Table 1, the average level of the equilibrium price, $a_F$, at different levels of $\lambda_U$ is very close to 0.925. We therefore interpret the square root of the payout variance as a rough proxy for return volatility.
and Zhang (2012) both argue that the Campbell et al. (2001) finding does not hold once an additional decade of data has been added. The current evidence is most consistent with the interpretation that there has been no trend in idiosyncratic volatility since the 1960’s. At the aggregate level, Campbell et al. (2001, Figure 1) and Brandt et al. (2010, Figure 2) show the time series of realized index-level volatility in the US from the 1920s; this series appears to have a downward trend, though the papers do not perform formal tests. As further evidence, we calculated the realized volatility of the S&P 500 index using rolling, overlapping 252-business-day windows and estimated a time trend of \(-3.4\) basis points per business day. The t-statistic using Newey-West with auto lag selection standard errors is \(-2.8\). Our volatility observations range from January 3, 1929 to July 16, 2018. Over this 89.5 year period, we estimate that average annualized volatility of the S&P 500 fell by 7.67 percent \((252 \times 0.00034 \times 89.5)\), from the low 20’s to the low teens.

The findings of no trend in idiosyncratic volatility is consistent with our model’s predictions on the right side of Figure 6 if the informed fraction \(\lambda_M + \lambda_S\) and indirect index turnover increased over the course of the 20th century. However, the small increase in systematic volatility predicted by the model is not consistent with the historical record.

8 Concluding remarks

A tendency for markets to be more micro than macro efficient has been a recurring theme of many of our results. For example:

- From Propositions \(5.3\) and \(6.1\) we get \(\lambda^*_S > 0\) — there are always some micro-informed investors. It is, in fact, possible that there are only micro-informed investors, so \(\lambda^*_M = 0\). This occurs when information is too costly to satisfy \((e3)\).

- More generally, Corollary \(6.1\) shows that when information is costlier, a larger fraction of the informed investors choose micro information.

- Proposition \(5.4\) shows increasing risk aversion raises the value of micro information.

- Figure \(3\) shows that for our model calibration, the economy is micro-efficient, unless the number of uninformed investors is implausibly low.

---

\(34\) Our 0.01 idiosyncratic return variance in Figure 6 translates to a volatility of idiosyncratic returns of roughly 10% — because the average price level for both the index and individual stocks is close to 1. This is roughly consistent, if slightly higher, than the average idiosyncratic volatility reported by Campbell et al. (2001) and Brandt et al. (2010).

\(35\) Details are in the Internet Appendix.
Figure 6: The left chart of the panels shows the variance decomposition of returns for the index fund, $VR_M$, and for market-adjusted stock returns, $VR_S$. The right chart shows the return variances, $Vol_{syst}^2$ from (49) and $Vol_{idio}^2$ from (50). The figure shows only the interior equilibrium region (see Proposition 6.1) where both $\lambda_M$ and $\lambda_S$ are positive.
Finally, Jung and Shiller (2005) argue that micro efficiency is characterized by a larger portion of stock return variation being driven by cash flow news than is the case for index return variation. As Panel B of Figure 6 shows, this holds in our model.

A key driver of these results is that the micro-informed are the only investors who collect surplus from accommodating idiosyncratic supply shocks. This creates a strong incentive to collect micro information. We see from (34) that the benefit $J_M/J_U$ to being the first macro-informed investor is finite, but the benefit to being the first micro-informed is infinite: $J_S/J_U$ takes on its highest value — zero — at $\lambda_S = 0$. This endogenous monopoly on liquidity provision for idiosyncratic shocks pushes markets towards micro efficiency.

A Appendix

A.1 Proof of Theorem 2.1

The proof of the theorem relies on three lemmas. To compare expected utility under alternative information choices, we will use the following expressions (A.1)–(A.2) for expected utility, which follow directly from Proposition 3.1 of Admati and Pfleiderer (1987).

Lemma A.1. Let $\Psi^{(0,K)}$ denote the covariance matrix of $S_i - RPS_i$, $i = 1, \ldots, K$, and let

$$
\Psi^{(0,K)} = \begin{pmatrix}
\var[M - RP_F] & 0 \\
0 & \Psi^{(0,K)}
\end{pmatrix}.
$$

The squared expected utility of an informed investor who chooses information set $I^\iota_K$ is

$$
J^2 = e^{2\gamma c} \times \left\{ \begin{array}{ll}
\kappa |\Sigma^{(0,K)}|/|\Psi^{(0,K)}|, & \iota = 0; \\
\exp(-Q_F)(\kappa \var[M|P_F]/\delta_F)|\Sigma^{(\iota,K)}|/|\Psi^{(1,K)}|, & \iota = 1, 2,
\end{array} \right.
$$

where $Q_F = (E[M - RP_F])^2/\var[M - RP_F]$. For an uninformed investor who conditions on $P_F$ and invests only in the index fund, the squared expected utility is given by

$$
J^2_U = e^{-Q_F} \frac{\var[M|P_F]}{\var[M - RP_F]}.
$$

Investors can evaluate (A.1)–(A.2) to make their information choices without first observing signals. Combining (A.1) with (5) and (6), the expected utility an investor attains by choosing information $I^\iota_K$ is given by

$$
J^2 = e^{2\gamma c} \times \left\{ \begin{array}{ll}
\kappa |\Sigma^{(0,K)}|/|\Psi^{(0,K)}|, & \iota = 0; \\
\exp(-Q_F)(\kappa \var[M|P_F]/\delta_F)|\Sigma^{(\iota,K)}|/|\Psi^{(1,K)}|, & \iota = 1, 2,
\end{array} \right.
$$
For \( \nu = 1 \), the expression follows from writing the ratio of determinants in (A.1) as

\[
\frac{\hat{\Sigma}^{(1,K)}}{\Psi^{(1,K)}} = \frac{\text{var}[M|P_F] \cdot |\hat{\Sigma}^{(0,K)}|}{|\Psi^{(1,K)}|},
\]

and then applying (6). The expressions in (A.3) hold as equalities when an investor uses the full capacity \( \kappa \), which is always possible and individually optimal if \( K \geq 1 \) or \( \nu = 2 \), so we will assume this condition holds. Interpret the case \( \nu = 0, K = 0 \) as the option not to invest, in which case the agent effectively consumes the acquired capacity.

The following lemma evaluates the determinants in (A.3).

**Lemma A.2.** For any \( K = 1, \ldots, N-1 \),

\[
\frac{|\Sigma^{(0,K)}|}{|\Psi^{(0,K)}|} = \left( \frac{\text{var}[S_i]}{\text{var}[S_i - RP_S_i]} \right)^K.
\]  

(A.4)

Also, \(|\Sigma^{(1,K)}| = \text{var}[M] \cdot |\Sigma^{(0,K)}| \) and \(|\Psi^{(1,K)}| = \text{var}[M - RP_F] \cdot |\Psi^{(0,K)}|\).  

(A.5)

**Proof.** Let \( G_K \) be the \( K \times K \) matrix with all diagonal entries equal to 1 and all off-diagonal entries equal to \(-1/(N-1)\). It follows from (2) that \( \Sigma^{(0,K)} = \sigma^2_S G_K \). In light of (4),

\[
\sum_{i=1}^{N} (S_i - RP_S_i) = \sum_{i=1}^{N} S_i - R \sum_{i=1}^{N} (P_i - P_F) = 0,
\]

Under (e1), it follows that, for \( i \neq j \),

\[
\text{cov}[S_i - RP_S_i, S_j - RP_S_j] = -\text{var}[S_i - RP_S_i]/(N-1),
\]

so the \( S_i - RP_S_i \) have the same correlation structure as the \( S_i \) themselves. In other words, \( \Psi^{(0,K)} = \text{var}[S_i - RP_S_i] G_K \), and then

\[
\frac{|\Sigma^{(0,K)}|}{|\Psi^{(0,K)}|} = \frac{|\sigma^2_S G_K|}{|\text{var}[S_i - RP_S_i] G_K|} = \frac{\sigma^2_S G_K}{|\text{var}[S_i - RP_S_i]| G_K},
\]

which yields (A.4). The block structure of \( \Sigma^{(1,K)} \) and \( \Psi^{(1,K)} \) yields (A.5). \( \square \)

**Lemma A.3.** Under the conditions of Theorem 2.1, we have

\[
\frac{\text{var}[M|P_F]}{\text{var}[M]} < \delta_F \quad \text{and} \quad \frac{\text{var}[S_i]}{\text{var}[S_i - RP_S_i]} > 1, \quad i = 1, \ldots, N,
\]  

(A.6)

and

\[
e^{-Q_F} \frac{\text{var}[M]}{\text{var}[M - RP_F]} \leq 1.
\]  

(A.7)

The first inequality confirms that making inferences from the price of the index fund consumes less information processing capacity than would be expected from the variance
reduction achieved; see the discussion surrounding \( \text{[7]} \). The reverse inequality would imply a “penalty” in conditioning on the index fund price, a scenario we see as uninteresting and rule out with the conditions in the theorem. The second inequality in \( \text{[A.6]} \) suggests that individual stock prices are, in a sense, sufficiently informative about fundamentals and not overwhelmed by noise trading. This second condition ensures that being micro informed is not strictly preferable to being uninformed — it is effectively a limit on the benefit of actively trading in individual stocks. As we will see in the proof of Theorem 2.1, this result also makes it suboptimal to have more than one stock in an investor’s consideration set. The inequality in \( \text{[A.7]} \) confirms that there is a benefit to investing in the index fund.

**Proof of Lemma A.3.** If investors strictly preferred becoming macro-informed over remaining uninformed and investing in the index fund, then \( \text{(e4)} \) would be violated. Thus, in any equilibrium satisfying \( \text{(e4)} \), an uninformed investor weakly prefers investing in the index fund over becoming macro-informed. We therefore have

\[
J^2_U \leq e^{2\gamma_c} e^{-Q_F} \frac{\text{var}[M|m, P_F]}{\text{var}[M - RP_F]} \equiv J^2_M,
\]

where \( m = m(f_M) \) is the signal in \( \text{[8]} \) acquired by an informed investor who allocates all capacity to learning about \( M \). Combining this inequality with \( \text{(A.2)} \), recalling from \( \text{(10)} \) that \( \text{var}[M|m, P_F]/\text{var}[M] = \kappa \), we get, by \( \text{(e3)} \),

\[
\frac{\text{var}[M|P_F]}{\text{var}[M]} \leq e^{2\gamma_c \kappa} < \delta_F.
\]

We can similarly compare \( J^2_U \) with the option of becoming micro-informed to get

\[
J^2_U \leq e^{2\gamma_c} e^{-Q_F} \frac{\text{var}[M|P_F]}{\text{var}[M - RP_F]} \frac{\text{var}[S_i|s_i, P_{S_i}]}{\text{var}[S_i - RP_{S_i}]} \equiv J^2_S,
\]

where \( s_i = s_i(f_S) \) is the signal in \( \text{[9]} \) acquired by an informed investor who allocates all capacity to \( \{P_F, (s_i, P_{S_i})\} \). Recalling from \( \text{(10)} \) that \( \delta_F \text{var}[S_i|s_i, P_{S_i}]/\text{var}[S_i] = \kappa \), this inequality reduces to

\[
1 \leq e^{2\gamma_c (\kappa/\delta_F)} \frac{\text{var}[S_i]}{\text{var}[S_i - RP_{S_i}]}.
\]

As \( e^{2\gamma_c \kappa} < \delta_F \) under \( \text{(e3)} \), the second inequality in \( \text{[A.6]} \) must hold.

Suppose no informed investors choose to learn more about \( M \) than its price, meaning that no investor chooses an information set of the type \( \mathcal{I}_K^{(2)} \). Then \( \text{(e2)} \) implies \( \text{var}[M|P_F] = \text{var}[M] \), which would contradict \( \text{[A.6]} \) because \( \delta_F < 1 \). It follows that some informed investor weakly prefers an information set \( \mathcal{I}_K^{(2)} \) over an information \( \mathcal{I}_K^{(0)} \) consisting of more precise signals about the same stocks and no information about \( M \). This preference implies

\[
e^{2\gamma_c} e^{-Q_F} \frac{\text{var}[M]}{\text{var}[M - RP_F]} \left( \frac{\text{var}[S_i]}{\text{var}[S_i - RP_{S_i}]} \right)^K \leq e^{2\gamma_c} \left( \frac{\text{var}[S_i]}{\text{var}[S_i - RP_{S_i}]} \right)^K \kappa,
\]

42
Combining Lemma A.2 and Lemma A.3, we get

\[ \text{Var} \]

showing that consideration set, in which case the investor cannot do better than dominated.

\[ \text{types of information sets} \]

**Proof of Theorem 2.1.** We need to show that for an informed investor, information choices other than \( \{m, P_F\} \) or \( \{s_i, P_{S_i}, P_F\} \) are suboptimal. We show this by considering the three types of information sets \( \mathcal{I}^{(i)}_K \), \( i = 0, 1, 2 \). For each case, we evaluate expected utility using \( \text{(A.3)} \) and show that any information choice other than \( \{m, P_F\} \) or \( \{s_i, P_{S_i}, P_F\} \) is strictly dominated.

**Case of \( i = 2 \).** We may take \( K \geq 1 \), since otherwise only the index fund is in the consideration set, in which case the investor cannot do better than \( \{m\} \) because of the capacity constraint. Then

\[
J^2 \geq e^{2\gamma_c} \exp(-Q_F) \frac{\kappa}{\var\[\Psi(1,K)\]} \cdot \text{Var}[M], \quad \text{using } \text{(A.3)}; \\
= e^{2\gamma_c} \exp(-Q_F) \kappa \frac{\var\[\Sigma(0,K)\] \cdot \var[M - RP_F]}{\var\[\Psi(0,K)\]}, \quad \text{using } \text{(A.5)}; \\
= e^{2\gamma_c} \exp(-Q_F) \left( \frac{\var[M|m, P_F]}{\var[M - RP_F]} \right) \frac{\var\[\Sigma(0,K)\] \cdot \var[M - RP_F]}{\var\[\Psi(0,K)\]}, \quad \text{using the first equality in } \text{(10)}; \\
= J^2 \frac{\var\[\Sigma(0,K)\] \cdot \var[M - RP_F]}{\var\[\Psi(0,K)\]}, \quad \text{using } \text{(A.8)}.
\]

Combining Lemma \( \text{(A.2)} \) and Lemma \( \text{(A.3)} \) we get \( \var\[\Sigma(0,K)\] \cdot \var\[\Psi(0,K)\] > 1 \), from which we conclude that \( J^2 > J^2 \). Thus, \( \mathcal{I}^{(2)}_K \) is strictly dominated by \( \{m\} \).

**Case of \( i = 1 \).** Start with \( K = 0 \). We have \( \var[S_i|s_i, P_{S_i}] = \var[S_i - RP_{S_i}|s_i, P_{S_i}] < \var[S_i - RP_{S_i}] \), and therefore from \( \text{(A.1)} \)

\[
J^2 = e^{2\gamma_c} \frac{\exp(-Q_F) \frac{\var[M|P_F]}{\var[M - RP_F]}}{\var\[S_i|s_i, P_{S_i}\]} > e^{2\gamma_c} \frac{\exp(-Q_F) \frac{\var[M|P_F]}{\var[M - RP_F]}}{\var\[S_i - RP_{S_i}\]} \frac{\var[S_i|s_i, P_{S_i}]}{\var\[S_i - RP_{S_i}\]} = J^2
\]

showing that \( \mathcal{I}^{(1)}_1 \) is preferred over \( \mathcal{I}^{(1)}_0 \). With \( K \geq 2 \), the squared expected utility is

\[
J^2 \geq e^{2\gamma_c} \exp(-Q_F) \frac{\delta_F\var[S_i|s_i, P_{S_i}]}{\var\[S_i\]} \left( \frac{\var[M|P_F]}{\var[M - RP_F]} \right) \frac{\var\[\Sigma(0,K)\] \cdot \var[M - RP_F]}{\var\[\Psi(0,K)\]} \frac{\var[S_i - RP_{S_i}]}{\var[S_i]}, \quad \text{using } \text{(10)} \text{ and } \text{(A.5)}; \\
= e^{2\gamma_c} \frac{\exp(-Q_F) \frac{\var[S_i|s_i, P_{S_i}]}{\var\[S_i\]} \left( \var[M|P_F] \right) \var\[\Sigma(0,K)\] \cdot \var[M - RP_F]}{\var\[\Psi(0,K)\] \var\[S_i - RP_{S_i}\]} \frac{\var[S_i - RP_{S_i}]}{\var\[S_i\]}, \quad \text{using } \text{(A.4)}; \\
= J^2 \frac{\var\[S_i - RP_{S_i}\]}{\var\[S_i\]},
\]

With \( K \geq 2 \), multiplying \( \text{(A.4)} \) by \( \var[S_i - RP_{S_i}] / \var[S_i] \) continues to yield an expression that is greater than 1. It follows that \( J^2 > J^2 \), so \( \mathcal{I}^{(1)}_K \), \( K \geq 2 \), is strictly dominated.
Case of \( t = 0 \). A nonempty information set requires \( K \geq 1 \), and then
\[
J^2 \geq e^{2\gamma c} \frac{K}{\Psi(0,K)}, \quad \text{using} \ (A.3);
\]
\[
\geq e^{2\gamma c - Q_F} \frac{\var[ M ]}{\var[ M - RP_F ]} \frac{K}{\Psi(0,K)}, \quad \text{using} \ (A.7)
\]
\[
eq e^{2\gamma c - Q_F} \frac{\var[ M | m, P_F ]}{\var[ M | P_F ]} \frac{K}{\Psi(0,K)}, \quad \text{using} \ (10)
\]
\[
= J^2_M \frac{\Sigma(0,K)}{\Psi(0,K)} \quad \text{>} \quad J^2_M, \quad \text{using} \ (A.8) \text{ and } (A.4).
\]

**No-deviation.** We have thus shown that either \( \{ m \} \) (macro-informed) or \( \{ s_i, P_S \} \) (micro-informed) dominates every other information set. If the proportion of macro-informed or micro-informed investors were zero, \((e2)\) would contradict one of the inequalities in \((A.6)\). The last assertion in the theorem follows. \( \square \)

### A.2 Solution of the constrained model

**Proof of Proposition A.1.** The analysis is simplified if we allow micro-informed agents to invest in the index fund and in a hedged security paying \( u_i - u_F = S_i \), with price \( P_S = P_i - P_F \). If we let \( \tilde{q}_F \) and \( \tilde{q}_S \) denote the demands in this case, the demands in the original securities are given by \( q_i^* = \tilde{q}_S \) and \( q_F^* = \tilde{q}_F - \beta \tilde{q}_S \). Write \( \mathcal{I}_i = \{ P_F, P, s_i \} \). By standard arguments, the modified demands are given by

\[
\begin{bmatrix}
\tilde{q}_F \\
\tilde{q}_S 
\end{bmatrix}
= \frac{1}{\gamma \var[ M ]} \begin{bmatrix} M \\ S_i \end{bmatrix} \mathcal{I}_i^{-1} \left( E \left[ \begin{bmatrix} M \\ S_i \end{bmatrix} | \mathcal{I}_i \right] - R \begin{bmatrix} P_F \\ P_i - P_F \end{bmatrix} \right).
\]

Now
\[
\var[ M | S_i ] \mathcal{I}_i = \left( \var[ M | \mathcal{I}_i ] \var[ S_i | \mathcal{I}_i ] \right) = \left( \var[ M | P_F ] \frac{(1 - f_S)\sigma^2_S}{1} \right), \quad (A.10)
\]

and
\[
E \left[ \begin{bmatrix} M \\ S_i \end{bmatrix} | \mathcal{I}_i \right] = \left[ E[ M | P_F ] \right]_{s_i}.
\]

Thus, \( \tilde{q}_F = q_F^U \), with \( q_F^U \) as given in \((20)\), and
\[
\tilde{q}_S = \frac{s_i - R(P_i - P_F)}{\gamma (1 - f_S)\sigma^2_S}.
\]

As \( q_i^* = \tilde{q}_S \), \((22)\) follows, and then \( q_F^* = \tilde{q}_F - \tilde{q}_S = q_F^U - q_i \) completes the proof. \( \square \)

**Proof of Proposition A.2.** The price \( P_F \) can be derived from first principles, but we can simplify the derivation by reducing it to the setting of Grossman and Stiglitz (1980). The informed \((19)\) and uninformed \((20)\) demands for the index fund and the market clearing

...
condition (24) reduce to the demands in equations (8) and (8') of Grossman and Stiglitz (1980) and their market clearing condition (9), once we take $\lambda = \lambda_M$ and $1 - \lambda_M = \lambda_U + \lambda_S$. The coefficients of the price $P_F$ in (17) can therefore be deduced from the price in their equation (A10). Theorem 1 of Grossman-Stiglitz gives an expression for $P_F$ in the form $\alpha_1 + \alpha_2 w_\lambda$, for constants $\alpha_1$ and $\alpha_2 > 0$, where, in our notation,

$$w_\lambda = m - \frac{\gamma(1 - f_M)\sigma_M^2}{\lambda_M}(X_F - \bar{X}_F).$$

Comparison with (17) yields the expression for $c_F/b_F$ in (25). From the coefficient of their $\theta$ (our $m$) in (A10) of Grossman-Stiglitz, we get

$$b_F = \frac{1}{R} \frac{\lambda_M}{(1-f_M)\sigma_M^2} + \frac{1-\lambda_M}{\text{var}[M|w_\lambda]} \frac{f_M}{\text{var}[w_\lambda]}.$$

Moreover,

$$\text{var}[w_\lambda] = (1 - f_M)\sigma_M^2 + \frac{\gamma^2(1 - f_M)^2\sigma_M^4}{\lambda_M^2} \sigma_{X_F}^2,$$

and $\text{var}[M|w_\lambda] = \text{var}[M|P_F]$. To evaluate $\text{var}[M|P_F]$, note that the only unknown term in (21) is $K_F b_F$, which we can now evaluate using (25) to get

$$K_F b_F = \frac{b_F^2 f_M \sigma_M^2}{b_F^2 f_M \sigma_M^2 + c_F^2 \sigma_{X_F}^2} = \frac{f_M \sigma_M^2}{f_M \sigma_M^2 + \frac{\gamma^2(1 - f_M)^2\sigma_M^4}{\lambda_M^2} \sigma_{X_F}^2}.$$ 

This yields an explicit expression for $\text{var}[M|P_F]$ which in turn yields an explicit expression for $b_F$ through (A.11). An expression for $c_F$ then follows using (25). Finally, to evaluate the constant term $a_F$, we can again match coefficients with the expression in (A10) of Grossman-Stiglitz. Alternatively, we can evaluate their (A10) at (using their notation) $\theta = E\theta^*$ and $x = E x^*$, which, in our notation yields

$$a_F = \frac{\bar{m}}{R} \frac{X_F}{R} \left[ \frac{1 - \lambda_M}{\gamma \text{var}[M|P_F]} + \frac{\lambda_M}{\gamma(1 - f_M)\sigma_M^2} \right]^{-1}.$$

Equation (26) follows directly from (22) and (16). In light of (11) and (13), condition (4) is satisfied if and only if $\xi = 0$.

### A.3 Attention equilibrium

**Proof of Proposition 5.1.** We use the expressions for $J_U$, $J_M$, and $J_S$ introduced in (A.2), (A.8), and (A.9), recalling that expected utility is negative. With prices given by Proposition 4.2 $\text{var}[M|P_F, m] = \text{var}[M|m] = (1 - f_M)\sigma_M^2$, so

$$J_M/J_U = e^{\gamma c} \left( \frac{\text{var}[M|P_F]}{(1 - f_M)\sigma_M^2} \right)^{-1/2}.$$
Combining (21) and (28) yields
\[ \text{var}[M|P_F] = f_M\sigma_M^2(1 - \rho_F^2) + (1 - f_M)\sigma_M^2, \]
from which (31) follows. Similarly, \[ \text{var}[S_i|P_{S_i}, s_i] = \text{var}[S_i|s_i] = (1 - f_S)\sigma_S^2, \]
so
\[ \frac{J_S}{J_U} = e^{\gamma c} \left( \frac{\text{var}[S_i - RP_{S_i}]}{(1 - f_S)\sigma_S^2} \right)^{-1/2}. \]
Using first (26) and then (30), we get
\[ \text{var}[S_i - RP_{S_i}] = (1 - f_S)\sigma_S^2 + \frac{\gamma^2(1 - f_S)^2\sigma_X^2}{\lambda_S^2}\sigma_S^2 = (1 - f_S)\sigma_S^2 + f_S\sigma_S^2 \left( \frac{1}{\rho_S^2} - 1 \right), \]
from which (32) follows.

\[ \textbf{Proof of Proposition 5.3.} \]
As noted in (34), \( J_S/J_U \) approaches zero as \( \lambda_S \) decreases to zero (and \( \lambda_M \) increases to \( 1 - \lambda_U \)). We know from (31) that \( J_M/J_U > 0 \) for all \( \lambda_M \); in fact, from (34) we know that \( J_M/J_U \geq \sqrt{1 - f_M} \). It follows from the strict monotonicity of \( J_M/J_U \) and \( J_S/J_U \) (Proposition 4.3) that either \( J_M/J_U > J_S/J_U \) for all \( \lambda_M \in [0, 1 - \lambda_U] \) or the two curves cross at exactly one \( \lambda_M \) in \( [0, 1 - \lambda_U] \). In the first case, all informed agents prefer to be micro-informed than macro-informed, so the only equilibrium is \( \lambda^*_M = 0 \).

In the second case, the unique point of intersection defines the equilibrium proportion \( \lambda^*_M \), as explained in the discussion of Figure 2. We therefore examine at which \( \lambda_M \) (if any) we have \( J_M/J_U = J_S/J_U \). We can equate (31) and (32) by setting
\[ \frac{1 - f_M}{f_M} \frac{1}{1 - \rho_F^2} = \frac{1 - f_S}{f_S} \frac{\rho_S^2}{1 - \rho_S^2}. \]
Using the expressions for \( \rho_F^2 \) and \( \rho_S^2 \) in (29) and (30), this equation becomes
\[ \frac{1 - f_M}{f_M} \lambda_M^2 + \frac{\lambda_M^2}{\gamma^2(1 - f_M)\sigma_M^2\sigma_X^2} = \frac{(1 - \lambda_U - \lambda_M)^2}{\gamma^2(1 - f_S)\sigma_S^2\sigma_X^2}. \]
Thus, \( \lambda_M \) satisfies a quadratic equation, which, with some algebraic simplification, can be put in the form \( A\lambda_M^2 + B\lambda_M + C = 0 \), where
\[ A = 1 - \varphi, \quad B = -2(1 - \lambda_U), \quad C = (1 - \lambda_U)^2 - \alpha \gamma^2, \quad \text{(A.13)} \]
with \( \varphi \) and \( \alpha \) as defined in (36). One of the two roots of this equation is given by \( \tilde{\lambda}_M \).
Denote the other root by
\[ \eta = \frac{-B + \sqrt{B^2 - 4AC}}{2A}. \]
We claim that \( \eta \not\in [0, 1 - \lambda_U] \). We may assume \( A \neq 0 \), because \( \eta \to \infty \) as \( A \to 0 \) because \( B < 0 \). If \( A < 0 \) then either \( \eta \) is complex or \( \eta < 0 \), again because \( B < 0 \). If \( A > 0 \), then \( A < 1 \) because \( \varphi > 0 \). Then if \( \eta \) is real, it satisfies \( \eta \geq -B/2A > -B/2 = 1 - \lambda_U \).
Combining these observations, we conclude that either \( \tilde{\lambda}_M \in [0, 1 - \lambda_U) \) and the information equilibrium has \( \lambda^*_M = \tilde{\lambda}_M \), or else the equilibrium occurs at \( \lambda^*_M = 0 \).

**Proof of Proposition 5.4.** Differentiation of \( \tilde{\lambda}_M \) with respect to \( \gamma \) yields

\[
\frac{d\tilde{\lambda}_M}{d\gamma} = -\frac{1}{\alpha \sqrt{\varphi}(1 - \lambda_U)} \left[ 1 + \frac{1 - \varphi}{\varphi} \frac{\gamma^2 \alpha}{(1 - \lambda_U)^2} \right]^{-1/2}.
\]

At an interior equilibrium, \( \tilde{\lambda}_M \) is real, so the expression on the right is real and negative.

At an interior equilibrium, \( \lambda^*_M \) is the solution to \( A\lambda^2 + B\lambda + C = 0 \), with the coefficients given by (A.13). Differentiating with respect to some parameter (e.g., \( \sigma_S \sigma_X \)) yields

\[
\dot{\lambda} = -\frac{\dot{A}\lambda^2 + \dot{C}}{2A\lambda + \dot{B}}.
\]

We note that \( 2A\lambda + \dot{B} < 0 \) can be rewritten \( (1 - \varphi)\lambda < 1 - \lambda_U \) which is always true because \( \varphi > 0 \) and \( \lambda_M < 1 - \lambda_U \) since \( \lambda_S > 0 \). Therefore, \( \text{sgn}(\lambda) = \text{sgn}(\dot{A}\lambda^2 + \dot{C}) \). Differentiating with respect to \( \sigma_M \sigma_X \) yields \( \dot{A} < 0 \) and \( \dot{C} < 0 \), which implies \( \dot{\lambda} < 0 \); differentiating with respect to \( \sigma_M \sigma_F \) yields \( \dot{A} > 0 \) and \( \dot{C} = 0 \), which implies \( \dot{\lambda} > 0 \).

**Proof of Proposition 5.5.** Using the expression for \( \rho^2_S \) in (30), we get

\[
\frac{f_S}{1 - f_S} \left( \frac{1}{\rho^2_S} - 1 \right) = \frac{\gamma^2(1 - f_S)\sigma^2_S\sigma^2_X}{\lambda^2_S}.
\]

This expression is strictly decreasing in \( f_S \), so \( J_S/J_U \) in (32) is strictly increasing in \( f_S \).

**Proof of Proposition 5.6.** We see from (31) that the derivative of \( J_M/J_U \) is negative precisely if the derivative of

\[
\frac{1 - f_M}{f_M} \frac{1}{1 - \rho^2_F} = \frac{1}{\gamma^2(1 - f_M)\sigma^2_M\sigma^2_X/\lambda^2_M + 1/f_M - 1}
\]

is negative, using the expression for \( \rho^2_F \) in (29). Differentiation yields

\[
\frac{1}{\gamma^2(1 - f_M)^2\sigma^2_M\sigma^2_X/\lambda^2_M} - \frac{1}{f^2_M} = \frac{1}{f_M} \left( \frac{1}{1 - \rho^2_F} - 1 \right) - \frac{1}{f^2_M},
\]

which is negative precisely if the first condition in (38) holds. The equivalence of the second condition in (38) follows from (29).

**A.4 Information equilibrium**

**Proof of Proposition 6.1.** We first show that (45) defines an information equilibrium at each \( c > 0 \), then verify uniqueness. For all three cases in (45), the specified \( \lambda_M, \lambda_S, \)
and \( \lambda_U \) are nonnegative and sum to 1, so it suffices to verify (39). For \( \zeta \leq c < \bar{c} \), we have \( J_M = J_S = J_U \) by construction, so the condition holds. For \( c \geq \bar{c} \), we again have \( J_S/J_U = 1 \) by construction. With \( \lambda_M = 0 \), we have \( \rho_s^2 = 0 \), and \( J_M/J_U \) in (31) evaluates to \( \exp(\gamma c)\sqrt{1-f_M} \geq \exp(\gamma \bar{c})\sqrt{1-f_M} = 1 \), so \( J_U/J_M \leq 1 \). Combining the two ratios we get \( J_S/J_M \leq 1 \). Thus, (39) holds.

For \( c < \zeta \), we consider two cases. First suppose case (i) of Proposition 5.3 holds at \( \zeta \). By definition, \( 1 - \lambda_M(\zeta) - \lambda_S(\zeta) = 0 \) and \( J_M/J_U = J_S/J_U \) at \( \lambda_M(\zeta), \lambda_S(\zeta), 0 \), so \( \lambda_M(\zeta) = \lambda_M^\ast(0) \) and \( \lambda_S(\zeta) = 1 - \lambda_M^\ast(0) \), by the definition of \( \lambda_M^\ast \). Because \( \lambda_M(c) \) and \( \lambda_S(c) \) are strictly decreasing in \( c \), they are strictly greater than \( \lambda_M^\ast(0) \) and \( 1 - \lambda_M^\ast(0) \). Decreasing \( \lambda_M \) decreases \( \rho_s^2 \), which decreases \( J_M/J_U \) in (31), and decreasing \( \lambda_S \) similarly decreases \( J_S/J_U \). By construction, \( J_M/J_U = J_S/J_U = 1 \) at \( (\lambda_M(c), \lambda_S(c), 1 - \lambda_M(c) - \lambda_S(c)) \), even for \( c < \zeta \), so at \( (\lambda_M^\ast(0), 1 - \lambda_M^\ast(0), 0) \) we have \( J_M/J_U < 1 \), \( J_S/J_U < 1 \), and \( J_M/J_S = 1 \), confirming (39).

Now suppose case (ii) of Proposition 5.3 holds at \( \zeta \); this includes the possibility that \( \bar{c} \leq \zeta \). Then \( \lambda_M^\ast(0) = \lambda_M(\zeta) = 0 \), and (15) specifies \( \lambda_M = 0 \) for all \( c < \zeta \). By the monotonicity argument used in case (i), \( J_S/J_U < 1 \) at all \( c < \zeta \). Moreover, Proposition 5.3(ii) entails \( J_S/J_M \leq 1 \), so this also holds for all \( c < \zeta \), and therefore (39) holds.

We now turn to uniqueness. At any \( c \), once we determine which proportions are strictly positive, the equilibrium is determined: if \( \lambda_U = 0 \), the other two proportions are determined by Proposition 5.3, if all three proportions are positive, they must satisfy \( J_M/J_U = J_S/J_U = 1 \) and must therefore be given by (42)–(43); if \( \lambda_M = 0 \) and \( \lambda_U > 0 \), the proportions are determined by the requirement that \( J_S/J_U = 1 \). We know from Proposition 5.3(iii) that \( \lambda_S > 0 \), so these are the only combinations we need to consider.

It therefore suffices to show that at any \( c \), the set of agents with positive proportions is uniquely determined. Suppose we try to introduce uninformed agents into an equilibrium from which they are absent. If we start with \( \lambda_M > 0 \) (and necessarily \( \lambda_S > 0 \)) then \( J_M/J_U \leq 1 \) and \( J_S/J_U \leq 1 \). Increasing \( \lambda_U \) requires decreasing either \( \lambda_M \) or \( \lambda_S \) and therefore decreasing either \( J_M/J_U \) or \( J_S/J_U \), precluding \( \lambda_U > 0 \), in light of (39). If \( \lambda_M = 0 \), the decrease must be in \( \lambda_S \) and the same argument applies. Suppose we try to introduce macro-informed agents into an equilibrium with only micro-informed and uninformed agents. The presence of uninformed agents requires \( J_M/J_U \geq 1 \). Increasing \( \lambda_M \) would increase \( J_M/J_U \), precluding \( \lambda_M > 0 \). Starting from an equilibrium with \( \lambda_S = 1 \) and increasing \( \lambda_M \) while leaving \( \lambda_U = 0 \) fixed is also infeasible because the value of \( \lambda_U \) determines the value of \( \lambda_M \) and \( \lambda_S \) through Proposition 5.3.

Proof of Corollary 6.1 (i) It suffices to consider the range \( \zeta \leq c \leq \bar{c} \) with \( \zeta < \bar{c} \), because \( \Pi_M \) is constant on \( (0, \zeta] \) and identically zero on \( [\bar{c}, \infty) \). It follows from (42) and (43) that

\[
\lambda_S^2(c) = \frac{\gamma^2(1 - f_S)^2\sigma^2\sigma_X^2}{f_S^2T^2M} \left( \frac{\lambda_M^2(c)f_M}{\gamma^2(1 - f_M)^2\sigma^2\sigma_X^2} + 1 \right) \equiv a\lambda_M^2(c) + b, \quad a, b > 0.
\]

Because \( \lambda_M(c) \) is strictly decreasing in \( c \), dividing both sides by \( \lambda_M^2(c) \) shows that \( \lambda_S^2(c)/\lambda_M^2(c) \) is strictly increasing in \( c \), hence \( \lambda_M(c)/(\lambda_M(c) + \lambda_S(c)) \) is strictly decreasing in \( c \). (ii) Follows from (15). (iii) We know from (29) and (30) that \( \rho_F^2 \) and \( \rho_S^2 \) are increasing in \( \lambda_M \) and
\[ \lambda_S, \text{ respectively}, \text{ so monotonicity of price efficiency follows from monotonicity in } (45). \]

### A.5 Variance ratios in Kacperczyk et al. (2016)

In this appendix, we briefly discuss the calculation of variance ratios in Kacperczyk et al. (2016), using notation from their paper, with references to relevant page numbers. The vector of (idiosyncratic and index) returns is given by

\[ \tilde{f} - r\tilde{p} = \Gamma^{-1}\mu + (I - B)z - Cx - A, \]

where \( z \) is a vector of independent factors, and \( x \) is a vector of independent supply shocks. The matrix \( I - B \) diagonal with entries \( \bar{\sigma}_i/\sigma_i \leq 1 \) (p.605), with

\[ \bar{\sigma}_i^{-1} = \sigma_i^{-1} + \bar{K}_i + \frac{\bar{K}_i^2}{\rho^2\sigma_x}. \]  

(A.14)

Here, \( \sigma_i \) is the prior variance of the payoff of security \( i \), \( \bar{\sigma}_i \) is the posterior variance, \( \sigma_x \) is a supply variance, \( \rho \) measures risk aversion, and \( \bar{K}_i \) is the total investor attention allocated to security \( i \). The cashflow variance in the return on the \( i \)th asset is therefore given by

\[ (\bar{\sigma}_i/\sigma_i)^2\text{var}[z_i] = \bar{\sigma}_i^2/\sigma_i. \]

The total variance of the return is (p. 606, equation 29),

\[ V_{ii} = \bar{\sigma}_i[1 + (\rho^2\sigma_x + \bar{K}_i)\sigma_i]. \]

The fraction of the total variance given by cashflow news is then

\[ \frac{\bar{\sigma}_i^2}{V_{ii}\sigma_i} = \frac{\bar{\sigma}_i[1 + (\rho^2\sigma_x + \bar{K}_i)\sigma_i]}{\bar{\sigma}_i[1 + (\rho^2\sigma_x + \bar{K}_i)\sigma_i] - \left(\frac{\sigma_i}{\bar{\sigma}_i} + (\rho^2\sigma_x + \bar{K}_i)\sigma_i\right)^{-1}.} \]

Using (A.14), this becomes

\[ \left(\sigma_i[\sigma_i^{-1} + \bar{K}_i + \frac{\bar{K}_i^2}{\rho^2\sigma_x}] + (\rho^2\sigma_x + \bar{K}_i)\sigma_i\right)^{-1}. \]  

(A.15)

As the proportion of investors who are informed increases, the total attention \( \bar{K}_i \) allocated to an asset does not decrease. (This property is similar to Proposition 1 in Kacperczyk et al. 2016, and we provide a proof in the Internet Appendix.) It follows that (A.15) is decreasing in the fraction of informed. But if the proportion of return variance due to cashflow variance is decreasing, then so is the ratio of cashflow variance to discount rate variance. Thus, VR decreases for all assets, in contrast to our results in Figure 6.

Figure 7 illustrates another contrast between our model and that of Kacperczyk et al. (2016). Recall that in our setting investors endogenously choose to specialize in micro or macro information, and, unless the cost of becoming informed is very high, we have both types of investors in equilibrium. In Kacperczyk et al. (2016), investors allocate all attention to the riskiest asset until they have sufficiently reduced its posterior variance, at which point they also start to allocate attention to the next riskiest asset.
With the calibration parameters of Table 1, the macro factor has greater variance than each micro factor, so investors initially learn only about the macro factor. In Figure 7, which uses our calibrated parameters and \( N = 100 \) stocks, this initially brings down the macro posterior variance (left) and the macro variance ratio (right) but leaves the corresponding micro quantities unchanged.

In the figure, the capacity on the horizontal axis is on an arbitrary scale of zero to 100. To interpret the figures, consider that the left panel shows \( 1 - f_M \) and \( 1 - f_S \). Thus, the figure says that \( f_S \) will remain at zero until investors have learned enough about the macro factor to increase \( f_M \) to about 0.76. A market with \( f_M \geq 0.76 \) or \( f_S \approx 0 \) would differ notably from the empirical results discussed in Section 7.1. The right panel of the figure similarly shows that the micro VR will remain very high until the macro VR has been reduced by a factor of almost 20.

![Figure 7: Ratios of posterior to prior variance (left) and ratios of cash flow variance to discount rate variance (right) as functions of information capacity in the model of Kacperczyk et al. (2016), using Table 1 parameters with \( \ell = 1 \) and \( N = 100 \) stocks.](image)

These patterns suggest that at many plausible parameters, the model of Kacperczyk et al. (2016) predicts that informed investors learn only about the macro factor, unless \( f_M \) is very high. This contrasts with the equilibrium in our model, in which we always find micro-informed investors and typically find macro-informed investors as well. The model of Kacperczyk et al. (2016) may be more descriptive of the attention allocation problem when several factors are of similar importance and have similar supply variances — factors representing industry sectors, for example, or value and growth factors, or stocks, bonds, and commodities. In our setting, a single market factor carries much greater risk than all other individual factors.

\[^{36}\text{Kacperczyk et al. (2016) assume that supply shocks for all factors have the same variance, but this restriction does not seem to be necessary for their results, so we use our calibrated values } \sigma_X^2, \text{ and } \sigma_X^2/N. \text{ We take the mean idiosyncratic supply shock to be zero. The capacity level at which investors begin allocating some attention to micro information does not depend on } N. \text{ See Internet Appendix for details.}\n
\[^{37}\text{We have truncated the vertical scale for legibility; at lower capacity levels the micro VR is roughly } 30,000. \text{ With our notation, the variance ratios at zero capacity are } 1/(\gamma^2\sigma_X^2\sigma_M^2) \text{ and } 1/(\gamma^2(\sigma_X^2/N^2)\sigma_M^2). \text{ The figure shows results for } \ell = 1. \text{ The macro-micro contrast is even greater with } \ell = 2.\]
References


1 Effect of $f_M$ and $f_S$ on attention equilibrium

Recall that for a fixed $\lambda_U$, the equilibrium $\lambda_M^*$ (proportion of macro-informed) is determined by the condition $J_M/J_U = J_S/J_U$, in the case of an interior equilibrium. Because $J_M/J_U$ does not depend on $f_S$, as $f_S$ rises and $J_S/J_U$ increases, $J_M/J_U$ can increase only if $\lambda_M$ increases (from Proposition 5.2). An interior equilibrium $\lambda_M^*$ must therefore increase with $f_S$. As the benefit of being micro-informed falls due to more precise micro information, the fraction of informed investors that focus on macro information grows. The left panel of Figure 4 from the paper demonstrates this adjustment. For every $\lambda_M$, a higher $f_S$ makes the micro-informed worse off, which pushes the equilibrium number of macro-informed higher.

Similarly, since $J_S/J_U$ does not depend on $f_M$ but decreases in $\lambda_M$ (i.e., the micro-informed are better off as there are fewer micro-informed), if $J_M/J_U$ decreases (increases) in $f_M$, then $\lambda_M^*$ must increase (decrease) in $f_M$. The right panel of Figure 4 from the paper illustrates this phenomenon. In the figure, the equilibrium $\lambda_M^*$ is sufficiently small so that macro precision makes the macro-informed better off. As macro precision $f_M$ increases, for a range of sufficiently small $\lambda_M$, the macro-informed become better off, which increases $\lambda_M^*$ (i.e., decreases the number of micro-informed, thus making the remaining micro-informed better off). Had the equilibrium $\lambda_M$ been sufficiently high, the effect would have had the opposite sign.

*Glasserman: Columbia Business School, pg20@columbia.edu. Mamaysky: Columbia Business School, hm2646@columbia.edu.
The preceding arguments establish the following result:

**Proposition 1.1** (Effect of information precision on equilibrium). *In the case of an interior equilibrium with $\lambda^*_M > 0$, the number of macro-informed increases as the micro signal becomes more precise:*

$$\frac{d\lambda^*_M}{df_S} > 0.$$  

*Condition [38] is necessary and sufficient for the number of macro-informed to increase as the macro signal becomes more precise:*

$$\frac{d\lambda^*_M}{df_M} > 0 \ (< 0) \text{ if and only if } \rho_F^2 < \frac{1}{1 + f_M} \left( > \frac{1}{1 + f_M} \right).$$

## 2 Calibration

### 2.1 Supply shocks and turnover

In our model, equal weighted index turnover is given by $1/N \sum_{i=1}^{N} |X_F - \bar{X}_F + X_i|$. Using standard results, we get

$$\frac{1}{N} \sum_{i=1}^{N} |X_F - \bar{X}_F + X_i| = \sqrt{\frac{2}{\pi}} \times \sqrt{\sigma^2_{X_F} + \sigma^2_X}.$$  

Direct turnover for stock $i$ is

$$\frac{1}{N} |X_F - \bar{X}_F + X_i|$$

and direct index turnover is

$$\frac{1}{N} \sum_{j=1}^{N} |X_F - \bar{X}_F + X_j|.$$  

The conditions we want to satisfy are

$$E \left[ \frac{1}{N} \sum_{j=1}^{N} |X_F - \bar{X}_F + X_j| \right] = 0.76$$

and

$$\rho^2(|X_F - \bar{X}_F + X_i|, \sum_{j=1}^{N} |X_F - \bar{X}_F + X_j|) = 0.47.$$
Writing \( Y_j = X_F - \bar{X}_F + X_j \), we can express these conditions as

\[
E|Y_j| = 0.76
\]

and

\[
\rho^2(|Y_i|, \sum_j |Y_j|) = 0.47. \tag{1}
\]

As \( \text{var}[Y_j] = \sigma_{X_F}^2 + \sigma_X^2 \), the first of these conditions gives

\[
\sqrt{\frac{2}{\pi}} \sqrt{\sigma_{X_F}^2 + \sigma_X^2} = 0.76 \Rightarrow \sigma_{X_F}^2 + \sigma_X^2 = 0.9073. \tag{2}
\]

For the second equation, we have

\[
\rho(|\sum_j |Y_j||, \sum_i |Y_i|) = \frac{\text{cov}[\sum_j |Y_j||, \sum_i |Y_i|]}{\sigma(\sum_j |Y_j|)\sigma(|Y_i|)}
\]

\[
= \frac{\text{cov}[\sum_{j \neq i} |Y_j||, |Y_i|] + \text{var}[|Y_i|]}{\sigma(\sum_j |Y_j|)\sigma(|Y_i|)}
\]

\[
= \frac{(N - 1)\rho(|Y_i||, |Y_j|)\sigma^2(|Y_i|) + \sigma^2(|Y_i|)}{\sigma(\sum_j |Y_j|)\sigma(|Y_i|)}
\]

\[
= \frac{(N - 1)\rho(|Y_i||, |Y_j|)\sigma^2(|Y_i|) + \sigma^2(|Y_i|)}{(N + (N - 1)\rho(|Y_i||, |Y_j|))^{1/2}}
\]

\[
\to \rho(|Y_i||, |Y_j|)^{1/2}, \quad \text{as } N \to \infty.
\]

We think of the number of stocks \( N \) as large, and passing to the limit simplifies the calculation. With this simplification, \( [1] \) becomes

\[
\rho(|Y_i||, |Y_j|) = 0.47.
\]

From formula (1.1.0) in Kamat (1958), p.26, we find that if \( (\xi_1, \xi_2) \) are bivariate normal with correlation \( \rho \), then

\[
\rho(|\xi_1||, |\xi_2|) = \frac{2}{\pi - 2} (\rho \arcsin(\rho) + \sqrt{1 - \rho^2} - 1) \equiv \text{Kamat}(\rho).
\]

Thus, we need

\[
\rho(Y_i, Y_j) = \text{Kamat}^{-1}(0.47) = 0.714452.
\]

3
But

\[ \rho(Y_i, Y_j) = \frac{\sigma_{Y_i}^2 \sigma_{Y_j}^2}{\sigma_{Y_i}^2 + \sigma_Y^2}. \]  

(3)

Combining this with (2) we get

\[ \sigma_{X_F}^2 = 0.714452 \times 0.9073 = 0.6482, \quad \sigma_X^2 = 0.9073 - 0.6482 = 0.2591, \]

and then

\[ \sigma_{X_F} = 0.8051, \quad \sigma_X = 0.5090. \]

In this derivation, we have ignored the correlation \(-1/(N-1)\) between idiosyncratic supply shocks \(X_i\) and \(X_j\). Those correlations vanish if we take \(N \to \infty\) (as we do above) and are negligible for large but finite \(N\).

### 2.2 ETF and futures trading volume

Table 1 shows the annualized trading volume of the top ETFs and equity index futures on a representative trading day.

<table>
<thead>
<tr>
<th>Name</th>
<th>Share Volume (daily)</th>
<th>Share Price</th>
<th>$ Traded per Year (trlns)</th>
<th>Name</th>
<th>Volume</th>
<th>Multiplier</th>
<th>Price</th>
<th>$ Annual Volume (trlns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPDR S&amp;P 500 ETF TRUST</td>
<td>70,441,296</td>
<td>278.76</td>
<td>4.95</td>
<td>S&amp;P 500 FUTURE Sep18</td>
<td>2,057</td>
<td>250</td>
<td>2,791.00</td>
<td>0.36</td>
</tr>
<tr>
<td>INVESCO QQQ TRUST SERIES 1</td>
<td>30,667,280</td>
<td>176.33</td>
<td>1.36</td>
<td>S&amp;P 500 FUTURE Dec18</td>
<td>102</td>
<td>250</td>
<td>2,797.20</td>
<td>0.02</td>
</tr>
<tr>
<td>iSHARES RUSSELL 2000 ETF</td>
<td>17,498,722</td>
<td>166.99</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iSHARES MSCI EMERGING MARKET</td>
<td>64,918,536</td>
<td>45.95</td>
<td>0.75</td>
<td>S&amp;P500 EMINI FUT Sep18</td>
<td>106,000</td>
<td>50</td>
<td>2,791.50</td>
<td>6.65</td>
</tr>
<tr>
<td>FINANCIAL SELECT SECTOR SPDR</td>
<td>49,738,832</td>
<td>28.07</td>
<td>0.35</td>
<td>S&amp;P500 EMINI FUT Dec18</td>
<td>1,066</td>
<td>50</td>
<td>2,794.75</td>
<td>0.04</td>
</tr>
<tr>
<td>iSHARES MSCI EAFE ETF</td>
<td>21,618,260</td>
<td>70.43</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPATH S&amp;P 500 VIX S/T FU ETN</td>
<td>28,464,014</td>
<td>31.84</td>
<td>0.23</td>
<td>DJIA MINI e-CBOT Sep18</td>
<td>15,982</td>
<td>5</td>
<td>25,316.00</td>
<td>0.51</td>
</tr>
<tr>
<td>SPDR DJIA TRUST</td>
<td>3,857,796</td>
<td>253.06</td>
<td>0.25</td>
<td>DJIA MINI e-CBOT Dec18</td>
<td>13</td>
<td>5</td>
<td>25,332.00</td>
<td>0.00</td>
</tr>
<tr>
<td>iSHARES CHINA LARGE-CAP ETF</td>
<td>18,915,848</td>
<td>47.39</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iSHARES CORE S&amp;P 500 ETF</td>
<td>3,017,272</td>
<td>280.73</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iSHARES IBOXX USD HIGH YIELD</td>
<td>13,330,054</td>
<td>85.91</td>
<td>0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TECHNOLOGY SELECT SECT SPDR</td>
<td>12,426,255</td>
<td>72.07</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROSHARES ULTRAPRO QQQ</td>
<td>12,553,372</td>
<td>62.10</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENERGY SELECT SECTOR SPDR</td>
<td>15,883,723</td>
<td>76.47</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INDUSTRIAL SELECT SECT SPDR</td>
<td>10,715,862</td>
<td>76.04</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: On June 13, 2018 we look at the top 15 most heavily traded ETFs (according to ETF.com), as well as the first two S&P500, S&P500 e-mini and Dow futures. The average daily trading volume over the prior 30 days for these 21 instruments represented an annualized dollar turnover of $18.25 trillion. Compared to the market capitalization of the Russell 3000 index of $30.49 trillion, this represents an annual turnover of 60%.
2.3 $M$ as present value of dividends

Say monthly dividends follow

$$D_{t+1} = \mu_D + \rho D_t + \epsilon_{t+1},$$

with $\mu_D = 1 - \rho$ so the long run dividend level is 1 and $\text{var}(\epsilon) = \sigma^2$. Let us define the $N$ period dividend at time $t + 1$ as

$$M \equiv D_{t+12} + \beta D_{t+13} + \cdots + \beta^{N-1} D_{t+1+N}. \tag{4}$$

We think of this as the present value of the future $N$ dividends starting at time $t+12$. We also assume that the first dividend follows the process $D_{12} = \mu_D + \rho D_0 + \epsilon_{12}$. If $D_0 = 1$, we can show that

$$E_t[M] = \frac{1 - \beta^N}{1 - \beta}.$$

The variance is given by

$$\text{var}_t(M) = \frac{\sigma^2}{(1 - \beta \rho)^2} \left( \frac{1 - \beta^{2N}}{1 - \beta^2} - 2(\beta \rho)^N \frac{1 - (\beta/\rho)^N}{1 - \beta/\rho} + (\beta \rho)^{2N} \frac{1 - (1/\rho)^{2N}}{1 - (1/\rho)^2} \right).$$

Since we are normalizing the dividend level to 1, we are interested in the volatility per unit of expected dividend as $N$ grows. Figure 1 shows the expectation, variance, volatility and normalized volatility

$$\text{volatility of normalized dividend} = \frac{\sqrt{\text{var}_t(M)}}{E_t[M]}$$

as a function of $N$.

2.4 Calibration of $f_M$

Tables 2 and 3 show results of regressing

$$\overline{CF}[t, t+x] - CF[t]$$

on day $t$ explanatory variables where $CF[t]$ is either last twelve month earnings or dividends of the S&P500 index on day $t$ and $\overline{CF}[t, t+x]$ is the average $CF[t]$ over years $t+1, t+2, \cdots, t+x$. The regressions are run with overlapping daily data starting in
January 31, 1990. Data are obtained from Bloomberg.

2.5 Properties of equilibrium

Figures 2 and 4 show equilibrium quantities for the model calibration in Section 7.1 of the paper for $\ell = 1, 2$ respectively. Figure 5 shows the number of macro and micro informed investors for the $\ell = 1, 2$ versions of the model calibration.

2.6 Trend in realized S&P500 volatility

Figure 6 shows the time trend in annual S&P500 volatility, using daily overlapping observations. Data are obtained from Bloomberg. The Newey-West t-statistic uses automatic lag selection.

3 Proof that attention in Kacperczyk et al. (2016) is weakly increasing in number of informed

We use the following notation, based on Kacperczyk et al. (2016):

$$\bar{\sigma}_i = \bar{\sigma}_i(\chi, k) = \left( \frac{1}{\sigma_i} + \chi k + \frac{\chi^2 k^2}{\rho^2 \sigma_x} \right)^{-1}$$

$$\lambda_i = \lambda_i(\chi, k) = \bar{\sigma}_i[1 + \left( \rho^2 \sigma_x + \chi k \right) \bar{\sigma}_i] + \rho^2 \bar{x}_i^2 \bar{\sigma}_i^2$$

The “true” $\lambda$ functions depend on $\chi$ and $k$ only through their product, but it will be useful to keep these arguments separate. We know from Kacperczyk et al. (2016) that $\lambda_i$ is strictly decreasing in $\chi$ if $k > 0$ and strictly decreasing in $k$ if $\chi > 0$.

Hold the total capacity $K$ fixed and vary $\chi$. Let $\{K_i(\chi)\}$ be equilibrium attention allocations at $\chi$.

**Proposition 3.1.** Each $\chi K_i(\chi)$ is increasing in $\chi$. Consequently, each variance ratio is decreasing in $\chi$.

**Proof.** The second statement follows from the first using (A.15), which shows that variance ratios are decreasing in $K_i$.

Write $M_\chi$ for the set of assets that attain the maximal $\lambda$: $i \in M_\chi$ iff $\lambda_i(\chi, K_i(\chi)) = \max_k \lambda_k(\chi, K_k(\chi))$. Let $0 < x < y$ be two values of $\chi$. To argue by contradiction, suppose
that for some asset \( i \),

\[ xK_i(x) > yK_i(y). \tag{5} \]

This requires \( K_i(x) > 0 \), which implies \( i \in M_x \). By the strict monotonicity of \( \lambda_i \), (5) implies \( \lambda_i(x, K_i(x)) < \lambda_i(y, K_i(y)) \). But then

\[
\begin{align*}
\lambda_j(x, K_j(x)) & \leq \max_k \lambda_k(x, K_k(x)) \\
& = \lambda_i(x, K_i(x)) \\
& < \lambda_i(y, K_i(y)) \\
& \leq \max_k \lambda_k(y, K_k(y)) = \lambda_j(y, K_j(y)), \quad \text{for all } j \in M_y.
\end{align*}
\]

By strict monotonicity of \( \lambda \), this implies that

\[ xK_j(x) > yK_j(y), \quad \text{for all } j \in M_y. \]

In equilibrium, the full capacity \( K \) is allocated, so

\[ xK = x \sum_k K_k(x) \geq x \sum_{j \in M_y} K_j(x) > y \sum_{j \in M_y} K_j(y) = yK, \]

which contradicts the assumption that \( x < y \). \( \square \)

4 Graphing variance ratios in the model of Kacperczyk et al. (2016)

In order to graph variance ratios in the model of Kacperczyk et al. (2016), as we do in Figure 7, we need to solve the attention allocation problem in their paper, which we refer to as KVNV.

From KVNV (14) and p.605, item 4, we have, at \( \bar{K}_i = k \),

\[ \lambda_i(k) = \bar{\sigma}_i(k)[1 + (\rho^2 \sigma_x + k)\bar{\sigma}_i(k)] + \rho^2 \bar{x}_i^2 \sigma_i^2(k), \]

and

\[ \bar{\sigma}_i(k) = \left( \sigma_i^{-1} + k + \frac{k^2}{\rho^2 \sigma_x} \right)^{-1}. \]
We want to apply these expressions with $i = M$ or $i = S$, with
\[ \sigma_i = \sigma^2_M, \sigma^2_S; \quad \rho = \gamma; \quad \sigma_x = \sigma^2_{X_F}, \sigma^2_{X_i}; \quad \bar{x}_i = \bar{X}_F, \bar{X}_S = 0. \]

**Solution**

If we assume a symmetric equilibrium, then in KVNV’s (12)–(14) we can take $K = \bar{K}$ and $K_i = \bar{K}_i$ and drop the $j$s. In other words, we are allocating the economy-wide capacity $\bar{K}$.

As we increase the capacity $\bar{K}$, we proceed as follows.

1. At zero capacity, we have $\bar{\sigma}_i(0) = \sigma_i$, and, at our calibration
   \[ \lambda_M(0) = \sigma^2_M[1 + \gamma^2 \sigma^2_{X_F} \sigma^2_M] + \gamma^2 \bar{X}_F^2 \sigma^2_M > \lambda_S(0) = \sigma^2_S[1 + \gamma^2 \sigma^2_{X_i} \sigma^2_S]. \]
   Thus, all capacity is initially allocated to $M$.

2. This continues until we reach the point $k^*$ at which $\lambda_M(k^*) = \lambda_S(0)$: for $\bar{K} \in [0, k^*)$, the allocation is $\bar{K}_M = \bar{K}$ and $\bar{K}_S = 0$.

3. For $\bar{K} \geq k^*$, we will allocate capacity to $M$ and to all $N$ stocks in order to make all $\lambda$s equal. Because all stocks have the same parameters, they will get the same allocation. So, for each $\bar{K} \geq k^*$, we need to find $k \leq \bar{K}$ such that
   \[ \lambda_M(\bar{K} - k) = \lambda_S(k/N), \]
   which allocates $\bar{K}_S = k/N$ to each stock and $\bar{K}_M = \bar{K} - k$ to $M$.

Once we have the optimal allocations $\bar{K}_M$ and $\bar{K}_S$, we can evaluate $\lambda_M(\bar{K}_M)$ and $\lambda_S(\bar{K}_S)$ and the posterior variances $\bar{\sigma}_i(\bar{K}_i)$, and from these we can calculate $VR_M$ and $VR_S$.

**References**


Figure 1: Expectation, variance, volatility and volatility per unit expectation of the $N$ month discounted dividend process given in equation (4) with starting dividend $D_0 = 1.$
Equilibrium summary $\gamma = 5.5$, $\sigma_{M}^2 = 0.01323$, $\sigma_{S}^2 = 0.004408$, $f_{M} = 0.47$, $f_{S} = 0.2$, $\sigma_{Xr}^2 = 0.6482$, $\sigma_{X}^2 = 0.2591$

Figure 2: Properties of equilibrium with $\ell = 1$. 
Figure 3: Properties of equilibrium with $\ell = 2$.  

Equilibrium summary $\gamma = 5.5$, $\sigma_M^2 = 0.01323$, $\sigma_S^2 = 0.004408$, $f_M = 0.47$, $f_S = 0.2$, $\sigma_{X_F}^2 = 2.593$, $\sigma_X^2 = 0.2591$
Equilibrium summary $\gamma = 5.5$, $\sigma_M^2 = 0.01323$, $\sigma_S^2 = 0.004408$, $f_M = 0.47$, $f_S = 0.2$, $\sigma_{X_F}^2 = 5.834$, $\sigma_X^2 = 0.2591$

Figure 4: Properties of equilibrium with $\ell = 3$. 
Figure 5: Number of informed in equilibrium.
Figure 6: S&P500 volatility, daily observations of lagged annual volatility. The time trend coefficient is -3.4 basis points per day. The Newey-West t-statistic for the time trend uses auto lag selection. Volatility observations range from January 3, 1929 to July 16, 2018.
Table 2: Earnings regression.

<table>
<thead>
<tr>
<th></th>
<th>Forward average EPS</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1 yr)</td>
<td>(2 yrs)</td>
<td>(3 yrs)</td>
<td>(4 yrs)</td>
</tr>
<tr>
<td>E₁dEPS</td>
<td>1.837**</td>
<td>1.688</td>
<td>0.976</td>
<td>0.870</td>
<td>0.968</td>
</tr>
<tr>
<td></td>
<td>(0.901)</td>
<td>(1.053)</td>
<td>(1.060)</td>
<td>(1.154)</td>
<td>(1.230)</td>
</tr>
<tr>
<td>E₂dEPS</td>
<td>−1.067</td>
<td>−0.911</td>
<td>−0.131</td>
<td>0.007</td>
<td>−0.122</td>
</tr>
<tr>
<td></td>
<td>(0.709)</td>
<td>(0.799)</td>
<td>(0.852)</td>
<td>(0.965)</td>
<td>(1.056)</td>
</tr>
<tr>
<td>dEPS₁l₁₂m</td>
<td>0.317</td>
<td>0.374**</td>
<td>0.380**</td>
<td>0.374**</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td>(0.205)</td>
<td>(0.157)</td>
<td>(0.156)</td>
<td>(0.180)</td>
<td>(0.236)</td>
</tr>
<tr>
<td>Price_Book</td>
<td>−0.001</td>
<td>−0.004</td>
<td>−0.007*</td>
<td>−0.008**</td>
<td>−0.006**</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>EBIT_Margin</td>
<td>−0.004</td>
<td>−0.007</td>
<td>−0.008</td>
<td>−0.010*</td>
<td>−0.013**</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.060</td>
<td>0.113</td>
<td>0.116</td>
<td>0.146*</td>
<td>0.184***</td>
</tr>
<tr>
<td></td>
<td>(0.092)</td>
<td>(0.098)</td>
<td>(0.093)</td>
<td>(0.076)</td>
<td>(0.068)</td>
</tr>
<tr>
<td>Observations</td>
<td>6,577</td>
<td>6,326</td>
<td>6,075</td>
<td>5,824</td>
<td>5,573</td>
</tr>
<tr>
<td>R²</td>
<td>0.303</td>
<td>0.373</td>
<td>0.472</td>
<td>0.583</td>
<td>0.689</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.302</td>
<td>0.372</td>
<td>0.471</td>
<td>0.582</td>
<td>0.688</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01
### Table 3: Dividend regressions.

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable: Forward average dividend per share</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1 yr)</td>
</tr>
<tr>
<td>E_dEPS</td>
<td>0.225*</td>
</tr>
<tr>
<td></td>
<td>(0.128)</td>
</tr>
<tr>
<td>E1_dEPS</td>
<td>−0.089</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
</tr>
<tr>
<td>dEPS_l12m</td>
<td>0.050**</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
</tr>
<tr>
<td>Price_Book</td>
<td>−0.002***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>EBIT_Margin</td>
<td>0.002***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.012</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
</tr>
<tr>
<td>Observations</td>
<td>6,577</td>
</tr>
<tr>
<td>R²</td>
<td>0.639</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.639</td>
</tr>
</tbody>
</table>

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