On the effects of restricting short-term investment*

Nicolas Crouzet, Ian Dew-Becker, and Charles G. Nathanson

February 19, 2019

Abstract

We study the effects of policies proposed for addressing “short-termism” in financial markets. We examine a noisy rational expectations model in which investors’ exposures and information about fundamentals endogenously vary across horizons. In this environment, taxing or outlawing short-term investment has no negative effect on the information in prices about long-term fundamentals. However, such a policy reduces the profits and utility of short- and long-term investors. Changing policies on the release of short-term information can help long-term investors – an objective of some policymakers – at the expense of short-term investors, but it also makes prices less informative and increases costs of speculation.

For decades economists and policymakers have expressed concern about the potentially negative effects of “short-termism” among investors in financial markets. Research has argued that short-term investors may increase the volatility and reduce the informativeness of asset prices (Froot, Scharfstein, and Stein (1992)), exacerbate fire sales and crashes (Cella, Ellul, and Giannetti (2013)), inefficiently incentivize managers to focus on short-term projects (Shleifer and Vishny (1990)), or reduce incentives of other investors to acquire information (Baldauf and Mollner (2017); Weller (2017)), making prices less informative overall.

Those who take the view that short-termism is bad for financial markets or the economy as a whole have proposed a broad array of policies to encourage long-term investment. One of the oldest

---

*Crouzet: Northwestern University. Dew-Becker: Northwestern University and NBER. Nathanson: Northwestern University. An earlier version of this paper was circulated under the title “Multi-frequency trade”. We appreciate helpful comments from the editor, Stijn van Nieuwerburgh, two anonymous referees, and Hengjie Ai, Bradyn Breon-Drish, Alex Chinco, Cecilia Parlatore, Ioanid Rosu, Laura Veldkamp, and seminar participants at Northwestern, Minnesota Macro Asset Pricing, Adam Smith, the Finance Cavalcade, Red Rock, Yale SOM, and FIRS.*
is the tax on transactions of Tobin (1978). Some policies directly depend on holding periods, such as US tax treatment of capital gains and dividends, the SEC’s most recent proxy access rules, the proposed Long-Term Stock Exchange, linking corporate voting rights to tenure, and the proposal of Bolton and Samama (2013) for corporations to explicitly reward long-term investors. Budish, Cramton, and Shim (2015) propose to eliminate trade at the very highest frequencies through frequent batch auctions, and there have also been proposals to limit quarterly financial reports and earnings guidance in the US, following similar changes in the UK, e.g. by Dimon and Buffett (2018). A number of these policies were endorsed in a letter from 2009 signed by leaders in business, finance, and law.

This paper theoretically evaluates the effect of policies targeting short-termism on price informativeness and investor outcomes. Unlike the previous literature, we consider a simple and very general setting with investors who are ex ante identical and then may endogenously specialize into different horizons. While there is some recent work on the consequences of various limits on information gathering ability and there have been empirical analyses of high-frequency traders, we are not aware of any other work that directly studies the effects of restrictions on short- and long-term strategies on price informativeness and investor profits in a general setting.

The model is meant to be as simple and as general as possible. Two key features it must have are that investors choose among investment strategies at different horizons, and that they choose how much information to acquire about fundamentals across horizons. We study a version of the noisy rational expectations model developed in Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016). Whereas that paper studies investment in a cross-section of assets, we argue here that investment policies over time can be thought of as a choice of exposures on many different future dates. Each of those dates represents a different “asset”, and the returns on those assets will be correlated across dates. The model in this paper is notable for allowing an arbitrarily long horizon (as opposed to...
two or three periods), with turnover at any frequency.

It is important to note that the model is not fully dynamic – all trade happens on date 0, so investors cannot rebalance in response to news or the realization of fundamentals, even though they might desire to. The model takes a dynamic problem, with information flowing and investment choices being made over time, and compresses it into a single time period, along the lines of the classic Arrow–Debreu type analysis, but without a complete set of state-contingent contracts. Dynamic market equilibria are difficult or impossible to solve, and we do not contribute to that area.\(^7\) The paper’s focus is instead on the choice of short- versus long-term investment strategies and information acquisition. Short-term investors arise naturally in the model as agents whose exposures to fundamentals fluctuate rapidly across dates due to the type of information they have acquired. The relevant concept of short- versus long-term here ranges between days and years – the model is not designed to analyze technical features of higher frequency trading, like market microstructure effects or exchange fragmentation.

There are at least three potential reasons a policymaker might want to regulate investment strategies. First, if price informativeness at long horizons is more important for economic decisions like physical investment, then long-term information acquisition might be encouraged. Second, policymakers might have a general bias toward long-term investors, perhaps because they are more likely to be people saving for retirement. Finally, policymakers might want to limit the losses of retail investors who make poor investment decisions. As is common in the literature on price efficiency, we do not explicitly model these regulatory motives.\(^8\) Instead, we use the model to examine how restrictions on investment policies affect price informativeness and the profits and utility of the various investors in order to help inform the policy debate.

\(^7\) Work on the infinite regress problem typically assumes that investors have only single-period objectives and does not for a choice of information across horizons. See Makarov and Rytochkov (2012), Kasa, Walker, and Whiteman (2013), and Rondina and Walker (2017). Recent work also examines dynamic models with strategic trade (with similar restrictions regarding horizons) – see Vayanos (1999, 2001), Ostrovsky (2012), Banerjee and Breen-Drisch (2016), Foucault, Hombert, and Rosu (2016), Du and Zhu (2017), and Dugast and Foucault (2017).

\(^8\) Bond, Edmans, and Goldstein (2012) review he literature on the value of price efficiency and identify two spillovers. First, information that stock prices reveal guides real activity through investment decisions (Dow and Gorton (1997); Kurlat and Veldkamp (2015)) and the decisions of outside investors and regulators to intervene in a firm’s activities (Bond, Goldstein, and Prescott (2009); Bond and Goldstein (2015)). Second, price informativeness allows shareholders to tie manager compensation to equity prices, thus improving the real efficiency of management activities (Fishman and Hagerty (1989); Holmström and Tirole (1993); Farboodi and Veldkamp (2017)).
The paper examines a number of specific policies, including direct restrictions on investment strategies, taxes on transactions, and taxing or subsidizing information acquisition. We first show that when sophisticated agents are restricted from investing and trading at some frequency, prices become uninformative at that frequency. So if a policy were implemented saying that investors could no longer maintain positions for less than a month, variation in prices at frequencies less than a month would become uninformative for fundamentals, and instead be driven purely by liquidity demand. High-frequency price volatility and mean reversion would also rise.

However, there is no spillover across horizons. A short-term restriction or transaction tax does not reduce price informativeness or increase return volatility at longer horizons, so prices would remain informative at frequencies lower than a month (in an extension of the model, informativeness can even rise). This separability across horizons follows from a statistical result showing that there is a robust independence across frequencies in stationary models, along with a separability in mean-variance (or constant absolute risk aversion) preferences.

The next question is how investment restrictions affect investor outcomes. An increase in short-term investment (e.g. due to a change in technology that makes short-term information acquisition or trading cheaper) turns out to make long-term investors worse off, essentially taking away some of the long-term investors’ trading opportunities. But restricting short-term investment does not transfer profits back to long-term investors; instead it simply eliminates those profits, making both short- and long-term investors worse off.

In the context of the model, the way to tilt markets in favor of long-term investors – if that is one’s goal – is to make acquisition of short-term information more expensive for investors. There have been numerous recent proposals to do just that, for example by limiting quarterly earnings guidance. The model in this paper is well suited to analyze such policies, and we show that they can shift the equilibrium toward long-term investors, increasing their average profits and utility (though the direction of this result depends on how one models information releases).

Finally, the paper examines the impact of the various policies on the profits of noise traders. Intuitively, the noise traders are constantly making mistakes, potentially affecting prices. There are two ways to protect them from those mistakes: stop them from trading, or reduce the losses they take on each trade. Stopping them from trading is in principle simple – just close asset
markets – but then one loses the information contained in prices, along with any gains from trade.

More interestingly, the paper shows that a better alternative is to subsidize or otherwise encourage information acquisition, which causes prices to become more informative and less responsive to noise trader (perhaps speculative) demand. Such a policy can, in the limit, drive noise trader losses to zero, while simultaneously making prices more useful for economic decisions and reducing the excess volatility caused by noise trader speculation. However, it is the opposite of the policy that we showed helps the long-term investors. Furthermore, it is important to temper the results on noise traders with the knowledge that there is no single canonical model of noise traders. The paper examines robustness to an alternative formulation driven by time-varying hedging demand and shows that welfare predictions are more difficult to make, though the predictions for price informativeness and return volatility are similar.

Overall, then, we obtain three main results about policies aimed at short-termism:


2. Restricting short-term investment hurts both short- and long-term investors, but helps noise traders.

3. Taxing or restricting the availability of short-term information helps long-term investors, hurts short-term investors and noise traders, and reduces short-term price efficiency. Subsidizing information or mandating greater disclosure by firms does the opposite.

On net, then, we would argue that mandatory information releases or subsidizing information acquisition are the most natural policies to address short-termism, as they both reduce speculative effects on prices and improve price efficiency. They do, however, come with costs to long-term investors, and also run against recent proposals to reduce quarterly reporting.

The answers to the questions of how restrictions on trade affect price informativeness and welfare are not obvious ex ante. One view is that there might be some sort of separation across frequencies, so that restrictions in one realm do not affect outcomes in another. On the other hand, investors obviously interact – they trade with each other – so it would be surprising if policies targeting a particular type of investor did not act to benefit others. What we find is a mix of the two: market characteristics at high frequencies can affect the profits and utility of long-term
investors – the model is not entirely separable across frequencies in that sense – but they do not affect low-frequency price informativeness in our baseline case. Furthermore, there is a tension between helping long-term investors, helping noise traders, and maintaining price informativeness. No single policy helps all the groups at the same time due to a zero-sum aspect of the model, and policies that may be attractive to certain investors can come with negative side effects for agents outside the model – e.g. executives, or policymakers like the FOMC – who might make decisions based on asset prices.

The remainder of the paper is organized as follows. Sections 1 and 2 lay out the model and its solution. Section 3 examines the effects for price volatility and informativeness of restrictions on investment at different horizons, while section 4 examines the impacts of such policies on the profits and welfare of different investors. Section 4 also examines the impact of restrictions on information releases such as earnings announcements. Section 5 discusses the robustness of our results to a number of alternative assumptions, and section 6 concludes.

1 The model

1.1 Market structure

Time is denoted by $t \in \{-1, 0, 1, \ldots, T\}$, with $T$ even and large. There is a fundamentals process $D_t$, on which investors trade forward contracts, with realizations on all dates except $-1$ and $0$. The time series is stacked into a vector $D \equiv [D_1, D_2, \ldots, D_T]'$ (versions of variables without time subscripts denote vectors) and is unconditionally distributed as

$$D \sim N(0, \Sigma_D). \quad (1)$$

For our benchmark results, we focus on the case where fundamentals are stationary. Stationarity implies that $\Sigma_D$ is constant along its diagonals, and we further assume that the eigenvalues of $\Sigma_D$ are finite and bounded away from zero (which is satisfied by standard ARMA processes).

There is a set of futures claims on realizations of the fundamental. When we say that the model features a choice of investment across dates or horizons, we mean that investors will choose portfolio allocations across the futures contracts, which then yield exposures to the realization of
fundamentals on different dates in the future.

A concrete example of a process $D_t$ is the price of crude oil: oil prices follow some stochastic process and investors trade futures on oil at many maturities. $D_t$ could also be the dividend on a stock, in which case the futures would be claims on dividends on individual dates. The analysis of futures is an abstraction for the sake of the theory, though we note that dividend futures are in fact traded (Binsbergen and Koijen (2017)). While the concept of a futures market on the fundamentals will be a useful analytic tool, we will also price portfolios of futures. Equity, for example, is a claim to the stream of fundamentals over time. Holding any given combination of futures claims on the fundamental is equivalent to holding futures contracts on equity claims.

The model does not allow for stochastic volatility, nonlinearity, or other changes in the higher moments of $D_t$ over time. The fact that we study the level of fundamentals, rather than their log, is also a restriction shared by CARA–Normal specifications (e.g. Grossman and Stiglitz (1980)). The restrictions, along with those implicit in the preferences below, mean that the model is useful primarily for qualitative analysis.

1.2 Information structure

There is a unit mass of “sophisticated” or rational investors, indexed by $i \in [0, 1]$, who have rational expectations, conditioning on both prices and private signals. The signals an agent observes are a collection $\{Y_{i,t}\}_{t=1}^{T}$ observed on date 0 with

$$Y_{i,t} = D_t + \varepsilon_{i,t}, \varepsilon_i \sim N(0, \Sigma_i),$$

where $\Sigma_i^{-1}$ is investor $i$’s signal precision matrix (which will be chosen endogenously). Through $Y_{i,t}$, investors can learn about fundamentals on all dates between 1 and $T$. $\varepsilon_{i,t}$ is a stationary error process in the sense that $\text{Cov}(\varepsilon_{i,t}, \varepsilon_{i,t+j})$ depends on $j$ but not $t$. Since $\text{Var}(\varepsilon_{i,t})$ is the same for all $t$, all dates are equally difficult to learn about and no particular date is given special prominence in the model. Investors must choose an information policy that treats all dates symmetrically.

The signal structure generates one of our desired model features, which is that investors can

---

9The model could accommodate predictable changes in volatility, such as intra-day patterns and volatility around scheduled announcements, through time-change methods as in Ané and Geman (2000), and Geman, Madan, and Yor (2001).
choose to learn about fundamentals across different dates in the future. When the errors are positively correlated across dates, the signals are relatively less useful for forecasting trends in fundamentals since the errors also have persistent trends. Conversely, when errors are negatively correlated across dates, the signals are less useful for forecasting transitory variation and provide more accurate information about moving averages. What types of fluctuations investors are informed about will determine their investment strategies.

1.3 Investment objective

On date 0, there is a market for forward claims on fundamentals on all dates in the future. Investor i’s demand for a date-\( t \) forward conditional on the set of prices and signals is denoted \( Q_{i,t} \). Investors have mean-variance utility over terminal wealth:\(^{10}\)

\[
U_{0,i} = \max_{\{Q_{i,t}\}} T^{-1} E_{0,i} \left[ \sum_{t=1}^{T} \beta^t Q_{i,t} (D_t - P_t) \right] - \frac{1}{2} (\rho T)^{-1} \text{Var}_{0,i} \left[ \sum_{t=1}^{T} \beta^t Q_{i,t} (D_t - P_t) \right],
\]

(3)

where \( 0 < \beta \leq 1 \) is the discount factor, \( E_{0,i} \) and \( \text{Var}_{0,i} \) are the expectation and variance operators conditional on agent i’s date-0 information set, \( \{P, Y_i\} \), and \( \rho \) is risk-bearing capacity per unit of time. Investors have identical preferences – they can follow different strategies with different rates of portfolio turnover, but they all want to earn the highest possible returns, with the least amount of risk, in the shortest time. The sense in which the model maps into the colloquial use of the term “short-termism” is that agents in the model may choose to follow investment strategies featuring very rapid changes in their positions across dates. Short- and long-term investors are distinguished by how long they maintain positions, not by their objective.\(^{11}\)

The key restriction here is that signals are acquired and trade occurs on date 0. In general settings there is no known closed-form solution to even the partial-equilibrium dynamic portfolio choice problem, let alone to the full market equilibrium. We therefore use a relatively minimal static model which eliminates those problems by assumption. The model nevertheless has the two characteristics that we stated we desire in the introduction: it allows for investment strategies that place

\(^{10}\)To see why this is over terminal wealth, note that when the profits from each futures claim, \( D_t - P_t \), are reinvested at the riskless rate \( \beta^{-1} \), terminal wealth, \( W_{T,i} \), is \( \sum_{t=1}^{T} \beta^{-T} Q_{i,t} (D_t - P_t) \), which is simply \( \beta^{-T} \) times the argument of the expectation and variance in the preferences. For motivation, see Dumas and Luciano (1991).

\(^{11}\)The model can only accommodate mean-variance or constant absolute risk aversion preferences and remain tractable.
different weight on fundamentals on different dates in the future, and it allows investors to make a choice about how precise their signals are for different types of fluctuations in fundamentals.\footnote{The two key differences from a fully dynamic model are that agents cannot condition on past realizations of fundamentals and that there is not a full set of state-contingent contracts. The former restriction will bind more weakly when agents make decisions primarily based on private signals rather than the realization of fundamentals. The latter restriction could potentially lead to time inconsistency, depending on how one assumes agents update information sets and preferences over time.}

The time discounting in (3) has the effect of making dates farther in the future less important in the objective of the investors. Define

\[ \tilde{Q}_{i,t} \equiv \beta^t Q_{i,t} \]  

(4)
to be agent \( i \)'s discounted demand. In what follows, the \( \tilde{Q}_{i,t} \) will be stationary processes. That means that \( Q_{i,t} = \beta^{-t} \tilde{Q}_{i,t} \) will generally grow in magnitude with maturity \( t \), though only to a relatively small extent for typical values of \( \beta \) and horizons on the order of 10–20 years.

1.4 Noise trader demand

In order to keep prices from being fully revealing, we assume there is uninformed demand from a set of noise trader who have irrational expectations. Their beliefs depend on a signal, \( Z_t \), that is in reality uncorrelated with fundamentals and represents a sentiment shock. Appendix A shows their demand, denoted \( N_t \), is then

\[ \tilde{N}_t = Z_t - kP_t, \]  

(5)

where \( \tilde{N}_t \equiv \beta^t N_t. \)

(6)
k is a coefficient determining the sensitivity of noise trader demand to prices, which depends on their risk aversion and how precise they believe their signals to be. In the benchmark case where \( D_t \) is stationary in levels, we assume that \( Z_t \) is also stationary in levels, which yields a useful symmetry between fundamentals, supply, and the signals, in that they are all stationary processes.
1.5 Asset market equilibrium in the time domain

We begin by solving for the market equilibrium on date 0 that takes the agents’ signal precisions, \( \Sigma_i^{-1} \), as given.

**Definition 1** For any given set of individual precisions \( \{\Sigma_i\}_{i \in [0,1]} \), a date-0 asset market equilibrium is a set of demand functions, \( \{Q_i(P,Y_i)\}_{i \in [0,1]} \), and a price vector \( P \), such that investors maximize utility and all markets clear: \( \int_i Q_i,di + N_t = 0 \) for all \( t \geq 1 \).

Investors submit demand curves for each futures contract and the equilibrium price vector, \( P \), is the one that clears all markets. The structure of the time-0 equilibrium is mathematically that of Admati (1985):

\[
P = A_1D + A_2Z, \quad \text{(7)}
\]

where

\[
A_1 \equiv I - (\rho^2 \Sigma_{avg}^{-1} \Sigma_Z^{-1} \Sigma_{avg}^{-1} + \Sigma_{avg}^{-1} + \rho^{-1}k + \Sigma_D^{-1})^{-1} (\rho^{-1}k + \Sigma_D^{-1}), \quad \text{(8)}
\]

\[
A_2 \equiv \rho^{-1}A_1 \Sigma_{avg}^{-1}, \quad \text{(9)}
\]

\[
\Sigma_{avg}^{-1} \equiv \int_i \Sigma_i^{-1} di. \quad \text{(10)}
\]

As Admati (1985) discusses, this equilibrium is not particularly illuminating since standard intuitions, including the idea that increases in demand should raise prices, do not hold. Prices of futures maturing on any particular date depend on fundamentals and demand for all other maturities, since the matrices \( A_1 \) and \( A_2 \) are not diagonal except in knife-edge cases. Interpreting the equilibrium requires interpreting complicated products of matrix inverses. The following section shows that the equilibrium can be solved by hand nearly exactly when it is rewritten in terms of frequencies.

2 Frequency domain interpretation

2.1 Frequency portfolios

The basic difficulty of the model is that fundamentals, noise trader demand, and signal errors are all correlated across dates. For any one of those three processes, we could use a standard orthogonal (eigen-) decomposition to yield a set of independent components. But in general three
time series with different correlation properties across dates will not have the same orthogonal
decomposition. A central result from time series analysis, though, is that a particular frequency
transform asymptotically orthogonalizes all standard stationary time series processes.

Such a transformation represents simply analyzing the prices of particular portfolios of futures
instead of the futures themselves. It must satisfy three requirements. First, the transformation
should be full rank, so that the set of portfolios allows an investor to obtain the same payoffs as
the futures themselves. Second, the transformed portfolios should be independent of each other.
And third, since we are studying trade at different frequencies, it would be nice if the portfolios
also had a frequency interpretation.

There are many different conceptions of fluctuations at different frequencies. One might imagine
step functions switching between +1 and -1 at different rates. For reasons we will see below, using
sines and cosines will be most natural in our setting.

Formally, the portfolio weights are represented as vectors of the form

\[ c_h = \sqrt{\frac{2}{T}} \left( \cos \left( \omega_h (t - 1) \right) \right)_{t=1}^{T}, \]

\[ s_h = \sqrt{\frac{2}{T}} \left( \sin \left( \omega_h (t - 1) \right) \right)_{t=1}^{T}, \]

where \( \omega_h \equiv \frac{2\pi h}{T}, \) (13)

for different values of the integer \( h \in \{0, 1, \ldots, T/2\}. \) \( c_0 \) is the lowest frequency portfolio, with the
same weight on all dates, while \( c_{T/2} \) is the highest frequency, with weights switching each period
between \( \pm 1. \)

Figure 1 plots the weights for a pair of those portfolios. The x-axis represents dates and the
y-axis is the weight of the portfolio on each date. The weights vary smoothly over time, with the
rate at which they change signs depending on the frequency \( \omega_h. \)

Economically, the idea is to think about the investment problem as being one of choosing
exposure to different types of fluctuations in fundamentals. A long-term investor can be thought of
as one whose exposure to fundamentals changes little over time, while a short-term investor holds
a portfolio whose weights change more frequently and by larger amounts.

Our claim is that studying the frequency portfolios is more natural than studying individual
futures claims. Investors do not typically acquire exposure to fundamentals on only a single date. Rather, they have exposures on multiple dates, and the portfolios we study are one way to express that. While investors will also obviously not hold a portfolio that takes the exact form of a cosine, any portfolio can be expressed as a sum of cyclical components. An investor whose portfolio loadings change frequently will have a portfolio whose weights are relatively larger on the high-frequency components, which figure 1 shows generate rapid changes in loadings.

2.2 Properties of the frequency transformation

The portfolio weights can be combined into a matrix, $\Lambda$, which implements the frequency transformation,

$$\Lambda \equiv \left[ \frac{1}{\sqrt{2}} c_0, c_1, \bar{s}_1, s_2, \ldots, c_{T-1}, \frac{1}{\sqrt{2}} \bar{s}_{T-1}, \frac{1}{\sqrt{2}} c_T \right].$$

($s_0$ and $s_{T/2}$ do not appear since they are identically equal to zero; the $1/\sqrt{2}$ scaling for $c_0$ and $c_{T/2}$ gives them the same norms as the other vectors).

We use lower-case letters to denote frequency-domain objects. So whereas $\tilde{Q}_i$ is investor $i$’s vector of discounted allocations to the various futures, $\tilde{q}_i$ is their vector of discounted allocations to the frequency portfolios, with

$$\tilde{Q}_i = \Lambda \tilde{q}_i.$$ 

(15)

In what follows, the index $j = 1, ..., T$ identifies columns of $\Lambda$. The $j$th column of $\Lambda$ is a vector that fluctuates at frequency $\omega_{\lfloor \frac{j}{2} \rfloor} = 2\pi \lfloor \frac{j}{2} \rfloor / T$, where $\lfloor \cdot \rfloor$ is the integer floor operator.$^{13}$ So there are two vectors, a sine and a cosine, for each characteristic frequency, with the exceptions of $j = 1$ (frequency 0, the lowest possible) and $j = T$ (frequency $\frac{T}{2}$, the highest possible).

Note also that $\Lambda$ has the property that $\Lambda^{-1} = \Lambda'$, so that frequency-domain vectors can be obtained through

$$\tilde{q}_i = \Lambda' \tilde{Q}_i.$$ 

(16)

In the same way that $\tilde{q}_i$ represents weights on frequency-specific portfolios, $d \equiv \Lambda' D$ is a representation of the realization of fundamentals written in terms of frequencies instead of time. The first element of $d$, for example, is proportional to the realized sample mean of $D$.

$^{13}$ $\lfloor x \rfloor$ is the largest integer that is less than or equal to $x$. 

12
As a simple example, consider the case with $T = 2$. The low-frequency or long-term component of dividends is then $d_0 = (D_1 + D_2)/\sqrt{2}$ and the high-frequency or transitory component is $d_1 = (D_1 - D_2)/\sqrt{2}$. Agents invest in the low-frequency component $d_0$ by buying an equal amount of the claims on $D_1$ and $D_2$ and they trade the high-frequency component $d_1$ by buying offsetting amounts of the claims on $D_1$ and $D_2$. A short-term investment in this case is one where the sign of the exposure to fundamentals changes, while the long-term investment has a fixed position.

The most important feature of the frequency transformation is that it approximately diagonalizes the variance matrices.

**Definition 2** For an $n \times n$ matrix $A$ with elements $a_{l,m}$, the weak matrix norm is

$$|A| \equiv \left( \frac{1}{n} \sum_{l=1}^{n} \sum_{m=1}^{n} a_{l,m}^2 \right)^{1/2}.$$  \hfill (17)

If $|A - B|$ is small, then the elements of $A$ and $B$ are close in mean square.

The frequency transformation leads us to study the spectral densities of the various time series:

**Definition 3** The spectrum at frequency $\omega$ of a stationary time series $X_t$ is

$$f_X(\omega) \equiv \sigma_{X,0} + 2 \sum_{t=1}^{\infty} \sigma_{X,t} \cos(\omega t),$$ \hfill (18)

where $\sigma_{X,t} = \text{cov}(X_s, X_{s-t})$. \hfill (19)

The spectrum, or spectral density, is used widely in time series analysis. The usual interpretation is that it represents a variance decomposition. $f_X(\omega)$ measures the part of the variance of $X_t$ associated with fluctuations at frequency $\omega$, which is formalized as follows.

**Lemma 1** For any stationary time series $\{X_t\}^T_{t=1}$, with frequency representation $x \equiv X', \Sigma_x \equiv X'X$, the elements of the vector $x$ are approximately uncorrelated in the sense that the covariance matrix of $x$, $\Sigma_x \equiv X'X$, is nearly diagonal,

$$|\Sigma_x - \text{diag}(f_X)| \leq bT^{-1/2},$$ \hfill (20)
and $x$ converges in distribution to

$$x \rightarrow_d N(0, \text{diag}(f_X)), \quad (21)$$

where $b$ is a constant that depends on the autocorrelations of $X$, and $\text{diag}(f_X)$ denotes a matrix with the vector $\{f_X(\omega_{|j/2|})\}_{j=1}^T$ on the main diagonal and zeros elsewhere.

**Proof.** These are textbook results (e.g. Brockwell and Davis (1991) and Gray (2006)). Online appendix 1 provides a derivation of the inequality (20) specific to our case. The convergence in distribution follows from Brillinger (1981), theorem 4.4.1. □

Lemma 1 says that $\Lambda$ approximately diagonalizes all stationary covariance matrices. So the frequency-specific components of fundamentals, prices, and noise trader demand are all (approximately) independent when analyzed in terms of frequencies. That is, $d = \Lambda'D$, $y_i = \Lambda'Y_i$, and $z = \Lambda'Z$ all have asymptotically diagonal variance matrices. That independence will substantially simplify our analysis, and it is a special property of the sines and cosines, as opposed to other conceptions of frequencies. The various primitive restrictions on the model, including mean-variance preferences and stationarity, are required in order to be able to take advantage of this diagonalization result. \textsuperscript{16}

### 2.3 Market equilibrium in the frequency domain

Instead of solving jointly for the prices of all futures, the approximate diagonalization result allows us to solve a series of parallel scalar problems, one for each frequency. Intuitively, since the frequency-specific portfolios have returns that are nearly uncorrelated with each other, the investors’

\textsuperscript{14}Specifically, $b = 4 \left( \sum_{j=1}^\infty |\sigma_{X,j}| \right)$.

\textsuperscript{15}A requirement of this lemma, which we impose for all the stationary processes studied in the paper, is that the autocovariances are summable in the sense that $\sum_{j=1}^\infty |\sigma_{X,j}|$ is finite (which holds for finite-order stationary ARMA processes, for example). Trigonometric transforms of stationary time series converge in distribution under more general conditions, though. See Shumway and Stoffer (2011), Brillinger (1981), and Shao and Wu (2007).

\textsuperscript{16}Finally, note that infill asymptotics, where $T$ grows by making the length of a time period shorter, are not sufficient for lemma 1 to hold. $T$ must be large relative to the range of autocorrelation of the process $X$. So, for example, if fundamentals have nontrivial autocorrelations over a horizon of a year, $T$ should be substantially larger than a year. If one shifts from annual to monthly data, then $T$ should rise by a factor of 12 for the approximations to be equally accurate. Additionally, $T$ should be as long as the investors’ actual horizons. That is, if they are investing for retirement, it would be on the other of decades (if this is a model with geometric time discounting, then perhaps we should have $T \rightarrow +\infty$; we use a finite $T$ in order to avoid the nontrivial technical challenges of dealing with infinitely large matrices). Overall, $T$ should be long enough both for the approximations involved in the frequency transformation to be correct, and to accurately represent the time horizon of individual investors.
utility can be written approximately as a sum of mean-variance optimizations\footnote{This follows from lemma 1 combined with the fact that $\Lambda'\Lambda = I$, so that $Q_i' D = Q_i' \Lambda' \Lambda D = q_i' d$.}

\[ U_{0,i} \approx \max_{\{q_{i,j}\}} \left\{ E_0, i \left( \tilde{q}_{i,j} (d_j - p_j) \right) - \frac{1}{2} \rho^{-1} \text{Var}_0, i \left( \tilde{q}_{i,j} (d_j - p_j) \right) \right\}. \quad (22) \]

In what follows, we solve the model using the approximation for $U_{0,i}$, and then show that it converges to the true solution from Admati (1985). When utility is completely separable across frequencies, there is an equilibrium frequency by frequency:

**Solution 1** Under the approximations $d \sim N(0, \text{diag}(fD))$ and $z \sim N(0, \text{diag}(fZ))$, the prices of the frequency-specific portfolios, $p_j$, satisfy, for all $j$

\[ p_j = a_{1,j} d_j + a_{2,j} z_j \quad (23) \]

\[ a_{1,j} \equiv 1 - \frac{\rho^{-1} k + f_{D,j}^{-1}}{\left( \rho f_{\text{avg},j}^{-1} \right)^2 f_{Z,j}^{-1} + f_{\text{avg},j}^{-1} + f_{D,j}^{-1} + \rho^{-1} k} \quad (24) \]

\[ a_{2,j} \equiv \frac{a_{1,j}}{\rho f_{\text{avg},j}^{-1}} \quad (25) \]

where $f_{\text{avg},j}^{-1} \equiv \int f_{i,j}^{-1} di$ is the average precision of the agents’ signals at frequency $j$.

**Proof.** See appendix B. ■

The price of the frequency-$j$ portfolio depends only on fundamentals and supply at that frequency due to the independence across frequencies. As usual, the informativeness of prices, $\text{Var} [d_j \mid p_j]$ can be shown to increase in the precision of the signals that investors obtain, while the impact of noise trader demand on prices is decreasing in signal precision and risk tolerance.

These solutions for the prices are standard results for scalar markets. What is different here is simply that the agents chose exposures across frequencies, rather than across dates; $p_j$ is the price of a portfolio whose exposure to fundamentals fluctuates over time at frequency $2\pi \lfloor j/2 \rfloor / T$. Both prices and demands at frequency $j$ depend only on signals and supply at frequency $j$ – the problem is completely separable across frequencies.

While solution 1 is an approximation, online appendix 2.2 shows that the error for the coefficients $a_1$ and $a_2$ and the prices is of order $T^{-1/2}$. In other words, the standard time-domain solution for
stationary time series processes becomes arbitrarily close to a simple set of parallel scalar problems in the frequency domain for large $T$.

In what follows, we assume that $ka_{2,j} < 1$ for all $j$, which ensures that $z$ represents a positive demand shock in equilibrium. The restriction is that noise trader demand not be too sensitive to prices; in the literature $k$ is usually equal to zero.

2.4 Optimal information choice in the frequency domain

The analysis so far takes the precision of the signals as fixed. Following Van Nieuwerburgh and Veldkamp (2009) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016), we allow investors to choose their signal precisions, $\Sigma_i^{-1}$ to maximize the expectation of their mean-variance objective (3) subject to an information cost,$^{18}$

$$
\max \{f_{i,j}\} E_{-1} \left[ U_{i,0} \mid \Sigma_i^{-1} \right] - \frac{\psi}{2T} \text{tr} \left( \Sigma_i^{-1} \right),
$$

where $E_{-1}$ is the expectation operator on date $-1$, i.e. prior to the realization of signals and prices (as distinguished from $E_{i,0}$, which conditions on $P$ and $Y_i$), and $\psi$ is the per-period cost of information. Total information here is measured by the trace operator $\text{tr} \left( \Sigma_i^{-1} \right)$. Note that while Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) focus on the case where investors have a fixed budget of precision, we are studying the dual problem in which information comes at a constant marginal cost. This can be thought of as a case where an investment firm can choose how many analysts to hire at a fixed wage, with total precision scaling linearly with the number of analysts.$^{19}$

Given the optimal demands, an agent’s expected utility is linear in the precision they obtain at each frequency.

Lemma 2 Each informed investor’s expected utility at time $-1$ may be written as a function of

---

$^{18}$The preferences can equivalently be written in terms of utility over terminal wealth, $W_{T,i}$. Specifically, maximization of $E_{-1} \left[ -\rho^{-1}T^{-1} \log E_{0,i} \left[ \exp (-\rho W_{T,i}) \right] \mid \Sigma_i^{-1} \right]$, where $E_{0,i}$ conditions on priors, agent $i$’s signals, and prices, is equivalent to maximization of (26) since $U_{0,i} = \rho^{-1}T^{-1} \log E_{0,i} \left[ \exp (-\rho W_{T,i}) \right]$.

$^{19}$The constraint model corresponds to a world where firms cannot expand the number of analysts that they employ, just shift them among tasks (frequencies). The cost model that we focus on represents a world where firms are free to hire more analysts from an elastic supply. This is more relevant if the financial sector does not account for most of the employment of the people capable of doing research.
their own signal precisions, \( f_{i,j}^{-1} \), and the average across other investors, \( f_{\text{avg},j}^{-1} \equiv \int_i f_{i,j}^{-1} \text{d}i \), with

\[
E_{-1} [U_{0,i} \mid \{f_{i,j}\}] = \frac{1}{2T} \sum_{j=1}^{T} \lambda_j \left( f_{\text{avg},j}^{-1} \right) f_{i,j}^{-1} + \text{constant}, \tag{27}
\]

where the constant does not depend on investor \( i \)'s precision, and the functions \( \lambda_j \) satisfy \( \lambda_j (x) > 0 \) and \( \lambda_j' (x) < 0 \) for all \( x \geq 0 \) and all \( j \).

**Proof.** See online appendix 2.3. ■

The terms \( \lambda_j \left( f_{\text{avg},j}^{-1} \right) \) represent the marginal utility of precision at each frequency, which the individual investor takes as given. Since expected utility and the information cost are both linear in the set of precisions that agent \( i \) chooses, \( \{f_{i,j}^{-1}\} \), it immediately follows that agents purchase signals at whatever subset of frequencies has \( \lambda_j \left( f_{\text{avg},j}^{-1} \right) \geq \psi \).

**Solution 2** Information is allocated so that

\[
f_{\text{avg},j}^{-1} = \begin{cases} 
\lambda_j^{-1} (\psi) & \text{if } \lambda_j (0) \geq \psi, \\
0 & \text{otherwise}.
\end{cases}
\tag{28}
\]

Because attention cannot be negative, when \( \lambda_j (0) \leq \psi \), no attention is allocated to frequency \( j \). Otherwise, attention is allocated so that its marginal benefit and its marginal cost are equated. This result does not pin down precisely how any specific investor's attention is allocated; in this class of models, when information cost are not strictly convex, only the aggregate allocation of attention across frequencies is determinate. For the purposes of studying price informativeness, though, characterizing this aggregate allocation is all that is necessary. Solution 2 is the water-filling equilibrium of Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016).

At this point there are still no investors who are explicitly “short-term” or “long-term”. Investors can follow many different strategies, with different mixes of short- and long-term focus. Even without any specialization to particular strategies, though, we now have sufficient structure to analyze the price effects of restrictions on the strategies that investors may follow.
3 The consequences of restricting investment frequencies for prices

This section focuses on the effects on prices of restrictions on the frequencies at which investment strategies can operate. It examines a particularly stark restriction that simply outlaws certain strategies. Specifically, we assume that investors are restricted to setting $\bar{q}_{i,j} = 0$ for $j$ in some set $\mathcal{R}$. We leave the noise traders unconstrained, assuming that, like retail investors, they face different regulations from large and sophisticated institutions.

This restriction on exposures to the frequency portfolios reduces the number of degrees of freedom that an investor has in making choices. To see this intuitively, consider a model in which each time period is an hour, and $T$ is a year, or 1625 trading hours. A restriction that investors cannot invest at a frequency higher than a day (6.5 hours) would mean that they would go from a strategy with 1625 degrees of freedom to one with only 250. A pension that sets a portfolio once a quarter would have only four degrees of freedom. In that sense, then, a frequency restriction is similar to a shift from a continuous market to one with infrequent batch auctions, as in Budish, Cramton, and Shim (2015).

Derivations of the results in the remainder of this section are in appendix C.

3.1 Results

3.1.1 Price informativeness across frequencies

In terms of frequencies, there is a complete separation in response to an investment restriction: prices become uninformative at restricted frequencies, while remaining unaffected at unrestricted frequencies.

Result 1 When investment by sophisticated investors is restricted at a set of frequencies $\mathcal{R}$, prices satisfy

$$p_j = \begin{cases} k^{-1}z_j & \text{for } j \in \mathcal{R} \\ a_{1,j}d_j + a_{2,j}z_j & \text{otherwise} \end{cases},$$

(29)

where $a_{1,j}$ and $a_{2,j}$ are the same as those defined in solution 1.
Intuitively, when sophisticated investors are restricted at a particular frequency, prices depend only on sentiment, since the agents with information cannot express their opinions. Moreover, the market becomes illiquid, and it is cleared purely through prices rather than quantities. On the other hand, since the solution for information acquisition at a frequency \( j \) does not depend on anything about any other frequency, the information acquired at a frequency \( j \notin \mathcal{R} \) is unaffected by the policy. As a result, we have the following corollary.

**Corollary 1.1** When investors are restricted from holding portfolios with weights that fluctuate at some set of frequencies \( j \in \mathcal{R} \), then prices at those frequencies, \( p_j \), become completely uninformative about dividends. The informativeness of prices for \( j \notin \mathcal{R} \) about dividends is unchanged. More formally, \( \text{Var} \left[ d_j \mid p_j \right] \) for \( j \notin \mathcal{R} \) is unaffected by the restriction. For \( j \in \mathcal{R} \), \( \text{Var} \left[ d_j \mid p_j \right] = \text{Var} \left[ d_j \right] \).

So a policy that eliminates short-term investment, e.g. by requiring holding periods of some minimum length, reduces the informativeness of prices for the short-term or transitory components of fundamentals, but has no effect on price informativeness in the long-run.

### 3.1.2 Price informativeness across dates

The fact that prices remain equally informative at some frequencies does not mean that they remain equally informative for any particular date. Dates and frequencies are linked through a standard result:

\[
\text{Var} \left( D_t \mid P \right) = \frac{1}{T} \sum_{j=1}^{T} \text{Var} \left[ d_j \mid p_j \right].
\]  

(30)

The variance of an estimate of fundamentals conditional on prices at a particular date is equal to the average of the variances across all frequencies. We then have the following:

**Corollary 1.2** Investment restrictions reduce price informativeness for fundamentals on all dates by equal amounts, and by an amount that weakly increases with the number of frequencies that are restricted.

If a person is making decisions based on estimates of fundamentals from prices and they are worried that prices are contaminated by high-frequency noise due to a restriction on short-term
investment, a natural response would be to examine an average of fundamentals and prices over time (across maturities of futures contracts).

**Corollary 1.3** The informativeness of prices for the sum of fundamentals depends only on informativeness at the lowest frequency:

\[
Var \left( T^{-1} \sum_{t=1}^{T} D_t \mid P \right) = Var \left[ T^{-1/2} d_0 \mid p_0 \right],
\]

where \( d_0 \) is the lowest frequency portfolio – with equal weight each date – and \( p_0 \) is its price.

Corollary 1.3 follows immediately from the definition of \( d_0 \) and the independence across frequencies in the solution. It shows that the informativeness of prices for moving averages of fundamentals depends only on the very lowest frequency. So even if prices have little or no information at high frequencies – \( Var [d_j \mid p_j] \) is high for large \( j \) – there need not be any degradation of information about averages of fundamentals over multiple periods, as they depend primarily on precision at lower frequencies (smaller values of \( j \)).

More concretely, going back to our example of oil futures, when investors are not allowed to choose exposure to the high-frequency portfolios, prices become noisier, making it more difficult to obtain an accurate forecast of the spot price of oil at some specific moment in the future. But if one is interested in the average of spot oil prices over a year, the model predicts that prices remain informative under restrictions on short-term strategies.

Thus, any restriction on investment reduces price informativeness for any particular date. But when short-term investment is restricted, there is little change in the behavior of moving averages of prices. So if a manager is making investment decisions based on fundamentals only at a particular moment, then that decision will be hindered by the policy since prices now have more noise. But if decisions are made based on averages of fundamentals over longer periods, the model predicts that there need not be adverse consequences.

Finally, it is natural to examine the informativeness of differences in prices across dates. As an example, we can consider the variance of the first difference of fundamentals.

**Corollary 1.4** The variance of an estimate of the change in fundamentals across dates conditional
on observing the vector of prices is

$$\text{Var}[D_t - D_{t-1} | P] = \sum_{j=1}^{T} 2 \left(1 - \cos \left(\omega \frac{j}{2}\right)\right) \text{Var}[d_j | p_j].$$

(32)

The function $2(1 - \cos(\omega))$ is equal to 0 at $\omega = 0$ and rises smoothly to 4 at the highest frequency, $\omega = \pi$. So period-by-period changes in fundamentals are driven primarily by high-frequency variation. As a result, policies restricting short-term investment will tend to have relatively large effects on the informativeness of prices for changes in fundamentals, as opposed to their limited effects on moving averages.\(^{20}\)

3.1.3 Return volatility

**Corollary 1.5** Given an information policy $f_{avg,j}^{-1}$, the variance of returns at frequency $j$, $r_j \equiv d_j - p_j$, is

$$\text{Var}(r_j) = \begin{cases} 
 f_{D,j} + \frac{f_{z,j}}{k^2} & \text{for } j \in R \\
 \min(\psi, \lambda_j(0)) & \text{otherwise}
\end{cases}.$$  

(33)

Moreover, the variance of returns at restricted frequencies satisfies $\text{Var}(r_j) > f_{D,j} + \frac{f_{z,j}}{(k+\rho f_{D,j})^2}$, which is the variance that returns would have at the same frequency if investment were unrestricted but agents were uninformed.

The volatility of returns at a restricted frequency is higher than it would be if the sophisticated investors were allowed to trade, even if they gathered no information. When uninformed active investors have risk-bearing capacity ($\rho > 0$), they absorb some of the exogenous demand by simply trading against prices, buying when prices are below their means and selling when they are above. The greater is the risk-bearing capacity, the smaller is the effect of sentiment volatility on return volatility. Thus, the restriction affects return volatilities through its effects on both liquidity provision and information acquisition.

\(^{20}\) Conversely, reductions in price informativeness at low frequencies have relatively large effects on moving averages and small effects on changes. This result can be relevant in situations where long-term investment is restricted, as in the case of a trading desk that cannot have exposure to cycles lasting longer than a day (e.g. Brock and Kleidon (1992) and Menkveld (2013)).
Restricting sophisticated investors from following short-term strategies in this model can thus substantially raise asset return volatility in the short-run – it can lead to, for example, large day-to-day fluctuations in prices (though those fluctuations in prices are, literally, variations in prices across maturities for different futures contracts on date 0). When unrestricted, sophisticated traders smooth prices across maturities, intermediating between excess demand on one day and excess supply in the next. When they are restricted from holding positions in futures that fluctuate from day to day, they can no longer provide that intermediation service, and short-term volatility increases. Restricting short-term investment increases transitory price volatility in this manner.

3.2 The pricing of equity

Equity is a claim on the entire future stream of fundamentals, so in the model we define it to be a claim that pays $D_t$ on each date $t$. Since the payoff of an equity claim is simply the sum of fundamentals, the date-1 equity claim has a payoff of exactly $d_0$. Corollary 1.3 then says that the absolute level of the price of equity remains equally informative under a short-run investment restriction as in the unrestricted case. That result is natural: if only short-run investment is restricted, then long-run investors, who simply buy and hold equity, are unaffected and can continue to maintain price efficiency.

However, that does not mean that equity prices are unaffected by the restriction. In particular, while the level of equity prices on an individual date remains equally informative, changes in equity prices over time are not. In particular, note that

$$P_{t}^{equity} - P_{t+1}^{equity} = P_t,$$

where $P_{t}^{equity} = \sum_{j=0}^{\infty} P_{t+j}$ is the price of equity on date $t$. The difference between equity prices between dates $t$ and $t+1$ is exactly equal to the price of the single-period dividend claim on date $t$. That is because a strategy that holds equity on date $t$ but then immediately sells it on date $t + 1$ only actually has exposure to fundamentals on date $t$.

So when restrictions on short-term investment make the prices of the individual futures claims less informative, they also make changes in the value of equity over time less informative. The results above for the informativeness of individual futures claims map directly into informativeness
of differences in equity prices across dates.

### 3.3 Numerical example

We now examine a numerical example to illustrate the predictions of the model for the behavior of investor positions, prices, and returns, both with and without restrictions on investment. Details of the analysis and further results are reported in online appendix 3.

The length of a time period is set to a week.\(^{21}\) The spectrum of fundamentals, \(f_D\), is calibrated to match the features of dividend growth for the CRSP total market index. Since dividends are nonstationary in the data, the numerical calibration assumes that \(\Delta D_t\) is stationary, so that the individual futures are claims to dividend growth.\(^{22}\) Online appendix 5 shows that the frequency analysis and theoretical results are essentially identical in this case to the baseline.\(^{23}\) The top-left panel of figure 2 plots the calibrated spectrum for dividend growth, \(f_D\). Empirically, there is substantial persistence in dividend growth, which causes \(f_D\) to peak at low frequencies.\(^{24}\)

The top-right panel of figure 2 plots the variance of returns on the dividend claims at each frequency, both with and without a restriction on investment at frequencies corresponding to cycles lasting less than one month (\(\omega \geq 2\pi/4\)), which could be thought of as similar to a tax on very short-term capital gains. Consistent with the results above, the variance of returns rises substantially at the restricted frequencies.

In addition to the benchmark case where each frequency is equally difficult to learn about, we also consider an alternative specification for the information cost in which the cost of precision increases as the frequency falls. Formally, in the benchmark specification, the total cost of information is \(\sum_j \psi f_{i,j}^{-1}\), and the alternative uses the generalization \(\sum_j \psi_j f_{i,j}^{-1}\), with \(\psi_j \propto (\omega_j + \omega_1)^{-1}\).\(^{25}\) That specification has two uses. First, it illustrates what would happen if a regulator taxed or

---

\(^{21}\) As discussed above, the model is not intended to match sub-second scale features of financial markets, like limit order books and exchange fragmentation. It could be plausibly applied to a daily or perhaps hourly frequency. Here we choose a week because that is the highest frequency at which aggregate economic indicators are released (specifically, initial claims for unemployment).

\(^{22}\) Technically, the spectrum, \(f_D\), is fit to the change in log dividends in calculating our calibration, but in the analysis that follows, we take \(f_D\) as applying to the first difference of the level of dividends.

\(^{23}\) See also section 5 for a discussion of this case.

\(^{24}\) We set \(\rho = 57.8\), \(k = 0.2\), and \(f_Z\) to be 1/8th of the smallest value of \(f_D\). The qualitative features of the model, as demonstrated in the results above, are not sensitive to those choices.

\(^{25}\) The average cost of information is set in this specification so that total information acquisition is equal to the baseline case, just shifted to higher frequencies.
subsidized information acquisition differentially across frequencies. Second, it will help match the empirical behavior of dividend strip variances.

The top-right panel shows that the consequence of that change is to cause the variance of returns to rise at low frequencies. The bottom-left panel of figure 2 plots the average precision of the signals obtained by investors at each frequency. Under the investment restriction, the precision goes to zero, since the information becomes useless. When information costs vary across frequency, so does information acquisition, and approximately inversely to the cost.

Finally, the bottom-right panel of figure 2 plots annualized Sharpe ratios of dividend strips at maturities of 1 to 7 years along with the equity claim (i.e. the claim to all dividend strips to maturity $T$) under the three different information policies. The dividend strips are modeled as claims to the level of dividends on a given date in the future. Since it is $\Delta D$ that is stationary here, a claim to $D_t$ is equal to a claim to $\sum_{s=1}^{t} \Delta D_s$. We assume that there is a unit supply of equity, which induces positive average returns on claims to dividends. Because $\Delta D_1$ affects the level of dividends on every date in the future, while $\Delta D_T$ affects only the level of dividends on date $T$, there is effectively greater supply of the shorter-maturity dividend claims, meaning that they earn higher returns in equilibrium, consistent with the findings of Binsbergen and Koijen (2017) and inducing downward-sloping Sharpe ratios.

In the benchmark case where investors acquire information at all frequencies, returns have the same variance at all frequencies and horizons, which is inconsistent with the data in Binsbergen and Koijen (2017). The cost specification that increases at low frequencies causes the variance curve to slope upward strongly with frequency, generating more strongly downward-sloping Sharpe ratios, both of which are consistent with the results reported by Binsbergen and Koijen (2017). That result is obtained because the variances of the dividend strips depend on lower frequencies when their maturities are longer. Figures A.2 and A.3 in the online appendix report further results and compare the model to the data reported by Binsbergen and Koijen (2017).

The previous section argues that while a restriction on short-term investment does not affect the informativeness of the level of equity prices on date 1, it does affect the informativeness of

---

differences across dates. Table 1 reports informativeness for both the level and various changes in equity prices over time. For the level, informativeness is measured as the increase in precision from observing prices,

$$\log \left( \frac{\text{var} \left[ \sum_{t=1}^{T} D_t \mid P_{1}^{\text{equity}} \right]^{-1}}{\text{var} \left[ \sum_{t=1}^{T} D_t \right]^{-1}} \right)$$

(35)

Similarly, for the $k$-period change in equity prices, we report

$$\log \left( \frac{\text{var} \left[ \sum_{t=1}^{T} D_t - \sum_{t=k+1}^{T} D_t \mid P_{k+1}^{\text{equity}} - P_{1}^{\text{equity}} \right]^{-1}}{\text{var} \left[ \sum_{t=1}^{T} D_t - \sum_{t=k+1}^{T} D_t \right]^{-1}} \right)$$

(36)

These measures of price informativeness map to the empirical measures of Bai, Philippon, and Savov (2016), who measure price informativeness across horizons based on the fraction of the variation in earnings explained by stock prices (see also Dávila and Parlatore (2018)).

Table 1 shows that the level of equity prices is no less efficient under the short-term investment restriction while the difference in equity prices between the first and second weeks is substantially less efficient. As the length of the difference gets longer, so that it focuses on lower frequencies, the efficiency rises back toward the baseline. Finally, looking at a second difference, which measures the change in price growth across two periods, isolating higher frequencies, the short-term restriction again has measurable effects (see corollary 1.4).27

4 Investor outcomes

This section studies the consequences of investment restrictions on investor profits and utility. We obtain two main results, which initially appear to be in conflict:

1. A rise in short-term investment, due to a decline in the cost of high-frequency information or trading, reduces the profits and utility of long-term investors.

2. However, restricting short-term investment in the way studied in section 3 only serves to further reduce the profits and utility of long-term investors.

So while long-term investors are worse off when short-term investment rises, cutting off short-

\[\text{corollary 1.4}^2\text{The case where low frequencies are more costly to learn about, e.g. because of a tax on low-frequency information acquisition or a subsidy to high-frequency acquisition, leads to precisely the opposite effects.}\]
term investment strategies – the ability to rapidly turn over portfolios – neither restores the old equilibrium, nor does it make the long-term investors better off. To achieve those goals, policies that change the cost of information acquisition are more suitable, as we will show.

The last part of the section examines the implications of the possible policy responses for noise traders, finding that noise traders are best off when prices are most informative.

4.1 Who are short- and long-term sophisticated investors?

We define a short-term investor as one whose portfolio is driven relatively more by high-frequency fluctuations, while a long-term investor holds a portfolio that is driven relatively more by low-frequency fluctuations. That definition can be formalized by a variance decomposition, using the facts

$$Var\left(\tilde{Q}_{i,t}\right) = \sum_{j=1}^{T} Var\left(\tilde{q}_{i,j}\right) \quad (37)$$

and

$$\frac{d}{df_{i,j}} [Var(\tilde{q}_{i,j})] > 0 \quad (38)$$

The component of the variance of $\tilde{Q}_{i,t}$ that is driven by fluctuations at frequency $j$, $Var(\tilde{q}_{i,j})$, is increasing in the precision of the signals agent $i$ acquires at frequency $j$ ($f_{i,j}^{-1}$). So if two investors have the same total variance of their positions, $Var\left(\tilde{Q}_{1,t}\right) = Var\left(\tilde{Q}_{2,t}\right)$, but one of them has higher-precision signals at high frequencies, i.e. $f_{1,j}^{-1} > f_{2,j}^{-1}$ for $j$ above some cutoff, then variation in that investor’s position is driven relatively more by high-frequency components.

For two investors with positions that have the same unconditional variance, the short-term investor – whose fluctuations happen relatively faster – is the one with relatively more precise signals about the transitory or high-frequency features of fundamentals. That is, short-term investors have short-term/high-frequency information, and long-term investors have long-term/low-frequency information. In the context of the example above, the short-term investor would be investor 1. As an extreme case – which is a simplification of the world for the sake of theoretical clarity – we will take short-term investors as people whose signals have positive precision only for $j$ above some cutoff $j_{HF}$, and long-term investors have signals with positive precision only for $j$ below some $j_{LF}$ with $j_{HF} > j_{LF}$.
4.2 Investor profits and utility

Result 2 Let \( R \equiv D - P \) be the vector of returns in the time domain. Investor \( i \)'s average discounted profits are

\[
E_{-1} [\tilde{Q}_i^t R_t] = \sum_{j=1}^{T} (1 - ka_{2,j}) (-E_{-1} [z_j r_j]) + ka_{1,j} E_{-1} [r_j d_j] + \rho (f_{i,j}^{-1} - f_{avg,j}^{-1}) \text{Var}_{-1} [r_j]
\] (39)

and expected profits at each frequency are nonnegative,

\[
E_{-1} [\tilde{q}_{i,j} r_j] \geq 0 \text{ for all } i, j
\] (40)

with equality only if \( f_{i,j}^{-1} = 0 \) and \( f_{D,j}^{-1} = \rho f_{avg,j}^{-1} f_{Z,j}^{-1} k \) (i.e. in a knife-edge case). Finally, the average earnings of noise traders are

\[
E \left[ \sum_{t=1}^{T} \tilde{N}_t R_t \right] = \sum_j \left( (1 - ka_{2,j}) z_j - ka_{1,j} d_j \right) (1 - a_{2,j} z_j - a_{1,j} d_j)
\] (41)

Each investor’s expected discounted profits depend on three terms. The first represents the profits earned from noise traders. \( E [z_j r_j] = -a_{2,j} f_{Z,j}^{-1} < 0 \) since the sophisticated investors imperfectly accommodate their demand. When the noise traders have high demand (that is, when \( z \) is high), they drive prices up and expected returns down. The sophisticated investors earn profits from trading with that demand.28

The second term represents the profits that the informed investors earn by buying from the noise traders when they have positive signals on average. The coefficient \( ka_{1,j} \) represents the slope of the supply curve that the informed investors face.

Finally, the third term in (39) represents a reallocation of profits from the less to the more informed sophisticated investors. An investor who has highly precise signals about fundamentals

\[28\] Note here that we are referring to flow profits, which do not include the cost of information acquisition that investors pay on date \(-1\). We do this partly because flow profits are more readily measurable than the potential fixed costs of setting up information acquisition technologies, and also because flow profits are still relevant in the case where investors face a constraint on information instead of a cost, or where the cost is in terms of utility units instead of money.
at frequency $j$ can accurately distinguish periods when prices are high due to strong fundamentals from those when prices are high due to high sentiment. That allows them to earn relatively more profits on average than an uninformed investor.

That said, an uninformed sophisticated investor does not earn negative expected profits at any frequency, even with $f_{i,j}^{-1} = 0$. There are always, except in a knife-edge case, profits to be earned by trading with noise traders. Intuitively, this result follows from the separability of the problem across frequencies. The reason that an investor must always earn nonnegative expected profits is that if at some frequency $j$ they did not, then they could simply set $q_{i,j} = 0$, ensuring profits of zero and hence higher utility. An uninformed investor forecasts returns based only on prices, so that knife-edge case represents the condition under which prices alone have no forecasting power for returns, and they set $q_{i,j} = 0$ in all states.

Result 2 yields two key insights. First, all investors, no matter their information, have the ability to earn profits at all frequencies through liquidity provision. Second, all else equal, investors who are informed about a particular frequency earn the most money from investing at these frequencies. Short-term investors – those with relatively more information about high-frequency fundamentals – earn relatively higher returns at high frequencies, while long-term investors earn relatively higher returns at low frequencies.

Finally, the decomposition of earnings for noise traders, equation (42), comes from the fact that the returns noise traders earn must be exactly the opposite of what the informed investors earn on average (i.e. equation (39) with $f_{i,j}^{-1} = f_{\text{avg},j}^{-1}$). Average noise trader earnings are quadratic in the coefficients determining prices, $a_{1,j}$ and $a_{2,j}$. That is caused by the interaction of two effects. First, when expected returns are more responsive to their demand shocks ($a_{2,j}$ is large) or to fundamentals ($1 - a_{1,j}$ is large), then expected returns vary more, giving more potential for losses. However, variation in prices inhibits their trading since they have downward sloping demand curves, with slope $k$. So when $1 - ka_{2,j}$ is small or $ka_{1,j}$ is small, losses are smaller.

### 4.3 The effects of an increase in short-term investment

This section studies the consequences of a decline in the cost of acquiring information at high frequencies for short- and long-term investors, as well as potential policy responses.
4.3.1 Effects on short- and long-term investors

Under the specification of the model where the total cost of information is $\sum_j \psi_j f_{i,j}^{-1}$ (where the baseline is the special case of $\psi_j = \psi$ for all $j$), the equilibrium condition for information acquisition is

$$f_{avg,j}^{-1} = \begin{cases} 
\lambda_j^{-1} (\psi_j) & \text{if } \lambda_j (0) \geq \psi_j, \\
0 & \text{otherwise.}
\end{cases} \quad (43)$$

(See online appendix 8.1 for a derivation of the other theoretical results in this case.) We examine the effects of a marginal reduction in $\psi_j$ for $j > j_{HF}$ starting from some the point $\psi_j = \lambda_j (0)$ — i.e. exactly where reducing information costs will lead to an initial increase in information acquisition. The investors who acquire information at those frequencies (setting $f_{i,j}^{-1} > 0$ for $j > j_{HF}$) are then the short-term investors, while those who do not, leaving $f_{i,j}^{-1} = 0$ for $j > j_{HF}$, are the long-term investors.29

**Corollary 2.1** Starting from a $\psi_j$ such that no investors acquire information at frequency $j$, a decline in $\psi_j$ that leads to an increase in the equilibrium $f_{avg,j}^{-1}$ reduces profits and utility of an investor for whom $f_{i,j}^{-1}$ remains unchanged. Specifically,

$$\frac{d}{d\psi_j} E^{-1} [\tilde{q}_{LF,j} T_j] \bigg|_{\psi_j = \lambda_j (0)^-} > 0 \quad (44)$$

$$\frac{d}{d\psi_j} E^{-1} \left[ \sum_t \tilde{Q}_{LF,t} (D_t - P_t) \right] \bigg|_{\psi_j = \lambda_j (0)^-} > 0 \quad (45)$$

$$\frac{d}{d\psi_j} E^{-1} [U_{LF,0}] \bigg|_{\psi_j = \lambda_j (0)^-} > 0 \quad (46)$$

where the notation $\psi_j = \lambda_j (0)^-$ indicates the derivative is taken to the left and the LF subscripts denote positions and utility of a long-term investor who keeps $f_{i,j}^{-1} = 0$ at the affected frequency. Concretely, in an economy initially populated only by long-term investors who gather no short-term information, a decline in the cost of short-term information increases $f_{avg,j}^{-1}$ for $j > j_{HF}$ and therefore reduces the expected profits at those frequencies, total expected profits, and the utility of

29 We assume here that only a fraction of investors begin acquiring high-frequency information, but another valid equilibrium is one in which all investors do — the model does not pin down the cross-sectional distribution. However, it is not difficult to generate specialization endogenously, e.g. through second-order differences in information acquisition skill across frequencies.
long-term investors.

The source of that result is the fact that investors with low-frequency information may still invest in the short-run (i.e. have exposures that change from day to day). Suppose, for example, that not only does $f_{LF,j}^{-1} = 0$ for high $j$, but also that $f_{avg,j}^{-1}$ does – nobody has short-term information. In that setting obviously any sophisticated investor will be happy to accommodate transitory fluctuations in noise trader demand. More concretely, an investor who has information that the long-term value of a stock is $50 will be willing to provide liquidity in the short-run, buying when the price is below $50 and selling when the price is higher. That liquidity provision will have high-frequency components when liquidity demand (noise trader demand) has high-frequency components (i.e. $f_{z,j} > 0$ for $j > j_{HF}$). That is, if there are short-run variations in sentiment, then there will be short-run variation in the low-frequency investor’s position.

The increase in short-term investment hurts those with low-frequency information because those with high-frequency information are better at providing short-term liquidity. Result 2 and corollary 2.1 formalize that idea and shows that how short-term investors hurt long-term investors – by crowding out their ability to provide liquidity. It is critical to note, though, that result 2 still shows that the increase in high-frequency investment never drives the profits earned by long-term investors below zero, even at high frequencies.

Furthermore, none of this is suboptimal from the perspective of the long-term investors – at the equilibrium, all investors are indifferent between acquiring information and not at any frequency where $\lambda_j(0) > \psi_j$, so their utility is not increased by acquiring more information at high frequencies. Moreover, that indifference also means that the short-term investors – those who actually create the increase in $f_{avg,j}^{-1}$ – also see declines in expected utility. The decline in $\psi_j$, by increasing information acquisition, makes prices more efficient, leaving less scope for investors to predict returns and earn profits.

The results also do not change the incentives of low-frequency investors to acquire information at low frequencies. While they lose money from a decrease in liquidity provision at high frequencies, their choices at low frequencies are unaffected. Therefore, a decline in the cost of high-frequency information has no effect on price informativeness at low frequencies. There is also nothing special about analyzing a shift in $\psi_j$ at high frequencies – the economic results are the same if the cost of
information changes at any frequency.

Nevertheless, there is something of an arms race here in that investor profits are decreasing in the information acquired by other investors. When the cost of high-frequency information falls, somebody will acquire more information, and the investor who does not will earn lower trading profits going forward. That said, making the arms race idea fully formal would require modeling a speed tournament or some sort of imperfect competition, so the link is only stylized.

Finally, we also note that the reduction in information costs and increase in short-term investment has positive effects on the overall market:

**Corollary 2.2** A reduction in $\psi_j$ that increases $f_{avg,j}^{-1}$ for $j > j_{HF}$, increases price informativeness and reduces return volatility at those frequencies. That is, for any frequency

$$\frac{d}{d\psi_j} \text{Var} [d_j | p_j] \bigg|_{\psi_j=\lambda_j(0)^-} \geq 0$$  \hspace{1cm} (47)

$$\frac{d}{d\psi_j} \text{Var} [r_j] \bigg|_{\psi_j=\lambda_j(0)^-} \geq 0$$  \hspace{1cm} (48)

So while long-term investors may be hurt by the reduction in information costs at high frequencies, to a regulator whose goal is simply to maximize price informativeness or minimize return volatility, short-term information acquisition and investment is beneficial.

### 4.3.2 Policy responses

If a decline in the cost of high-frequency information or trading hurts the incumbent long-term investors, a natural question to the incumbents might be how to restore the old equilibrium. We consider three responses that have been proposed: restricting or eliminating short-term investment, taxing transactions (or variation in positions), and limiting the availability of short-term information.

First, consider a restriction on short-term/high-frequency investment that says that no sophisticated investor may set $q_{i,j} \neq 0$ for $j$ above some cutoff, as in the previous section (and as in a daily batch auction). Result 2 implies that such a restriction would, rather than restoring the profits and utility of the long-term investors, actually reduce them further. This follows from the fact that restricting investment eliminates the terms in the summation for $j$ above the cutoff, which
are all nonnegative. While short-term investors make liquidity provision at high frequencies more difficult, outlawing short-term investment simply makes it impossible. So eliminating short-term investment does not restore the old equilibrium – it actually compounds the effect of the increase in short-term investment.

**Corollary 2.3** Limiting short-term investment with a policy that restricts all sophisticated investors from holding \( q_{i,j} \neq 0 \) for \( j > j_{HF} \) weakly reduces the profits and expected utility of all sophisticated investors.

Imposing a tax on changes in positions, specifically, a tax on \((Q_{i,t} - Q_{i,t-1})^2\), will have similar effects to a restriction on short-term investment in that the tax is most costly for short-term strategies with high turnover. Online appendix 6 formalizes that intuition.

The final policy response would be to somehow limit the acquisition of high-frequency information. In the context of the model, that would represent a restriction on the ability of investors to learn about period-to-period variation in fundamentals, for example by making it more costly to acquire high-frequency information. The most obvious response to a decline in the cost at frequency \( j \) is to directly impose a tax that exactly reverses the decline.

In the context of the model, a restriction on information acquisition could in fact exactly restore the equilibrium that exists in the absence of the short-term investors. Since the long-term investors do not acquire high-frequency information, the restriction has no direct effect on them. In terms of the results above, the reason that short-term investors harm long-term investors in the model is that they increase \( f_{\text{avg},j}^{-1} \) for high values of \( j \). A policy that makes short-term information more expensive does the opposite, reducing \( f_{\text{avg},j}^{-1} \) and shifting the market back to the previous equilibrium.

A specific example of a policy that could make it more costly for investors to acquire high-frequency information might be a reduction in the information that firms freely release. For example, there have been suggestions to change financial reporting requirements so that less short-run information is revealed proposed by the CFA institute (Schacht et al. (2007)) and Brookings Institution (Pozen (2014)). In the UK quarterly earnings reports are no longer mandatory (Gigler et al. (2014)). When firms stop reporting quarterly earnings, or providing short-term earnings guidance, they are in a sense making information acquisition more expensive – instead of simply reading
and interpreting announcements, investors now must do research to try to measure short-term performance.

To be clear, the claim here is not that markets should be tilted in the direction of long-term investors. Restricting information can help long-term investors in some cases, but it would also have potential negative externalities from reduced price informativeness that have been studied in the literature, making investment decisions worse, making monitoring of firms more difficult, and limiting firms’ ability to tie managerial pay to performance (Bond, Edmans, and Goldstein (2012)). Furthermore, the next section shows that restricting information in the model, even though it helps one class of investors, will hurt others.

4.4 Outcomes for noise traders

The formalization of noise traders used here is that they are investors whose demand depends on an uninformative signal that they erroneously believe forecasts fundamentals. Under that interpretation, a natural objective of a policymaker might be to set policies to try to reduce the losses of these investors and to keep their speculative demand from affecting prices and creating volatility (that is the motivation of the transaction tax in Tobin (1978)).

The decomposition of average earnings of noise traders, equation (42), shows that there are two ways to drive their losses to zero. One is for prices to be completely informative, with $a_{1,j} = 1$ and $a_{2,j} = 0$ (i.e. $p_j = d_j$) in equation (42). That case is ideal in that noise traders have no losses and prices are maximally useful as signals for making decisions. Noise trader losses are zero in this case because informed investors have perfectly elastic demand curves, and will trade any quantity since they know the price is exactly equal to fundamentals. The second way to reduce noise trader losses to zero is to drive $a_{1,j}$ to zero and $a_{2,j} = k^{-1}$. In that case, prices are completely uninformative, and they move in such a way that there is no trade. This achieves the goal of minimizing noise trader losses, but at the cost of eliminating all information from asset markets.\(^\text{30}\)

The two policies examined above – restricting trade and restricting information – both drive in the direction of the second way to reduce noise trader losses. However, they differ in their effects

\(^{30}\)Note, though, that noise trader profits are non-monotonic in both $a_1$ and $a_2$, so while we can draw conclusions about the extreme cases of $a_1$ and $a_2$ equal to 0 or 1, the exact response of profits to interior values of those coefficients is parameter dependent.
on price informativeness. We next consider each policy separately.

4.4.1 The consequences of restricting investment strategies

Restricting all investment by the informed investors at a given frequency eliminates all information from prices, but it also means that the noise traders have nobody to trade with, so their losses are identically zero. Similarly, restricting information, by reducing \( f_{\text{avg},j}^{-1} \) to zero, sets \( a_{1,j} \) to zero, so that the noise traders have no losses because sophisticated investors have no information (the second part of equation (42)). We also have

\[
\begin{aligned}
  f_{\text{avg},j}^{-1} = 0 \Rightarrow 1 - ka_{2,j} &= \frac{f_{D}^{-1}}{\rho^{-1}k + f_{D}^{-1}} \\
  \end{aligned}
\]

The noise traders will still lose money to the informed investors in general, through the first term in equation (42). As the amount of fundamental uncertainty grows, though — \( f_{D}^{-1} \) shrinks — the losses eventually fall to zero.

So unlike above, for the purpose of protecting noise traders, instead of long-horizon investors, the trading restriction is more effective than the information restriction. The information restriction does not in general reduce the losses of the noise traders to zero, while the trade restriction does. Either policy is only second-best, though, in the sense that they help noise traders by reducing the informativeness of prices and increasing price volatility.

The policy of restricting investment is most natural at frequencies where \( f_{Z,j} \) is large and \( f_{D,j} \) is small. At such a frequency, the information loss from restricting investment is relatively small — in fact it could be zero if \( \lambda_j(0) \) is sufficiently small (since there would also be no information acquisition in the absence of the restriction) — and the benefit is relatively large, since it increases in \( f_Z \) (equation (42)). So restrictions on investment make the most sense at frequencies with little variation in fundamentals but substantial variation in sentiment or noise trader demand.

**Corollary 2.4** At any frequency where \( f_{Z,j} \) is sufficiently large or \( f_{D,j} \) sufficiently small that \( \lambda_j(0) \leq \psi \) (recall that \( \lambda_j(0) \) represents the marginal value of acquiring information when \( f_{\text{avg},j}^{-1} = 0 \)), there is no information acquisition in equilibrium and prices are completely uninformative. At those frequencies, restricting trade by mandating that \( q_{i,j} = 0 \) reduces the losses of noise traders to
zero and has no effect on price efficiency, since prices are already uninformative.

A common view is that there is relatively little important economic news at high frequencies since economic decisions such as physical investment depend on relatively long-term expectations. In such a case, one would think that \( f_{D,j} \) is small at high frequencies. The results here then show that it would be natural to restrict high-frequency investment since there is no information loss and the effects of noise trader demand or speculation are eliminated. The model here formalizes that common intuition.

4.4.2 The consequences of subsidizing information

On the other hand, if the goal was to reduce noise trader losses without any reduction in price informativeness, then the ideal policy would be one that increases \( f_{avg,j}^{-1} \). Specifically, as the quantity of information that investors acquire becomes infinite (\( f_{avg,j}^{-1} \to \infty \)), prices become completely informative, in that \( a_{1,j} \to 1 \) and \( a_{2,j} \to 0 \), and noise traders have zero average losses,

\[
\lim_{f_{avg,j}^{-1} \to \infty} E[n_{j}r_{j}] = 0 \quad (50)
\]

Increases in \( f_{avg,j}^{-1} \) could be encouraged by subsidizing or otherwise encouraging information production by investors (e.g. a tax credit for research). In the context of the discussion in the previous section, this corresponds to actively trying to reduce the \( \psi_{j} \) that investors face (ideally to zero, if the goal is to send \( f_{avg,j}^{-1} \) to infinity). Certainly an information subsidy would not be costless to implement, and whether its benefits outweigh the costs is theoretically ambiguous.\(^{31}\)

In the model, there is tension among the outcomes of short-term investors, long-term investors, and noise traders. Long-term investors benefit from reductions in \( f_{avg,j}^{-1} \) at high frequencies, but that comes at the cost of reducing price informativeness and hurting noise traders and short-term investors. Noise traders benefit from increasing \( f_{avg,j}^{-1} \) (when \( f_{avg,j}^{-1} \) is sufficiently large, at least), or mandating greater disclosure about fundamentals, but that hurts the informed investors in general, since their trading opportunities shrink. These results follow from the zero-sum property of the

\(^{31}\)A closely related policy, and one that might be cheaper than a subsidy for research, is to mandate greater information production by firms, such as more frequent or thorough earnings announcements. Online appendix 4 examines a version of the model in which a public signal is revealed on date 0, and the next section discusses the results.
payoffs in the game that the investors are playing. For policymakers wishing to promote price efficiency, encouraging greater information acquisition and higher $f^{-1}_{avg,j}$ is ideal.

5 Robustness and alternative specifications

The appendix reports results for a number of perturbations to the assumptions in the analysis discussed so far. We summarize them here.

First, the baseline analysis assumes that fundamentals are stationary in levels. Online appendix 5 shows that the results are nearly identical when fundamentals and noise trader demand are instead stationary in first differences. In particular, all the frequency portfolios take precisely the same form as in the level-stationary case, except for the very lowest frequency. That means that the analysis above also goes through identically except for a small change at that single frequency.

Second, there are alternative specifications for the information acquisition problem that investors face. In the first, investors have a fixed budget of information instead of access to unlimited amounts of information at a constant cost (see online appendix 8.2). In that case, since the total supply of information is constant, restrictions on trade cause reallocations across frequencies; for example, while restricting high-frequency investment always reduces high-frequency informativeness, in this case it can also increase the informativeness of prices at low frequencies. This leads to ambiguity about the effects of such restrictions on price informativeness in the time domain and on investor profits and utility.

The second alternative information specification is to assume that the cost of information depends on the entropy of the signals, rather than their precision. Online appendix 8.3 shows that the main results continue to hold in that case and that the entropy specification also helps match the behavior of dividend strip returns.

The third potential modification of the model is to treat the noise trader demand as being driven by liquidity shocks rather than sentiment shocks. While the basic equilibrium is identical in this case to the baseline, some of the results on investor utility change. With rational liquidity demand, there are gains from trade that do not exist in the zero-sum baseline model. Online appendix 9 reports results for this case.

Finally, in the analysis of information restrictions, we suggested that a decrease in mandatory
reporting might constitute an increase in the cost of information. An alternative interpretation is that it would represent a decline in the precision of a public signal that all investors can freely observe. Online appendix 4 examines that case and shows that an increase in public disclosures can reverse the declines in utility and profits following a decline in the cost of information. The effects of disclosure therefore depend on modeling choices, which need to be evaluated empirically. The model is about costly information acquisition and processing, so it is somewhat inconsistent with the general approach to assume that it is costless for investors to interpret, for example, financial statements. Nevertheless, the basic pattern of the effects is the same, in that both a tax on information acquisition, raising $\psi_j$, and an increase in the precision of public signals raises the profits of low-frequency investors and noise traders and reduces the profits of high-frequency investors. A public signal has the added advantage of increasing price informativeness.

Online appendix 6 examines an extension of the model with quadratic trading costs. Those costs can be viewed as either technological or representing a transaction tax implemented by a regulator. It shows that an increase in trading costs affects high frequencies more strongly than low frequencies, causing investors to acquire less information at high frequencies and reduce high-frequency investment.

6 Conclusion

The aim of this paper is to understand the effects of policies aimed at reducing “short-termism” in financial markets. It develops results on the effects on price informativeness and investor welfare of restrictions on investment and information acquisition at different frequencies. In order to study those questions, we develop a model in which investors can make meaningful decisions about the horizon of their investment strategies, and in which they face endogenous information choices.

We obtain three main results in the baseline specification of the model:


2. Restricting short-term investment hurts both short- and long-term investors, but helps noise traders.

3. Taxing or restricting the availability of short-term information helps long-term investors,
hurts short-term investors and noise traders, and reduces short-term price efficiency.

The first result is a natural consequence of the statistical independence of the model across frequencies. The second result shows that, while lower costs of acquiring high-frequency information reduce the utility and profits of long-term investors, restricting short-term investment in response does not make long-term investors better off. A buy-and-hold investor is able to provide the market short-term liquidity – a person with a price target of $50 should be willing to accommodate transitory demand shocks that drive the price above their target. But when high frequency information is cheaper, more investors chose to focus on short-term strategies. These short-term investors are also better at providing short-term liquidity than long-term investors; this is what makes long-term investors worse off. However, eliminating all short-term investment does not solve the problem. In fact, it makes it worse by eliminating the earnings from liquidity provision for all investors. However, the results for noise traders are reversed – they benefit from restrictions on investment and are hurt by limits on information.

Finally, the third result shows that information policies have distributional effects. If one’s goal is to both maximize price informativeness and limit the impact of speculation by noise traders, subsidizing information acquisition can potentially (if the subsidy is sufficiently strong) solve both of those problems. However, since there is not a single accepted model of noise trading, the third result is relatively more delicate. We also examine an alternative specification in which noise traders are replaced by investors with time-varying hedging demand. In that case, it is more difficult to obtain clear predictions for welfare, but the first two main results continue to hold. Furthermore, price efficiency may have positive externalities that are not modeled here, as discussed in Bond, Edmans, and Goldstein (2012).

We do not make normative claims about what the right objective is. There are many externalities not considered here. For example, price informativeness is important to many agents in the economy who are not represented in our model. We also have a specific model of noise traders as irrational agents, but the role of noise trader demand in facilitating trade can also be played by agents who simply have exogenous liquidity needs, in which case the optimal policy response would more clearly tilt towards information subsidies. It is also not obvious whether short- or long-term investors should necessarily be supported. The goal of the paper is not to resolve the question of
which policy is best, but rather simply to provide a general analysis of the effects of the various policies.

References


A Noise trader demand

We assume that noise traders have preferences similar to those of sophisticates, but they have different information. They receive signals about fundamentals, and believe that the signals are informative, although the signals are actually random. The signals are also perfectly correlated across the noise traders, so that they do not wash out in the aggregate. They can be therefore thought of as common sentiment shocks among noise traders. Furthermore, the noise traders assume that prices contain no information about fundamentals.

The noise traders optimize

$$
\max_{\{N_t\}_{t=1}^T} \frac{T^{-1}}{2} \sum_{t=1}^T \beta^t N_t E_{0,N} [D_t - P_t] - \frac{1}{2} (\rho T)^{-1} \text{Var}_{0,N} \left[ \sum_{t=1}^T \beta^t N_t (D_t - P_t) \right]
$$

where $N_t$ is the demand of the noise traders and $E_{0,N}$ and $\text{Var}_{0,N}$ are their expectation and variance operators conditional on their signals.

We model the noise traders as being Bayesians who simply misunderstand the informativeness of their signals, and ignore prices. Their prior belief, before receiving signals, is that

$$D \sim N \left(0, \Sigma_N^{prior} \right).$$
They then receive signals that they believe (incorrectly) are of the form

\[ S \sim N \left( D, \Sigma_N^{signal} \right). \]  

The usual Bayesian update then yields the distribution of \( D \) conditional on \( S \),

\[ D \mid S \sim N \left( \Sigma_N \left( \Sigma_N^{signal} \right)^{-1} S, \Sigma_N \right) \]  

where \( \Sigma_N \equiv \left( \Sigma_N^{signal} \right)^{-1} + \left( \Sigma_N^{prior} \right)^{-1} \).  

So we have

\[ E_{0,N}[D] = \Sigma_N \left( \Sigma_N^{signal} \right)^{-1} S \]  

\[ \text{Var}_{0,N}[D] = \Sigma_N \]  

Define \( \tilde{N}_t \equiv \beta^t N_t \) and \( \tilde{N} = [N_1, \ldots, N_T]' \). The optimization problem then becomes

\[ \max_{\tilde{N}} T^{-1} \tilde{N}' \left( \Sigma_N \left( \Sigma_N^{signal} \right)^{-1} S - P \right) - \frac{1}{2} (\rho T)^{-1} \tilde{N}' \Sigma_N \tilde{N}. \]  

This has the solution:

\[ \tilde{N} = \rho^{-1} \Sigma_N^{-1} \left( \Sigma_N \left( \Sigma_N^{signal} \right)^{-1} S - P \right) \]  

\[ = \rho^{-1} \left( \left( \Sigma_N^{signal} \right)^{-1} S - \Sigma_N^{-1} P \right). \]  

For the sake of simplicity, we assume that \( \Sigma_N = k^{-1} I \), where \( I \) is the identity matrix and \( k \) is a parameter. (This can be obtained, for instance, by assuming that \( \Sigma_N^{signal} = \Sigma_N^{prior} = 2kI \). We then have

\[ \tilde{N} = \rho^{-1} \left( \Sigma_N^{signal} \right)^{-1} S - kP, \]  

so that the vector \( Z = (Z_1, \ldots, Z_T)' \) from the main text is:

\[ Z \equiv \rho^{-1} \left( \Sigma_N^{signal} \right)^{-1} S, \]
and the true variance of $S$, $\Sigma_S$, can always be chosen to yield any particular $\Sigma_Z \equiv \text{Var}(Z)$ by setting

$$\Sigma_S = \rho^2 \Sigma_N^{signal} \Sigma_Z \Sigma_N^{signal}.$$ (63)

**B Derivation of solution 1**

To save notation, we suppress the $j$ subscripts indicating frequencies in this section when they are not necessary for clarity. So in this section $f_D$, for example, is a scalar representing the spectral density of fundamentals at some arbitrary frequency (rather than vectors, which is what the unsubscripted variables represent in the main text).

In this section we solve a general version of the model that allows for a constant component of the supply, denoted $s$. This can be thought of as the mean aggregate supply of the underlying. The main results implicitly set $s = 0$, but the analysis of equity returns uses nonzero $s$. We assume that the noise traders’ demand curve depends on prices relative to their mean, so that supply does not enter. This is without loss of generality as it is simply a normalization.

### B.1 Statistical inference

We guess that prices take the form

$$p = a_1 d + a_2 z + a_3 s$$ (64)

where $s$ is nonstochastic. Standard analysis then yields that

$$E[d \mid y, p] = \tau^{-1}_i \left( f_i^{-1} y_i + \frac{a_1}{a_2} f_Z^{-1} (p - a_3 s) \right)$$ (65)

The online appendix reports a full derivation.
B.2 Demand and equilibrium

The agent’s utility function is (where variables without subscripts here indicate vectors),

\[
U_i = \max_{\{q_{i,t}\}} \rho^{-1} E_{0,i} \left[ T^{-1} \tilde{Q}_i^t (D - P) \right] - \frac{1}{2} \rho^{-2} \text{Var}_{0,i} \left[ T^{-1/2} \tilde{Q}_i^t (D - P) \right] \tag{66}
\]

\[
\begin{align*}
&= \max_{\{q_{i,t}\}} \rho^{-1} E_{0,i} \left[ T^{-1} \tilde{q}_i^t (d - p) \right] - \frac{1}{2} \rho^{-2} \text{Var}_{0,i} \left[ T^{-1/2} \tilde{q}_i^t (d - p) \right] \tag{67} \\
&= \max_{\{q_{i,t}\}} \rho^{-1} T^{-1} \sum_j \tilde{q}_{i,j} E_{0,i} \left[ (d_j - p_j) \right] - \frac{1}{2} \rho^{-2} T^{-1} \sum_j \tilde{q}_{i,j}^2 \text{Var}_{0,i} \left[ d_j - p_j \right], \tag{68}
\end{align*}
\]

where the last line follows by imposing the asymptotic independence of \( d \) across frequencies (we analyze the error induced by that approximation below). The utility function is thus entirely separable across frequencies, with the optimization problem for each \( \tilde{q}_{i,j} \) independent from all others.

The remainder of the equilibrium computation then follows standard steps, fully reported in the online appendix.

C Results on price informativeness with restricted frequencies

C.1 Result 1 and corollaries 1.1 and 1.5

When there are no active investors and just exogenous supply, we have that \( 0 = z_j + kp_j \) and so:

\[
\begin{align*}
p_j &= k^{-1} z_j, \tag{69} \\
r_j &= d_j - k^{-1} z_j. \tag{70}
\end{align*}
\]

Because of the separability of information choices across frequencies, the coefficients \( a_{1,j} \) and \( a_{2,j} \) are unchanged at all other frequencies. Moreover, it is clear that \( \text{Var}(d_j|p_j) = \text{Var}(d_j) \) at the restricted frequencies, since prices now only carry information about supply, which is uncorrelated with dividends.

Note that for any \( j \in \mathcal{R} \),

\[
\text{Var}(r_j) = f_{D,j} + \frac{f_{Z,j}}{k^2}, \tag{71}
\]

47
Additionally, if investors were allowed to hold exposure at those frequencies, but the endogenously chose not to allocate any attention to the frequency, the return volatility would be:

$$\text{Var}_{\text{unrestr.}}(r_j) = \lambda_j(0) = f_{D,j} + \frac{f_{Z,j}}{(k + \rho f_{D,j}^{-1})^2} < \text{Var}(r_j). \quad (72)$$

### C.2 Corollary 1.2 and result 1.4

Under the diagonal approximation, we have:

$$D | P \sim N(\bar{D}, \Lambda \text{diag}(\tau_0^{-1}) \Lambda') \quad (73)$$

where $\tau_0$ is a vector of frequency-specific precisions conditional on prices, as of time 0. Given the independence of prices across frequencies, the $j$-th element of $\tau_0$ is:

$$\tau_{0,j}^{-1} = \text{Var}(d_j | p_j). \quad (74)$$

Using this expression, we can compute:

$$\text{Var}(D_t | P) = 1_t' \Lambda \text{diag}(\tau_0^{-1}) \Lambda' 1_t \quad (75)$$

$$= (\Lambda' 1_t)' \text{diag}(\tau_0^{-1}) (\Lambda' 1_t) \quad (76)$$

$$= \sum_j \lambda_{t,j}^2 \text{Var}(d_j | p_j) \quad (77)$$

$$= \lambda_{t,0}^2 \text{Var}(d_0 | p_0) + \lambda_{t,T}^2 \text{Var}(d_T | p_T) + \sum_{k=1}^{T/2-1} (\lambda_{t,2k}^2 + \lambda_{t,2k+1}^2) \text{Var}(d_k | p_k) \quad (78)$$

where $1_t$ is a vector equal to 1 in its $t$-th element and zero elsewhere, and $\lambda_{t,j}$ is the $t, j$ element of $\Lambda$. The last line follows from the fact that the spectrum has $f_{X,2k} = f_{X,2k+1}$ for $0 < k < T/2 - 1$. Furthermore, note that for $0 < k < T/2 - 1$,

$$\lambda_{t,2k}^2 + \lambda_{t,2k+1}^2 = \frac{2}{T} \cos(\omega_k (t - 1))^2 + \frac{2}{T} \sin(\omega_k (t - 1))^2 \quad (79)$$

$$= \frac{2}{T} \quad (80)$$
which yields equation (30). Result 3 immediately follows from this expression and the fact that 
\[ \lambda_{t,0}^2 = \lambda_{t,T}^2 = \frac{1}{T}. \]

Result 1.4 uses the fact that
\[ \text{Var}(D_t - D_{t-1} | P) = (\lambda_{t,1} - \lambda_{t-1,1})^2 \tau_{0,1}^{-1} + (\lambda_{t,T} - \lambda_{t-1,T})^2 \tau_{0,T}^{-1} \]
\[ + \sum_{k=1}^{T/2-1} [\left(\lambda_{t,2k} - \lambda_{t-1,2k}\right)^2 + (\lambda_{t,2k+1} - \lambda_{t-1,2k+1})^2] \tau_{0,k}^{-1}, \]
and the fact that \((\cos(x) - \cos(y))^2 + (\sin(x) - \sin(y))^2 = 4 \sin \left(\frac{1}{2}(x - y)\right)^2 = 2 (1 - \cos(x - y))\).

**D Result 2**

Expression (37) in the main text follows from the steps used in appendix C.2. Recall from (??) that, omitting the \(j\) notation,
\[ \tilde{q}_i = \rho \left( f_i^{-1} y_i + \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) \right) \]
\[ = \rho f_i^{-1} \varepsilon_i + \rho \left( f_i^{-1} + \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) a_1 \right) d + \rho \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) a_2 z \]
Recall also that:
\[ \tau_i = \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} + f_D^{-1} + f_i^{-1}, \]
so that:
\[ \tilde{q}_i = \rho \left( \tau_i - \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} - f_D^{-1} \right) \varepsilon_i + \rho \left( f_i^{-1} + \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) a_1 \right) d + \rho \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) a_2 z \]
Moreover,
\[ f_i^{-1} - a_1 \tau_i + \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} = \tau_i - \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} - f_D^{-1} - a_1 \tau_i + \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} \]
\[ = (1 - a_1) \tau_i - f_D^{-1}. \]
Therefore

\[ \rho^{-1} \tilde{q}_i = \left( \tau_i - \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} - f_D^{-1} \right) \varepsilon_i + ((1 - a_1) \tau_i - f_D^{-1}) d + \left( \frac{a_1}{a_2} f_Z^{-1} - a_2 \tau_i \right) z, \quad (89) \]

so that

\[ \rho^{-2} \text{Var} (\tilde{q}_i) = \left( \tau_i - \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} - f_D^{-1} \right) + ((1 - a_1) \tau_i - f_D^{-1})^2 f_D + \left( \frac{a_1}{a_2} f_Z^{-1} - a_2 \tau_i \right)^2 f_Z. \quad (90) \]

(where the first term uses the fact that Var \( f_i^{-1} \varepsilon_i = f_i^{-1} \)). The derivative of this expression with respect to \( \tau_i \) is:

\[
\left. \rho^{-2} \frac{\partial \text{Var} (\tilde{q}_i)}{\partial \tau_i} \right|_{\tau_i = \tau} = 2 \tau_i ((1 - a_1)^2 f_D + a_2^2 f_Z) - 1 \\
\geq 2 \left( f_D^{-1} + \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} \right) ((1 - a_1)^2 f_D + a_2^2 f_Z) - 1 \\
= 2 \left( (1 - a_1)^2 + a_2^2 f_Z f_D^{-1} + (1 - a_1)^2 \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} f_D + a_1^2 \right) - 1 \\
= 2 \left( 1 - 2a_1(1 - a_1) + a_2^2 f_Z f_D^{-1} + (1 - a_1)^2 \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} f_D \right) - 1 \quad (91) \\
= 2 \left( -2a_1(1 - a_1) + a_2^2 f_Z f_D^{-1} + (1 - a_1)^2 \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} f_D \right) + 1 \\
= 2 \left( (1 - a_1) \left( \frac{a_1}{a_2} \right) (f_Z^{-1} f_D)^{\frac{1}{2}} - a_2 (f_Z^{-1} f_D)^{-\frac{1}{2}} \right)^2 + 1 \\
> 0,
\]

where to go from the first to the second line, we used the fact that \( \tau_i \geq \left( \frac{a_1}{a_2} \right)^2 f_Z^{-1} + f_D^{-1} \), and where we also used the fact that \( a_1 \leq 1 \). Since \( \tau_i \) is a monotonic transformation of \( f_i^{-1} \), this establishes equation (38) from the main text.

For result 2, first note that \( E_{-1} \left[ \tilde{Q}'_i R \right] = E_{-1} \left[ \tilde{q}'_i A' \Lambda r \right] = E_{-1} \left[ \tilde{q}_i r \right] = \sum_j E_{-1} \left[ \tilde{q}_{i,j} r_j \right] \), where the last equality follows from the diagonal approximation. Moreover, straightforward but tedious
algebra shows that:

\[
\begin{align*}
  f^{-1}_i + \left( \frac{a_1}{a_2} f_{Z}^{-1} - \tau_i \right) a_1 &= \rho(f^{-1}_i - f^{-1}_{avg})(1 - a_1) + k a_1, \\
  \left( \frac{a_1}{a_2} f_{Z}^{-1} - \tau_i \right) a_2 &= -\rho(f^{-1}_i - f^{-1}_{avg}) a_2 + (k a_2 - 1).
\end{align*}
\] (92) (93)

We can use these expressions, and the fact that \( r = (1 - a_1)d - a_2 z \) to re-write \( \tilde{q}_i \) as:

\[
\tilde{q}_i = \rho f^{-1}_i \varepsilon_i + \rho (f^{-1}_i - f^{-1}_{avg}) r + k a_1 d + (k a_2 - 1) z.
\] (94)

Therefore,

\[
E_{-1} [\tilde{q}_i r] = \rho (f^{-1}_i - f^{-1}_{avg}) \text{Var} (r) + k a_1 E_{-1} [rd] + (k a_2 - 1) E_{-1} [rz],
\] (95)

which is the decomposition from result 2.

The result that expected profits are nonnegative is a simple consequence of the investors' objective:

\[
\max_{\{\tilde{q}_{i,j}\}} \rho^{-1} T^{-1} \sum_j \left( E_{0,i} [\tilde{q}_{i,j} (d_j - p_j)] - \frac{1}{2} \rho^{-2} T^{-1} \sum_j \text{Var}_{0,i} [\tilde{q}_{i,j} (d_j - p_j)] \right)
\] (96)

Since the variance is linear in \( \tilde{q}_{i,j}^2 \), if \( E_{0,i} [\tilde{q}_{i,j} r_j] < 0 \), utility can always be increased by setting \( \tilde{q}_{i,j} = 0 \) (or, even more, by reversing the sign of \( \tilde{q}_{i,j} \)). In order for \( E_{-1} [\tilde{q}_{i,j} r_j] = 0 \), it must be the case that \( \text{Var}_{-1,i} [E_{0,i} [d_j - p_j]] = 0 \), since any deviation of \( E_{0,i} [d_j - p_j] \) will cause the investor to optimally take a nonzero position. We have, from above,

\[
\begin{align*}
  a_1 &= \frac{\tau_{avg} - f^{-1}_{D}}{\tau_{avg} + \rho^{-1} k} = \frac{(\rho f^{-1}_{avg})^2 f^{-1}_{Z} + f^{-1}_{avg}}{(\rho f^{-1}_{avg})^2 f^{-1}_{Z} + f^{-1}_{avg} + f^{-1}_{D} + \rho^{-1} k}, \\
  a_2 &= \frac{a_1}{\rho f^{-1}_{avg}}, \\
  \tau_{avg} &= \left( \rho f^{-1}_{avg} \right)^2 f^{-1}_{Z} + f^{-1}_{avg} + f^{-1}_{D}.
\end{align*}
\] (97) (98) (99)
The expression for \( a_2 \) is invalid in the case when \( f_{\text{avg}}^{-1} = 0 \). In that case, we have

\[
E[d \mid y_i, p] = \tau_i^{-1} \left( f_i^{-1} y_i + \frac{a_1}{a_Z^2} f_{\text{avg}}^{-1} p \right) \tag{100}
\]

\[
E[d - p \mid y_i, p] = \tau_i^{-1} f_i^{-1} y_i + \left( \tau_i^{-1} \frac{a_1}{a_Z^2} f_{\text{avg}}^{-1} - 1 \right) (a_1 d + a_2 z) \tag{101}
\]

\[
\text{Var} [E[d - p \mid y_i, p]] = \left( \tau_i^{-1} f_i^{-1} + \left( \tau_i^{-1} \frac{a_1}{a_Z^2} f_{\text{avg}}^{-1} - 1 \right) a_1 \right)^2 f_{\text{avg}}^{-1} + \left( \tau_i^{-1} \frac{a_1}{a_Z^2} f_{\text{avg}}^{-1} - 1 \right)^2 a_2 f \tag{102}
\]

Now first we must have \( \tau_i^{-1} \frac{a_1}{a_Z^2} f_{\text{avg}}^{-1} - 1 = 0 \) in order for the third term to be zero. But if that is true, then for the first term to be zero we must have \( f_i^{-1} = 0 \) (since \( \tau_i^{-1} \) is always positive). Combining \( f_i^{-1} = 0 \) with \( \tau_i^{-1} \frac{a_1}{a_Z^2} f_{\text{avg}}^{-1} - 1 = 0 \), we obtain

\[
f_{\text{avg}}^{-1} = \rho f_{\text{avg}}^{-1} f_{\text{avg}}^{-1} k. \tag{103}
\]
Table 1. Information losses relative to benchmark

<table>
<thead>
<tr>
<th></th>
<th>Short-term rest.</th>
<th>Low-freq. cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity price</td>
<td>0.00</td>
<td>4.06</td>
</tr>
<tr>
<td>1-week difference</td>
<td>1.59</td>
<td>0.00</td>
</tr>
<tr>
<td>1-month difference</td>
<td>0.46</td>
<td>1.08</td>
</tr>
<tr>
<td>1-year difference</td>
<td>0.05</td>
<td>2.89</td>
</tr>
<tr>
<td>Monthly second difference</td>
<td>0.61</td>
<td>-0.11</td>
</tr>
</tbody>
</table>

Note: log differences in precision of functions of equity prices for fundamentals between the baseline model and the short-term restriction or alternative cost function that is high at low frequencies. The first row is for the level of the equity price on date 1. The second is for the difference in the price of equity in week 1 and week 2 (which isolates just the dividend in the first week). The third and fourth rows are for the difference between the price of equity on the first date and one month and one year later, respectively. Finally, the bottom row is the efficiency of $P_{equity}^9 - 2P_{equity}^5 + P_{equity}^1$, which measures the change in price growth across the first and second months.
Figure 1: Portfolio weights

Notes: Portfolio weights for the cosing frequency portfolios $c_1$ and $c_{10}$, as defined in the main text. The horizontal axis is time, or the maturity of the corresponding futures contract. The vertical axis is the weight which each portfolio puts on that futures contract.
Figure 2: Numerical example

Notes: The four panels plot results for the numerical example. The frequency-specific cost case is where the cost of information at frequency $\omega_j$ is proportional to $(\omega_j + \omega_1)^{-1}$. The investment restriction says that sophisticated investors cannot hold portfolios that fluctuate at frequencies corresponding to cycles lasting one month or less. The Sharpe ratios for the dividend strip returns are calculated based on per-period returns as discussed in appendix ?? and then annualized.
1 Proof of lemma 1

The proof here follows “Time Series Analysis” lecture notes of Suhasini Subba Rao. The broad idea of the proof is as follows. Let \( \Sigma \) be any matrix of the form:

\[
\Sigma = \begin{pmatrix}
\sigma_0 & \sigma_1 & \ldots & \sigma_{T-1} \\
\sigma_1 & \sigma_0 & \sigma_1 & \ldots & \sigma_{T-2} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\sigma_{T-1} & \ldots & \ldots & \sigma_0
\end{pmatrix}
\]

(1)

where \( x_0 > 0 \). Matrices of this type contain all the variance-covariance matrices of order \( T \) of arbitrary weakly stationary processes. The lemma follows from “approximating” \( \Sigma \) by the circulant matrix:

\[
\Sigma_{circ} = circ(\sigma_{circ}) \quad , \quad \sigma \equiv (\sigma_0, \sigma_1 + \sigma_{T-1}, \sigma_2 + \sigma_{T-2}, \ldots, \sigma_{T-2} + \sigma_2, \sigma_{T-1} + \sigma_1)',
\]

(2)
where, for any real vector \( \{ x_i \}_{i=0}^{T-1} \),

\[
circ(x) \equiv \begin{pmatrix} x_0 & \cdots & x_{T-1} \\ x_{T-1} & x_0 & \cdots & x_{T-2} \\ \vdots & & \ddots & \vdots \\ x_1 & \cdots & x_0 \end{pmatrix}.
\] (3)

In order to obtain this approximation, we first need the following result.

**Appendix lemma 1** For any matrix \( \Sigma \) of the form given above, and associated circulant matrix \( \Sigma_{\text{circ}} \), the family of vectors \( \Lambda \) defined in the main text exactly diagonalizes \( \Sigma_{\text{circ}} \):

\[
\Sigma_{\text{circ}} \Lambda = \Lambda \text{diag} \left( \{ f_\Sigma (\omega_{[j/2]}) \}_{j=1}^{T} \right),
\] (4)

where each distinct eigenvalue in \( \{ f_\Sigma (\omega_{[j/2]}) \}_{j=1}^{T} \) is given by:

\[
f_\Sigma (\omega_h) = \sigma_0 + 2 \sum_{t=1}^{T-1} \sigma_t \cos(\omega_h t), \quad \omega_h \equiv 2\pi h/T,
\] (5)

for some \( h = 0, \ldots, \frac{T}{2} \).

Given that \( \Lambda \) is orthonormal,

\[
\Lambda' \Sigma_{\text{circ}} \Lambda = \text{diag} \left( f_\Sigma \right).
\] (6)

The approximate diagonalization of the matrix \( \Sigma \) consists in writing:

\[
\Lambda' \Sigma \Lambda = \text{diag} \left( f_\Sigma \right) + R_\Sigma,
\] (7)

where the \( T \times T \) matrix \( R_\Sigma \) is given by:

\[
R_\Sigma \equiv \Lambda' (\Sigma - \Sigma_{\text{circ}}) \Lambda.
\] (8)

This is an approximation in the sense that \( R_\Sigma \) is generically small. Specifically, it is of order \( T^{-1} \) element-wise. The following lemma proves the first result stated in lemma 1 of the main text.
Appendix lemma 2  For any $T \geq 2$, we have:

$$|R_\Sigma| \leq \frac{4}{\sqrt{T}} \sum_{j=1}^{T-1} |j\sigma_j|,$$

(9)

where $|M|$ denotes the weak matrix norm, as in the main text.

**Proof.** Define $\Delta \Sigma = \Sigma_{\text{circ}} - \Sigma$. First note that since:

$$\Sigma^{(i,j)} = \left\{ \begin{array}{ll} \sigma_0 & \text{if } i = j \\ \sigma_{|i-j|} & \text{otherwise} \end{array} \right. ,$$

(10)

$$\Sigma_{\text{circ}}^{(i,j)} = \left\{ \begin{array}{ll} \sigma_0 & \text{if } i = j \\ \sigma_{|i-j|} + \sigma_{T-|i-j|} & \text{otherwise} \end{array} \right. ,$$

(11)

we have:

$$\Delta \Sigma^{(i,j)} = \left\{ \begin{array}{ll} 0 & \text{if } i = j \\ \sigma_{T-|i-j|} & \text{otherwise} \end{array} \right. ,$$

(12)

where $\Sigma^{(i,j)}$ is the $(i,j)$ element of $\Sigma$. This means that the matrix $\Delta \Sigma$ has constant and symmetric diagonals. Moreover, the first subdiagonals both contain $\sigma_{T-1}$, the second contain $\sigma_{T-2}$, and so on. That is,

$$\Delta \Sigma = \begin{pmatrix}
0 & \sigma_{T-1} & \sigma_{T-2} & \sigma_2 & \sigma_1 \\
\sigma_{T-1} & \ddots & \ddots & \ddots & \ddots \\
\sigma_{T-2} & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
\sigma_2 & \ddots & \ddots & \ddots & \sigma_{T-1} \\
\sigma_1 & \sigma_2 & \sigma_{T-2} & \sigma_{T-1} & 0
\end{pmatrix}$$

(13)

Therefore,

$$\sum_{i=1}^{T} \sum_{j=1}^{T} |\Delta \sigma_{i,j}| = 2 \sum_{j=1}^{T-1} |j\sigma_j|.$$  

(14)
Let $\lambda_k$ denote the $k$-th column of the matrix $\Lambda$. For any $(l, m) \in [1, T]^2$, we have:

$$
\left| R^{(l,m)}_\Sigma \right| = \left| \lambda_i' \Delta \Sigma \lambda_m \right|
$$

$$
= \left| \sum_{i=1}^{T} \sum_{j=1}^{T} \lambda_{i,l} \lambda_{j,m} \Delta \sigma_{i,j} \right|
$$

$$
\leq \sum_{i=1}^{T} \sum_{j=1}^{T} \left| \lambda_{i,l} \right| \left| \lambda_{j,m} \right| \left| \Delta \sigma_{i,j} \right|
$$

$$
\leq \sum_{i=1}^{T} \sum_{j=1}^{T} \frac{\sqrt{2}}{\sqrt{T}} \frac{\sqrt{2}}{\sqrt{T}} \left| \Delta \sigma_{i,j} \right|
$$

$$
= \frac{4}{T} \sum_{j=1}^{T-1} |j\sigma_j|.
$$

This implies that:

$$
\| R_\Sigma \|_\infty \leq \frac{4}{T} \sum_{j=1}^{T-1} |j\sigma_j|,
$$

where $\| . \|_\infty$ is the element-wise max norm. The inequality for the weak norm follows from the fact that the weak norm and the element-wise max norm satisfy $\| . \| \leq \sqrt{T} \| . \|_\infty$. ■

# 2 Results on the frequency solution

## 2.1 Derivation of solution 1

To save notation, we suppress the $j$ subscripts indicating frequencies in this section when they are not necessary for clarity. So in this section $f_D$, for example, is a scalar representing the spectral density of fundamentals at some arbitrary frequency (rather than vectors, which is what the unsubscripted variables represent in the main text).

In this section we solve a general version of the model that allows for a constant component of the supply, denoted $s$. This can be thought of as the mean aggregate supply of the underlying. The main results implicitly set $s = 0$, but the analysis of equity returns uses nonzero $s$. We assume that the noise traders’ demand curve depends on prices relative to their mean, so that supply does not enter. This is without loss of generality as it is simply a normalization.
2.1.1 Statistical inference

We guess that prices take the form

\[ p = a_1 d + a_2 z + a_3 s \]  

(17)

where \( s \) is nonstochastic. The joint distribution of fundamentals, signals, and prices is then

\[
\begin{bmatrix}
    d \\
    y_i \\
    p - a_3 s
\end{bmatrix}
\sim N
\left(0, \begin{bmatrix}
    f_D & f_D & a_1 f_D \\
    f_D & f_D + f_i & a_1 f_D \\
    a_1 f_D & a_1 f_D & a_1^2 f_D + a_2^2 f_Z
\end{bmatrix}\right)
\]

(18)

The expectation of fundamentals conditional on the signal and price is

\[
E[d | y_i, p] = \left[ f_D, a_1 f_D \right] \left[ f_D + f_i a_1 f_D a_2^2 f_D + a_2^2 f_Z \right]^{-1} \begin{bmatrix} y_i \\ p - a_3 s \end{bmatrix}
= [1, a_1] \left[ 1 + f_i f_D^{-1} a_1 a_1 a_1^2 + a_2^2 f_Z f_D^{-1} \right]^{-1} \begin{bmatrix} y_i \\ p - a_3 s \end{bmatrix}
\]

(19)

(20)

and the variance satisfies

\[
\tau_i \equiv Var[d | y_i, p]^{-1} = f_D^{-1} \left(1 - \begin{bmatrix} 1 & a_1 \end{bmatrix} \begin{bmatrix} 1 + f_i f_D^{-1} a_1 a_1 a_1^2 + a_2^2 f_Z f_D^{-1} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ a_1 \end{bmatrix} \right)^{-1}
= \frac{a_1^2 f_Z^{-1} + f_i^{-1} + f_D^{-1}}{a_2^2}
\]

(21)

(22)

We use the notation \( \tau \) to denote a posterior precision, while \( f^{-1} \) denotes a prior precision of one of the basic variables of the model. The above then implies that

\[
E[d | y_i, p] = \tau_i^{-1} \left(f_i^{-1} y_i + \frac{a_1}{a_2} a_1 f_Z^{-1} (p - a_3 s)\right)
\]

(23)
2.1.2 Demand and equilibrium

The agent’s utility function is (where variables without subscripts here indicate vectors),

\[ U_i = \max_{\{Q_{i,t}\}} \rho^{-1}E_{0,i} \left[ T^{-1}q^*_i (D - P) \right] - \frac{1}{2} \rho^{-2} \text{Var}_{0,i} \left[ T^{-1/2}q^*_i (D - P) \right] \]  \tag{24}

\[ = \max_{\{Q_{i,t}\}} \rho^{-1}E_{0,i} \left[ T^{-1}q^*_i (d - p) \right] - \frac{1}{2} \rho^{-2} \text{Var}_{0,i} \left[ T^{-1/2}q^*_i (d - p) \right] \]  \tag{25}

\[ = \max_{\{Q_{i,t}\}} \rho^{-1}T^{-1} \sum_j \tilde{q}_{i,j}E_{0,i} \left[ (d_j - p_j) \right] - \frac{1}{2} \rho^{-2}T^{-1} \sum_j \tilde{q}_{i,j}^2 \text{Var}_{0,i} \left[ d_j - p_j \right] \]  \tag{26}

where the last line follows by imposing the asymptotic independence of \( d \) across frequencies (we analyze the error induced by that approximation below). The utility function is thus entirely separable across frequencies, with the optimization problem for each \( \tilde{q}_{i,j} \) independent from all others.

Taking the first-order condition associated with the last line above for a single frequency (with \( \tilde{q}_i, d, \) etc. again representing scalars, for any \( j \)), we obtain

\[ \tilde{q}_i = \rho \tau_i E [d - p \mid y_i, p] \]

\[ = \rho \left( f_i^{-1}y_i + \frac{a_1}{a_2} f^{-1}_z (p - a_3 s) - \tau_i p \right) \]  \tag{27}

\[ = \rho \left( f_i^{-1}y_i + \frac{a_1}{a_2} f^{-1}_z (a_1 d + a_2 z) - \tau_i (a_1 d + a_2 z + a_3 s) \right) \]  \tag{28}

\[ = \rho \left( f_i^{-1}y_i + \frac{a_1}{a_2} f^{-1}_z (a_1 d + a_2 z) - \tau_i (a_1 d + a_2 z + a_3 s) \right) \]  \tag{29}

Summing up all demands and inserting the guess for the price yields

\[ -z + k(a_1 d + a_2 z) + s = \int_i \rho \left( f_i^{-1}y_i + \left( \frac{a_1}{a_2} f^{-1}_z \right) (a_1 d + a_2 z) - \tau_i a_3 s \right) di \]  \tag{30}

\[ = \int_i \rho \left( f_i^{-1}d + \left( \frac{a_1}{a_2} f^{-1}_z \right) (a_1 d + a_2 z) - \tau_i a_3 s \right) di, \]  \tag{31}

where the second line uses the law of large numbers. Matching coefficients on \( d, z, \) and \( s \) then yields

\[ \int_i \rho \left( \frac{a_1}{a_2} f^{-1}_z \right) di = -a_2^{-1} (1 - ka_2) \]  \tag{32}

\[ \int_i \rho f_i^{-1}a_1^{-1} + \rho \left( \frac{a_1}{a_2} f^{-1}_z \right) di = k \]  \tag{33}
and therefore

\[ k - \int_i \rho f_i^{-1} a_1^{-1} = a_2^{-1} (k a_2 - 1) \]  
\[ \int_i \rho f_i^{-1} = \frac{a_1}{a_2} \]

Now define aggregate precision to be

\[ f_{avg}^{-1} \equiv \int_i f_i^{-1} di \]

We then have

\[ \tau_i = \frac{a_1^2 f_i^{-1}}{a_2} + f_i^{-1} + f_D^{-1} \]
\[ \tau_{avg} \equiv \int \tau_idi = (\rho f_{avg}^{-1})^2 f_Z^{-1} + f_{avg}^{-1} + f_D^{-1} \]

Inserting the expression for \( \tau_i \) into (32) yields

\[ a_1 = \frac{\tau_{avg} - f_D^{-1}}{\tau_{avg} + \rho^{-1}k} \]  
\[ a_2 = \frac{a_1}{\rho f_{avg}} \]  
\[ a_3 = -\frac{1}{\rho \tau_{avg}} \]

The expression for \( a_1 \) can be written more explicitly as:

\[ a_1 = \frac{\tau_{avg} - f_D^{-1}}{\tau_{avg} + \rho^{-1}k} = \frac{a_1^2 f_Z^{-1} + f_{avg}^{-1} + f_D^{-1} + \rho^{-1}k - \rho^{-1}k - f_D^{-1}}{a_1^2 f_Z^{-1} + f_{avg}^{-1} + f_D^{-1} + \rho^{-1}k} \]
\[ = 1 - \frac{\rho^{-1}k + f_D^{-1}}{(\rho f_{avg})^{-2} f_Z^{-1} + f_{avg}^{-1} + \rho^{-1}k + f_D^{-1}}. \]

The expression for \( a_2 \) is invalid in the case when \( f_{avg}^{-1} = 0 \). In that case, we have

\[ a_2 = \frac{1}{\rho f_D^{-1} + k}. \]
2.2 Bounding the approximation error

This section considers the case where supply is set to zero, so that $s = 0$. It derives the following result.

**Proposition 1** The difference between solution 1 and the exact Admati (1985) solution is small in the sense that

$$|A_1 - \Lambda \text{diag} (a_1) N'| \leq c_1 T^{-1/2}$$  \hspace{1cm} (46)

$$|A_2 - \Lambda \text{diag} (a_2) N'| \leq c_2 T^{-1/2}$$  \hspace{1cm} (47)

for constants $c_1$ and $c_2$. Furthermore, the variances of the approximation error for prices and quantities are bounded by:

$$|\text{Var} (\Lambda p - P)| \leq c_P T^{-1/2}$$  \hspace{1cm} (48)

$$|\text{Var} (\Lambda \tilde{q}_i - \tilde{Q}_i)| \leq c_Q T^{-1/2}$$  \hspace{1cm} (49)

for constants $c_P$ and $c_Q$.

We use the notation $\tilde{O}$ to mean that, for any matrices $A$ and $B$,

$$|A - B| = \tilde{O} \left( T^{-1/2} \right) \iff |A - B| \leq b T^{-1/2}$$  \hspace{1cm} (50)

for some constant $b$ and for all $T$. This is a stronger statement than typical big-$O$ notation in that it holds for all $T$, as opposed to holding only for some sufficiently large $T$. Standard properties of norms yield the following result. If $|A - B| = \tilde{O} \left( T^{-1/2} \right)$ and $|C - D| = \tilde{O} \left( T^{-1/2} \right)$, then

$$|cA - cB| = \tilde{O} \left( T^{-1/2} \right)$$  \hspace{1cm} (51)

$$|A^{-1} - B^{-1}| = \tilde{O} \left( T^{-1/2} \right)$$  \hspace{1cm} (52)

$$|(A + C) - (B + D)| = \tilde{O} \left( T^{-1/2} \right)$$  \hspace{1cm} (53)

$$|AC - BD| = \tilde{O} \left( T^{-1/2} \right)$$  \hspace{1cm} (54)
In other words, convergence in weak norm carries through under addition, multiplication, and inversion. Following the time domain solution (8), $A_1$ and $A_2$ can be expressed as a function of the Toeplitz matrices $\Sigma_D$, $\Sigma_Z$ and $\Sigma_{avg}$ using those operations. It follows that $|A_1 - \Lambda \text{diag} (a_1) \Lambda'| \leq c_1 T^{-\frac{1}{2}}$ for some constant $c_1$, and the same holds for $A_2$ for some constant $c_2$.

For the variance of prices, we define

$$R_1 \equiv A_1 - \Lambda \text{diag} (a_1) \Lambda',$$  
$$R_2 \equiv A_2 - \Lambda \text{diag} (a_2) \Lambda'.$$  

In what follows, we use the strong norm $\|\cdot\|$, defined as:

$$\|A\| = \max_{x'x = a} (x'A'Ax)^{\frac{1}{2}}.$$  

Finally, we use the following property of the weak norm: for any two square matrices $A, B$ of size $T \times T$,

$$|AB| \leq \sqrt{T} |A| |B|.$$  

The proof for this inequality is standard and reported at the end of this section. We then have the following:

$$|\text{Var} [P - \Lambda p]| = |\text{Var} [(A_1 - \Lambda a_1 \Lambda')D + (A_2 - \Lambda a_2 \Lambda')Z]|$$  
$$\leq |R_1 \Sigma_D R_1'| + |R_2 \Sigma_Z R_2'|$$  
$$\leq \sqrt{T} (|R_1 \Sigma_D||R_1| + |R_2 \Sigma_Z||R_2|)$$  
$$\leq \sqrt{T} \left(\|\Sigma_D\| |R_1|^2 + \|\Sigma_Z\| |R_2|^2\right)$$  
$$\leq \sqrt{T} K (|R_1|^2 + |R_2|^2).$$  

The second line follows from the triangle inequality. The third line comes from property (58). The fourth line uses the fact that for any two square matrices $G, H$, $\|GH\| \leq \|G\| \|H\|$; for a proof, see Gray (2006), lemma 2.3. The last line follows from the assumption that the eigenvalues of $\Sigma_D$ and $\Sigma_Z$ are bounded. Indeed, since $\Sigma_D$ and $\Sigma_Z$ are symmetric and real, they are Hermitian; following
Gray (2006), eq. (2.16), we then have \( \| \Sigma_Z \| = \max_t |\alpha_{Z,t}| \) and \( \| \Sigma_D \| = \max_t |\alpha_{D,t}| \), where \( \alpha_{X,t} \) denotes the eigenvalues of the matrix \( X \).

Given that \( |R_1| \leq c_1 T^{-\frac{1}{2}} \) and \( |R_2| \leq c_2 T^{-\frac{1}{2}} \), this implies:

\[
|Var [P - \Lambda p]| \leq K \sqrt{T} (c_1^2 + c_2^2) T^{-1} = c_1 T^{-\frac{1}{2}}. \tag{64}
\]

\[
|Var [\tilde{Q} - \Lambda \tilde{q}]|.
\]

A similar proof establishes the result for \( |Var [\tilde{Q} - \Lambda \tilde{q}]| \).

To prove inequality (58), note that:

\[
|AB|^2 = \frac{1}{T} \sum_{m,n} \left( \sum_{t=1}^{T} a_{mt} b_{tn} \right)^2
\leq \frac{1}{T} \sum_{m,n} \left( \sum_{t=1}^{T} a_{mt}^2 \right) \left( \sum_{t=1}^{T} b_{tn}^2 \right)
= T \left( \frac{1}{T} \sum_{m,t} a_{mt}^2 \right) \left( \frac{1}{T} \sum_{n,t} b_{nt}^2 \right)
= T |A|^2 |B|^2, \tag{66}
\]

so that \( |AB| \leq \sqrt{T} |A| |B| \). In this sequence of inequalities, going from the second to the third line uses the Cauchy-Schwarz inequality.

**2.3 Proof of lemma 2**

First, since the trace operator is invariant under rotations,

\[
tr \left( \Sigma_i^{-1} \right) = \sum_j f_{i,j}^{-1}. \tag{67}
\]

The information constraint is linear in the frequency-specific precisions. Investors also face a technical constraint that the elements of \( f_{i,j} \) corresponding to paired sines and cosines must have the same value. That is, if \( \lfloor j/2 \rfloor = \lfloor k/2 \rfloor \), then \( f_{i,j} = f_{i,k} \); this condition is necessary for \( \epsilon_{i,t} \) to be stationary.
Inserting the optimal value of \( q_{i,j} \) into the utility function, we obtain

\[
E_{-1} \left[ U_{i,0} \right] = \frac{1}{2} E \left[ T^{-1} \sum_j \tau_{i,j} E \left[ d_j - p_j \mid y_{i,j}, p_j \right]^2 \right]
\]  

(68)

\( U_{i,0} \) is utility conditional on an observed set of signals and prices. \( E_{-1} \left[ U_{i,0} \right] \) is then the expectation taken over the distributions of prices and signals.

\( Var \left[ E \left[ d_j - p_j \mid y_{i,j}, p_j \right] \right] \) is the variance of the part of the return on portfolio \( j \) explained by \( y_{i,j} \) and \( p_j \), while \( \tau_{i,j}^{-1} \) is the residual variance. The law of total variance says

\[
Var \left[ E \left[ d_j - p_j \mid y_{i,j}, p_j \right] \right] = Var \left[ E \left[ d_j - p_j \mid y_{i,j}, p_j \right] \right] + E \left[ Var \left[ d_j - p_j \mid y_{i,j}, p_j \right] \right]
\]

(69)

where the second term on the right-hand side is just \( \tau_{i,j}^{-1} \) and the first term is \( E \left[ E \left[ d_j - p_j \mid y_{i,j}, p_j \right]^2 \right] \) since everything has zero mean. The unconditional variance of returns is

\[
Var(r_j) = Var \left[ d_j - p_j \right] = (1 - a_{1,j})^2 f_{D,j} + \frac{a_{1,j}^2}{\rho f_{avg,j}^{-1}} f_{Z,j}.
\]

(70)

So then

\[
E_{-1} \left[ U_{i,0} \right] = \frac{1}{2} T^{-1} \sum_j \left[ \left( (1 - a_{1,j})^2 f_{D,j} + \frac{a_{1,j}^2}{\rho f_{avg,j}^{-1}} f_{Z,j} \right) \tau_{i,j} - 1 \right].
\]

(71)

We thus obtain the result that agent \( i \)'s expected utility is linear in the precision of the signals that they receive (since \( \tau_{i,j} \) is linear in \( f_{i,j}^{-1} \); see equation 38). Now define

\[
\lambda_j \left( f_{avg,j}^{-1} \right) = (1 - a_{1,j})^2 f_{D,j} + \left( \frac{a_{1,j}}{\rho f_{avg,j}^{-1}} \right)^2 f_{Z,j} = Var(r_j).
\]

(72)

From equations (39)-(40), when \( f_{avg,j}^{-1} > 0 \), \( \lambda_j \) can be re-written as:

\[
\lambda_j \left( f_{avg,j}^{-1} \right) = \frac{f_{D,j} \left( f_{D,j}^{-1} + \rho^{-1}k \right)^2 + (\rho f_{avg,j}^{-1})^2 f_{Z,j}^{-1} + f_{Z,j} f_{avg,j}^{-1}}{\left( (\rho f_{avg,j}^{-1})^2 f_{Z,j}^{-1} + f_{D,j}^{-1} + \rho^{-1}k + f_{avg,j}^{-1} \right)^2},
\]

(73)
which can be further decomposed as:

\[
\lambda_j\left(f_{\text{avg},j}^{-1}\right) = \frac{1}{\left((\rho f_{\text{avg},j}^{-1})^2 f_{Z,j}^{-1}f_{D,j}^{-1} + \rho^{-1}k + f_{\text{avg},j}^{-1}\right)^2} + \frac{1}{\left((\rho f_{\text{avg},j}^{-1})^2 f_{Z,j}^{-1}f_{D,j}^{-1} + \rho^{-1}k + f_{\text{avg},j}^{-1}\right)^2}
\]

(74)

Each of these three terms is decreasing in \( f_{\text{avg},j}^{-1} \), so that the function \( \lambda_j(\cdot) \) is decreasing.

\[
E_{-1}[U_{i,0}] = \frac{1}{2} T^{-1} \sum_j \left[ \frac{a_1^2}{a_2^2} f_{Z}^{-1} + f_{D}^{-1} \right]
\]

(75)

### 2.4 Corollary 2.1

Assume that long-term investors are initially uninformed about the frequency; then \( f_i^{-1} = 0 \), for all \( i \) so:

\[
\tau_i = \left(\frac{a_1}{a_2}\right)^2 f_{Z}^{-1} + f_{D}^{-1}
\]

(76)

Using expression (89) from the published appendix, we then have

\[
\rho^{-1} q_{LF,i} = \left(1 - a_1\right) \left(\frac{a_1}{a_2}\right)^2 f_{Z}^{-1} - a_1 f_{D}^{-1} + \left(\frac{a_1(1-a_1)}{a_2} f_{Z}^{-1} - a_2 f_{D}^{-1}\right) z.
\]

(77)

Given that \( r = (1 - a_1)d - a_2z \) and that \( z \) and \( d \) are independent,

\[
\rho^{-1} E_{-1}[q_{LF,i|r}] = \left(1 - a_1\right) \left(\frac{a_1}{a_2}\right)^2 f_{Z}^{-1} - a_1 f_{D}^{-1} \left(1 - a_1\right) f_{D} - \left(\frac{a_1(1-a_1)}{a_2} f_{Z}^{-1} - a_2 f_{D}^{-1}\right) a_2 f_{Z} + \left(1 - a_1\right)^2 \left(\frac{a_1}{a_2}\right)^2 f_{Z}^{-1} d + 2a_1(1 - a_1) + a_2^2 f_{Z} f_{D}^{-1} - \left(1 - a_1\right) \left(\frac{a_1}{a_2}\right) \left(f_{Z}^{-1} f_{D} \right)^{\frac{1}{2}} - a_2 \left(f_{Z} f_{D}^{-1}\right)^{\frac{1}{2}} f_{Z} f_{D}^{-1}
\]

(78)

\[
(\left(f_{Z}^{-1} f_{D} \right) \left(1 - a_1\right) \left(\frac{a_1}{a_2}\right) - a_2 f_{Z} f_{D}^{-1})^{\frac{1}{2}} \]

(79)

From the equilibrium condition for \( f_{\text{avg},j}^{-1} \) stated in the text, a marginal reduction in \( \alpha_j \) at \( \alpha_j = \lambda_j(0)/\psi \) leads to a marginal increase in \( f_{\text{avg},j}^{-1} \), so the signs of the derivatives with respect to \( \alpha_j \) are simply the reverse of the signs of the derivatives with respect to \( f_{\text{avg},j}^{-1} \). We now calculate derivatives with respect to \( f_{\text{avg},j}^{-1} \).
For any $f_{avg}^{-1} > 0$, where $a_1/a_2 = \rho f_{avg}^{-1}$, the derivative of this expression with respect to $f_{avg}^{-1}$ is

$$
\rho^{-1} \frac{dE_{-1} [\bar{q}_{LF,t}]}{df_{avg}} = 2 \left( (1 - a_1) \left( \frac{a_1}{a_2} \right) (f_{Z}^{-1} f_D)^\frac{1}{2} - a_2 (f_{Z}^{-1} f_D)^\frac{1}{2} \right)
\times \left\{ \rho \left[ (1 - a_1) (f_{Z}^{-1} f_D)^\frac{1}{2} - a_1 (f_{Z}^{-1} f_D)^\frac{1}{2} \right] - \left[ (f_{Z}^{-1} f_D)^\frac{1}{2} + (f_{Z}^{-1} f_D)^\frac{1}{2} \right] \rho \frac{\partial a_1}{\partial f_{avg}} f_{avg}^{-1} \right\}
$$

(80)

Moreover, when $f_{avg}^{-1} > 0$,

$$
\frac{\partial a_1}{\partial f_{avg}} f_{avg}^{-1} = a_1 (1 - a_1) + (1 - a_1) \frac{(\rho f_{avg}^{-1})^2 f_{Z}^{-1}}{(\rho f_{avg}^{-1})^2 f_{Z}^{-1} + f_{avg}^{-1} + f_{D}^{-1} + \rho^{-1} k}.
$$

(81)

The following limits follow from the discussion in Appendix 2.1.2:

$$
\lim_{f_{avg}^{-1} \to 0^+} a_1 = 0, \quad \lim_{f_{avg}^{-1} \to 0^+} a_2 = \frac{1}{\rho f_{D}^{-1} + k}.
$$

(82)

Using these limits and the expressions just derived, we arrive at

$$
\lim_{f_{avg}^{-1} \to 0^+} \frac{\partial E_{-1} [\bar{q}_{LF,t}]}{df_{avg}^{-1}} = -2\rho \frac{(f_{Z}^{-1} f_D)^\frac{1}{2} (f_{Z}^{-1} f_D)^\frac{1}{2}}{f_{D}^{-1} + \rho^{-1} k} < 0.
$$

(83)

Re-introducing the notation $j$, for the frequency at which entry takes place, we then have

$$
\frac{d}{df_{avg,j}^{-1}} E_{-1} \left[ \sum_t \tilde{Q}_{LF,t} (D_t - P_t) \right] = \frac{d}{df_{avg,j}^{-1}} \sum_k E_{-1} [\tilde{q}_{LF,k} r_k] = \frac{d}{df_{avg,j}^{-1}} E_{-1} [\tilde{q}_{LF,j} r_j] < 0;
$$

(84)

that is, all the effect of entry on total profits is concentrated on frequency $j$, where entry reduces profits, as just established.

For the last result, we again use the frequency separability,

$$
\frac{d}{df_{avg,j}^{-1}} E_{-1} [U_{LF,0}] = \frac{d}{df_{avg,j}^{-1}} E_{-1} [u_{LF,0,j}],
$$

(85)

where

$$
E_{-1} [u_{LF,0,j}] = \frac{1}{2} T^{-1} \left[ \left( (1 - a_{1,j})^2 f_{D,j} + a_{2,j}^2 f_{Z,j} \right) \tau_{i,j} - 1 \right]
$$

(86)

is the component of utility from fluctuations at at frequency $j$. This latter definition uses expression (71), derived in Appendix 2.3. Omitting the $j$ notation for clarity, the derivative of this expression
with respect to $f_{\text{avg}}$ assuming that $f_i^{-1} = 0$ is:

$$2T \frac{dE_{-1}[u_{LF,0}]}{df_{\text{avg}}^{-1}} = ((1 - a_1)^2 f_D + a_1^2 (\rho f_{\text{avg}}^{-1})^2 f_Z) 2 \rho^2 f_Z^{-1} f_{\text{avg}}^{-1}$$

$$+ \left( -2(1 - a_1) \frac{\partial a_1}{\partial f_{\text{avg}}} f_D + 2 a_1 \frac{\partial a_1}{\partial f_{\text{avg}}} (\rho f_{\text{avg}}^{-1}) f_Z + 2 a_1^2 \rho^2 f_Z f_{\text{avg}}^{-1} \right) \left( (\rho f_{\text{avg}}^{-1})^2 f_Z^{-1} + f_D^{-1} \right)$$

(87)

Given that:

$$\lim_{f_{\text{avg}}^{-1} \to 0^+} a_1 = 0,$$

(88)

the only term in this expression for which the limit may not be 0 as $f_{\text{avg}}^{-1} \to 0^+$ is:

$$-2(1 - a_1) \frac{\partial a_1}{\partial f_{\text{avg}}} f_D + 2 a_1 \frac{\partial a_1}{\partial f_{\text{avg}}} \rho f_{\text{avg}}^{-1} f_Z.$$

(89)

However, given equation (81), we have that:

$$\lim_{f_{\text{avg}}^{-1} \to 0^+} \frac{\partial a_1}{\partial f_{\text{avg}}} f_{\text{avg}}^{-1} = 0,$$

(90)

and so the second term in (89) goes to 0 as $f_{\text{avg}}^{-1} \to 0^+$. For the second term, note that, using (81) we have that:

$$\frac{\partial a_1}{\partial f_{\text{avg}}} = \frac{a_1}{f_{\text{avg}}^{-1}} + o(1) = \frac{1 + (\rho f_{\text{avg}}^{-1}) f_Z^{-1}}{(\rho f_{\text{avg}}^{-1})^2 f_Z^{-1} + f_D^{-1} + f_{\text{avg}}^{-1} + \rho^{-1} k} + o(1).$$

(91)

Therefore,

$$\lim_{f_{\text{avg}}^{-1} \to 0^+} 2T \frac{dE_{-1}[u_{LF,0}]}{df_{\text{avg}}^{-1}} = -2 \frac{f_D}{f_D^{-1} + \rho^{-1} k} = -2 f_D a_2 < 0,$$

(92)

which proves the last statement of corollary 2.1.

### 2.5 Corollary 2.2

The second inequality follows immediately from the facts proved above that $\frac{d}{df_{\text{avg},j}} \lambda_j \left(f_{\text{avg},j}^{-1}\right) < 0$ and $\lambda_j \left(f_{\text{avg},j}^{-1}\right) = \text{Var} \ (r_j)$. The first inequality follows from the fact proved above that

$$\text{Var} \ [d_j | p_j] = (\rho f_{\text{avg}}^{-1})^2 f_Z^{-1} + f_D^{-1}$$

(93)
3 Results for the calibration

3.1 Calibration

Our goal is to calibrate the model to be consistent with the behavior of aggregate stock market dividends at the annual frequency, and also use information about other major economic time series to provide reasonable values for the spectrum at higher frequencies. The reason that we use the annual frequency for dividends is that there are seasonal effects within the year, in that most dividends are paid quarterly, but in different months.

We first calculate the spectrum of annual dividend growth by calculating the periodogram – the squared Fourier transform – of annual data obtained from CRSP.

To obtain information about high frequencies, we use weekly initial unemployment claims. That series is obviously somewhat removed from dividends, but has the advantage of being perhaps the only economic indicator that is available at such high frequencies. It is used, for example, by the Federal Reserve Bank of Philadelphia’s real-time business conditions index. It is strongly cyclical, and closely related to the unemployment rate, so we use it is as general measure of economic activity. It also has the advantage that its sample periodogram has a highly similar shape to that of dividend growth at the frequencies where they overlap.

To estimate the spectrum we first shift the level of the periodogram for initial claims so that it has the same mean as that of dividend growth at the frequencies where they overlap.

We estimate the true spectrum as a latent variable. We model it with a Gaussian prior for its log such that the covariance between any pair of frequencies is proportional to $\exp(-\phi |\omega_1 - \omega_2|)$, where $\phi$ is a parameter determining the smoothness of the estimated spectrum. The factor of proportionality, denoted $\sigma^2_P$, is the prior variance for the level of the log spectrum (we use the log because the log periodogram is homoskedastic).

The two periodograms yield a pair of samples, $\{X_1, F_1\}$ and $\{X_2, F_2\}$, where the $X$ vectors are the sample Fourier frequencies and the $F$ vectors are the values of the log periodogram. Those two samples are stacked into a pair of large vectors, $\tilde{X}$ and $\tilde{F}$. The prior covariance matrix is then $\Sigma$, where the $i, j$ entry is $\sigma^2_P \exp(-\phi |\tilde{X}_i - \tilde{X}_j|)$. Denote the estimate of the true spectrum as $\hat{F} + b$, where $b$ is a constant.
Technically, the log periodogram is not normal – it is distributed as the log of a $\chi^2/2$. We treat it as normal for simplicity, following a quasi-maximum likelihood approach. Denote the estimated spectrum with the vector $\hat{F}$. Then the quasi-log-likelihood, taking into account the prior and the data likelihood, is

$$-\hat{F} \Sigma^{-1} \hat{F} - \left( \hat{F} + b - \hat{F} \right) \Sigma^{-1} \text{samp} \left( \hat{F} + b \right)$$

(94)

where 1 is a vector of 1’s and $\Sigma_{\text{samp}}$ is the variance matrix of the log spectrum. This is, given basic properties of the periodogram, the variance of a $\chi^2/2$ (see, e.g., Brillinger (1981)).

The first-order condition for $b$ is

$$0 = 1' \Sigma^{-1} \text{samp} \left( \hat{F} + b \right)$$

(95)

$$b = (1' \Sigma^{-1} \text{samp} 1)^{-1} 1' \Sigma^{-1} \text{samp} \left( \hat{F} - \hat{F} \right)$$

(96)

Inserting that into the optimization, the first-order condition for $\hat{F}$ is

$$\max_{\hat{F}} -\hat{F} \Sigma^{-1} \hat{F} - \left( \hat{F} + 1 (1' \Sigma^{-1} \text{samp} 1)^{-1} 1' \Sigma^{-1} \text{samp} \left( \hat{F} - \hat{F} \right) - \hat{F} \right) \Sigma^{-1} \text{samp} \left( \hat{F} + 1 (1' \Sigma^{-1} \text{samp} 1)^{-1} 1' \Sigma^{-1} \text{samp} \left( \hat{F} - \hat{F} \right) - \hat{F} \right)$$

(97)

Yielding

$$\hat{F} = (\Sigma^{-1} + V)^{-1} V \hat{F}$$

(98)

where $V \equiv \left( I + 1 (1' \Sigma^{-1} \text{samp} 1)^{-1} 1' \Sigma^{-1} \text{samp} \right) \Sigma^{-1} \text{samp} \left( I + 1 (1' \Sigma^{-1} \text{samp} 1)^{-1} 1' \Sigma^{-1} \text{samp} \right)$

(99)

Because the set of frequencies, $\hat{X}$, at which we have data is not the same as the set of frequencies in the numerical example, we linearly interpolate from $\hat{F} + b$ to obtain $f_D$.

### 3.2 Calculating returns on dividend strips and equity

As noted in the text, the numerical example uses the case where fundamentals (dividends, in this case) are stationary in first differences (see section 5 for the derivations in that case). Since the model is calibrated to the weekly frequency, a claim on the level of dividends at the end of the first year is a claim to $\sum_{t=1}^{52} \Delta D_t$, where $\Delta$ is the first-difference operator. The futures claims are in this case claims to $\Delta D_t$, with prices $P_t$. The price of the 1-year dividend strip is then $\sum_{t=1}^{52} P_t$. An
n-year dividend future, giving claims to the level of dividends at the end of year n is then a claim to 
\[ \sum_{t=1}^{52n} \Delta D_t \] with price \[ \sum_{t=1}^{52n} P_t \]. Equity is a claim to dividends in each period. It is straightforward 
to show that its final payoff is \[ \sum_{t=1}^{T} (T + 1 - t) \Delta D_t \].

A difficulty with interpreting the returns on these contracts is that they all have different 
maturities. Note that an individual futures return is a single-period return. The per-period return 
on a dividend strip is then just the average return on the individual futures,

\[
R_{period}^C = \frac{\sum_{t=1}^{52n} (\Delta D_t - P_t)}{52n}
\] (100)

The analogous calculation for equity is

\[
R_{period}^{Equity} = \frac{\sum_{t=1}^{T} (T + 1 - t) (\Delta D_t - P_t)}{\sum_{t=1}^{T} (T + 1 - t)}
\] (101)

We calculate per-period return variances similarly. Specifically,

\[
\sigma^2_{period,n} = \frac{\text{var} \left( \sum_{t=1}^{52n} (\Delta D_t - P_t) \right)}{52n}
\] (102)

\[
\sigma^2_{period,Equity} = \frac{\text{var} \left( \sum_{t=1}^{T} (T + 1 - t) (\Delta D_t - P_t) \right)}{\sum_{t=1}^{T} (T + 1 - t)}
\] (103)

These values are all multiplied by 52 to put them into annual terms.

To account for the positive average returns on dividend strips and equity, we give them a 
positive supply in the model (since the sophisticated investors bear that supply, they drive the 
price down and returns up). We assume that there is a unit supply of equity. Since equity has a 
payoff of \[ \sum_{t=1}^{T} (T + 1 - t) \Delta D_t \], that means that the supply of the claim to date-t dividend growth 
is \[ T + 1 - t \].
3.3 Variance of dividend strip returns

In the nonstationary model, the variance of dividend growth in a single period is

\[
\text{Var} (\Delta D_t / P_t) = 1_t \Lambda \text{diag} (f_R) \Lambda' 1_t
\]

(104)

\[
= (\Lambda' 1_t)' \text{diag} (f_R) (\Lambda' 1_t)
\]

(105)

\[
= \sum_j \lambda_{t,j}^2 \text{Var} (d_j - p_j)
\]

(106)

\[
= \lambda_{t,0}^2 \text{Var} (d_0 - p_0) + \lambda_{t,T/2}^2 \text{Var} (d_{\frac{T}{2}} - p_{\frac{T}{2}}) + \sum_{j=1}^{T/2-1} (\lambda_{t,2j}^2 + \lambda_{t,2j+1}^2) \text{Var} (d_j)
\]

(107)

where \( f_R \) is the spectrum of returns, \( f_R \equiv \text{Var} (d_j - p_j), 1_t \) is a vector equal to 1 in its \( t \)th element and zero elsewhere, and \( \lambda_{t,j} \) is the \( j \)th trigonometric transform evaluated at \( t \). This takes advantage of the fact that the variance at the cosine and sine associated with a given frequency must be the same. From here on, we write \( r_j \equiv d_j - p_j \).

More generally, then

\[
\text{Var} \left( \frac{1}{s} \sum_{m=0}^{s-1} R_{t+m} \right) = \frac{1}{s^2} \left( \sum_{m=0}^{s-1} 1_{t+m} \right)' \Lambda \text{diag} (f_R) \Lambda' \left( \sum_{m=0}^{s-1} 1_{t+m} \right)
\]

(108)

\[
= \frac{1}{s^2} \left( \sum_{m=0}^{s-1} \lambda_{t+m,0} \right)^2 f_{R,0} + \frac{1}{s^2} \left( \sum_{m=0}^{s-1} \lambda_{t+m,T/2} \right)^2 f_{R,T/2}
\]

(109)

\[
+ \frac{1}{s^2} \sum_{j=1}^{T/2-1} \left[ \left( \sum_{m=0}^{s-1} \lambda_{t+m,2j} \right)^2 + \left( \sum_{m=0}^{s-1} \lambda_{t+m,2j+1} \right)^2 \right] f_{R,j}
\]

(110)

For \( 0 < j < T/2 \)

\[
\left( \sum_{m=0}^{s-1} \lambda_{t+m,j} \right)^2 + \left( \sum_{m=0}^{s-1} \lambda_{t+m,j} \right)^2 = \sum_{m=0}^{s-1} \sum_{k=0}^{s-1} \frac{2}{T} \left[ \cos (2\pi j (t + m - 1) / T) \cos (2\pi j (t + k - 1) / T) + \sin (2\pi j (t + m - 1) / T) \sin (2\pi j (t + k - 1) / T) \right]
\]

(111)

Now note that

\[
2 \cos (x) \cos (y) + 2 \sin (x) \sin (y) = 2 \cos (x - y)
\]

(112)
So we have
\[
\left( \sum_{m=0}^{s-1} \lambda_{t+m,j} \right)^2 + \left( \sum_{m=0}^{s-1} \lambda_{t+m,j} \right)^2 = \frac{2}{T} \sum_{m=0}^{s-1} \sum_{k=0}^{s-1} \cos \left( \frac{2\pi j}{T} (m - k) \right) 
\]
(113)
\[
= 2 \frac{s}{T} \sum_{m=-s}^{s-1} \frac{s - |m|}{s} \cos \left( \frac{2\pi j}{T} m \right) 
\]
(114)
\[
= 2 \frac{s}{T} \mathcal{F}_s \left( \frac{2\pi j}{T} \right) 
\]
(115)
where \( \mathcal{F}_s \) denotes the \( s \)-th-order Fejér kernel. Note that when \( s = T \), the above immediately reduces to zero, since \( \cos (2\pi j) = 0 \). That is the desired result, as an average over all dates should be unaffected by fluctuations at any frequency except zero. For \( j = 0 \),
\[
\left( \sum_{m=0}^{s-1} f_{t+m,0} \right)^2 = \left( \sum_{m=0}^{s-1} \sqrt{1/T} \right)^2 
\]
(116)
\[
= \left( \frac{1}{T^{1/2}} \right)^2 
\]
(117)
\[
= \frac{s}{T} \mathcal{F}_s (0), 
\]
(118)
since \( \mathcal{F}_s (0) = s \) (technically, this holds as a limit: \( \lim_{x \to 0} \mathcal{F}_s (x) = s \)). For \( j = T/2 \),
\[
\left( \sum_{m=0}^{s-1} f_{t+m,T/2} \right)^2 = \frac{1}{T} \left( \sum_{m=1}^{s} (-1)^m \right)^2 = \begin{cases} \frac{1}{T} \text{ for odd } s \\ 0 \text{ otherwise} \end{cases} 
\]
(119)
\[
= \frac{s}{T} \frac{1}{s} \left( \frac{\sin (s\pi/2)}{\sin (\pi/2)} \right)^2 = \frac{s}{T} \mathcal{F}_s (\pi) 
\]
(120)
So we finally have that
\[
\text{Var} \left( \frac{1}{s} \sum_{m=0}^{s-1} R_{t+m} \right) = \frac{1}{sT} \left( \mathcal{F}_s (0) f_{R,0} + \sum_{j=1}^{T/2-1} \mathcal{F}_s (\omega_j) f_{R,2j} + \mathcal{F}_s (\pi) f_{R,T/2} \right) 
\]
(121)
In the case where fundamentals are difference-stationary, the return on a claim to the level of fundamentals on date \( s \) is exactly \( \sum_{t=1}^{s} R_t \).
3.4 Further numerical results

Figures A.1, A.2, and A.3 give further detail in addition to the results reported in the main text. Figure A.1 replicates figure 2, but replacing the case with frequency-specific information costs with a case where information flows are measured by their entropy rather than precision. Appendix 8.3 described the analysis for that case. Figures A.2 and A.3 report the mean, standard deviation, and Sharpe ratio of the dividend strips and equity in the model with frequency-specific information costs and entropy costs. They also report the average of the values reported for dividend strips across four markets in Binsbergen and Koijen (2017).

4 Public release of information

This section considers a simple extension of the model in which there is a public signal that is revealed on date 0. It has the same structure as the other signals in that it takes the form, at each frequency,

\[ \zeta = d + \varepsilon \zeta \quad (122) \]

\[ \text{Var}(\varepsilon \zeta) = f_\zeta \quad (123) \]

This section examines the effects of varying the precision of that signal, \( f_\zeta^{-1} \).

4.1 Statistical inference

We guess that prices take the form

\[ p = a_1 d + a_2 z + a_\zeta \zeta \quad (124) \]

\( (p - a_\zeta \zeta) / a_1 \) is a signal about the dividend with noise equal to \( (a_2 / a_1) z \), which has variance \( (a_2 / a_1)^2 f_z \). The posterior variance of dividends is then

\[ \tau_i = \frac{a_1^2}{a_2^2} f_z^{-1} + f_t^{-1} + f_\zeta^{-1} + f_D^{-1} \quad (125) \]
and the posterior mean is

\[
E[d - p \mid y_i, p - a_\zeta \zeta] = \tau_i^{-1} a_1^2 f_Z^{-1} (p - a_\zeta \zeta) a_1^{-1} + \tau_i^{-1} f_i^{-1} y_i + \tau_i^{-1} f_\zeta^{-1} \zeta - p
\]  

(126)

It will be useful later to calculate the variance of fundamentals conditional on just observing prices, which is

\[
Var (d \mid p) = \frac{a_2^2}{(a_1 + a_\zeta)^2} f_Z + \frac{a_\zeta^2}{(a_1 + a_\zeta)^2} f_\zeta
\]

(127)

4.2 Demand and equilibrium

Agent \(i\)'s demand is

\[
\tilde{q}_i = \rho \tau_i E[d - p \mid y_i, p]
\]

(128)

\[
= \rho \left( \frac{a_1^2}{a_2^2} f_Z^{-1} (p - a_\zeta \zeta) a_1^{-1} + f_i^{-1} y_i + f_\zeta^{-1} \zeta - \tau_i p \right)
\]

(129)

\[
= \rho \left( \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) p + f_i^{-1} y_i + \left( f_\zeta^{-1} - \frac{a_1}{a_2} a_\zeta f_Z^{-1} \right) \zeta \right)
\]

(130)

Summing up all demands and inserting the guess for the price yields

\[
-z + k (a_1 d + a_2 z + a_\zeta \zeta) = \rho \int (\frac{a_1}{a_2} f_Z^{-1} - \tau_{avg}) p + f_i^{-1} y_i + \left( f_\zeta^{-1} - \frac{a_1}{a_2} a_\zeta f_Z^{-1} \right) \zeta) \, di
\]

(131)

\[
= \rho \left( \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_{avg} \right) (a_1 d + a_2 z + a_\zeta \zeta) + f_{avg}^{-1} d + \left( f_\zeta^{-1} - \frac{a_1}{a_2} a_\zeta f_Z^{-1} \right) \zeta \right)
\]

(132)

where the second line uses the law of large numbers. Matching coefficients on \(d\), \(z\), and \(\zeta\) then yields

\[
k = \rho \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_{avg} \right) + \rho f_{avg}^{-1} a_1^{-1}
\]

(133)

\[
-a_2^{-1} (1 - ka_2) = \rho \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_{avg} \right)
\]

(134)

\[
k = \rho \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_{avg} \right) + \rho \left( a_\zeta^{-1} f_\zeta^{-1} - \frac{a_1}{a_2} f_Z^{-1} \right)
\]

(135)

\[
a_\zeta = f_\zeta^{-1} a_1 \left( f_{avg}^{-1} + \frac{a_2^2}{a_2^2} f_Z^{-1} \right)^{-1}
\]

(136)
\[
\frac{a_1}{a_2} = \rho f_{\text{avg}}^{-1}
\]  
(137)

$$a_1 = \frac{(\rho f_{\text{avg}}^{-1})^2 f_{Z}^{-1} + f_{\text{avg}}^{-1}}{\tau_{\text{avg}} + \rho^{-1} k}$$  
(138)

which implies

$$a_\zeta = \frac{f_{\zeta}^{-1}}{\tau_{\text{avg}} + \rho^{-1} k}$$  
(139)

### 4.3 Utility and profits

Utility, as before, is equal to the variance of returns multiplied by precision,

$$E_{-1} [u_{i,0}] = \lambda_j \left( f_{\text{avg}}^{-1}, f_{\zeta}^{-1} \right) \left( (\rho f_{\text{avg}}^{-1})^2 f_{Z}^{-1} + f_{i}^{-1} + f_{D}^{-1} + f_{\zeta}^{-1} \right) - 1$$  
(140)

where \( \lambda_j \left( f_{\text{avg}}^{-1}, f_{\zeta}^{-1} \right) \) is the variance of returns, and we write it as a function of \( f_{\zeta}^{-1} \) since that is a choice variable of a regulator in this case.

It is straightforward to show that average profits are also linear in \( \lambda_j \left( f_{\text{avg}}^{-1}, f_{\zeta}^{-1} \right) \left( (\rho f_{\text{avg}}^{-1})^2 f_{Z}^{-1} + f_{i}^{-1} + f_{D}^{-1} + f_{\zeta}^{-1} \right) \), so results on utility will map directly into results on profits (with appropriate adjustments for the cost of information).

### 4.4 Results

#### 4.4.1 Limits and noise trader profits

The main text considers a limit as the information of investors becomes infinite. Here, that would correspond to setting \( f_{\zeta}^{-1} \to \infty \). That immediately implies \( \tau_{\text{avg}} \to \infty, a_1 \to 0, a_2 \to 0, \) and \( a_\zeta \to 1 \). The analog to the first two limits from section 4.4.2 in this case is that prices are perfectly informative in the sense that they depend just on fundamentals, since when \( f_{\zeta}^{-1} \to \infty, \zeta = d, \) and hence \( p = d \).

Noise trader profits are,

$$E [(z - k ((a_1 + a_\zeta) d + a_2 z + a_\zeta \xi)) (1 - a_1 - a_\zeta) d - a_2 z - a_\zeta \xi)]$$  
(141)

$$= -a_2 f_{Z} - k (a_1 + a_\zeta) (1 - a_1 - a_\zeta) f_{D} + ka_2^2 f_{Z} + ka_2^2 f_{\zeta}$$  
(142)
So when $a_1 = 0$, $a_2 = 0$, and $a_\zeta = 1$, noise trader losses are zero, yielding the third limit from section 4.4.2. Note, again, that this is the opposite of average profits of informed investors, so when $f_\zeta^{-1} \to \infty$, the average profits of informed investors also go to zero.

### 4.4.2 Information acquisition, profits, and utility

The profits and utility of uninformed investors – the long-term investors in the example in the text, are linear in

\[
\lambda_j \left( f_{avg}^{-1}, f_\zeta^{-1} \right) \left( (\rho f_{avg}^{-1})^2 f_Z^{-1} + f_D^{-1} + f_\zeta^{-1} \right)
\]  

(143)

We are interested in changes in $f_\zeta$, which will affect $f_{avg}^{-1}$ in equilibrium. When there is any positive amount of information acquisition, we have $\lambda_j \left( f_{avg}^{-1}, f_\zeta^{-1} \right) = \psi_j$. Taking a total derivative with respect to $f_\zeta^{-1}$ (or just invoking the implicit function theorem) yields

\[
\frac{df_{avg}^{-1}}{df_\zeta^{-1}} = -\frac{\lambda_{j,2} \left( f_{avg}^{-1}, f_\zeta^{-1} \right)}{\lambda_{j,1} \left( f_{avg}^{-1}, f_\zeta^{-1} \right)}
\]  

(144)

where $\lambda_{j,k}$ denotes the derivative of $\lambda_j$ with respect to its $k$th argument.

The derivative of profits when information is being acquired ($\lambda_j \left( f_{avg}^{-1}, f_\zeta^{-1} \right) = \psi_j$) is then

\[
\frac{d}{df_\zeta^{-1}} \left[ \lambda_j \left( f_{avg}^{-1}, f_\zeta^{-1} \right) \left( (\rho f_{avg}^{-1})^2 f_Z^{-1} + f_D^{-1} + f_\zeta^{-1} \right) \right] = \psi_j \left( 2 \rho^2 f_{avg}^{-1} f_Z^{-1} \frac{df_{avg}^{-1}}{df_\zeta^{-1}} + 1 \right)
\]  

(145)

When information is not being acquired, $f_{avg}^{-1} = 0$, the derivative becomes

\[
\lambda \left( 0, f_\zeta^{-1} \right) + \left( f_D^{-1} + f_\zeta^{-1} \right) \lambda_{j,2} \left( 0, f_\zeta^{-1} \right)
\]  

(146)

We have the following results

1. Information acquisition is weakly decreasing in $f_\zeta^{-1}$.
   - This result is obtained by simply showing that $\lambda_{j,2} \left( f_{avg}^{-1}, f_\zeta^{-1} \right) < 0$.
2. The profits and utility of passive investors increase in disclosure when $f_{avg}^{-1} > 0$.
   - This involves simply confirming that (145) is positive.
3. The utility of all sophisticated investors increase in $f_\zeta^{-1}$ when $f_{avg}^{-1} > 0$. 

– This follows directly from the second result.

4. The average profits of sophisticated investors, and the losses of noise traders, decrease in $f^{-1}$ when $f^{-1}_{avg} > 0$.

– The derivative of the profits of the average investor, who has signal precision $f^{-1}_{avg}$, is

$$\psi_j \left( (2\rho^2 f^{-1}_{avg} f^{-1}_Z + 1) \frac{df^{-1}_{avg}}{df^{-1}_\zeta} + 1 \right)$$

(147)

which can be shown to be negative

5. The losses of noise traders and the utility of sophisticated investors converge to zero as $f^{-1}_{\zeta} \to \infty$.

– See the previous section.

6. The total precision for fundamentals in public information – prices and the public signal – increases in $f^{-1}_{\zeta}$.

– Public precision is $(\rho f^{-1}_{avg})^2 f^{-1}_Z + f^{-1}_D + f^{-1}_{\zeta}$. The derivative with respect to $f^{-1}_{\zeta}$ is

$$2\rho^2 f^{-1}_{avg} f^{-1}_Z \frac{df^{-1}_{avg}}{df^{-1}_\zeta} + 1$$

(148)

which is the same as the derivative used for the second result. When $f^{-1}_{avg} = 0$, the result holds trivially.

4.5 Numerical example

We consider a simple numerical example with $\rho = k = f^{-1}_D = f^{-1}_Z = 1$ and $\psi = 0.4757$ and examine how profits, utility, and price informativeness vary with $f^{-1}_{\zeta}$. The four panels of figure A.4 plot results from a numerical solution, with $f^{-1}_{\zeta}$ varying along the x-axis. Note that the scales are generally in logs.

The top-left panel plots the profits of the various agents. The dotted line is the expected profits for uninformed sophisticated investors. They initially benefit as information is released publicly since it reduces their informational disadvantage compared to more highly informed agents (at $f^{-1}_{\zeta} = 0$, their profits are not zero, just numerically very small). Eventually $f^{-1}_{\zeta}$ rises sufficiently high that $f^{-1}_{avg} = 0$. At that point, more precision for public signals just makes prices more infor-
mative and reduces the profits of all sophisticated investors.

The solid line in the top-left panel plots the average profits of a sophisticated investor with the average level of precision, $f_{avg}^{-1}$. Their profits fall as $f_{avg}^{-1}$ rises because they acquire less information – $f_{avg}$ falls. Since profits are zero sum, as their average profits fall, the (negative) average profits of the noise traders rise – they lose less money.

The bottom-left panel of figure A.4 plots $f_{avg}^{-1}$. It shows that increases in the precision of the public signal reduce incentives for agents to acquire information.

The top-right panel shows that utility initially increases with the public signal – agents are able to trade with the noise traders facing less risk (since they are better informed about fundamentals) without having to pay for private signals. Eventually, though, when there is sufficient information, prices become so efficient that profits and hence utility fall, eventually to the point where there are no profits to be earned.

Finally, the bottom-right panel of figure A.4 reports the information available to investors, either purely from prices or from combining prices and the public signal. In both cases, we see that they rise as the public signal becomes more precise.

5 Results when fundamentals are difference-stationary

In the main text, we assume that the level of fundamentals is stationary. Here we examine an extension in which fundamentals are stationary in terms of first differences and show that the results go through nearly identically, with the primary difference being in how the long-term portfolio is defined.

5.1 Informed investors under difference stationarity

We assume that $D_0$ is known to investors when making decisions, and without loss of generality normalize $D_0 = 0$. Define $\Delta$ to be the first difference operator so that

$$\Delta D_t = D_t - D_{t-1}$$

(149)
and define the vector $\Delta \mathbf{D} \equiv [\Delta D_1, \Delta D_2, \ldots \Delta D_T]'$. We assume that

$$\Delta \mathbf{D} \sim N(0, \Sigma_D).$$

(150)

For any given allocation to the futures contracts, there is an allocation to claims on $\Delta \mathbf{D}$ that gives an identical payoff. Specifically, an allocation $Q_i' \Delta \mathbf{D}$ can be exactly replicated by

$$Q_i' \Delta \mathbf{D} = Q_i' \mathbf{L}_1 \Delta \mathbf{D}$$

(151)

$$= (\mathbf{L}_1 Q_i)' \Delta \mathbf{D}$$

(152)

where $\mathbf{L}_1$ is a matrix that creates partial sums,

$$\mathbf{L}_1 = \begin{bmatrix}
1 & 0 & 0 & \cdots \\
1 & 1 & 0 & \\
1 & 1 & 1 & \\
\vdots & & & \\
\end{bmatrix}$$

(153)

So an allocation of $Q_i$ to the futures is equivalent to an allocation of $\mathbf{L}_1 Q_i$ to claims on the first differences of fundamentals, which we will call the growth rate futures. Define the notation

$$Q_{\Delta,i} \equiv \mathbf{L}_1 Q_i$$

(154)

Furthermore, the prices of the growth rate futures are simply the vector $\Delta \mathbf{P}$ (by the law of one price). We can therefore rewrite the optimization problem equivalently as

$$\max T^{-1} \sum_{t=1}^{T} \beta_t Q_{\Delta,i,t} E_{0,i} [\Delta D_t - \Delta P_t] - \frac{1}{2} \left( \rho T^{-1} \right) \text{Var}_{0,i} \left[ \sum_{t=1}^{T} \beta_t Q_{\Delta,i,t} (\Delta D_t - \Delta P_t) \right]$$

(155)

Now suppose for the moment that we are able to solve the entire model in terms of first differences (that is not obvious as we will need to ensure that noise trader demand is also difference stationary). So we have an allocation $Q_{\Delta D,i}$. An allocation to the first differences is then equivalent to an allocation of $(\mathbf{L}_1')^{-1} Q_{\Delta,i}$ to the levels (which follows trivially from the definition of $Q_{\Delta,i}$ in
Since our maintained assumption is that we will solve the model in first differences in the same way we did in the main text for levels, that means that we will continue to use the rotation $\Lambda$, but now in first differences. So the frequency domain allocations in terms of first differences will be

$$\tilde{Q}_{\Delta,i,t} = \Lambda \tilde{q}_{\Delta,i}$$ (156)

where $\tilde{Q}_{\Delta,i,t} \equiv Q_{\Delta,D,i,t}\beta^t$. $\tilde{q}_{\Delta,i}$ now represents the allocations to different frequencies of growth in fundamentals. The key question, then, is what that implies for the behavior of portfolios in terms of levels. We have

$$\tilde{Q}_i = (L_1')^{-1} \tilde{Q}_{\Delta,i}$$ (157)
$$= (L_1')^{-1} \Lambda \tilde{q}_{\Delta,i}$$ (158)

So in terms of levels, the basis vectors, instead of being $\Lambda$, are $(L_1')^{-1} \Lambda$.

For $(L_1')^{-1}$ we have

$$(L_1')^{-1} = \begin{bmatrix}
1 & -1 & 0 & \cdots & 0 \\
0 & 1 & -1 & \ddots & \\
0 & 0 & 1 & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & -1 \\
0 & \cdots & 0 & 0 & 1
\end{bmatrix}$$ (159)

So the way that $(L_1')^{-1}$ transforms a matrix is to take a forward difference of each column, and then retaining the value of the final row. A way to see the implications of that transformation is to approximate the finite differences of the sines and cosines as derivatives. The columns of $(L_1')^{-1} \Lambda$ are equal to $(L_1')^{-1} c_j$ and $(L_1')^{-1} s_j$, which can be written using standard trigonometric formulas.
as:

\[ (L'_1)^{-1} c_j \approx \begin{bmatrix} \frac{2 \sin \left( \frac{1}{2} \omega_j \right)}{\sqrt{\frac{T}{2}}} \left( \sin \left( \frac{\omega_j (t - \frac{1}{2})}{2} \right) \right)_{t=2}^T \\
\frac{\sqrt{\frac{T}{2}} \cos \left( \omega_j (T - 1) \right)}{} \end{bmatrix} \]  

(160)

\[ (L'_1)^{-1} s_j \approx \begin{bmatrix} \frac{-2 \sin \left( \frac{1}{2} \omega_j \right)}{\sqrt{\frac{T}{2}}} \left( \cos \left( \frac{\omega_j (t - \frac{1}{2})}{2} \right) \right)_{t=2}^T \\
\frac{\sqrt{\frac{T}{2}} \sin \left( \omega_j (T - 1) \right)}{} \end{bmatrix} \]  

(161)

The column \( c_j \) represents a portfolio in terms of the first differences of fundamentals with weights equal to a cosine fluctuating at frequency \( \omega_j \). \((L'_1)^{-1} c_j \) measures the loadings of that portfolio on claims to the level of fundamentals. These loadings also fluctuate at frequency \( \omega_j \), with the only difference being the replacement of the cosine with a sine function. (Intuitive, the loadings are approximately equal to the derivative of the columns of \( \Lambda \) with respect to time; taking derivatives does not affect the characteristic frequency of fluctuations.)

So consider a relatively short-term investor, whose portfolio weights are all close to zero except for a large value in the vector \( q_{\Delta,i} \) at some large value of \( j \). By assumption, that investor holds a portfolio whose loadings on the first differences of fundamentals fluctuate at frequency \( \omega_j \). What the approximations in (160–161) show, though, is that that investor’s positions measured in terms of the level of fundamentals (i.e. \( \tilde{Q}_i \)) has loadings that also fluctuate at frequency \( \omega_j \).

One subtlety is in the lowest-frequency portfolio, \((L'_1)^{-1} \left( \frac{1}{\sqrt{T}} c_0 \right) \). That portfolio puts equal weight on growth in fundamentals on all dates – it is a bet on the sample mean growth rate. In terms of levels, note that \((L'_1)^{-1} \left( \frac{1}{\sqrt{T}} c_0 \right) = \left[ 0, 0, 0, ..., \sqrt{2/T} \right] \). A person who wants to bet on the mean growth rate between dates 1 and \( T \) can do that by buying a claim to fundamentals only on date \( T \).

### 5.2 Noise traders under difference stationarity

Last, we need to show that noise trader demand will also take a form such that the entire model can be solved in terms of first differences (and then shifted back into levels for interpretation). First, as above, since the model expressed in first differences is just a linear transformation of the levels,
the noise traders’ optimization problem can be written in terms of first differences,

$$\max T^{-1} \sum_{t=1}^{T} \beta^t N_{\Delta,t} E_{0,N} [\Delta D_t - \Delta P_t] - \frac{1}{2} (\rho T^{-1}) \text{Var}_{0,N} \left[ \sum_{t=1}^{T} \beta^t N_{\Delta,t} (\Delta D_t - \Delta P_t) \right]$$  \hspace{1cm} (162)

where $N_{\Delta,t}$ is the demand of the noise traders for the claims on first differences.

We assume that the noise traders understand that fundamentals have a unit root and that they therefore have priors and signals that refer to the change in fundamentals. The analogs to (52) and (53) are then

\[
\Delta D \sim N \left( 0, \Sigma_{N\Delta}^{\text{prior}} \right) \\
S \sim N \left( \Delta D, \Sigma_{N\Delta}^{\text{signal}} \right)
\]  \hspace{1cm} (163) (164)

and the Bayesian update is

\[
\Delta D \mid S \sim N \left( \Sigma_{N\Delta} \left( \Sigma_{N\Delta}^{\text{signal}} \right)^{-1} S, \Sigma_{N\Delta} \right)
\]  \hspace{1cm} (165)

where \( \Sigma_{N\Delta} = \left( \left( \Sigma_{N\Delta}^{\text{signal}} \right)^{-1} + \left( \Sigma_{N\Delta}^{\text{prior}} \right)^{-1} \right)^{-1} \)  \hspace{1cm} (166)

6 Quadratic trading costs

The restriction that investors have exactly zero exposure at certain frequencies is a natural one to study in the model. But there are other ways of imposing limits on investors’ exposures across frequencies. This appendix examines the equilibrium when there are quadratic costs of trading. Relative to the frictionless benchmark, introducing these costs has analogous effects to the more abstract restriction \( q_{i,j} = 0 \) for \( j \in \mathcal{R} \). Changes in trading costs could be caused either by the imposition of a quadratic tax on shares traded (i.e. a particular form of a Tobin tax), or by changes in the trading technology. The proofs for this section follow in appendix 7.

The model does not literally have trade over time. However, the exposures that investors choose in the futures market can be replicated through a commitment to trade (at a fixed price) the fundamental on future dates. That is, define a date-\( t \) equity claim to be an asset that pays dividends equal to the fundamental on each date from \( t+1 \) to \( T \). Since the futures contracts involve
exchanging money only at maturity, the date-\(t\) cost of an equity claim is \(P_t^{\text{equity}} = \sum_{j=1}^{T-t} \beta^{-j} P_{t+j}\).

An investor’s exposure to fundamentals on date \(t\), \(Q_{i,t}\), can be acquired either by buying \(Q_{i,t}^{\text{EQ}}\) units of forwards on date 0 or by holding \(Q_{i,t}^{\text{EQ}}\) units of equity entering date \(t\). In the latter case, the volume of trade by investor \(i\) would be equal to the change in \(Q_{i,t}\) over time. That is, \(\Delta Q_{i,t}^{\text{EQ}} = \Delta Q_{i,t}\).

We assume that investors now maximize the following objective:

\[
U_{0,i} = \max_{\{Q_{i,t}\}} E_{0,i} \left[ T^{-1} \sum_{t=1}^{T} Q_{i,t} (D_t - P_t) - \frac{1}{2} cT^{-2} E_{0,i} [QV \{Q_t\}] - \frac{1}{2} bT^{-2} E_{0,i} \left[ \sum_{t=1}^{T} Q_{i,t}^{2} \right] \right],
\]

where \(b > 0\) is a cost of holding large positions in the assets, \(c \geq 0\) is a cost incurred from quadratic variation in positions, with quadratic variation defined as:

\[
QV \{Q_t\} \equiv \left[ \sum_{t=2}^{T} (Q_{i,t} - Q_{i,t-1})^2 + (Q_{i,1} - Q_{i,T})^2 \right].
\]

The term involving \(b\) in (167) replaces the aversion to variance in the benchmark setting. That change is made for the sake of tractability, but its economic consequences are minimal (see, e.g., Kasa, Walker, and Whiteman (2013)). We also set discount rates to zero here to maintain tractability.

Appendix 7 shows that:

\[
T^{-1}QV \{Q_t\} = 2 \sum_{j=1}^{T} \sin^2 \left( \frac{\omega_{|j/2|}}{2} \right) q_{i,j}^2.
\]

Note that we have defined quadratic variation as the sum of the squared changes in \(Q_{i,t}\) between \(t = 2\) and \(T\) plus \((Q_{i,1} - Q_{i,T})^2\). Without the final term, there would be no cost to investors of entering and exiting very large positions at the beginning and end of the investment period. This term helps account for that, and has the added benefit of yielding the simple closed-form expression in the frequency domain reported above. The right-hand side shows that the quadratic variation in the volume induced by an investor depends on their squared exposures at each frequency scaled by \(\sin^2 \left( \frac{\omega_{|j/2|}}{2} \right)\), which rises from 0 to 1 as \(j\) rises. Intuitively, when \(c > 0\), holding exposure to higher frequency fluctuations in fundamentals is more costly because it requires more frequent portfolio rebalancing.
The equilibrium of the model is described in detail in Appendix 7. Here, we highlight key results and explain how they relate to the previous results on restricting trade frequencies.

**Result 1** When \( c > 0 \), all else equal, investors’ equilibrium signal precision is higher at lower frequencies.

With the assumption of fixed quadratic trading costs, the marginal benefit of increasing precision at frequency \( j \) is given by:

\[
\frac{1}{2} (c \sin^2 (\omega_j/2) + b)^{-1} \text{Var} \left[ d_j \mid p_j, y_i, j \right]^2.
\]

(170)

In particular, it is declining with both the signal precision and the frequency of exposure. Given that the marginal cost of information is the same across frequencies, investors choose higher signal precisions at lower frequencies, all else equal.

The main result regarding the effect of the quadratic trading cost is the following.

**Result 2** A small increase in trading costs, when starting from zero, reduces information acquisition at all frequencies except frequency 0. The effect is larger at higher frequencies. As a corollary, the effect of an increase in trading costs on price informativeness is weaker at longer horizons.

The first part of this result suggests that if the goal is to reduce short-term investment, then a quadratic tax is a more blunt instrument than placing an explicit restriction on investment at targeted frequencies. A tax on volume affects all investors, regardless of the strategy that they follow. However, the second part of the result says that trading costs affect short-term strategies most strongly. The quadratic cost thus leads, endogenously, to the same changes in information acquisition studied in the main model; namely, the variance of dividends conditional on prices, \( \text{Var}(d_j|p_j) \), rises more at higher frequencies. The corollary regarding price informativeness refers to the fact that the variance of moving averages of the form:

\[
\text{Var} \left( \frac{1}{n} \sum_{m=0}^{n-1} D_{t+m} \mid P \right)
\]

(171)

increases less as a result of the increase in trading costs for longer horizons \( n \). In the extreme case of \( n = T \), which corresponds to the frequency 0 component of the signals, the increase in trading
costs has in fact no effect on equilibrium signal precision and thus price informativeness. This can be seen from the expression for the marginal benefit of signal precision above, which is independent of $c$ when $j = 0$.

Finally, to examine the effects of trading costs on noise trader profits, we have

**Result 3** Prices continue to take the form

$$p_j = a_{1,j}d_j + a_{2,j}z_j$$  \hspace{1cm} (172)

At all frequencies, increases in trading costs weakly reduce $a_{1,j}$ and strictly increase $a_{2,j}$ (except at frequency zero, where they have no effect).

Again, an increase in trading costs is broadly similar to a restriction on investment in the sense that it makes markets less liquid and prices less informative. By liquid what we mean is that an exogenous demand shock – an increase in $z_j$ – has a larger effect on prices when trading costs are larger. This policy can therefore reduce the losses of noise traders by reducing their overall trade with the informed investors, but again at the cost of less informative prices. As above, if one has evidence that $f_Z$ is large relative to $f_D$ at high frequencies, then this trade-off may be favorable. There is not much to learn about, so losing information has relatively low costs, and since the sentiment shocks are large, inhibiting them is particularly valuable.

Thus, overall, the message of the model with quadratic costs is consistent with the previous analysis. Increasing trading costs leads to less informed trading and the effect is tilted toward high frequencies; at lower frequencies, information acquisition decisions are less impacted. As a result, the effect of the increase on the informativeness of prices for fundamentals in the long run is limited.
7 Quadratic costs proofs

7.1 Frequency domain expressions for trading costs

Using $Q_i = \Lambda q_i$, each agent’s position at time $t$ can be written as

$$Q_{i,t} = \sum_j \begin{bmatrix} q_j \cos \left(2\pi j t/T \right) \\ + q_j \sin \left(2\pi j t/T \right) \end{bmatrix}.$$  

(173)

Trading costs are then written in terms of $(Q_{i,t} - Q_{i,t-1})^2$ as:

$$QV \{Q_i\} \equiv \sum_{t=2}^T (Q_{i,t} - Q_{i,t-1})^2 + (Q_{i,1} - Q_{i,T})^2.$$  

(174)

We can write that as

$$QV \{Q_i\} = (DQ)' (DQ)$$  

(175)

where $D$ is a matrix that generates first differences,

$$D \equiv \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \\ 1 & 0 & \cdots & 0 & -1 \end{bmatrix}.$$  

(176)

Using again the fact that $Q_i = \Lambda q_i$,

$$QV \{Q_i\} = q' \Lambda' D' D \Lambda q$$  

(177)

In what follows, we will need to evaluate the matrix $\Lambda' D' D \Lambda$. The $m,n$ element of that matrix is the inner product of the $m$ and $n$ columns of $D \Lambda$. Each column of $D \Lambda$ contains the first difference of the corresponding column of $\Lambda$, with the exception of the last element, $(D \Lambda)_{m,T}$, which is equal
to $\Lambda_{m,t} - \Lambda_{n,T}$. We have the following standard trigonometric results: for $m \neq n$:

$$
\sum_{t=1}^{T} (\cos (\omega_m t) - \cos (\omega_m (t-1))) (\cos (\omega_n t) - \cos (\omega_n (t-1))) = 0, \quad (178)
$$

$$
\sum_{t=1}^{T} (\cos (\omega_m t) - \cos (\omega_m (t-1))) (\sin (\omega_n t) - \sin (\omega_n (t-1))) = 0, \quad (179)
$$

$$
\sum_{t=1}^{T} (\sin (\omega_m t) - \sin (\omega_m (t-1))) (\sin (\omega_n t) - \sin (\omega_n (t-1))) = 0, \quad (180)
$$

where recall that $\omega_m = \frac{2\pi m}{T}$, and:

$$
\sum_{t=1}^{T} (\cos (\omega_m t) - \cos (\omega_m (t-1)))^2 = 2T \sin^2 (\omega_m / 2), \quad (181)
$$

$$
\sum_{t=1}^{T} (\sin (\omega_m t) - \sin (\omega_m (t-1)))^2 = 2T \sin^2 (\omega_m / 2), \quad (182)
$$

$$
\sum_{t=1}^{T} (\cos (\omega_m t) - \cos (\omega_m (t-1))) (\sin (\omega_m t) - \sin (\omega_m (t-1))) = 0. \quad (183)
$$

These results immediately imply that the off-diagonal elements of $\Lambda'D'D\Lambda$ are equal to zero and the $j$th element of the main diagonal is $2T \sin^2 (\omega_{j/2} / 2)$. We then have

$$
QV \{Q_i\} = q\Lambda'D'D\Lambda q
$$

$$
= \sum_{j=1}^{T} 2T \sin^2 (\omega_{(j/2) / 2}) q_{i,j}^2, \quad (185)
$$

Total holding costs can be written as:

$$
\sum_{t=1}^{T} Q_t^2 = \sum_{j=1}^{T} q_j^2, \quad (186)
$$

which is just Parseval’s theorem.
7.2 Equilibrium of the trading cost model

Throughout the analysis, unless it is necessary, we omit the index \( j \) of the particular frequency in order to simplify notation.

7.2.1 Investment and equilibrium

The first-order condition for frequency \( j \) is

\[
0 = E[d_j - p_j | y_{i,j}, p_j] - 2c \sin^2 \left( \omega_{t/2} / 2 \right) q_j - b q_j
\]

(187)

\[
q = \frac{E[d_j - p_j | y_{i,j}, p_j]}{\gamma_j}
\]

(188)

\[
= \gamma_j^{-1} \tau_i^{-1} \left( f_i^{-1} y_i + \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) p \right)
\]

(189)

where

\[
\gamma_j \equiv 2c \sin^2 \left( \omega_{t/2} / 2 \right) + b
\]

(190)

is the marginal cost of \( q_j \). We can then solve for the coefficients \( a_1 \) and \( a_2 \) as before.

Inserting the formula for the conditional expectation and integrating across investors yields

\[
\int_i \gamma_j^{-1} \tau_i^{-1} \left( f_i^{-1} y_i + \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) (a_1 d - a_2 z) \right) di = z_j
\]

(191)

\[
\int_i \gamma_j^{-1} \tau_i^{-1} \left( f_i^{-1} d + \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) (a_1 d - a_2 z) \right) di = z_j
\]

(192)

Matching coefficients then yields

\[
\int_i \gamma_j^{-1} \tau_i^{-1} \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) di = -a_2^{-1}
\]

(193)

\[
\int_i \gamma_j^{-1} \tau_i^{-1} \left( f_i^{-1} + \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_i \right) a_1 \right) di = 0
\]

(194)

Combining those two equations, we obtain

\[
\int_i \gamma_j^{-1} \tau_i^{-1} f_i^{-1} di = \frac{a_1}{a_2}
\]

(195)

Now put the definition of \( \tau_i \) into that equation for \( f_i^{-1} \)
\[ \int_i \gamma_j^{-1} \tau_i^{-1} \left( \tau_i \frac{a_1^2}{a_2^2} f_Z^{-1} - f_D^1 \right) \, di = \frac{a_1}{a_2} \] (196)

\[ \gamma_j^{-1} \int_i 1 - \left( \frac{a_1^2}{a_2^2} f_Z^{-1} - f_D^1 \right) \tau_i^{-1} \, di = \frac{a_1}{a_2} \] (197)

### 7.2.2 Expected utility

At any particular frequency,

\[ U_{i,j} = q_{i,j} E_{0,i} [d_j - p_j] - \frac{1}{2} q_{i,j}^2 2c \sin^2 (\omega_{i,j}/2) - \frac{1}{2} b q_{i,j}^2 \] (198)

\[ = \frac{1}{2} E [d_j - p_j \mid y_{i,j}, p_j]^2 \] (199)

Expected utility prior to observing signals is then

\[ EU_{i,j} = \frac{1}{2} E \left[ \frac{E [d_j - p_j \mid y_{i,j}, p_j]^2}{\gamma_j} \right] \] (200)

\[ E \left[ E [d_j - p_j \mid y_{i,j}, p_j]^2 \right] \] is the variance of the part of the return on portfolio \( j \) explained by \( y_{i,j} \) and \( p_j \), while \( \tau_{i,j} \) is the residual variance. We know from the law of total variance that

\[ Var [d_j - p_j] = Var [E [d_j - p_j \mid y_{i,j}, p_j]] + E [Var [d_j - p_j \mid y_{i,j}, p_j]] \] (201)

where the second term on the right-hand side is just \( \tau_{i,j}^{-1} \) and the first term is \( E \left[ E [d_j - p_j \mid y_{i,j}, p_j]^2 \right] \) since everything has zero mean. The unconditional variance of returns is simply

\[ Var [d_j - p_j] = Var [(1 - a_1) d_j + a_2 z_j] \] (202)

\[ = (1 - a_{1,j})^2 f_{D,j} + a_2^2 f_{Z,j} \] (203)

So then

\[ EU_{i,j} = \frac{1}{2} \frac{Var [d_j - p_j] - \tau_{i,j}^{-1}}{\gamma_j} \] (204)
What we end up with is that utility is decreasing in $\tau_{i,j}^{-1}$. That is,

$$EU_{i,j} = -\frac{1}{2} \frac{\tau_{i,j}^{-1}}{\gamma_j} + \text{constants.} \quad (205)$$

### 7.2.3 Information choice

With the linear cost on precision, agents maximize

$$-\frac{1}{2} \frac{\tau_{i,j}^{-1}}{\gamma_j} - \psi f_{i,j}^{-1} \quad (206)$$

$$= -\frac{1}{2} \left( \frac{a_1^2}{a_2^2} f_{Z,j}^{-1} + f_{i,j}^{-1} + f_{D,j}^{-1} \right)^{-1} \gamma_j^{-1} - \psi f_{i,j}^{-1} \quad (207)$$

The FOC for $f_{i,j}^{-1}$ is

$$\psi = \frac{1}{2} \tau_{i,j}^{-2} \gamma_j^{-1} \quad (208)$$

$$\tau_{i,j} = \frac{1}{\sqrt{2}} \psi^{-1/2} \gamma_j^{-1/2} \quad (209)$$

But $\tau$ has a lower bound of $\frac{a_1^2}{a_2^2} f_{Z}^{-1} + f_{D}^{-1}$, so it’s possible that this has no solution. That would be a state where agents do no learning. Formally,

$$\tau_{i,j} = \max \left( \frac{a_1^2}{a_2^2} f_{Z}^{-1} + f_{D}^{-1}, \frac{1}{\sqrt{2}} \psi^{-1/2} \gamma_j^{-1/2} \right) \quad (210)$$

Note that, unlike in the other model, the equilibrium is unique here – all agents individually face a concave problem with an interior solution.

**Frequencies with no learning**  Now using the result for $a_1/a_2$ from above, at the frequencies where nobody learns, $f_{i}^{-1} = 0$, we have

$$\frac{a_1}{a_2} = \int_i \gamma_j^{-1} \tau_i^{-1} f_i^{-1} \, di \quad (211)$$

$$= 0 \quad (212)$$
which then implies

\[ \tau_{i,j} = \max \left( f^{-1}_D, \frac{1}{\sqrt{2}} \psi^{-1/2} \gamma^{-1/2}_j \right) \]  \hfill (213)

To get \( a_2 \), we have

\[
\int_i (cj^2 + b) \tau_i^{-1} \left( \frac{a_1}{a_2} f^{-1}_Z - \tau_i \right) \, di = -a_2^{-1} \]  \hfill (214)
\[
\gamma_j = a_2 \]  \hfill (215)

So the sensitivity of the price to supply shocks is increasing in the cost of holding inventory, \( b \), and the trading costs, \( c \). It is also higher at higher frequencies – it is harder to temporarily push through supply than to do it persistently.

**Frequencies with learning** At the frequencies at which there is learning, where

\[ f^{-1}_D < \frac{1}{\sqrt{2}} \psi^{-1/2} \gamma^{-1/2}_j \]  \hfill (216)

we have, just by rewriting the \( \tau \) equation,

\[ f_i^{-1} = \tau_i - \frac{a_1}{a_2} f^{-1}_Z - f^{-1}_D \]  \hfill (217)

Using the second equation from above,

\[
\int_i \gamma_j^{-1} \tau_i^{-1} \left( \frac{a_1}{a_2} f^{-1}_Z - \tau_i \right) \, di = -a_2^{-1} \]  \hfill (218)
\[
\int_i \gamma_j^{-1} \tau_i^{-1} \left( \frac{a_1}{a_2} f^{-1}_Z - a_2 \tau_i \right) \, di = -1 \]  \hfill (219)
\[
\int_i \gamma_j^{-1} \left( \tau_i^{-1} \frac{a_1}{a_2} f^{-1}_Z - a_2 \right) \, di = -1 \]  \hfill (220)

Under the assumption of a symmetric strategy, this is

\[ \tau^{-1} \frac{a_1}{a_2} f^{-1}_Z - a_2 = -\gamma_j \]  \hfill (221)
\[ \frac{a_1}{a_2} = \tau f_Z (-\gamma_j + a_2) \]  \hfill (222)
Using the other equilibrium condition, we have

\[
\int_i \gamma^{-1}_j \tau^{-1}_i \left( \tau_i - \frac{a_2^2}{a_2^2} f_Z^{-1} - f_D^{-1} \right) \, di = \frac{a_1}{a_2} \tag{223}
\]

\[
\int_i \gamma^{-1}_j \left( 1 - \tau^{-1}_i \frac{a_1}{a_2} f_Z^{-1} \frac{a_1}{a_2} - \tau^{-1}_i f_D^{-1} \right) \, di = \frac{a_1}{a_2} \tag{224}
\]

\[
1 - (-\gamma_j + a_2) \frac{a_1}{a_2} - \tau^{-1}_i f_D^{-1} = (c_j^2 + b) \frac{a_1}{a_2} \tag{225}
\]

\[
1 - \tau^{-1}_i f_D^{-1} = a_1 \tag{226}
\]

Plugging in the formula for \( \tau_i \) when there is learning,

\[
1 - \sqrt{2} \psi^{1/2} \gamma^{1/2}_j f_D^{-1} = a_1. \tag{227}
\]

The expression for \( a_2 \) can be obtained from:

\[
\frac{a_1}{\tau f_Z} = (-\gamma_j + a_2) a_2. \tag{228}
\]

Since \( a_1/\tau f_Z > 0 \), we know that there is only one solution to this equation for \( a_2 > 0 \). The positive root is

\[
a_2 = \gamma_j + \sqrt{\gamma_j^2 + 4 \frac{a_1}{\tau f_Z}} \tag{229}
\]

### 8 Alternative information cost specifications

This section considers alternative specifications for information costs. In each case, we examine the robustness of all of the paper’s theoretical results. The following results hold regardless of the information cost structure:

- Corollaries 1.3 and 1.4 depend only on the properties of the frequency transformation

- Result 2, corollaries 2.3 and 2.4, and the limits for \( a_{1,j} \) and \( a_{2,j} \) under information subsidies depend only on the properties of the date-0 rational expectations equilibrium (REE).
8.1 Frequency-specific linear costs

This section reports results for the case where the total cost of information is $\sum_{j=1}^{T} \psi_j f_{i,j}^{-1}$ instead of $\sum_{j=1}^{T} \psi f_{i,j}^{-1}$. Expected utility is the same, so it is simple to show that the equilibrium information choices are

$$f_{\text{avg},j}^{-1} = \begin{cases} 0 & \text{if } \psi_j > \lambda_j(0) \\ \lambda_j^{-1}(\psi_j) & \text{if } \psi_j \leq \lambda_j(0) \end{cases}$$

where, as before $\lambda_j(0) = f_{D,j} + \frac{f_{Z,j}}{\rho f_{D,j} + k}$.

8.1.1 Result 1 and corollaries 1.1 and 1.2

These results rely on the fact that the equilibrium information choices are independent across frequencies. Since that holds in this case, corollaries 1.1 and 1.2 are unchanged.

8.1.2 Corollary 1.5

This result depends primarily on the date-0 REE. The only change is that at the unrestricted frequencies, the variance of returns is $\min(\psi_j, \lambda_j(0))$ – the cost now has a frequency index.

8.1.3 Corollary 2.1 and 2.2

These results are derived in published appendix D.

8.1.4 Corollary 2.5

This result again follows from the separability of the information choice across frequencies and continues to hold with $\psi$ replaced by $\psi_j$.

8.2 Information capacity constraint

This section examines the case where investors are constrained in the total information they can acquire, rather than facing a linear cost of adding more precision. These problems mathematically
are duals of each other, meaning that they coincide holding the parameters fixed. The comparative statics, however, are different in some cases.

8.2.1 Information cost structure

The constraint specification is

$$\max_{\{f_{i,j}\}} \sum_{j=1}^{T} \lambda_j \left( f_{avg,j}^{-1} \right) f_{i,j}^{-1} \quad \text{s.t.} \quad \sum_{j=1}^{T} \psi_j f_{i,j}^{-1} \leq C$$

for some $C$. Denoting the Lagrange multiplier on the information constraint by $\mu$, the equilibrium information choices are

$$f_{avg,j}^{-1} = \begin{cases} 0 & \text{if } \mu > \frac{\lambda_j(0)}{\psi_j} \\ \lambda_j^{-1} \left( \mu \psi_j \right) & \text{if } \mu \leq \frac{\lambda_j(0)}{\psi_j} \end{cases}$$

(232)

where $\mu$ is the solution to

$$\sum_{j \text{ s.t. } \mu \leq \frac{\lambda_j(0)}{\psi_j}} \lambda_j^{-1} \left( \mu \psi_j \right) = C.$$  

(233)

8.2.2 Result 1

The equation for prices at restricted frequencies continues to hold since it depends only on the date-0 REE. The values of $a_{1,j}$ and $a_{2,j}$ at unrestricted frequencies shift in response to the restriction, unlike in the baseline case, due to the lack of complete separability. $a_{1,j}$ weakly rises, while the effect on $a_{2,j}$ is ambiguous.

8.2.3 Corollary 1.1

It remains the case that prices at the restricted frequencies become completely uninformative. At the unrestricted frequencies price informativeness weakly increases, depending on whether attention is reallocated to those frequencies. Specifically, we have the following result.

Lemma 1 (Corollary 1.1, modified) When investors are restricted from holding portfolios with weights that fluctuate at some set of frequencies $j \in \mathcal{R}$, the prices at those frequencies, $p_j$, become completely uninformative about dividends. The informativeness of prices for $j \in \mathcal{R}$ about dividends
weakly increases. More formally, \( \text{Var}[d_j|p_j] \) for \( j \notin \mathcal{R} \) weakly increases following the restriction. For \( j \in \mathcal{R} \), \( \text{Var}[d_j|p_j] = \text{Var}[d_j] \).

**Proof.** First we have:

\[
\text{Var}[d_j|p_j] = \frac{1}{f_{D,j}^{-1} + (\rho f_{\text{avg},j}^{-1})f_{Z,j}^{-1}},
\]

so price informativeness is strictly increasing in \( f_{\text{avg},j}^{-1} \). Moreover, at any restricted frequency, \( f_{\text{avg},j}^{-1} = 0 \) so \( \text{Var}[d_j|p_j] = f_D = \text{Var}[d_j] \).

Let \( \mu_{\text{unr}} \) be the marginal value of capacity in the unrestricted case. If \( \{ j \text{ s.t. } \mu_{\text{unr}} \psi_j \leq \lambda_j(0) \} \cup \mathcal{R} = \emptyset \), then the restriction has no effect on information choices, and \( \mu_{\text{res}} = \mu_{\text{unr}} \), where \( \mu_{\text{res}} \) is the marginal value of capacity under the restriction.

Consider the case where \( \{ j \text{ s.t. } \mu_{\text{unr}} \psi_j \leq \lambda_j(0) \} \cap \mathcal{R} \neq \emptyset \). We next show that in that case, \( \mu_{\text{res}} < \mu_{\text{unr}} \).

Assume otherwise, i.e. \( \mu_{\text{res}} \geq \mu_{\text{unr}} \). Then \( \forall j, \lambda_j^{-1}(\mu_{\text{res}} \psi_j) \leq \lambda_j^{-1}(\mu_{\text{unr}} \psi_j) \). Moreover, if \( \lambda_j(0) \geq \mu_{\text{res}} \psi_j \), then \( \lambda_j(0) \geq \mu_{\text{unr}} \psi_j \). So:

\[
\sum_{j \text{ s.t. } \lambda_j(0) \geq \mu_{\text{res}} \psi_j, j \notin \mathcal{R}} \lambda_j^{-1}(\mu_{\text{res}} \psi_j) \leq \sum_{j \text{ s.t. } \lambda_j(0) \geq \mu_{\text{res}} \psi_j, j \notin \mathcal{R}} \lambda_j^{-1}(\mu_{\text{unr}} \psi_j) \leq \sum_{j \text{ s.t. } \lambda_j(0) \geq \mu_{\text{unr}} \psi_j, j \notin \mathcal{R}} \lambda_j^{-1}(\mu_{\text{unr}} \psi_j) \leq \sum_{j \text{ s.t. } \lambda_j(0) \geq \mu_{\text{unr}} \psi_j} \lambda_j^{-1}(\mu_{\text{unr}} \psi_j) = C.
\]

This contradicts optimality in the restricted case (the investors are not exhausting their information budget). Therefore \( \mu_{\text{res}} < \mu_{\text{unr}} \).

So the restriction implies that \( \mu_{\text{res}} \leq \mu_{\text{unr}} \), with equality if and only if no restricted frequencies where being learned about before the restriction. In turn, \( \mu_{\text{res}} \leq \mu_{\text{unr}} \) implies that learning at all unrestricted frequencies weakly increases, using the first-order condition (243).

So by contrast with the fixed marginal cost case, where learning is unchanged at unrestricted frequencies, here it goes up weakly, as attention is reallocated toward unrestricted frequencies.
8.2.4 Corollary 1.2

This result changes under the constraint. The properties of the frequency transformation yield

\[ Var(D_t|P) = \frac{1}{T} \sum_{j=1}^{T} Var[d_j|p_j] \]  

the effect of the restriction is now to increase \( Var[d_j|p_j] \) but to weakly reduce it at other frequencies. The net effect on the informativeness of the vector of prices then becomes ambiguous. It remains the case, though, that informativeness on all dates is affected equally.

8.2.5 Corollary 1.5

This result continues to hold but with the modification at the unrestricted frequencies of \( \min(\mu_{res}\psi_j, \lambda_j(0)) \), where \( \mu_{res} \) is the Lagrange multiplier in the constrained case. Note that since \( j_{avg,j}^{-1} \) weakly increases at the unrestricted frequencies, return variance weakly decreases at those frequencies.

8.2.6 Corollary 2.1

In the case of a constraint, a change in the cost of information acquisition at a particular has both an income and a substitution effect. The substitution effect will cause agents to shift attention from the frequencies whose costs have not fallen to those that fall. The income effect causes agents to (weakly) increase attention on all frequencies, since the constraint relaxes. The consequence is that the first part of the result,

\[ \frac{d}{d\psi_j} E^{-1}_t [\tilde{q}_{LF,jr}] \bigg|_{\psi_j=\lambda_j(0)^-} > 0 \]  

continues to hold, since it does not depend on anything about the other frequencies. However, the two other inequalities no longer hold.

This corollary gives the clearest motivation for the use of the cost specification instead of the constraint. The constraint specification means that a decline in information acquisition costs does not lead investors to acquire information (other than mechanically) since, by assumption, they cannot.
8.2.7 Corollary 2.5

This result is also unchanged except that “$\lambda_j(0) \leq \psi$” is replaced by “$\lambda_j(0) \leq \mu_{res}\psi_j$.”

8.3 Entropy cost for information

This section examines a specification where instead of the cost of information being measured in terms of total precision, it is measured in terms of the joint entropy of the prior and posterior, as in Sims (2003). Specifically, the information flow contained in the signals can be measured by the difference between the prior entropy, which is equal to $\frac{1}{2} \log |\Sigma_D|$ plus a constant, and the posterior entropy, $\frac{1}{2} \log (\Sigma_D^{-1} + \Sigma_i^{-1})^{-1}$. As in Kacperczyk, van Nieuwerburgh, and Veldkamp (2016), we exponentiate the entropy. Using the frequency transformation (and ignoring approximation error) and ignoring constants, total information flow is then measured by

$$\prod_{j=1}^{T} \left( f_{i,j}^{-1} + f_{D,j}^{-1} \right)$$

8.3.1 Information cost structure and equilibrium

The attention allocation problem with an entropy cost function can be written as:

$$\max \{ f_{i,j} \} \sum_{j=1}^{T} \lambda_j \left( f_{avg,j}^{-1} \right) f_{i,j}^{-1} - \kappa \prod_{j=1}^{T} \left( f_{i,j}^{-1} + f_{D,j}^{-1} \right)$$

Assume that there exists $j$ such that:

$$\lambda_j > \kappa \prod_{k \neq j} f_{D,k}^{-1}.$$  

Then the problem is unbounded (a value of $+\infty$ can be reached by setting $f_{i,k}^{-1} = 0$ for $k \neq j$ and $f_{i,j}^{-1} = +\infty$). Therefore, it must be the case in equilibrium that

$$\lambda_j \leq \kappa \prod_{k \neq j} f_{D,k}^{-1} \ \forall j.$$ 

It is straightforward to confirm that optimization requires that investors only allocate attention...
to a single frequency that achieves the maximum value across all frequencies of

$$\lambda_j - \kappa \prod_{k \neq j} f_{D,k}^{-1}$$

(241)

It is possible that there are multiple frequencies with this property. Regardless, each investor only
allocates attention to a single frequency. Define

$$\xi \equiv \kappa \prod_k f_{D,k}^{-1}$$

(242)

The equilibrium information choices are then given by:

$$f^{-1}_{\text{avg},j} = \begin{cases} 0 & \text{if } \frac{\lambda_j(0)}{f_{D,j}} < \xi \\ \lambda_j^{-1}(f_{D,j}\xi) & \text{if } \frac{\lambda_j(0)}{f_{D,j}} \geq \xi \end{cases}$$

(243)

As in the baseline case, the allocation of precision across investors is indeterminate, up to the fact
that investors must learn about at most one frequency.

This version of the model is very similar to the linear cost with heterogeneous frequency-specific
cost case, with $f_{D,j}$ playing the role of $\psi_j$. The model retains the linearity of utility with respect
to precision, and the information decisions remain completely separable across frequencies. Those
facts mean that all the results go through without any changes except corollary 2.1.

For corollary 2.1, we modify the entropy constraint to make the cost of precision frequency
dependent, as

$$\max_{\{f_i^{-1}\}} \sum_{j=1}^T \lambda_j \left( f_{\text{avg},j}^{-1} \right) f_{i,j}^{-1} - \kappa \prod_{j=1}^T \left( \psi_j f_{i,j}^{-1} + f_{D,j}^{-1} \right)$$

(244)

Then the equilibrium information allocation is

$$f^{-1}_{\text{avg},j} = \begin{cases} 0 & \text{if } \xi > \frac{\lambda_j(0)}{f_{D,j}\psi_j} \\ \lambda_j^{-1}(f_{D,j}\xi\psi_j) & \text{if } \xi \leq \frac{\lambda_j(0)}{f_{D,j}\psi_j} \end{cases}$$

(245)

At that point, the analysis from the baseline version applies. The only changes are that $\min(\psi, \lambda_j(0))$
must be replaced by $\min(f_{D,j}\xi, \lambda_j(0))$, where $\xi \equiv \kappa \prod_k f_{D,k}^{-1}$ in corollary 1.5 and $\lambda_j(0) \leq \psi$ should
be replaced by $\lambda_j(0) \leq f_{D_j} \xi$, with $\xi \equiv \kappa \prod_k f_{D_k}^{-1}$ in corollary 2.5.

## 9 Hedging model

This section provides the full derivation for an alternative model of “noise traders”. The key feature that the model needs in order for there to be trade – i.e. for prices to not be fully revealing – is that there must be shocks to demand for the fundamental that are uncorrelated with the realization of the fundamental. In the main text, those shocks are driven by uninformative signals that a subset of investors erroneously treat as informative. Here, we study an alternative case in which the demand shocks represent hedging demands from a subset of investors. The analysis is similar to that of Wang (1994), and extended to account for the information structure in this paper. See also Savov (2014) for related work on hedging demand and investment.

The analysis in this section takes place entirely in the frequency domain and applies to a representative frequency, so we drop the $j$ subscripts.

Suppose there is a set of investors who have a private technology that they can invest in. It has payoffs that are correlated with the fundamental, so that trading the fundamental is useful for hedging purposes. For simplicity, we assume that these investors do not have any other signals about fundamentals. We call these investors the hedgers.

The hedgers have investment opportunities that are imperfectly correlated. Each individual hedger, indexed by $h$, has an investment opportunity $z_h$, where the distribution of $z_h$ across the hedgers is $N(z, \sigma_z^2)$, with $z \sim N(0, f_Z)$. $z$ and $z_h$ are both random variables drawn on date 0, $z$ is not directly observed by any investor, while $z_h$ is observed by hedger $h$, but not by any other investors.

Investing a quantity $k_h$ in the project yields a random payoff of $k_h x_h$, where $x_h = z_h + d + \varepsilon_{x,h}$, with $\varepsilon_{x,h} \sim N(0, \sigma_x^2)$. The inclusion of $d$ as part of the payoff means that the agent can hedge the project by trading the fundamental. $z_h$ is the expected payoff in the absence of any other information, while $\varepsilon_{x,h}$ represents uninsurable risk that investor $h$ faces.

We guess that prices follow

$$p = a_1 d - a_2 z$$

(246)
The hedgers’ optimization is then over both their investment in the private opportunity, \( k_h \), and their investment in the fundamental, \( q_h \).

\[
\max_{k_h, q_h} E [k_h x_h + q_h (d - p) \mid p, z_h] - \frac{\rho_H^1}{2} \var [k_h x_h + q_h (d - p) \mid p, z_h] \tag{247}
\]

\[
= \max_{k_h, q_h} E [k_h x_h + q_h (E [d \mid p, z_h] - p)] - \frac{\rho_H^1}{2} \left( (k_h + q_h)^2 \var [d \mid p, z_h] + k_h^2 \sigma_x^2 \right) \tag{248}
\]

### 9.1 Beliefs

The optimization involves means and variances conditional on \( z_h \) and \( p \). It is possible to obtain them in general, but a useful simplification is to assume that hedgers only forecast \( d \) using \( p \), not their private investment opportunity, \( z_h \). That corresponds to the limiting case where \( \sigma_x \to \infty \), since then each hedger’s investment opportunity is minimally informative about the average investment opportunity. So in what follows all expectations and variances condition only on \( p \). We have

\[
\var (d \mid p) = f_D \left( 1 - \frac{a_1^2 f_D}{a_1^2 f_D + a_2^2 f_Z} \right) = f_D \left( \frac{f_Z}{a_2^2 f_D + f_Z} \right) \tag{249}
\]

\[
E [d \mid p] = \frac{a_1 f_D}{a_1^2 f_D + a_2^2 f_Z} p \tag{250}
\]

### 9.2 Optimization

Note that

\[
k_h x_h + q_h (d - p) = k_h z_h + (k_h + q_h) (d - p) + k_h \varepsilon_{x,h} + k_h p \tag{251}
\]

The optimization problem is then

\[
\max_{k_h, q_h} E [k_h x_h + q_h (d - p) \mid p] - \frac{\rho_H^1}{2} \left( (k_h + q_h)^2 \var [d \mid p] + k_h^2 \sigma_x^2 \right) \tag{252}
\]

\[
= \max_{k_h, q_h} k_h z_h + k_h p + E [(k_h + q_h) (d - p) \mid p] - \frac{\rho_H^1}{2} (k_h + q_h)^2 \var [(d - p) \mid p] - \frac{\rho_H^2}{2} k_h^2 \sigma_x^2 \tag{253}
\]
The two first-order conditions (FOCs) are

\[
\rho_H (z_h + E [d \mid p, z_h]) = (k_h + q_h) \text{var} (d \mid p, z_h) + k_h \sigma_x^2 \tag{254}
\]

\[
\rho_H E [d - p \mid p, z_h] = (k_h + q_h) \text{var} (d - p \mid p, z_h) \tag{255}
\]

Subtracting the second equation from the first yields

\[
k_h = \frac{\rho_H (z_h + p)}{\sigma_x^2} \tag{256}
\]

So, naturally, agents invest more in their private project when its expected return is higher or its risk is lower. \(z_h + p\) is the expected return on an investment that is long one unit of the private investment and short one unit of the fundamental, and \(\sigma_x^2\) is its variance, so this is the standard mean-variance optimal quantity invested.

Combining that result with the FOC for \(q_h\), we have

\[
q_h = \frac{\rho_H E [d - p \mid p, z_h]}{\text{var} (d - p \mid p, z_h)} - \frac{\rho_H (z_h + p)}{\sigma_x^2}
\]

\[
= \rho_H \left( a_1 f_D - a_1^2 f_D - a_2^2 f_Z \right) \frac{p}{f_D a_2^2 f_Z} - \frac{\rho_H (z_h + p)}{\sigma_x^2} \tag{258}
\]

\[
= \rho_H \left( a_1^2 f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) p - \rho_H \sigma_x^{-2} z_h \tag{259}
\]

where \(\tau_H\) is the precision of the hedgers’ beliefs,

\[
\tau_H \equiv \text{var} (d \mid p)^{-1} = \frac{a_1^2}{a_2^2} f_Z^{-1} + f_D^{-1} \tag{260}
\]

### 9.3 Expected utility

From above, expected utility conditional on prices and \(z_h\) is

\[
k_h z_h + k_h p + E [(k_h + q_h) (d - p) \mid p] - \frac{\rho_H^{-1}}{2} (k_h + q_h)^2 \text{var} [(d - p) \mid p, z_h] - \frac{\rho_H^{-1}}{2} k_h^2 \sigma_x^2 \tag{261}
\]
Multiplying the $k_h$ and $q_h$ FOCs by $k_h$ and $q_h$ then using them to substitute out the variances, then inserting the solutions for $k_h$ and $k_h + q_h$, utility becomes

$$
\frac{1}{2} (k_h z_h + E [k_h d + q_h (d - p) \mid p]) = \frac{1}{2} (k_h (z_h + p) + (k_h + q_h) E [d - p \mid p]) 
$$

(262)

$$
= \frac{1}{2} \left( \rho_H E [d - p \mid p]^2 \frac{\var (d - p \mid p)}{\sigma_x^2} + \rho_H (z_h + p)^2 \right) 
$$

(263)

The second term represents the utility gained from exposure to $x$, which is obtained by going long the private investment and short an equal amount of the fundamental, leaving pure exposure to $\varepsilon_{h,x}$. The first term is the usual utility gained from investing in the fundamental. These two investments are completely independent of each other.

The law of total variance, as in the main results, gives us

$$
\var [d - p] = \var [E [d - p \mid p]] + E [\var [d - p \mid p]] 
$$

(264)

$$
= E \left[ E [d - p \mid p]^2 \right] + \var (d - p \mid p) 
$$

(265)

which we can substitute in for the first term. For the second term, we have

$$
E \left[ (z_h + p)^2 \right] = \var (z_h + p) 
$$

(266)

$$
= \var ((z_h - z) + z + a_1 d - a_2 z) 
$$

(267)

$$
= \sigma_x^2 + (1 - a_2)^2 f_Z + a_1^2 f_D 
$$

(268)

Substituting back into the equation for expected utility,

$$
E \left[ \frac{1}{2} \left( \rho_H E [d - p \mid p]^2 \frac{\var (d - p \mid p)}{\sigma_x^2} + \rho_H (z_h + p)^2 \right) \right] = \frac{1}{2} \rho_H \left( \frac{\var [d - p] - \var (d - p \mid p)}{\var (d - p \mid p)} + \frac{\sigma_x^2 + (1 - a_2)^2 f_Z + a_1^2 f_D}{\sigma_x^2} \right) 
$$

(269)

$$
= \frac{1}{2} \rho_H \frac{(1 - a_1)^2 f_D + a_2^2 f_Z - \tau_H^{-1}}{\tau_H^{-1}} + \frac{1}{2} \rho_H \frac{\sigma_x^2 + (1 - a_2)^2 f_Z + a_1^2 f_D}{\sigma_x^2} 
$$

(270)

$$
= \frac{1}{2} \left\{ \tau_H (1 - a_1)^2 f_D + \sigma_x^{-2} a_1^2 \right\} + \frac{1}{2} \left\{ \tau_H a_2^2 + \sigma_x^{-2} (1 - a_2)^2 \right\} f_Z - 1 + \rho_H \sigma_x^{-2} \sigma_Z^2 
$$

(271)
9.4 Equilibrium

Now suppose there are unit masses of both the informed investors and the hedgers. This is without loss of generality as their influence can be controlled by shifting $\rho_H$ and $\rho$ (where the latter remains the risk tolerance of the sophisticated investors from the main analysis). The equilibrium condition is

$$0 = \rho_H \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) \left( a_1 d - a_2 z \right) - \rho_H \int_h z_h \sigma_x^{-2} dh + \int_i \rho \left( \int f_i^{-1} d + \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_{avg} \right) \left( a_1 d - a_2 z \right) \right) \frac{d}{di} (272)$$

Matching coefficients on $z$ and $d$ and using the law of large numbers so that $\int_h z_h dh = z$,

$$0 = \rho_H \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) + \rho_H \sigma_x^{-2} a_2^{-1} + \rho \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_{avg} \right) \quad (273)$$

$$0 = \rho_H \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) + \rho f_{avg} a_1^{-1} + \rho \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_{avg} \right) \quad (274)$$

Equating the right hand sides of those two equations yields

$$\frac{\rho f_{avg}}{\rho_H \sigma_x^{-2}} = \frac{a_1}{a_2} \quad (275)$$

Inserting that formula into the second equation yields

$$0 = \rho_H \left( a_1^{-1} \frac{a_1^2}{a_2^2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) + \rho f_{avg} a_1^{-1} + \rho a_1^{-1} \frac{a_1^2}{a_2} f_Z^{-1} - \rho \tau_{avg} \quad (276)$$

$$a_1 = \frac{(\rho_H + \rho) \left( \frac{\rho f_{avg}}{\rho_H \sigma_x^{-2}} \right)^2 f_Z^{-1} + \rho f_{avg}^{-1}}{\rho_H \left( \tau_H + \sigma_x^{-2} \right) + \rho \tau_{avg}} \quad (277)$$

9.5 Restricting speculators

The main text refers to the agents able to gather information as “sophisticates” as opposed to the unsophisticated noise traders. Here we describe them as speculators, who are making pure bets on the fundamental, as opposed to hedgers, who hold the fundamental (partly) to hedge their private investments.

The main text considers the experiment of restricting trading by the sophisticates. Here, if only
the hedgers can trade, and not the speculators, the market clearing condition is

\[ 0 = \rho_H \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) (a_1 d - a_2 z) - \rho_H z_i \sigma_x^{-2} \]  \quad (278)

Again matching coefficients,

\[ 0 = \rho_H \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) a_1 \]  \quad (279)

\[ 0 = -\rho_H \left( \frac{a_1}{a_2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) a_2 - \rho_H \sigma_x^{-2} \]  \quad (280)

This immediately implies \( a_1 = 0 \), and hence

\[ a_2 = \frac{\sigma_x^{-2}}{f_D^{-1} + \sigma_x^{-2}} \]  \quad (281)

So we again get that result, not surprisingly, that prices are uninformative when the investors who have access to information about fundamentals are no longer allowed to invest.

That does not mean, though, that the welfare benefits in this case go to zero, since the hedgers can still trade with each other. In fact, in both cases, they can always perfectly hedge the idiosyncratic part of \( z_i \), since prices depend only on aggregate \( z \) – each individual agent has no effect on prices.

Expected utility when speculators cannot trade is

\[ \frac{1}{2} \left\{ \frac{f_D^{-1} \sigma_x^{-2}}{f_D^{-1} + \sigma_x^{-2}} f_Z + \rho_H \sigma_x^{-2} \sigma_z^2 \right\} \]  \quad (282)

More generally, expected utility is

\[ EU_H = \frac{1}{2} \left\{ \begin{array}{l} \left( \tau_H (1 - a_1)^2 + \sigma_x^{-2} a_1^2 \right) f_D + \\
\left( \tau_H a_2^2 + \sigma_x^{-2} (1 - a_2)^2 \right) f_Z - 1 + \rho \sigma_x^{-2} \sigma_z^2 \end{array} \right\} \]  \quad (283)

### 9.6 Speculator profits

The formulas for utility and expected profits go through in this case unchanged since they depend just on the optimization of the speculators, taking \( a_1 \) and \( a_2 \) as given.
9.7 Results

This section describes how the results from the main text are affected by the replacement of the noise traders with hedgers.

9.7.1 Solution 1

There continues to be a linear solution, but in this case the coefficients are

\[
\begin{align*}
a_1 & = \frac{(\rho_H + \rho) \left( \frac{\rho f_{avg}^{-1}}{\rho_H f_x^{-1}} \right)^2 f_z^{-1} + \rho f_{avg}^{-1}}{\rho_H \left( \tau_O + \sigma_x^{-2} \right) + \rho f_{avg}} \\
a_2 & = a_1 \frac{\rho H \sigma_x^{-2}}{f_{avg}^{-1}}
\end{align*}
\]

(284)

(285)

9.7.2 Lemma 2

The derivation of this result depends only on the information structure and the existence of a linear equilibrium, so the utility of the speculators is the same here as in the main text.

9.7.3 Solution 2

The solution follows directly from the linearity of utility. As before, tedious algebra confirms that \( \lambda'(\cdot) < 0 \).

9.7.4 Result 1

The fact that the trade restrictions affect only targeted frequencies follows directly from the separability of the model across frequencies, so is unchanged here. The formula for prices in the case where speculators cannot trade is

\[
p_j = -\frac{\sigma_x^{-2}}{f_D^{-1} + \sigma_x^{-2} \tau j}
\]

(286)

9.7.5 Corollary 1.1

The lack of informativeness at restricted frequencies follows trivially from the pricing function at those frequencies. The lack of any change in informativeness at unrestricted frequencies follows from the fact that the pricing function at those frequencies is unaffected.
9.7.6 Corollaries 1.2–1.4

These results all are driven entirely by the properties of the frequency transformation and are therefore unaffected by the choice of noise traders versus hedgers.

9.7.7 Corollary 1.5

The volatility of returns in the absence of speculators is

\[ f_D + \left( \frac{\sigma_x^{-2}}{f_D^{-1} + \sigma_x^2} \right)^2 f_Z \]  

(287)

When speculators are present but uninformed, the pricing function is

\[ p = \frac{f_D \rho_H}{f_D \rho_H + (\rho + \rho_H) \sigma_x^2} \]  

\[ \text{var} \ (r) = f_D + \left( \frac{f_D \rho_H}{f_D \rho_H + (\rho + \rho_H) \sigma_x^2} \right)^2 f_Z \]  

(288)

(289)

Straightforward algebra shows that \( \frac{f_D \rho_H}{f_D \rho_H + (\rho + \rho_H) \sigma_x^2} < \frac{\sigma_x^{-2}}{f_D^{-1} + \sigma_x^2} \), which implies that return volatility is lower with uninformed speculators than with a trading restriction.

9.7.8 Result 2

The exact form of the formula for profits of speculators no longer holds. However, the nonnegativity does hold. The specific corollaries are more important and are discussed further below.

9.7.9 Corollary 2.1

\[ \frac{d}{d \psi_j} E_{-1} \left[ \tilde{q}_{LF,j} \tau_j \right] \bigg|_{\psi_j = \lambda_j(0)^-} > 0 \]  

(290)

\[ \frac{d}{d \psi_j} E_{-1} \left[ \sum_t \tilde{Q}_{LF,t} (D_t - P_t) \right] \bigg|_{\psi_j = \lambda_j(0)^-} > 0 \]  

(291)

\[ \frac{d}{d \psi_j} E_{-1} \left[ U_{LF,0} \right] \bigg|_{\psi_j = \lambda_j(0)^-} > 0 \]  

(292)

The problem faced by the sophisticated investors is the same in the sense that they continue
to acquire information to the point that \( \lambda_j \left( f^{-1}_{avg,j} \right) = \psi_j \), unless \( \lambda_j(0) \leq \psi_j \), in which case they acquire no information. A marginal decline in \( \psi_j \) then leads to a marginal increase in \( f^{-1}_{avg,j} \).

To obtain the derivative of \( E_{-1} [\tilde{q}_{LF,j} r_j] \), simply use the formulas for speculator profits from the main analysis. That result then immediately implies the derivative in the second line, due to the separability across frequencies. Similarly, it remains the case that speculator utility is equal to \( \sum_j Var(r_j) \tau_{i,j} \), and differentiation of \( Var(r_j) \tau_{i,j} \) with respect to \( f^{-1}_{avg,j} \) yields the desired result.

### 9.7.10 Corollary 2.2

From above, we have

\[
\text{var} [d_j | p_j] = \left( \frac{\rho f^{-1}_{avg,j} f^{-1}_{Z,j} + f^{-1}_{D,j}}{\rho X^2 f^{-1}_{Z,j} + f^{-1}_{D,j}} \right)^{-1},
\]

which is obviously decreasing in \( f^{-1}_{avg,j} \). It is also possible to confirm that return volatility is decreasing in \( f^{-1}_{avg,j} \).

### 9.7.11 Corollary 2.3

This result follows from the fact that each frequency independently contributes nonnegatively to the profits and utility of speculators, so it continues to hold here.

Similarly, the profits and utility of the hedgers must weakly fall under an investment restriction since they always have the option of not investing at any particular frequency.

### 9.7.12 Corollary 2.4

The formula for the earnings of noise traders does not apply to the hedgers. Moreover, their earnings are not simply the negative of those of the speculators since they also have their private investment opportunity.

It remains the case that at any frequency where \( \lambda_j(0) \leq \psi \), there is no information acquisition in equilibrium. That immediately implies that restricting speculators from trading still has no impact on price informativeness, since prices are uninformative in any case.

However, there is an important change in the result for utility. We now have, in the case where
prices are already uninformative

\[ \text{when } f^{-1}_{avg} = 0, \quad \frac{d}{d\rho} EU_H > 0 \quad (294) \]

That is, when the speculators are acting purely to provide insurance to the hedgers, any increase in their risk-bearing capacity increases the expected utility of the hedgers. So whereas in the case of noise traders, restricting trade at a frequency where no information was being acquired was beneficial, with hedgers it actually is socially harmful.

9.7.13 Section 4.4.2

It remains the case that

\[ \lim_{f_{avg}} a_1 = 1 \quad (295) \]
\[ \lim_{f_{avg}} a_2 = 0 \quad (296) \]

Furthermore, straightforward algebra shows that, writing the expected utility of the hedgers as a function of \( f^{-1}_{avg} \), we have

\[ \lim_{f_{avg} \to \infty} EU_H \left( f^{-1}_{avg} \right) > EU_H (0) \quad (297) \]

which shows that the hedgers are better off in a fully informative equilibrium than in the alternative uninformative case.
Figure A.1: Numerical example with entropy cost

Notes: See figure 2. The only change is that the frequency-specific information cost case is replaced by the case where information flow is measured by entropy. That setup is analyzed in appendix 8.3.
Notes: The three panels report per-period characteristics of dividend strip returns in the model with the frequency-specific cost specification along with empirical moments reported by Binsbergen and Koijen (2017) (averaged across the four markets they examine). All values are annualized.
Figure A.3: Details of dividend strips returns, entropy cost case

Notes: See figure A.2. This figure reports results for the case where information flow is measured by entropy instead of precision.
Figure A.4: Numerical example in public signal model

Notes: Numerical results for the model described in appendix 4 in which investors observe a public signal about fundamentals. The outputs are all for a single frequency, taking advantage of separability. The x-axis represents the precision of the public signal. Average information acquisition in the bottom-left panel is $f_{avg}^{-1}$. Utility in the top-right panel is net of information acquisition costs ($\psi_i f_i^{-1}$). Information flows in the bottom-right panel are the precision obtained from conditioning either on prices alone or prices and the public signal.