Online appendix for "On the effects of restricting short-term investment"

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

January 28, 2019

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 Σ be any matrix of the form:

$$\Sigma = \begin{pmatrix} \sigma_0 & \sigma_1 & ... & ... & \sigma_{T-1} \\ \sigma_1 & \sigma_0 & \sigma_1 & ... & \sigma_{T-2} \\ ... & ... & ... & ... & ... \\ \sigma_{T-1} & ... & ... & ... & \sigma_0 \end{pmatrix} \tag{1}$$

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" Σ by the circulant matrix:

$$\Sigma_{circ} = circ(\sigma_{circ}) \quad , \quad \sigma \equiv (\sigma_0, \sigma_1 + \sigma_{T-1}, \sigma_2 + \sigma_{T-2}, ..., \sigma_{T-2} + \sigma_2, \sigma_{T-1} + \sigma_1)', \tag{2}$$

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

$$circ(x) \equiv \begin{pmatrix} x_0 & \cdot & \cdot & x_{T-1} \\ x_{T-1} & x_0 & \cdot & x_{T-2} \\ \cdot & & & \\ x_1 & \cdot & \cdot & x_0 \end{pmatrix}. \tag{3}$$

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

Appendix lemma 1 *For any matrix Σ of the form given above, and associated circulant matrix Σ_{circ}, the family of vectors Λ defined in the main text exactly diagonalizes Σ_{circ}:*

$$\Sigma_{circ}\Lambda = \Lambda diag\left(\left\{f_\Sigma\left(\omega_{\lfloor j/2\rfloor}\right)\right\}_{j=1}^{T}\right), \tag{4}$$

where each distinct eigenvalue in $\left\{f_\Sigma\left(\omega_{\lfloor j/2\rfloor}\right)\right\}_{j=1}^{T}$ is given by:

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

for some $h = 0, ..., \frac{T}{2}$.

Given that Λ is orthonormal,

$$\Lambda'\Sigma_{circ}\Lambda = diag\left(f_\Sigma\right). \tag{6}$$

The approximate diagonalization of the matrix Σ consists in writing:

$$\Lambda'\Sigma\Lambda = diag\left(f_\Sigma\right) + R_\Sigma, \tag{7}$$

where the $T \times T$ matrix R_Σ is given by:

$$R_\Sigma \equiv \Lambda'\left(\Sigma - \Sigma_{circ}\right)\Lambda. \tag{8}$$

This is an approximation in the sense that R_Σ 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:*

$$|R_\Sigma| \leq \frac{4}{\sqrt{T}} \sum_{j=1}^{T-1} |j\sigma_j|, \tag{9}$$

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

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

$$\Sigma^{(i,j)} = \begin{cases} \sigma_0 & \text{if } i = j \\ \sigma_{|i-j|} & \text{otherwise} \end{cases}, \tag{10}$$

$$\Sigma_{circ}^{(i,j)} = \begin{cases} \sigma_0 & \text{if } i = j \\ \sigma_{|i-j|} + \sigma_{T-|i-j|} & \text{otherwise} \end{cases}, \tag{11}$$

we have:

$$\Delta\Sigma^{(i,j)} = \begin{cases} 0 & \text{if } i = j \\ \sigma_{T-|i-j|} & \text{otherwise} \end{cases} \tag{12}$$

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

$$\Delta\Sigma = \begin{pmatrix} 0 & \sigma_{T-1} & \sigma_{T-2} & & \sigma_2 & \sigma_1 \\ \sigma_{T-1} & \ddots & \ddots & \ddots & & \sigma_2 \\ \sigma_{T-2} & \ddots & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & \sigma_{T-2} \\ \sigma_2 & & \ddots & \ddots & \ddots & \sigma_{T-1} \\ \sigma_1 & \sigma_2 & & \sigma_{T-2} & \sigma_{T-1} & 0 \end{pmatrix} \tag{13}$$

Therefore,

$$\sum_{i=1}^{T} \sum_{j=1}^{T} |\Delta\sigma_{i,j}| = 2 \sum_{j=1}^{T-1} |j\sigma_j|. \tag{14}$$

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

$$\begin{aligned}
\left|R_{\Sigma}^{(l,m)}\right| &= \left|\lambda_l' \Delta\Sigma\lambda_m\right| \\
&= \left|\sum_{i=1}^{T}\sum_{j=1}^{T} \lambda_{i,l}\lambda_{j,m}\Delta\sigma_{i,j}\right| \\
&\leq \sum_{i=1}^{T}\sum_{j=1}^{T} |\lambda_{i,l}|\,|\lambda_{j,m}|\,|\Delta\sigma_{i,j}| \qquad (15)\\
&\leq \sum_{i=1}^{T}\sum_{j=1}^{T} \frac{\sqrt{2}}{\sqrt{T}}\frac{\sqrt{2}}{\sqrt{T}} |\Delta\sigma_{i,j}| \\
&= \tfrac{4}{T}\sum_{j=1}^{T-1} |j\sigma_j|.
\end{aligned}$$

This implies that:

$$||R_{\Sigma}||_{\infty} \leq \frac{4}{T}\sum_{j=1}^{T-1} |j\sigma_j|, \tag{16}$$

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

$$p = a_1 d + a_2 z + a_3 s \tag{17}$$

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

$$\begin{bmatrix} d \\ y_i \\ p - a_3 s \end{bmatrix} \sim N \left(0, \begin{bmatrix} f_D & f_D & a_1 f_D \\ f_D & f_D + f_i & a_1 f_D \\ a_1 f_D & a_1 f_D & a_1^2 f_D + a_2^2 f_Z \end{bmatrix} \right) \tag{18}$$

The expectation of fundamentals conditional on the signal and price is

$$\begin{aligned} E\left[d \mid y_i, p\right] &= \begin{bmatrix} f_D & a_1 f_D \end{bmatrix} \begin{bmatrix} f_D + f_i & a_1 f_D \\ a_1 f_D & a_1^2 f_D + a_2^2 f_Z \end{bmatrix}^{-1} \begin{bmatrix} y_i \\ p - a_3 s \end{bmatrix} && (19) \\ &= [1, a_1] \begin{bmatrix} 1 + f_i f_D^{-1} & a_1 \\ a_1 & a_1^2 + a_2^2 f_Z f_D^{-1} \end{bmatrix}^{-1} \begin{bmatrix} y_i \\ p - a_3 s \end{bmatrix} && (20) \end{aligned}$$

and the variance satisfies

$$\begin{aligned} \tau_i &\equiv Var\left[d \mid y_i, p\right]^{-1} = f_D^{-1} \left(1 - \begin{bmatrix} 1 & a_1 \end{bmatrix} \begin{bmatrix} 1 + f_i f_D^{-1} & a_1 \\ a_1 & a_1^2 + a_2^2 f_Z f_D^{-1} \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ a_1 \end{bmatrix} \right)^{-1} && (21) \\ &= \frac{a_1^2}{a_2^2} f_Z^{-1} + f_i^{-1} + f_D^{-1} && (22) \end{aligned}$$

We use the notation τ 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

$$E\left[d \mid y_i, p\right] = \tau_i^{-1} \left(f_i^{-1} y_i + \frac{a_1}{a_2^2} f_Z^{-1} \left(p - a_3 s\right) \right) \tag{23}$$

2.1.2 Demand and equilibrium

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

$$U_i = \max_{\{Q_{i,t}\}} \rho^{-1} E_{0,i} \left[T^{-1} \tilde{Q}_i' (D-P) \right] - \frac{1}{2} \rho^{-2} Var_{0,i} \left[T^{-1/2} \tilde{Q}_i' (D-P) \right] \tag{24}$$

$$= \max_{\{Q_{i,t}\}} \rho^{-1} E_{0,i} \left[T^{-1} \tilde{q}_i' (d-p) \right] - \frac{1}{2} \rho^{-2} Var_{0,i} \left[T^{-1/2} \tilde{q}_i' (d-p) \right] \tag{25}$$

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

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

$$\tilde{q}_i = \rho \tau_i E \left[d - p \mid y_i, p \right] \tag{27}$$

$$= \rho \left(f_i^{-1} y_i + \frac{a_1}{a_2^2} f_Z^{-1} (p - a_3 s) - \tau_i p \right) \tag{28}$$

$$= \rho \left(f_i^{-1} y_i + \frac{a_1}{a_2^2} f_Z^{-1} (a_1 d + a_2 z) - \tau_i (a_1 d + a_2 z + a_3 s) \right) \tag{29}$$

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

$$-z + k (a_1 d + a_2 z) + s = \int_i \rho \left(f_i^{-1} y_i + \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i \right) (a_1 d + a_2 z) - \tau_i a_3 s \right) di \tag{30}$$

$$= \int_i \rho \left(f_i^{-1} d + \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i \right) (a_1 d + a_2 z) - \tau_i a_3 s \right) di, \tag{31}$$

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

$$\int_i \rho \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i \right) di = -a_2^{-1} (1 - k a_2) \tag{32}$$

$$\int_i \rho f_i^{-1} a_1^{-1} + \rho \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_i \right) di = k \tag{33}$$

$$a_3 = \frac{-1}{\rho \int_i \tau_i di} \tag{34}$$

and therefore

$$k - \int_i \rho f_i^{-1} a_1^{-1} = a_2^{-1} (k a_2 - 1) \tag{35}$$

$$\int_i \rho f_i^{-1} = \frac{a_1}{a_2} \tag{36}$$

Now define aggregate precision to be

$$f_{avg}^{-1} \equiv \int_i f_i^{-1} di \tag{37}$$

We then have

$$\tau_i = \frac{a_1^2}{a_2^2} f_Z^{-1} + f_i^{-1} + f_D^{-1} \tag{38}$$

$$\tau_{avg} \equiv \int \tau_i di = \left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_{avg}^{-1} + f_D^{-1} \tag{39}$$

Inserting the expression for τ_i into (32) yields

$$a_1 = \frac{\tau_{avg} - f_D^{-1}}{\tau_{avg} + \rho^{-1} k} \tag{40}$$

$$a_2 = \frac{a_1}{\rho f_{avg}^{-1}} \tag{41}$$

$$a_3 = \frac{-1}{\rho \tau_{avg}} \tag{42}$$

The expression for a_1 can be written more explicitly as:

$$a_1 = \frac{\tau_{avg} - f_D^{-1}}{\tau_{avg} + \rho^{-1} k} = \frac{\frac{a_1^2}{a_2^2} f_Z^{-1} + f_{avg}^{-1} + f_D^{-1} + \rho^{-1} k - \rho^{-1} k - f_D^{-1}}{\frac{a_1^2}{a_2^2} f_Z^{-1} + f_{avg}^{-1} + f_D^{-1} + \rho^{-1} k} \tag{43}$$

$$= 1 - \frac{\rho^{-1} k + f_D^{-1}}{\left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_{avg}^{-1} + \rho^{-1} k + f_D^{-1}}. \tag{44}$$

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

$$a_2 = \frac{1}{\rho f_D^{-1} + k}. \tag{45}$$

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*

$$\left|A_1 - \Lambda diag\left(a_1\right)\Lambda'\right| \leq c_1 T^{-1/2} \tag{46}$$

$$\left|A_2 - \Lambda diag\left(a_2\right)\Lambda'\right| \leq c_2 T^{-1/2} \tag{47}$$

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

$$\left|Var\left(\Lambda p - P\right)\right| \leq c_P T^{-1/2} \tag{48}$$

$$\left|Var\left(\Lambda \tilde{q}_i - \tilde{Q}_i\right)\right| \leq c_Q T^{-1/2} \tag{49}$$

for constants c_P and c_Q.

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

$$|A - B| = \bar{O}\left(T^{-1/2}\right) \Longleftrightarrow |A - B| \leq bT^{-1/2} \tag{50}$$

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 $|A - B| = \bar{O}\left(T^{-1/2}\right)$ and $|C - D| = \bar{O}\left(T^{-1/2}\right)$, then

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

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

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

$$|AC - BD| = \bar{O}\left(T^{-1/2}\right) \tag{54}$$

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 Σ_D, Σ_Z and Σ_{avg} using those operations. it follows that $|A_1 - \Lambda diag\,(a_1)\,\Lambda'| \leq c_1 T^{-\frac{1}{2}}$ for some constant c_1, and the same holds for A_2 for some constant c_2.

For the variance of prices, we define

$$R_1 \equiv A_1 - \Lambda diag\,(a_1)\,\Lambda', \tag{55}$$

$$R_2 \equiv A_2 - \Lambda diag\,(a_2)\,\Lambda'. \tag{56}$$

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

$$\|A\| = \max_{x'x=a} \left(x'A'Ax\right)^{\frac{1}{2}}. \tag{57}$$

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

$$|AB| \leq \sqrt{T}\,|A|\,|B|\,. \tag{58}$$

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

$$\begin{aligned} |Var\,[P - \Lambda p]| &= \left|Var\left[(A_1 - \Lambda a_1 \Lambda')D + (A_2 - \Lambda a_2 \Lambda')Z\right]\right| && (59)\\ &\leq \left|R_1 \Sigma_D R_1'\right| + \left|R_2 \Sigma_Z R_2'\right| && (60)\\ &\leq \sqrt{T}\left(|R_1 \Sigma_D|\,|R_1| + |R_2 \Sigma_Z|\,|R_2|\right) && (61)\\ &\leq \sqrt{T}\left(\|\Sigma_D\|\,|R_1|^2 + \|\Sigma_Z\|\,|R_2|^2\right) && (62)\\ &\leq \sqrt{T}K\left(|R_1|^2 + |R_2|^2\right). && (63) \end{aligned}$$

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 Σ_D and Σ_Z are bounded. Indeed, since Σ_D and Σ_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:

$$\begin{aligned} |Var\left[P - \Lambda p\right]| &\leq K\sqrt{T}\left(c_1^2 + c_2^2\right) T^{-1} &(64)\\ &= c_P T^{-\frac{1}{2}}. &(65)\end{aligned}$$

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

To prove inequality (58), note that:

$$\begin{aligned} |AB|^2 &= \tfrac{1}{T}\sum_{m,n}\left(\sum_{t=1}^{T} a_{mt}b_{tn}\right)^2 \\ &\leq \tfrac{1}{T}\sum_{m,n}\left(\sum_{t=1}^{T} a_{mt}^2\right)\left(\sum_{t=1}^{T} b_{tn}^2\right) \\ &= \tfrac{1}{T}\left(\sum_{m,t} a_{mt}^2\right)\left(\sum_{n,t} b_{nt}^2\right) \\ &= T\left(\tfrac{1}{T}\left(\sum_{m,t} a_{mt}^2\right)\right)\left(\tfrac{1}{T}\left(\sum_{n,t} b_{nt}^2\right)\right) \\ &= T\,|A|^2\,|B|^2\,, \end{aligned} \tag{66}$$

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 $\varepsilon_{i,t}$ to be stationary.

Inserting the optimal value of $q_{i,j}$ into the utility function, we obtain

$$E_{-1}\left[U_{i,0}\right] \equiv \frac{1}{2} E\left[T^{-1} \sum_j \tau_{i,j} E\left[d_j - p_j \mid y_{i,j}, p_j\right]^2\right] \tag{68}$$

$U_{i,0}$ 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

$$Var\left[d_j - p_j\right] = Var\left[E\left[d_j - p_j \mid y_{i,j}, p_j\right]\right] + E\left[Var\left[d_j - p_j \mid y_{i,j}, p_j\right]\right] \tag{69}$$

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

$$Var(r_j) = Var\left[d_j - p_j\right] = \left(1 - a_{1,j}\right)^2 f_{D,j} + \frac{a_{1,j}^2}{\left(\rho f_{avg,j}^{-1}\right)^2} f_{Z,j}. \tag{70}$$

So then

$$E_{-1}\left[U_{i,0}\right] = \frac{1}{2} T^{-1} \sum_j \left[\left(\left(1 - a_{1,j}\right)^2 f_{D,j} + \frac{a_{1,j}^2}{(\rho f_{avg}^{-1})^2} f_{Z,j}\right) \tau_{i,j} - 1\right]. \tag{71}$$

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

$$\lambda_j\left(f_{avg,j}^{-1}\right) \equiv (1 - a_{1,j})^2 f_{D,j} + \left(\frac{a_{1,j}}{\rho f_{avg,j}^{-1}}\right)^2 f_{Z,j} = Var(r_j). \tag{72}$$

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

$$\lambda_j\left(f_{avg,j}^{-1}\right) = \frac{f_{D,j}\left(f_{D,j}^{-1} + \rho^{-1} k\right)^2 + (\rho f_{avg,j}^{-1})^2 f_{Z,j}^{-1} + f_{Z,j} \rho^{-2}}{\left((\rho f_{avg,j}^{-1})^2 f_{Z,j}^{-1} + f_{D,j}^{-1} + \rho^{-1} k + f_{avg,j}^{-1}\right)^2}, \tag{73}$$

which can be further decomposed as:

$$\lambda_j\left(f_{avg,j}^{-1}\right) = \frac{1}{\left((\rho f_{avg,j}^{-1})^2 f_{Z,j}^{-1}+f_{D,j}^{-1}+\rho^{-1}k+f_{avg,j}^{-1}\right)^2} + \frac{f_{Z,j}-\frac{f_{avg,j}^{-1}}{\rho}}{\left((\rho f_{avg,j}^{-1})^2 f_{Z,j}^{-1}+f_{D,j}^{-1}+\rho^{-1}k+f_{avg,j}^{-1}\right)^2} + \frac{\rho^{-1}k(1+f_{D,j}^{-1}\rho^{-1}k)}{\left((\rho f_{avg,j}^{-1})^2 f_{Z,j}^{-1}+f_{D,j}^{-1}+\rho^{-1}k+f_{avg,j}^{-1}\right)^2} \tag{74}$$

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

$$E_{-1}\left[U_{i,0}\right] = \frac{1}{2}T^{-1}\sum_j\left[\frac{a_1^2}{a_2^2}f_Z^{-1}+f_D^{-1}\right] \tag{75}$$

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:

$$\tau_i = \left(\frac{a_1}{a_2}\right)^2 f_Z^{-1}+f_D^{-1}. \tag{76}$$

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

$$\rho^{-1}\tilde{q}_{LF,i} = \left((1-a_1)\left(\frac{a_1}{a_2}\right)^2 f_Z^{-1}-a_1 f_D^{-1}\right)d+\left(\frac{a_1(1-a_1)}{a_2}f_Z^{-1}-a_2 f_D^{-1}\right)z. \tag{77}$$

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

$$\begin{aligned}\rho^{-1}E_{-1}\left[\tilde{q}_{LF,i}r\right] &= \left((1-a_1)\left(\tfrac{a_1}{a_2}\right)^2 f_Z^{-1}-a_1 f_D^{-1}\right)(1-a_1)f_D-\left(\tfrac{a_1(1-a_1)}{a_2}f_Z^{-1}-a_2 f_D^{-1}\right)a_2 f_Z\\ &= (1-a_1)^2\left(\tfrac{a_1}{a_2}\right)^2 f_Z^{-1}f_D-2a_1(1-a_1)+a_2^2 f_Z f_D^{-1}\\ &= \left((1-a_1)\left(\tfrac{a_1}{a_2}\right)(f_Z^{-1}f_D)^{\frac{1}{2}}-a_2(f_Z f_D^{-1})^{\frac{1}{2}}\right)^2\end{aligned} \tag{78}$$

$$(f_Z^{-1}f_D)\left((1-a_1)\left(\frac{a_1}{a_2}\right)-a_2 f_Z f_D^{-1}\right)^2 \tag{79}$$

From the equilibrium condition for $f_{avg,j}^{-1}$ stated in the text, a marginal reduction in α_j at $\alpha_j=\lambda_j(0)/\psi$ leads to a marginal increase in $f_{avg,j}^{-1}$, so the signs of the derivatives with respect to α_j are simply the reverse of the signs of the derivatives with respect to $f_{avg,j}^{-1}$. We now calculate derivatives with respect to $f_{avg,j}^{-1}$.

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

$$\begin{aligned}\rho^{-1}\frac{dE_{-1}[\tilde{q}_{LF,i}r]}{df_{avg}^{-1}} &= 2\left((1-a_1)\left(\frac{a_1}{a_2}\right)(f_Z^{-1}f_D)^{\frac{1}{2}} - a_2(f_Z f_D^{-1})^{\frac{1}{2}}\right) \\ &\times\left\{\rho\left[(1-a_1)(f_Z^{-1}f_D)^{\frac{1}{2}} - a_1(f_Z f_D^{-1})^{\frac{1}{2}}\right] - \left[(f_Z^{-1}f_D)^{\frac{1}{2}} + (f_Z f_D^{-1})^{\frac{1}{2}}\right]\rho\frac{\partial a_1}{\partial f_{avg}^{-1}}f_{avg}^{-1}\right\}\end{aligned} \tag{80}$$

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

$$\frac{\partial a_1}{\partial f_{avg}^{-1}}f_{avg}^{-1} = a_1(1-a_1) + (1-a_1)\frac{(\rho f_{avg}^{-1})^2 f_Z^{-1}}{(\rho f_{avg}^{-1})^2 f_Z^{-1} + f_{avg}^{-1} + f_D^{-1} + \rho^{-1}k}. \tag{81}$$

The following limits follow from the discussion in Appendix 2.1.2:

$$\lim_{f_{avg}^{-1}\to 0^+} a_1 = 0, \qquad \lim_{f_{avg}^{-1}\to 0^+} a_2 = \frac{1}{\rho f_D^{-1} + k}. \tag{82}$$

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

$$\lim_{f_{avg}^{-1}\to 0^+}\frac{\partial E_{-1}\left[\tilde{q}_{LF,i}r\right]}{\partial f_{avg}^{-1}} = -2\rho\frac{(f_Z f_D^{-1})^{\frac{1}{2}}(f_Z^{-1}f_D)^{\frac{1}{2}}}{f_D^{-1}+\rho^{-1}k} < 0. \tag{83}$$

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

$$\frac{d}{df_{avg,j}^{-1}}E_{-1}\left[\sum_t \tilde{Q}_{LF,t}\left(D_t - P_t\right)\right] = \frac{d}{df_{avg,j}^{-1}}\sum_k E_{-1}\left[\tilde{q}_{LF,k}r_k\right] = \frac{d}{df_{avg,j}^{-1}}E_{-1}\left[\tilde{q}_{LF,j}r_j\right] < 0; \tag{84}$$

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,

$$\frac{d}{df_{avg,j}^{-1}}E_{-1}\left[U_{LF,0}\right] = \frac{d}{df_{avg,j}^{-1}}E_{-1}\left[u_{LF,0,j}\right], \tag{85}$$

where

$$E_{-1}[u_{LF,0,j}] \equiv \frac{1}{2}T^{-1}\left[\left((1-a_{1,j})^2 f_{D,j} + a_{2,j}^2 f_{Z,j}\right)\tau_{i,j} - 1\right] \tag{86}$$

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_{avg} assuming that $f_i^{-1} = 0$ is:

$$\begin{aligned} 2T\tfrac{dE_{-1}[u_{LF,0}]}{df_{avg}^{-1}} &= \left((1-a_1)^2 f_D + a_1^2(\rho f_{avg}^{-1})^2 f_Z\right) 2\rho^2 f_Z^{-1} f_{avg}^{-1} \\ &+ \left(-2(1-a_1)\tfrac{\partial a_1}{\partial f_{avg}^{-1}} f_D + 2a_1 \tfrac{\partial a_1}{\partial f_{avg}^{-1}} \left(\rho f_{avg}^{-1}\right)^2 f_Z + 2a_1^2 \rho^2 f_Z f_{avg}^{-1}\right) \left(\left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_D^{-1}\right) \end{aligned} \tag{87}$$

Given that:

$$\lim_{f_{avg}^{-1} \to 0^+} a_1 = 0, \tag{88}$$

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

$$-2(1-a_1)\frac{\partial a_1}{\partial f_{avg}^{-1}} f_D + 2a_1 \frac{\partial a_1}{\partial f_{avg}^{-1}} \rho f_{avg}^{-1} f_Z. \tag{89}$$

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

$$\lim_{f_{avg}^{-1} \to 0^+} \frac{\partial a_1}{\partial f_{avg}^{-1}} f_{avg}^{-1} = 0, \tag{90}$$

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

$$\frac{\partial a_1}{\partial f_{avg}^{-1}} = \frac{a_1}{f_{avg}^{-1}} + o(1) = \frac{1 + (\rho f_{avg}^{-1}) f_Z^{-1}}{(\rho f_{avg}^{-1})^2 f_Z^{-1} + f_D^{-1} + f_{avg}^{-1} + \rho^{-1} k} + o(1). \tag{91}$$

Therefore,

$$\lim_{f_{avg}^{-1} \to 0^+} 2T \frac{dE_{-1}[u_{LF,0}]}{df_{avg}^{-1}} = -2\frac{f_D}{f_D^{-1} + \rho^{-1} k} = -2 f_D a_2 < 0, \tag{92}$$

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

$$Var\left[d_j \mid p_j\right] = \left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_D^{-1} \tag{93}$$

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 ϕ is a parameter determining the smoothness of the estimated spectrum. The factor of proportionality, denoted σ_P^2, 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, $\bar{X}$ and $\bar{F}$. The prior covariance matrix is then Σ, where the i, j entry is $\sigma_P^2 \exp\left(-\phi\left|\bar{X}_i - \bar{X}_j\right|\right)$. 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 $\chi_2^2/2$. 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

$$-\hat{F}\Sigma^{-1}\hat{F} - \left(\hat{F} + b - \bar{F}\right)\Sigma_{samp}^{-1}\left(\hat{F} + b1 - \bar{F}\right) \tag{94}$$

where 1 is a vector of 1's and Σ_{samp} is the variance matrix of the log spectrum. This is, given basic properties of the periodogram, the variance of a $\chi_2^2/2$ (see, e.g., Brillinger (1981)).

The first-order condition for b is

$$\begin{aligned} 0 &= 1'\Sigma_{samp}^{-1}\left(\hat{F} + b1 - \bar{F}\right) && (95) \\ b &= \left(1'\Sigma_{samp}^{-1}1\right)^{-1} 1'\Sigma_{samp}^{-1}\left(\hat{F} - \bar{F}\right) && (96) \end{aligned}$$

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

$$\max_{\hat{F}} -\hat{F}\Sigma^{-1}\hat{F} - \left(\hat{F} + 1\left(1'\Sigma_{samp}^{-1}1\right)^{-1} 1'\Sigma_{samp}^{-1}\left(\hat{F} - \bar{F}\right) - \bar{F}\right)\Sigma_{samp}^{-1}\left(\hat{F} + 1\left(1'\Sigma_{samp}^{-1}1\right)^{-1} 1'\Sigma_{samp}^{-1}\left(\hat{F} - \bar{F}\right) - \bar{F}\right) \tag{97}$$

Yielding

$$\hat{F} = \left(\Sigma^{-1} + V\right)^{-1} V\bar{F} \tag{98}$$

$$\text{where } V \equiv \left(I + 1\left(1'\Sigma_{samp}^{-1}1\right)^{-1} 1'\Sigma_{samp}^{-1}\right)\Sigma_{samp}^{-1}\left(I + 1\left(1'\Sigma_{samp}^{-1}1\right)^{-1} 1'\Sigma_{samp}^{-1}\right) \tag{99}$$

Because the set of frequencies, $\bar{X}$, 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 Δ is the first-difference operator. The futures claims are in this case claims to ΔD_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,

$$R_n^{period} = \frac{\sum_{t=1}^{52n} (\Delta D_t - P_t)}{52n} \tag{100}$$

The analogous calculation for equity is

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

We calculate per-period return variances similarly. Specifically,

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

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

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

$$
\begin{aligned}
Var\left(\Delta D_t - P_t\right) &= 1_t' \Lambda diag\left(f_R\right) \Lambda' 1_t && (104)\\
&= \left(\Lambda' 1_t\right)' diag\left(f_R\right) \left(\Lambda' 1_t\right) && (105)\\
&= \sum_j \lambda_{t,j}^2 Var(d_j - p_j) && (106)\\
&= \lambda_{t,0}^2 Var(d_0 - p_0) + \lambda_{t,T/2}^2 Var(d_{\frac{T}{2}} - p_{\frac{T}{2}}) + \sum_{j=1}^{T/2-1} \left(\lambda_{t,2j}^2 + \lambda_{t,2j+1}^2\right) Var(d_j - p_j) && (107)
\end{aligned}
$$

where f_R is the spectrum of returns, $f_R \equiv Var\left(d_j - p_j\right)$, 1_t is a vector equal to 1 in its tth element and zero elsewhere, and $\lambda_{t,j}$ is the jth 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

$$
\begin{aligned}
Var\left(\frac{1}{s}\sum_{m=0}^{s-1} R_{t+m}\right) &= \frac{1}{s^2}\left(\sum_{m=0}^{s-1} 1_{t+m}\right)' \Lambda diag\left(f_R\right) \Lambda' \left(\sum_{m=0}^{s-1} 1_{t+m}\right) && (108)\\
&= \frac{1}{s^2}\left(\sum_{m=0}^{s-1} \lambda_{t+m,0}\right)^2 f_{R,0} + \frac{1}{s^2}\left(\sum_{m=0}^{s-1} \lambda_{t+m,T/2}\right)^2 f_{R,T/2} && (109)\\
&\quad + \frac{1}{s^2}\sum_{j=1}^{T/2-1}\left[\left(\sum_{m=0}^{s-1} \lambda_{t+m,2j}\right)^2 + \left(\sum_{m=0}^{s-1} \lambda_{t+m,2j+1}\right)^2\right] f_{R,j} && (110)
\end{aligned}
$$

For $0 < j < T/2$

$$
\left(\sum_{m=0}^{s-1} \lambda_{t+m,j}\right)^2 + \left(\sum_{m=0}^{s-1} \lambda_{t+m,j}\right)^2 = \sum_{m=0}^{s-1}\sum_{k=0}^{s-1} \frac{2}{T}\left[\begin{array}{c} \cos\left(2\pi j\left(t+m-1\right)/T\right)\cos\left(2\pi j\left(t+k-1\right)/T\right) \\ +\sin\left(2\pi j\left(t+m-1\right)/T\right)\sin\left(2\pi j\left(t+k-1\right)/T\right)\end{array}\right] \quad (111)
$$

Now note that

$$
2\cos\left(x\right)\cos\left(y\right) + 2\sin\left(x\right)\sin\left(y\right) = 2\cos\left(x-y\right) \quad (112)
$$

So we have

$$\left(\sum_{m=0}^{s-1}\lambda_{t+m,j}\right)^2+\left(\sum_{m=0}^{s-1}\lambda_{t+m,j}\right)^2 = \frac{2}{T}\sum_{m=0}^{s-1}\sum_{k=0}^{s-1}\cos\left(\frac{2\pi j}{T}(m-k)\right) \tag{113}$$

$$= 2\frac{s}{T}\sum_{m=-(s-1)}^{s-1}\frac{s-|m|}{s}\cos\left(\frac{2\pi j}{T}m\right) \tag{114}$$

$$= 2\frac{s}{T}F_s\left(\frac{2\pi j}{T}\right) \tag{115}$$

where F_s denotes the sth-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 $j = 0$,

$$\left(\sum_{m=0}^{s-1}f_{t+m,0}\right)^2 = \left(\sum_{m=0}^{s-1}\sqrt{1/T}\right)^2 \tag{116}$$

$$= \left(s\frac{1}{T^{1/2}}\right)^2 \tag{117}$$

$$= \frac{s}{T}F_s(0), \tag{118}$$

since $F_s(0) = s$ (technically, this holds as a limit: $\lim_{x\to 0}F_s(x) = s$). For $j = T/2$,

$$\left(\sum_{m=0}^{s-1}f_{t+m,T/2}\right)^2 = \frac{1}{T}\left(\sum_{m=1}^{s}(-1)^m\right)^2 = \begin{cases}\frac{1}{T} \text{ for odd } s\\ 0 \text{ otherwise}\end{cases} \tag{119}$$

$$= \frac{s}{T}\frac{1}{s}\left(\frac{\sin(s\pi/2)}{\sin(\pi/2)}\right)^2 = \frac{s}{T}F_s(\pi) \tag{120}$$

So we finally have that

$$Var\left(\frac{1}{s}\sum_{m=0}^{s-1}R_{t+m}\right) = \frac{1}{sT}\left(F_s(0)f_{R,0}+\sum_{j=1}^{T/2-1}F_s(\omega_j)f_{R,2j}+F_s(\pi)f_{R,T/2}\right) \tag{121}$$

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,

$$\zeta = d + \varepsilon_\zeta \tag{122}$$

$$Var(\varepsilon_\zeta) = f_\zeta \tag{123}$$

This section examines the effects of varying the precision of that signal, f_ζ^{-1}.

4.1 Statistical inference

We guess that prices take the form

$$p = a_1 d + a_2 z + a_\zeta \zeta \tag{124}$$

$(p - a_\zeta \zeta)/a_1$ 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

$$\tau_i = \frac{a_1^2}{a_2^2} f_Z^{-1} + f_i^{-1} + f_\zeta^{-1} + f_D^{-1} \tag{125}$$

and the posterior mean is

$$E\left[d-p \mid y_i, p-a_\zeta\zeta\right] = \tau_i^{-1}\frac{a_1^2}{a_2^2}f_Z^{-1}\left(p-a_\zeta\zeta\right)a_1^{-1} + \tau_i^{-1}f_i^{-1}y_i + \tau_i^{-1}f_\zeta^{-1}\zeta - p \tag{126}$$

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

$$Var\left(d \mid p\right) = \frac{a_2^2}{\left(a_1+a_\zeta\right)^2}f_Z + \frac{a_\zeta^2}{\left(a_1+a_\zeta\right)^2}f_\zeta \tag{127}$$

4.2 Demand and equilibrium

Agent i's demand is

$$\begin{aligned}
\tilde{q}_i &= \rho\tau_i E\left[d-p \mid y_i, p\right] && (128)\\
&= \rho\left(\frac{a_1^2}{a_2^2}f_Z^{-1}\left(p-a_\zeta\zeta\right)a_1^{-1} + f_i^{-1}y_i + f_\zeta^{-1}\zeta - \tau_i p\right) && (129)\\
&= \rho\left(\left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_i\right)p + f_i^{-1}y_i + \left(f_\zeta^{-1} - \frac{a_1}{a_2^2}a_\zeta f_Z^{-1}\right)\zeta\right) && (130)
\end{aligned}$$

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

$$\begin{aligned}
-z + k\left(a_1 d + a_2 z + a_\zeta\zeta\right) &= \rho\int_i\left(\left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_i\right)p + f_i^{-1}y_i + \left(f_\zeta^{-1} - \frac{a_1}{a_2^2}a_\zeta f_Z^{-1}\right)\zeta\right)di && (131)\\
&= \rho\left(\left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_{avg}\right)\left(a_1 d + a_2 z + a_\zeta\zeta\right) + f_{avg}^{-1}d + \left(f_\zeta^{-1} - \frac{a_1}{a_2^2}a_\zeta f_Z^{-1}\right)\zeta\right) && (132)
\end{aligned}$$

where the second line uses the law of large numbers. Matching coefficients on d, z, and ζ then yields

$$\begin{aligned}
k &= \rho\left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_{avg}\right) + \rho f_{avg}^{-1}a_1^{-1} && (133)\\
-a_2^{-1}\left(1-ka_2\right) &= \rho\left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_{avg}\right) && (134)\\
k &= \rho\left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_{avg}\right) + \rho\left(a_\zeta^{-1}f_\zeta^{-1} - \frac{a_1}{a_2^2}f_Z^{-1}\right) && (135)
\end{aligned}$$

$$a_\zeta = f_\zeta^{-1}a_1\left(f_{avg}^{-1} + \frac{a_1^2}{a_2^2}f_Z^{-1}\right)^{-1} \tag{136}$$

$$\frac{a_1}{a_2} = \rho f_{avg}^{-1} \tag{137}$$

$$a_1 = \frac{\left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_{avg}^{-1}}{\tau_{avg} + \rho^{-1} k} \tag{138}$$

which implies

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

4.3 Utility and profits

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

$$E_{-1}\left[u_{i,0}\right] = \lambda_j \left(f_{avg}^{-1}, f_\zeta^{-1}\right) \left(\left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_i^{-1} + f_D^{-1} + f_\zeta^{-1}\right) - 1 \tag{140}$$

where $\lambda_j \left(f_{avg}^{-1}, f_\zeta^{-1}\right)$ is the variance of returns, and we write it as a function of f_ζ^{-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 $\lambda_j \left(f_{avg}^{-1}, f_\zeta^{-1}\right) \left(\left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_i^{-1} + f_D^{-1} + f_\zeta\right.$ 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 $\tau_{avg} \to \infty$, $a_1 \to 0$, $a_2 \to 0$, 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,

$$E\left[\left(z - k\left(\left(a_1 + a_\zeta\right) d + a_2 z + a_\zeta \varepsilon_\zeta\right)\right)\left(\left(1 - a_1 - a_\zeta\right) d - a_2 z - a_\zeta \varepsilon_\zeta\right)\right] \tag{141}$$

$$= -a_2 f_Z - k\left(a_1 + a_\zeta\right)\left(1 - a_1 - a_\zeta\right) f_D + k a_2^2 f_Z + k a_\zeta^2 f_\zeta \tag{142}$$

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

$$\lambda_j\left(f_{avg}^{-1}, f_\zeta^{-1}\right)\left(\left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_D^{-1} + f_\zeta^{-1}\right) \tag{143}$$

We are interested in changes in f_ζ, 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_ζ^{-1} (or just invoking the implicit function theorem) yields

$$\frac{df_{avg}^{-1}}{df_\zeta^{-1}} = -\frac{\lambda_{j,2}\left(f_{avg}^{-1}, f_\zeta^{-1}\right)}{\lambda_{j,1}\left(f_{avg}^{-1}, f_\zeta^{-1}\right)} \tag{144}$$

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

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

$$\frac{d}{df_\zeta^{-1}}\left[\lambda_j\left(f_{avg}^{-1}, f_\zeta^{-1}\right)\left(\left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_D^{-1} + f_\zeta^{-1}\right)\right] = \psi_j\left(2\rho^2 f_{avg}^{-1} f_Z^{-1}\frac{df_{avg}^{-1}}{df_\zeta^{-1}} + 1\right) \tag{145}$$

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

$$\lambda\left(0, f_\zeta^{-1}\right) + \left(f_D^{-1} + f_\zeta^{-1}\right)\lambda_{j,2}\left(0, f_\zeta^{-1}\right) \tag{146}$$

We have the following results

1. Information acquisition is weakly decreasing in f_ζ^{-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_ζ^{-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_ζ^{-1} when $f_{avg}^{-1} > 0$.

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

$$\psi_j \left(\left(2\rho^2 f_{avg}^{-1} f_Z^{-1} + 1\right) \frac{df_{avg}^{-1}}{df_\zeta^{-1}} + 1 \right) \tag{147}$$

which can be shown to be negative

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

– See the previous section.

6. The total precision for fundamentals in public information – prices and the public signal – increases in f_ζ^{-1}.

– Public precision is $\left(\rho f_{avg}^{-1}\right)^2 f_Z^{-1} + f_D^{-1} + f_\zeta^{-1}$. The derivative with respect to f_ζ^{-1} is

$$2\rho^2 f_{avg}^{-1} f_Z^{-1} \frac{df_{avg}^{-1}}{df_\zeta^{-1}} + 1 \tag{148}$$

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

4.5 Numerical example

We consider a simple numerical example with $\rho = k = f_D^{-1} = f_Z^{-1} = 1$ and $\psi = 0.4757$ and examine how profits, utility, and price informativeness vary with f_ζ^{-1}. The four panels of figure A.4 plot results from a numerical solution, with f_ζ^{-1} 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 $f_\zeta^{-1} = 0$, their profits are not zero, just numerically very small). Eventually f_ζ^{-1} rises sufficiently high that $f_{avg}^{-1} = 0$. At that point, more precision for public signals just makes prices more informative 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_{ζ}^{-1} rises because they acquire less information – f_{avg}^{-1} 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 Δ to be the first difference operator so that

$$\Delta D_t = D_t - D_{t-1} \tag{149}$$

and define the vector $\Delta D \equiv [\Delta D_1, \Delta D_2, ...\Delta D_T]'$. We assume that

$$\Delta D \sim N\left(0, \Sigma_D\right). \tag{150}$$

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

$$\begin{aligned} Q_i' D &= Q_i' L_1 \Delta D && (151) \\ &= \left(L_1' Q_i\right)' \Delta D && (152) \end{aligned}$$

where L_1 is a matrix that creates partial sums,

$$L_1 \equiv \begin{bmatrix} 1 & 0 & 0 & \cdots \\ 1 & 1 & 0 & \\ 1 & 1 & 1 & \\ \vdots & & & \ddots \end{bmatrix} \tag{153}$$

So an allocation of Q_i to the futures is equivalent to an allocation of $L_1' Q_i$ to claims on the first differences of fundamentals, which we will call the growth rate futures. Define the notation

$$Q_{\Delta,i} \equiv L_1' Q_i \tag{154}$$

Furthermore, the prices of the growth rate futures are simply the vector ΔP (by the law of one price). We can therefore rewrite the optimization problem equivalently as

$$\max T^{-1} \sum_{t=1}^{T} \beta^t Q_{\Delta,i,t} E_{0,i} \left[\Delta D_t - \Delta P_t\right] - \frac{1}{2} \left(\rho T^{-1}\right) Var_{0,i} \left[\sum_{t=1}^{T} \beta^t Q_{\Delta,i,t} \left(\Delta D_t - \Delta P_t\right)\right] \tag{155}$$

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 $\left(L_1'\right)^{-1} Q_{\Delta,i}$ to the levels (which follows trivially from the definition of $Q_{\Delta,i}$ in

(154)).

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 Λ, but now in first differences. So the frequency domain allocations in terms of first differences will be

$$\tilde{Q}_{\Delta D,i} = \Lambda \tilde{q}_{\Delta,i} \tag{156}$$

where $\tilde{Q}_{\Delta D,i,t} \equiv Q_{\Delta D,i,t}\beta^t$. $\tilde{q}_{\Delta,i}$ 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

$$\tilde{Q}_i = \left(L_1'\right)^{-1} \tilde{Q}_{\Delta,i} \tag{157}$$

$$= \left(L_1'\right)^{-1} \Lambda \tilde{q}_{\Delta,i} \tag{158}$$

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

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

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

So the way that $\left(L_1'\right)^{-1}$ 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 $\left(L_1'\right)^{-1} \Lambda$ are equal to $\left(L_1'\right)^{-1} c_j$ and $\left(L_1'\right)^{-1} s_j$, which can be written using standard trigonometric formulas as:

$$\left(L_1'\right)^{-1} c_j \approx \begin{bmatrix} 2\sin\left(\frac{1}{2}\omega_j\right)\sqrt{\frac{2}{T}}\left\{\sin\left(\omega_j\left(t-\frac{1}{2}\right)\right)\right\}_{t=2}^{T} \\ \sqrt{\frac{2}{T}}\cos\left(\omega_j(T-1)\right) \end{bmatrix} \tag{160}$$

$$\left(L_1'\right)^{-1} s_j \approx \begin{bmatrix} -2\sin\left(\frac{1}{2}\omega_j\right)\sqrt{\frac{2}{T}}\left\{\cos\left(\omega_j\left(t-\frac{1}{2}\right)\right)\right\}_{t=2}^{T} \\ \sqrt{\frac{2}{T}}\sin\left(\omega_j(T-1)\right) \end{bmatrix} \tag{161}$$

The column c_j represents a portfolio in terms of the first differences of fundamentals with weights equal to a cosine fluctuating at frequency ω_j. $\left(L_1'\right)^{-1} c_j$ measures the loadings of that portfolio on claims to the level of fundamentals. These loadings also fluctuate at frequency ω_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 Λ 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 ω_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 ω_j.

One subtlety is in the lowest-frequency portfolio, $(L_1'^{-1}\left(\frac{1}{\sqrt{2}}c_0\right)$. 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 $\left(L_1'\right)^{-1}\left(\frac{1}{\sqrt{2}}c_0\right) = \left[0,0,0,...,\sqrt{2/T}\right]$. 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.[1]

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,

[1] The highest frequency portfolio, $(L_1'^{-1}\left(\frac{1}{\sqrt{2}}c_{\frac{T}{2}}\right)$, is given by $1/\sqrt{T}\,(2,-2,...,2,1)'$, and therefore fluctuates at the highest sample frequency.

the noise traders' optimization problem can be written in terms of first differences,

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

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

$$\Delta D \sim N\left(0, \Sigma_{N\Delta}^{prior}\right) \tag{163}$$

$$S \sim N\left(\Delta D, \Sigma_{N\Delta}^{signal}\right) \tag{164}$$

and the Bayesian update is

$$\Delta D \mid S \sim N\left(\Sigma_{N\Delta} \left(\Sigma_{N\Delta}^{signal}\right)^{-1} S, \Sigma_{N\Delta}\right) \tag{165}$$

$$\text{where } \Sigma_{N\Delta} \equiv \left(\left(\Sigma_{N\Delta}^{signal}\right)^{-1} + \left(\Sigma_{N\Delta}^{prior}\right)^{-1} \right)^{-1} \tag{166}$$

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 $P_t^{equity} = \sum_{j=1}^{T-t} \beta^{-j} P_{t+j}$. An investor's exposure to fundamentals on date t, $Q_{i,t}$ can be acquired either by buying $Q_{i,t}$ units of forwards on date 0 or by holding $Q_{i,t}^{EQ}$ 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, $\Delta Q_{i,t}^{EQ} = \Delta Q_{i,t}$.

We assume that investors now maximize the following objective:

$$U_{0,i} = \max_{\{Q_{i,t}\}} E_{0,i} \left[T^{-1} \sum_{t=1}^{T} Q_{i,t} \left(D_t - P_t \right) \right] - \frac{1}{2} c T^{-2} E_{0,i} \left[QV \left\{ Q_i \right\} \right] - \frac{1}{2} b T^{-2} E_{0,i} \left[\sum_{t=1}^{T} Q_{i,t}^2 \right], \quad (167)$$

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:

$$QV \left\{ Q_i \right\} \equiv \left[\sum_{t=2}^{T} \left(Q_{i,t} - Q_{i,t-1} \right)^2 + \left(Q_{i,1} - Q_{i,T} \right)^2 \right]. \quad (168)$$

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:

$$T^{-1} QV \left\{ Q_i \right\} = 2 \sum_{j=1}^{T} \sin^2 \left(\omega_{\lfloor j/2 \rfloor} / 2 \right) q_{i,j}^2. \quad (169)$$

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 $\sin^2 \left(\omega_{\lfloor j/2 \rfloor} / 2 \right)$, 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:

$$\frac{1}{2}(c\sin^2\left(\omega_{\lfloor j/2\rfloor}/2\right)+b)^{-1}Var\left[d_j \mid p_j, y_{i,j}\right]^2. \tag{170}$$

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, $Var(d_j|p_j)$, rises more at higher frequencies. The corollary regarding price informativeness refers to the fact that the variance of moving averages of the form:

$$Var\left(\frac{1}{n}\sum_{m=0}^{n-1}D_{t+m}\mid P\right) \tag{171}$$

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*

$$p_j = a_{1,j} d_j + a_{2,j} z_j \tag{172}$$

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

$$Q_{i,t} = \sum_j \begin{bmatrix} q_j \cos(2\pi jt/T) \\ +q_{j'} \sin(2\pi jt/T) \end{bmatrix}. \tag{173}$$

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

$$QV\{Q_i\} \equiv \sum_{t=2}^{T} (Q_{i,t} - Q_{i,t-1})^2 + (Q_{i,1} - Q_{i,T})^2. \tag{174}$$

We can write that as

$$QV\{Q_i\} = (DQ)'(DQ) \tag{175}$$

where D is a matrix that generates first differences,

$$D \equiv \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \\ 1 & 0 & \cdots & 0 & -1 \end{bmatrix}. \tag{176}$$

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

$$QV\{Q_i\} = q'\Lambda' D' D \Lambda q \tag{177}$$

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 Λ, 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 $m \neq n$:

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

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

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

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

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

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

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

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

We then have

$$QV\left\{Q_i\right\} = q\Lambda' D' D\Lambda q \tag{184}$$

$$= \sum_{j=1}^{T} 2T\sin^2\left(\omega_{\lfloor j/2 \rfloor}/2\right) q_{i,j}^2 \tag{185}$$

Total holding costs can be written as:

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

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

$$
\begin{aligned}
0 &= E\left[d_j - p_j \mid y_{i,j}, p_j\right] - 2c\sin^2\left(\omega_{\lfloor j/2\rfloor}/2\right) q_j - bq_j && (187)\\
q &= \frac{E\left[d_j - p_j \mid y_{i,j}, p_j\right]}{\gamma_j} && (188)\\
&= \gamma_j^{-1}\tau_i^{-1}\left(f_i^{-1}y_i + \left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_i\right)p\right) && (189)
\end{aligned}
$$

where

$$\gamma_j \equiv 2c\sin^2\left(\omega_{\lfloor j/2\rfloor}/2\right) + b \qquad (190)$$

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

$$
\begin{aligned}
\int_i \gamma_j^{-1}\tau_i^{-1}\left(f_i^{-1}y_i + \left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_i\right)(a_1 d - a_2 z)\right)di &= z_j && (191)\\
\int_i \gamma_j^{-1}\tau_i^{-1}\left(f_i^{-1}d + \left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_i\right)(a_1 d - a_2 z)\right)di &= z_j && (192)
\end{aligned}
$$

Matching coefficients then yields

$$
\begin{aligned}
\int_i \gamma_j^{-1}\tau_i^{-1}\left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_i\right)di &= -a_2^{-1} && (193)\\
\int_i \gamma_j^{-1}\tau_i^{-1}\left(f_i^{-1} + \left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_i\right)a_1\right)di &= 0 && (194)
\end{aligned}
$$

Combining those two equations, we obtain

$$\int_i \gamma_j^{-1}\tau_i^{-1}f_i^{-1}di = \frac{a_1}{a_2} \qquad (195)$$

Now put the definition of τ_i into that equation for f_i^{-1}

$$\int_i \gamma_j^{-1}\tau_i^{-1}\left(\tau_i - \frac{a_1^2}{a_2^2} f_Z^{-1} - f_D^{-1}\right) di \quad = \quad \frac{a_1}{a_2} \tag{196}$$

$$\gamma_j^{-1}\int_i 1 - \left(\frac{a_1^2}{a_2^2} f_Z^{-1} - f_D^{-1}\right)\tau_i^{-1} di \quad = \quad \frac{a_1}{a_2} \tag{197}$$

7.2.2 Expected utility

At any particular frequency,

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

$$= \quad \frac{1}{2}\frac{E\left[d_j - p_j \mid y_{i,j}, p_j\right]^2}{\gamma_j} \tag{199}$$

Expected utility prior to observing signals is then

$$EU_{i,j} \equiv \frac{1}{2} E\left[\frac{E\left[d_j - p_j \mid y_{i,j}, p_j\right]^2}{\gamma_j}\right] \tag{200}$$

$E\left[E\left[d_j - p_j \mid y_{i,j}, p_j\right]^2\right]$ 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

$$Var\left[d_j - p_j\right] = Var\left[E\left[d_j - p_j \mid y_{i,j}, p_j\right]\right] + E\left[Var\left[d_j - p_j \mid y_{i,j}, p_j\right]\right] \tag{201}$$

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 simply

$$Var\left[d_j - p_j\right] \quad = \quad Var\left[\left(1 - a_1\right) d_j + a_2 z_j\right] \tag{202}$$

$$= \quad \left(1 - a_{1,j}\right)^2 f_{D,j} + a_2^2 f_{Z,j} \tag{203}$$

So then

$$EU_{i,j} = \frac{1}{2}\frac{Var\left[d_j - p_j\right] - \tau_{i,j}^{-1}}{\gamma_j} \tag{204}$$

What we end up with is that utility is decreasing in $\tau_{i,j}^{-1}$. That is,

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

7.2.3 Information choice

With the linear cost on precision, agents maximize

$$\begin{aligned} & -\frac{1}{2}\frac{\tau_{i,j}^{-1}}{\gamma_j} - \psi f_{i,j}^{-1} && (206)\\ = & -\frac{1}{2}\left(\frac{a_1^2}{a_2^2} f_{Z,j}^{-1} + f_{i,j}^{-1} + f_{D,j}^{-1}\right)^{-1} \gamma_j^{-1} - \psi f_{i,j}^{-1} && (207)\end{aligned}$$

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

$$\begin{aligned} \psi &= \frac{1}{2}\tau_{i,j}^{-2}\gamma_j^{-1} && (208)\\ \tau_{i,j} &= \frac{1}{\sqrt{2}}\psi^{-1/2}\gamma_j^{-1/2} && (209)\end{aligned}$$

But τ 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,

$$\tau_{i,j} = \max\left(\frac{a_1^2}{a_2^2} f_Z^{-1} + f_D^{-1}, \frac{1}{\sqrt{2}}\psi^{-1/2}\gamma_j^{-1/2}\right) \tag{210}$$

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

$$\begin{aligned} \frac{a_1}{a_2} &= \int_i \gamma_j^{-1}\tau_i^{-1} f_i^{-1} di && (211)\\ &= 0 && (212)\end{aligned}$$

which then implies

$$\tau_{i,j} = \max\left(f_D^{-1}, \frac{1}{\sqrt{2}}\psi^{-1/2}\gamma_j^{-1/2}\right) \tag{213}$$

To get a_2, we have

$$\int_i \left(cj^2 + b\right)\tau_i^{-1}\left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_i\right)di = -a_2^{-1} \tag{214}$$

$$\gamma_j = a_2 \tag{215}$$

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

$$f_D^{-1} < \frac{1}{\sqrt{2}}\psi^{-1/2}\gamma_j^{-1/2} \tag{216}$$

we have, just by rewriting the τ equation,

$$f_i^{-1} = \tau_i - \frac{a_1^2}{a_2^2}f_Z^{-1} - f_D^{-1} \tag{217}$$

Using the second equation from above,

$$\int_i \gamma_j^{-1}\tau_i^{-1}\left(\frac{a_1}{a_2^2}f_Z^{-1} - \tau_i\right)di = -a_2^{-1} \tag{218}$$

$$\int_i \gamma_j^{-1}\tau_i^{-1}\left(\frac{a_1}{a_2}f_Z^{-1} - a_2\tau_i\right)di = -1 \tag{219}$$

$$\int_i \gamma_j^{-1}\left(\tau_i^{-1}\frac{a_1}{a_2}f_Z^{-1} - a_2\right)di = -1 \tag{220}$$

Under the assumption of a symmetric strategy, this is

$$\tau^{-1}\frac{a_1}{a_2}f_Z^{-1} - a_2 = -\gamma_j \tag{221}$$

$$\frac{a_1}{a_2} = \tau f_Z\left(-\gamma_j + a_2\right) \tag{222}$$

Using the other equilibrium condition, we have

$$\int_i \gamma_j^{-1}\tau_i^{-1}\left(\tau_i - \frac{a_1^2}{a_2^2} f_Z^{-1} - f_D^{-1}\right) di = \frac{a_1}{a_2} \tag{223}$$

$$\int_i \gamma_j^{-1}\left(1 - \tau_i^{-1}\frac{a_1}{a_2} f_Z^{-1}\frac{a_1}{a_2} - \tau_i^{-1} f_D^{-1}\right) di = \frac{a_1}{a_2} \tag{224}$$

$$1 - \left(-\gamma_j + a_2\right)\frac{a_1}{a_2} - \tau_i^{-1} f_D^{-1} = \left(cj^2 + b\right)\frac{a_1}{a_2} \tag{225}$$

$$1 - \tau_i^{-1} f_D^{-1} = a_1 \tag{226}$$

Plugging in the formula for τ_i when there is learning,

$$1 - \sqrt{2}\psi^{1/2}\gamma_j^{1/2} f_D^{-1} = a_1. \tag{227}$$

The expression for a_2 can be obtained from:

$$\frac{a_1}{\tau f_Z} = \left(-\gamma_j + a_2\right) a_2. \tag{228}$$

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

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

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

$$f_{avg,j}^{-1} = \begin{cases} 0 & \text{if } \psi_j > \lambda_j(0) \\ \lambda_j^{-1}\left(\psi_j\right) & \text{if } \psi_j \leq \lambda_j(0) \end{cases} \tag{230}$$

where, as before $\lambda_j(0) = f_{D,j} + \frac{f_{Z,j}}{\left(\rho f_{D,j}^{-1} + k\right)^2}$.

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 $\min\left(\psi_j, \lambda_j(0)\right)$ – 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 ψ replaced by ψ_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

$$\max_{\{f_{i,j}^{-1}\}} \sum_{j=1}^{T} \lambda_j \left(f_{avg,j}^{-1}\right) f_{i,j}^{-1} \quad \text{s.t.} \quad \sum_{j=1}^{T} \psi_j f_{i,j}^{-1} \leq C \tag{231}$$

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

$$f_{avg,j}^{-1} = \begin{cases} 0 & \text{if } \mu > \frac{\lambda_j(0)}{\psi_j} \\ \lambda_j^{-1}\left(\mu\psi_j\right) & \text{if } \mu \leq \frac{\lambda_j(0)}{\psi_j} \end{cases}, \tag{232}$$

where μ is the solution to

$$\sum_{j \text{ s.t. } \mu \leq \frac{\lambda_j(0)}{\psi_j}} \lambda_j^{-1}\left(\mu\psi_j\right) = C. \tag{233}$$

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, $Var[d_j|p_j]$ for $j \notin \mathcal{R}$ weakly increases following the restriction. For $j \in \mathcal{R}$, $Var[d_j|p_j] = Var[d_j]$.*

Proof. First we have:

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

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

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

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

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

$$\begin{aligned}
\sum_{j \text{ s.t. } \lambda_j(0) \geq \mu_{res}\psi_j, j \notin R} \lambda_j^{-1}\left(\mu_{res}\alpha_j\right) &\leq \sum_{j \text{ s.t. } \lambda_j(0) \geq \mu_{res}\psi_j, j \notin R} \lambda_j^{-1}\left(\mu_{unr}\psi_j\right) \\
&\leq \sum_{j \text{ s.t. } \lambda_j(0) \geq \mu_{unr}\psi_j, j \notin R} \lambda_j^{-1}\left(\mu_{unr}\psi_j\right) \\
&< \sum_{j \text{ s.t. } \lambda_j(0) \geq \mu_{unr}\psi_j} \lambda_j^{-1}\left(\mu_{unr}\psi_j\right) \\
&= C.
\end{aligned} \tag{234}$$

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

So the restriction implies that $\mu_{res} \leq \mu_{unr}$, with equality if and only if no restricted frequencies where being learned about before the restriction. In turn, $\mu_{res} \leq \mu_{unr}$ 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

$$Var(D_t|P) = \frac{1}{T}\sum_{j=1}^{T} Var[d_j|p_j] \tag{235}$$

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 μ_{res} is the Lagrange multiplier in the constrained case. Note that since $f_{avg,j}^{-1}$ 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,

$$\left.\frac{d}{d\psi_j}E_{-1}\left[\tilde{q}_{LF,j}r_j\right]\right|_{\psi_j=\lambda_j(0)^-} > 0 \tag{236}$$

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, $\frac{1}{2}\log\left|\left(\Sigma_D^{-1}+\Sigma_i^{-1}\right)^{-1}\right|$. 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

$$\prod_{j=1}\left(f_{i,j}^{-1}+f_{D,j}^{-1}\right) \tag{237}$$

8.3.1 Information cost structure and equilibrium

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

$$\max_{\{f_{i,j}^{-1}\}}\sum_{j=1}^{T}\lambda_j\left(f_{avg,j}^{-1}\right)f_{i,j}^{-1}-\kappa\prod_{j=1}^{T}\left(f_{i,j}^{-1}+f_{D,j}^{-1}\right) \tag{238}$$

Assume that there exists j such that:

$$\lambda_j > \kappa\prod_{k\neq j}f_{D,k}^{-1}. \tag{239}$$

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

$$\lambda_j \leq \kappa\prod_{k\neq j}f_{D,k}^{-1} \quad \forall j. \tag{240}$$

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

$$\lambda_j - \kappa \prod_{k \neq j} f_{D,k}^{-1} \tag{241}$$

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

$$\xi \equiv \kappa \prod_{k} f_{D,k}^{-1} \tag{242}$$

The equilibrium information choices are then given by:

$$f_{avg,j}^{-1} = \begin{cases} 0 & \text{if } \frac{\lambda_j(0)}{f_{D,j}} < \xi \\ \lambda_j^{-1}\left(f_{D,j}\xi\right) & \text{if } \frac{\lambda_j(0)}{f_{D,j}} \geq \xi \end{cases}. \tag{243}$$

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 ψ_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

$$\max_{\{f_{i,j}^{-1}\}} \sum_{j=1}^{T} \lambda_j\left(f_{avg,j}^{-1}\right) f_{i,j}^{-1} - \kappa \prod_{j=1}^{T}\left(\psi_j f_{i,j}^{-1} + f_{D,j}^{-1}\right) \tag{244}$$

Then the equilibrium information allocation is

$$f_{avg,j}^{-1} = \begin{cases} 0 & \text{if } \xi > \frac{\lambda_j(0)}{f_{D,j}\psi_j} \\ \lambda_j^{-1}\left(f_{D,j}\xi\psi_j\right) & \text{if } \xi \leq \frac{\lambda_j(0)}{f_{D,j}\psi_j} \end{cases}, \tag{245}$$

At that point, the analysis from the baseline version applies. The only changes are that $\min\left(\psi, \lambda_j(0)\right)$ must be replaced by $\min\left(f_{D,j}\xi, \lambda_j(0)\right)$, 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 $N\left(z, \sigma_z^2\right)$, with $z \sim N\left(0, f_Z\right)$. 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 $\varepsilon_{x,h} \sim N\left(0, \sigma_x^2\right)$. 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

$$p = a_1 d - a_2 z \tag{246}$$

The hedgers' optimization is then over both their investment in the private opportunity, k_h, and their investment in the fundamental, q_h.

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

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

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 $\sigma_z \to \infty$, 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

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

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

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

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

$$= \max_{k_h,q_h} k_h z_h + k_h p + E\left[(k_h + q_h)(d-p) \mid p\right] - \frac{\rho_H^{-1}}{2} (k_h + q_h)^2 \operatorname{var}\left[(d-p) \mid p\right] - \frac{\rho_H^{-1}}{2} k_h^2 \sigma_x^2 \tag{253}$$

The two first-order conditions (FOCs) are

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

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

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 σ_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

$$q_h = \frac{\rho_H E[d - p \mid p, z_h]}{\operatorname{var}(d - p \mid p, z_h)} - \frac{\rho_H (z_h + p)}{\sigma_x^2} \tag{257}$$

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

$$= \rho_H \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) p - \rho_H \sigma_x^{-2} z_h \tag{259}$$

where τ_H is the precision of the hedgers' beliefs,

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

9.3 Expected utility

From above, expected utility conditional on prices and z_h is

$$k_h z_h + k_h p + E[(k_h + q_h)(d - p) \mid p] - \frac{\rho_H^{-1}}{2} (k_h + q_h)^2 \operatorname{var}[(d - p) \mid p, z_h] - \frac{\rho_H^{-1}}{2} k_h^2 \sigma_x^2 \tag{261}$$

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

$$\frac{1}{2}\left(k_h z_h + E\left[k_h d + q_h (d-p) \mid p\right]\right) = \frac{1}{2}\left(k_h (z_h + p) + (k_h + q_h) E\left[d - p \mid p\right]\right) \tag{262}$$

$$= \frac{1}{2}\left(\frac{\rho_H E\left[d-p \mid p\right]^2}{\operatorname{var}(d-p \mid p)} + \frac{\rho_H (z_h + p)^2}{\sigma_x^2}\right) \tag{263}$$

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

$$\operatorname{var}[d-p] = \operatorname{var}\left[E\left[d-p \mid p\right]\right] + E\left[\operatorname{var}\left[d-p \mid p\right]\right] \tag{264}$$

$$= E\left[E\left[d-p \mid p\right]^2\right] + \operatorname{var}(d-p \mid p) \tag{265}$$

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

$$E\left[(z_h + p)^2\right] = \operatorname{var}(z_h + p) \tag{266}$$

$$= \operatorname{var}((z_h - z) + z + a_1 d - a_2 z) \tag{267}$$

$$= \sigma_z^2 + (1-a_2)^2 f_Z + a_1^2 f_D \tag{268}$$

Substituting back into the equation for expected utility,

$$E\left[\frac{1}{2}\left(\frac{\rho_H E\left[d-p \mid p\right]^2}{\operatorname{var}(d-p \mid p)} + \frac{\rho_H (z_h+p)^2}{\sigma_x^2}\right)\right] = \frac{1}{2}\rho_H\left(\frac{\operatorname{var}[d-p] - \operatorname{var}(d-p \mid p)}{\operatorname{var}(d-p \mid p)} + \frac{\sigma_z^2 + (1-a_2)^2 f_Z + a_1^2 f_D}{\sigma_x^2}\right) \tag{269}$$

$$= \frac{1}{2}\rho_H \frac{(1-a_1)^2 f_D + a_2^2 f_Z - \tau_H^{-1}}{\tau_H^{-1}} + \frac{1}{2}\rho_H \frac{\sigma_z^2 + (1-a_2)^2 f_Z + a_1^2 f_D}{\sigma_x^2} \tag{270}$$

$$= \frac{1}{2}\left\{\begin{array}{c}\left(\tau_H (1-a_1)^2 + \sigma_x^{-2} a_1^2\right) f_D + \\ \left(\tau_H a_2^2 + \sigma_x^{-2}(1-a_2)^2\right) f_Z - 1 + \rho_H \sigma_x^{-2}\sigma_z^2\end{array}\right\} \tag{271}$$

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 ρ_H and ρ (where the latter remains the risk tolerance of the sophisticated investors from the main analysis). The equilibrium condition is

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

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

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

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

Equating the right hand sides of those two equations yields

$$\frac{\rho f_{avg}^{-1}}{\rho_H \sigma_x^{-2}} = \frac{a_1}{a_2} \tag{275}$$

Inserting that formula into the second equation yields

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

$$a_1 = \frac{(\rho_H + \rho) \left(\frac{\rho f_{avg}^{-1}}{\rho_H \sigma_x^{-2}} \right)^2 f_Z^{-1} + \rho f_{avg}^{-1}}{\rho_H \left(\tau_H + \sigma_x^{-2} \right) + \rho \tau_{avg}} \tag{277}$$

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

$$0 = \rho_H \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) (a_1 d - a_2 z) - \rho_H z_i \sigma_x^{-2} \tag{278}$$

Again matching coefficients,

$$0 = \rho_H \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) a_1 \tag{279}$$

$$0 = -\rho_H \left(\frac{a_1}{a_2^2} f_Z^{-1} - \tau_H - \sigma_x^{-2} \right) a_2 - \rho_H \sigma_x^{-2} \tag{280}$$

This immediately implies $a_1 = 0$, and hence

$$a_2 = \frac{\sigma_x^{-2}}{f_D^{-1} + \sigma_x^{-2}} \tag{281}$$

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

$$\frac{1}{2} \left\{ \frac{f_D^{-1} \sigma_x^{-2}}{f_D^{-1} + \sigma_x^{-2}} f_Z + \rho_H \sigma_x^{-2} \sigma_z^2 \right\} \tag{282}$$

More generally, expected utility is

$$EU_H \equiv \frac{1}{2} \left\{ \begin{array}{c} \left(\tau_H (1 - a_1)^2 + \sigma_x^{-2} a_1^2 \right) f_D + \\ \left(\tau_H a_2^2 + \sigma_x^{-2} (1 - a_2)^2 \right) f_Z - 1 + \rho \sigma_x^{-2} \sigma_z^2 \end{array} \right\} \tag{283}$$

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

$$a_1 = \frac{(\rho_H + \rho)\left(\frac{\rho f_{avg}^{-1}}{\rho_H f_x^{-1}}\right)^2 f_Z^{-1} + \rho f_{avg}^{-1}}{\rho_H\left(\tau_O + \sigma_x^{-2}\right) + \rho\tau_{avg}} \tag{284}$$

$$a_2 = a_1 \frac{\rho_H \sigma_x^{-2}}{\rho f_{avg}^{-1}} \tag{285}$$

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

$$p_j = -\frac{\sigma_x^{-2}}{f_D^{-1} + \sigma_x^{-2}} z_j \tag{286}$$

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

$$f_D + \left(\frac{\sigma_x^{-2}}{f_D^{-1} + \sigma_x^{-2}}\right)^2 f_Z \tag{287}$$

When speculators are present but uninformed, the pricing function is

$$\begin{aligned} p &= \frac{f_D \rho_H}{f_D \rho_H + (\rho + \rho_H)\,\sigma_x^2} z & (288) \\ \operatorname{var}(r) &= f_D + \left(\frac{f_D \rho_H}{f_D \rho_H + (\rho + \rho_H)\,\sigma_x^2}\right)^2 f_Z & (289) \end{aligned}$$

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

$$\begin{aligned} \frac{d}{d\psi_j} E_{-1}\left[\tilde{q}_{LF,j} r_j\right]\Big|_{\psi_j = \lambda_j(0)^-} &> 0 & (290) \\ \frac{d}{d\psi_j} E_{-1}\left[\sum_t \tilde{Q}_{LF,t}\left(D_t - P_t\right)\right]\Bigg|_{\psi_j = \lambda_j(0)^-} &> 0 & (291) \\ \frac{d}{d\psi_j} E_{-1}\left[U_{LF,0}\right]\Big|_{\psi_j = \lambda_j(0)^-} &> 0 & (292) \end{aligned}$$

The problem faced by the sophisticated investors is the same in the sense that they continue to acquire information to the point that $\lambda_j\left(f_{avg,j}^{-1}\right) = \psi_j$, unless $\lambda_j(0) \leq \psi_j$, in which case they acquire no information. A marginal decline in ψ_j then leads to a marginal increase in $f_{avg,j}^{-1}$.

To obtain the derivative of $E_{-1}\left[\tilde{q}_{LF,j} r_j\right]$, 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 $f_{avg,j}^{-1}$ yields the desired result.

9.7.10 Corollary 2.2

From above, we have

$$\mathrm{var}\left[d_j \mid p_j\right] = \left(\frac{\rho f_{avg,j}^{-1}}{\rho_H \sigma_x^{-2}} f_{Z,j}^{-1} + f_{D,j}^{-1}\right)^{-1}, \tag{293}$$

which is obviously decreasing in f_{avg}^{-1}. It is also possible to confirm that return volatility is decreasing in $f_{avg,j}^{-1}$.

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

$$\text{when } f_{avg}^{-1} = 0, \ \frac{d}{d\rho} EU_H > 0 \tag{294}$$

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

$$\lim_{f_{avg}^{-1}} a_1 = 1 \tag{295}$$

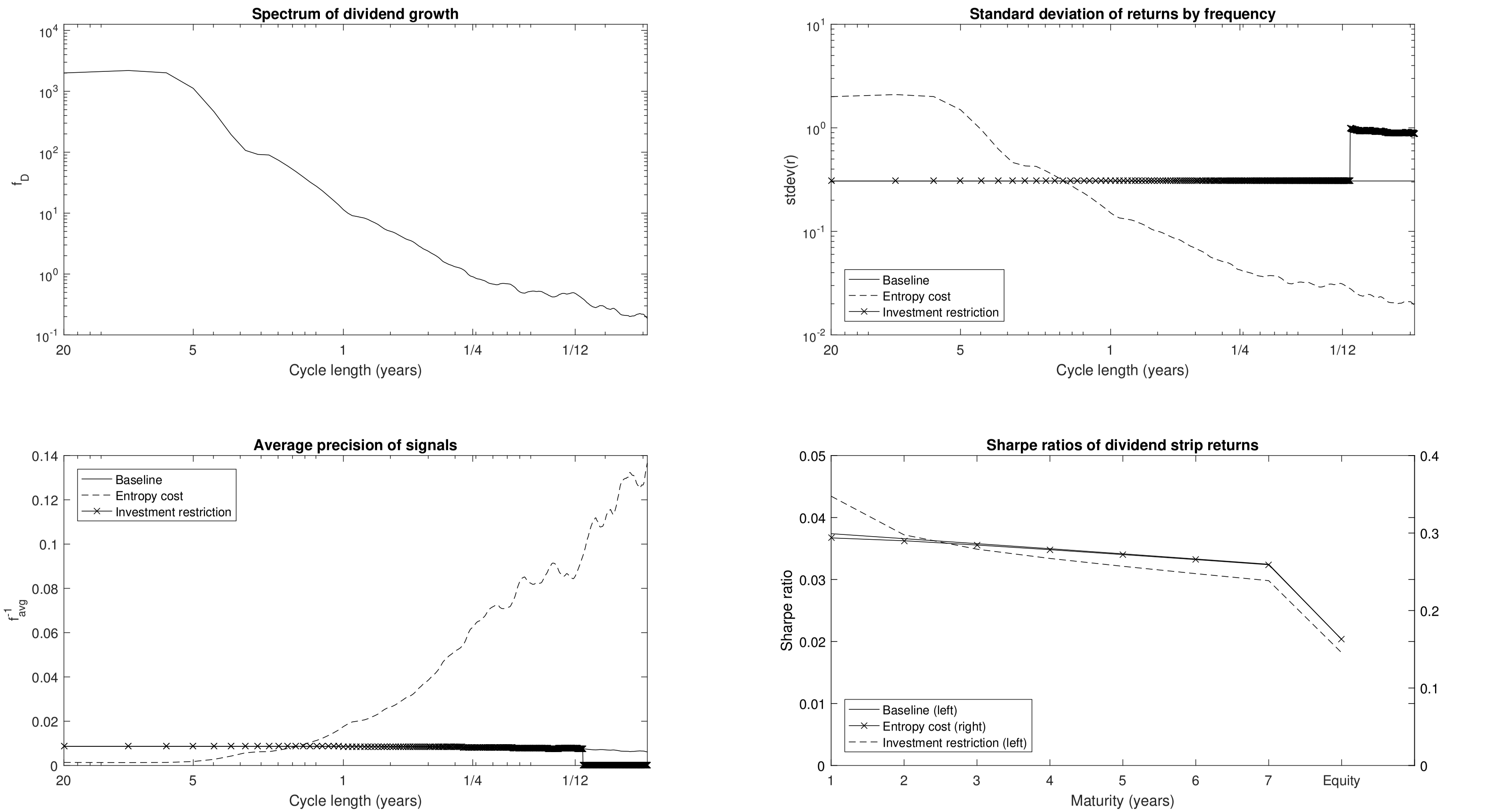
$$\lim_{f_{avg}^{-1}} a_2 = 0 \tag{296}$$

Furthermore, straightforward algebra shows that, writing the expected utility of the hedgers as a function of f_{avg}^{-1}, we have

$$\lim_{f_{avg}^{-1} \to \infty} EU_H \left(f_{avg}^{-1} \right) > EU_H \left(0 \right) \tag{297}$$

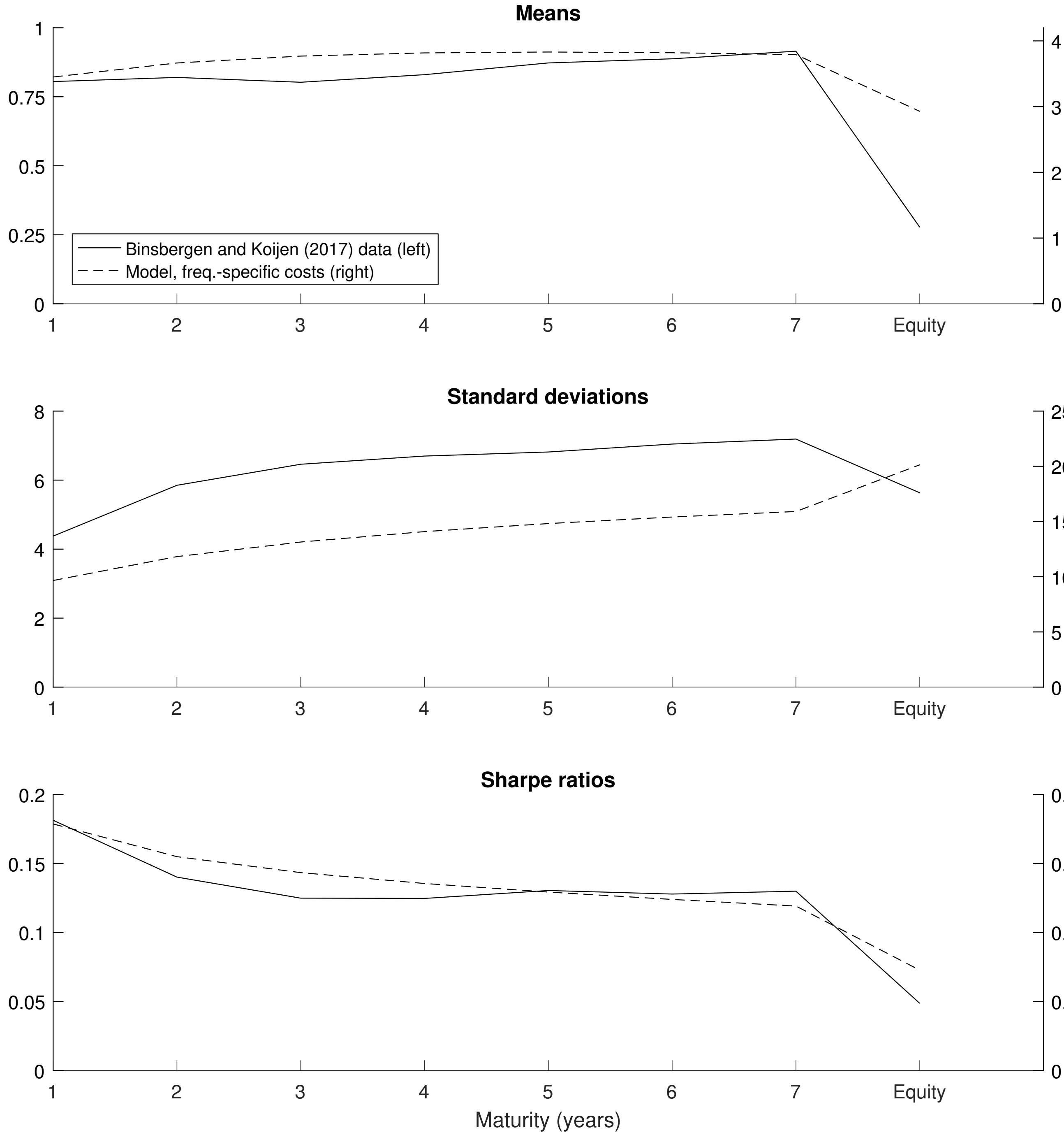
which shows that the hedgers are better off in a fully informative equilibrium than in the alternative uninformative case.

Figure A.1: Numerical example with entropy cost



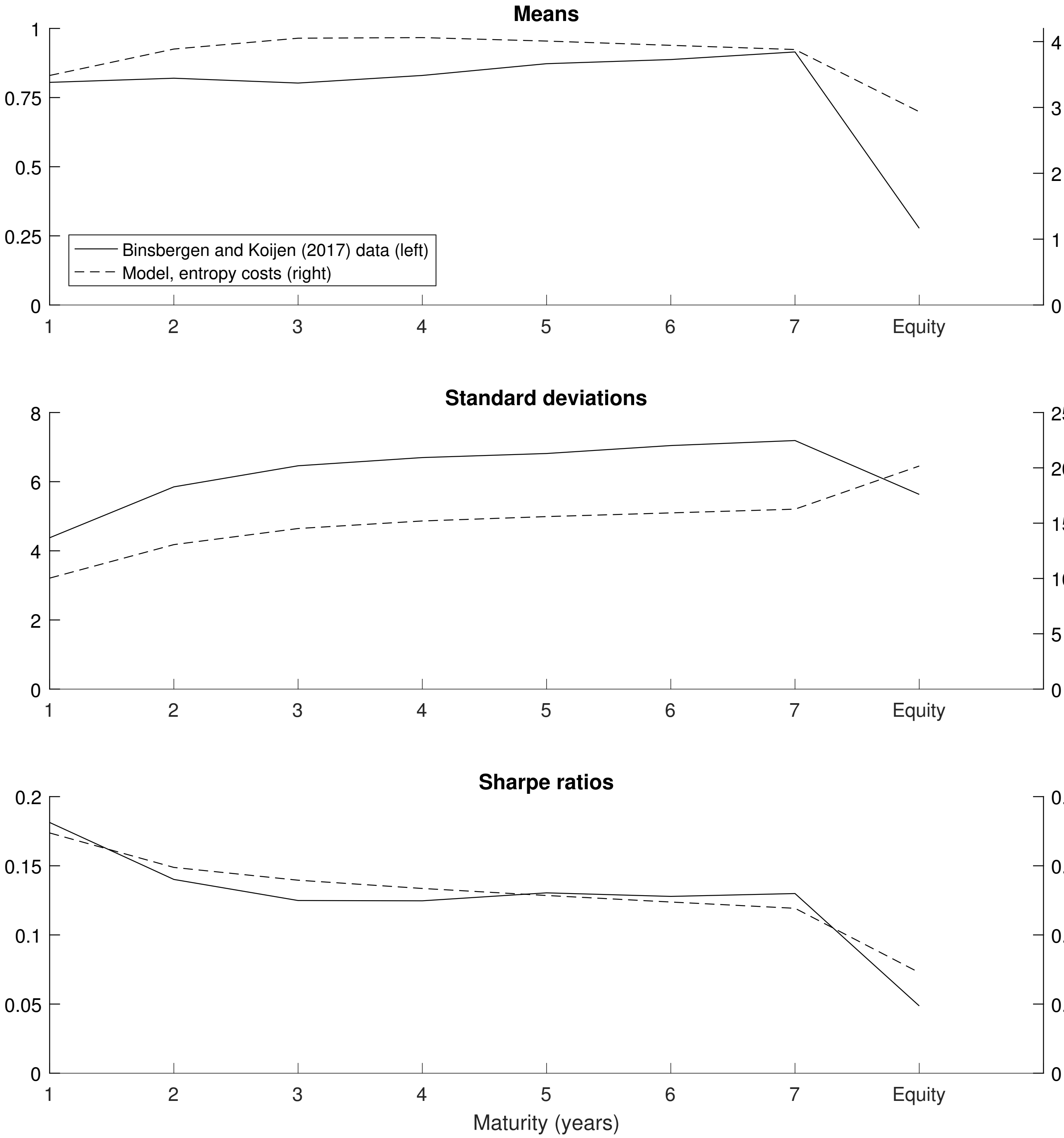
Notes: See figure 2. The only change is that the frequency-specific information cost case is replaced by the case where information flow is measured by entropy. That setup is analyzed in appendix 8.3.

Figure A.2: Details of dividend strips returns, frequency-specific information cost case



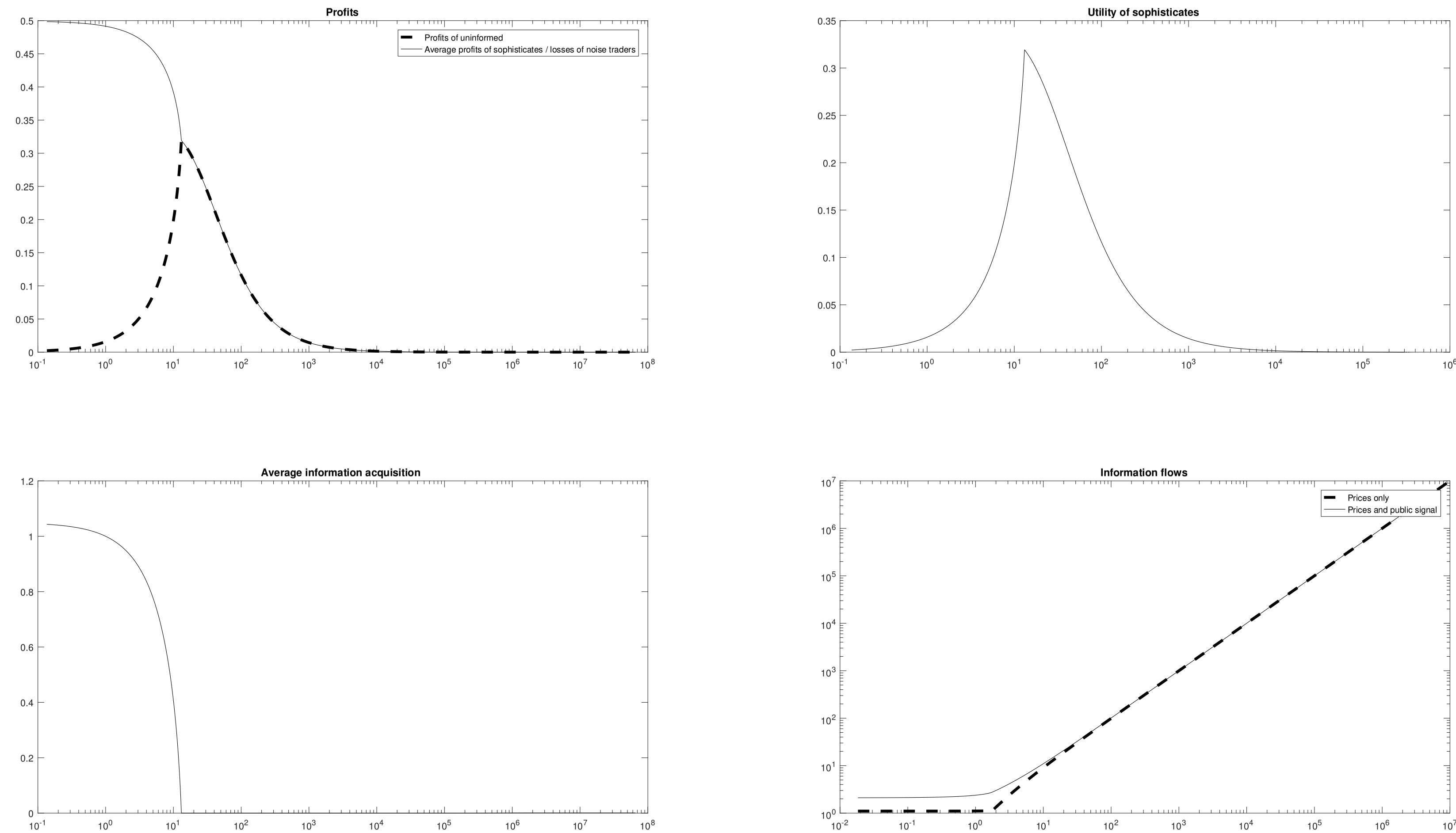
Notes: The three panels report per-period characteristics of dividend strip returns in the model with the frequency-specific cost specification along with empirical moments reported by Binsbergen and Koijen (2017) (averaged across the four markets they examine). All values are annualized.

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 $(psi_j f_i^{-1})$. Information flows in the bottom-right panel are the precision obtained from conditioning either on prices alone or prices and the public signal.