

APPENDIX TO:
International Equity Flows and Returns:
A Quantitative Equilibrium Approach

Rui Albuquerque
Boston University School of Management and CEPR

Gregory H. Bauer
Bank of Canada

Martin Schneider
New York University and Federal Reserve Bank of Minneapolis

May 12, 2006

Appendix

A Proof of Theorem 1

The proof of the theorem proceeds in four steps. First, we show that, if prices take the form in the Theorem, then the vector ϕ_t^i captures investor i 's payoff-relevant information. In other words, $\phi_t^U = \hat{\mathbf{F}}_t^U$ is a sufficient statistic for forecasting stock returns given U-investors' information set \mathcal{I}_t^U , whereas $\phi_t^S = (\hat{\mathbf{F}}_t^{S'}, \hat{\mathbf{F}}_t^{U'})'$ is a sufficient statistic for forecasting all future stock *and* private returns given S-investors' information set \mathcal{I}_t^S . Second, we use the Kalman filter to obtain difference equations for the evolution of the individual state vectors ϕ_t^i and verify that these difference equations have a stationary solution. Third, we solve for optimal portfolio allocations as a function of ϕ_t^i and current price. Finally, we derive market clearing prices and verify that they indeed take the form given in the Theorem. In the numerical exercises below, finding an equilibrium then reduces to solving a nonlinear system of equations in the price coefficients $(\bar{\pi}, \boldsymbol{\pi}_S, \boldsymbol{\pi}_U)$.

In equilibrium the local equity asset price depends on the beliefs of S-investors and U-investors of the underlying factors:

$$P_t = \bar{\pi} + \boldsymbol{\pi}'_S \hat{\mathbf{F}}_t^S + \boldsymbol{\pi}'_U \hat{\mathbf{F}}_t^U. \quad (1)$$

Consider first the agents' payoff-relevant information. Suppose the information sets \mathcal{I}_t^i contain only normal random variables. This implies normality of the conditional expectations $\hat{\mathbf{F}}_t^i$ and, if the price satisfies (1), also of all per-share returns. It follows that $\phi_t^S = (\hat{\mathbf{F}}_t^{S'}, \hat{\mathbf{F}}_t^{U'})'$ is a sufficient statistic for forecasting all future returns, given the information set \mathcal{I}_t^S . Now, $\phi_t^U = \hat{\mathbf{F}}_t^U$ is a sufficient statistic for forecasting returns given the information set \mathcal{I}_t^U . This includes one-step-ahead returns, since the current price can be written as a function of $\hat{\mathbf{F}}_t^U$. Indeed, U-investors know $\hat{\mathbf{F}}_t^U$, so that observing the price is the same as observing the signal

$$P_t - \bar{\pi} - \boldsymbol{\pi}'_U \hat{\mathbf{F}}_t^U = \boldsymbol{\pi}'_S \hat{\mathbf{F}}_t^S.$$

But then $\boldsymbol{\pi}'_S \hat{\mathbf{F}}_t^U = E[\boldsymbol{\pi}'_S \hat{\mathbf{F}}_t^S | \mathcal{I}_t^U] = \boldsymbol{\pi}'_S \hat{\mathbf{F}}_t^S$, and we can write the price as $P_t = \bar{\pi} + (\boldsymbol{\pi}'_S + \boldsymbol{\pi}'_U) \hat{\mathbf{F}}_t^U$. It follows that the state vector ϕ_t^U captures the payoff-relevant information of U-investors' consumption-savings and portfolio choice problem.

Since all state variables are normal and homoskedastic, the evolution of investors' beliefs can be described by tracking conditional expectations, using the Kalman filter. S-investors learn about the state of the business cycle by observing dividends

and returns on their private opportunities. They do not learn from the price, since they already know $\hat{\mathbf{F}}_t^S$ and hence $\hat{\mathbf{F}}_t^U$. We collect their relevant observables in a vector $\mathbf{y}_t^S = (D_t - \bar{D}, R_t^B - \bar{R}^B - \eta_B F_{t-1}^B)'$ that can be represented as

$$\mathbf{y}_t^S = \mathbf{M}^{ySF} \mathbf{F}_{t-1} + \mathbf{M}^{yS\varepsilon} \varepsilon_t. \quad (2)$$

With the dynamics of the vector \mathbf{F}_t and equation (2) we obtain a state-space system. S-investors' conditional expectation of the state vector, $\hat{\mathbf{F}}_t^S$, then takes the form

$$\begin{aligned} \hat{\mathbf{F}}_t^S &= \boldsymbol{\rho} \hat{\mathbf{F}}_{t-1}^S + \mathbf{K}^S \left(\mathbf{y}_t^S - \mathbf{M}^{ySF} \hat{\mathbf{F}}_{t-1}^S \right) \\ &= \boldsymbol{\rho} \hat{\mathbf{F}}_{t-1}^S + \hat{\boldsymbol{\varepsilon}}_t^S, \end{aligned} \quad (3)$$

where \mathbf{K}^S is a steady-state Kalman gain matrix. The matrix $\mathbf{M}^{yS\varepsilon}$ allows errors in the observation equation (2) to be correlated with errors in the state equation (3).

U-investors obtain valuable information from dividends as well as from the signal contained in prices, so that $\mathbf{y}_t^U = (D_t - \bar{D}, \boldsymbol{\pi}'_S \hat{\mathbf{F}}_t^S)$. These variables¹ can be represented using $\hat{\mathbf{F}}_t^S$:

$$\mathbf{y}_t^U = \mathbf{M}^{yUF} \hat{\mathbf{F}}_{t-1}^S + \mathbf{M}^{yU\varepsilon} \hat{\boldsymbol{\varepsilon}}_t^S. \quad (4)$$

Equations (3) and (4) form the state-space system of U-investors. Their conditional expectation, and hence their state variable ϕ_t^U , can be written as

$$\begin{aligned} \hat{\mathbf{F}}_t^U &= \boldsymbol{\rho} \hat{\mathbf{F}}_{t-1}^U + \mathbf{K}^U \left(\mathbf{y}_t^U - \mathbf{M}^{yUF} \hat{\mathbf{F}}_{t-1}^U \right) \\ &= (\boldsymbol{\rho} - \mathbf{K}^U \mathbf{M}^{yUF}) \hat{\mathbf{F}}_{t-1}^U + \mathbf{K}^U \mathbf{M}^{yUF} \hat{\mathbf{F}}_{t-1}^S + \mathbf{K}^U \mathbf{M}^{yU\varepsilon} \hat{\boldsymbol{\varepsilon}}_t^S. \end{aligned} \quad (5)$$

Note that all that U-investors are doing is trying to forecast the forecast of the S-investors. Letting $\hat{\mathbf{u}}_t^U = \mathbf{y}_t^U - \mathbf{M}^{yUF} \hat{\mathbf{F}}_{t-1}^U$, the Kalman filter on U-investors' problem yields

$$E_t [\hat{\mathbf{u}}_t^U \hat{\mathbf{u}}_t^U] = \mathbf{M}^{yUF} E_t \left[\left(\mathbf{F}_{t-1} - \hat{\mathbf{F}}_{t-1}^U \right) \left(\mathbf{F}_{t-1} - \hat{\mathbf{F}}_{t-1}^U \right)' \right] \mathbf{M}^{yUF'} + \mathbf{M}^{yU\varepsilon} E [\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'] \mathbf{M}^{yU\varepsilon'}. \quad (6)$$

The law of motion of $\phi_t^U = \hat{\mathbf{F}}_t^U$ is then

$$\phi_{t+1}^U = \boldsymbol{\Phi}^U \phi_t^U + \mathbf{M}^{\phi\varepsilon U} \boldsymbol{\varepsilon}_{t+1}^U.$$

Repeating the same process for S-investors' conditional forecasts $\hat{\mathbf{F}}_t^S$ we have $\phi_t^S = (\hat{\mathbf{F}}_t^S, \hat{\mathbf{F}}_t^{U'})'$:

$$\phi_{t+1}^S = \boldsymbol{\Phi}^S \phi_t^S + \mathbf{M}^{\phi\varepsilon S} \boldsymbol{\varepsilon}_{t+1}^S. \quad (7)$$

¹The matrices are $\mathbf{M}^{yUF} = \begin{pmatrix} \boldsymbol{\pi}'_S \boldsymbol{\rho} \\ \mathbf{M}_1^{ySF} \end{pmatrix}$, where \mathbf{M}_1^{ySF} is the first line of \mathbf{M}^{ySF} and $\mathbf{M}^{yU\varepsilon} = \begin{pmatrix} \boldsymbol{\pi}'_S \\ \mathbf{e}_1 \end{pmatrix}$, where \mathbf{e}_1 is the first unit vector.

Let us turn to the decision problem of both investors. Write returns as

$$\mathbf{R}_{t+1}^i = \bar{\mathbf{R}}^i + \mathbf{M}^{R\phi^i} \phi_t^i + \mathbf{M}^{R\varepsilon\phi^i} \varepsilon_{t+1}^{\phi^i}$$

for each investor. Guess that investors' i value function is of the form

$$V(w_t^i; \phi_t^i) = -\exp\left[-\kappa^i - \tilde{\gamma} w_t^i - \mathbf{u}'_i \phi_t^i - \frac{1}{2} \phi_t^{i'} \mathbf{U}_i \phi_t^i\right].$$

Define $\Omega^i = \left(\mathbf{M}^{\phi\varepsilon i'} \mathbf{U}_i \mathbf{M}^{\phi\varepsilon i} + (\Sigma_{\phi\phi}^i)^{-1}\right)^{-1}$, where $\Sigma_{\phi\phi}^i = E\left[\varepsilon_t^{\phi^i} \varepsilon_t^{\phi i'}\right]$. We have that (superscript i dropped for simplicity)

$$\begin{aligned} E_t V(w_{t+1}, \phi_{t+1}) &= -\frac{(\det \Sigma_{\phi\phi})^{-\frac{1}{2}}}{(\det \Omega)^{-\frac{1}{2}}} \exp\left(-\kappa - \tilde{\gamma} (R_f(w_t - c_t) + \psi'_t (\bar{\mathbf{R}} + \mathbf{M}^{R\phi} \phi_t))\right) \\ &\quad \times \exp\left(-\frac{1}{2} \phi_t' \Phi' U \Phi \phi_t - \mathbf{u}' \Phi \phi_t + \frac{1}{2} \tilde{\omega} \Omega \tilde{\omega}'\right), \end{aligned}$$

where $\tilde{\omega} = \tilde{\gamma} \psi'_t \mathbf{M}^{R\varepsilon\phi} + (\phi_t' \Phi' \mathbf{U} + \mathbf{u}') \mathbf{M}^{\phi\varepsilon}$. Solving for the optimal portfolio, we obtain

$$\begin{aligned} \psi_t &= \tilde{\gamma}^{-1} (\mathbf{M}^{\psi\psi})^{-1} (\bar{\mathbf{M}}^\psi + \mathbf{M}^{\psi\phi} \phi_t) \\ &= \bar{\psi} + \Psi \phi_t, \end{aligned}$$

where the matrices are given by $\bar{\mathbf{M}}^{\psi'} = \bar{\mathbf{R}}' - \mathbf{u}' \mathbf{M}^{\phi\varepsilon} \Omega \mathbf{M}^{R\varepsilon\phi'}$, $\mathbf{M}^{\psi\psi} = \mathbf{M}^{R\varepsilon\phi} \Omega \mathbf{M}^{R\varepsilon\phi'}$, and $\mathbf{M}^{\phi\psi'} = \mathbf{M}^{R\phi'} - \Phi' \mathbf{U} \mathbf{M}^{\phi\varepsilon} \Omega \mathbf{M}^{R\varepsilon\phi'}$. The first term in matrix Ψ gives the myopic demand of the investor (i.e., $\tilde{\gamma}^{-1} (\mathbf{M}^{\psi\psi})^{-1} \mathbf{M}^{R\phi}$), and the second term gives the hedging demand of the investor (i.e., $-\tilde{\gamma}^{-1} (\mathbf{M}^{\psi\psi})^{-1} \mathbf{M}^{R\varepsilon\phi} \Omega \mathbf{M}^{\phi\varepsilon'} \mathbf{U} \Phi$).

>From the value function $V(w_t^i; \phi_t^i)$ we see that risk-averse investors care not only about fluctuations in wealth but also about changes in beliefs, captured by the state vector ϕ_t^i . The quadratic term reflects investors' taste for "unusual" investment opportunities. Intuition for this effect can be obtained by thinking about the case of one state variable. \mathbf{U}_i is then a positive number and continuation utility is higher the further ϕ_t^i is from its mean of zero. Since ϕ_t^i is payoff-relevant, it drives expected returns at some time in the future. An unusual value signals that above-average expected returns will be available by going either long or short.

We can now describe in detail the coefficients of the optimal portfolio policy $\psi_t^i = \bar{\psi}^i + \Psi^i \phi_t^i$. We have

$$\begin{aligned} \psi_t^i &= \tilde{\gamma}^{-1} \tilde{\Sigma}_{R^i R^i}^{-1} E^i(\mathbf{R}_{t+1}^i | \phi_t^i) \\ &\quad - \tilde{\gamma}^{-1} \tilde{\Sigma}_{R^i R^i}^{-1} Cov\left(\left(\mathbf{u}'_i + E^i[\phi_{t+1}^i | \phi_t^i]' \mathbf{U}_i\right) \phi_{t+1}^i, \mathbf{R}_{t+1}^i | \phi_t^i\right), \\ &= \tilde{\gamma}^{-1} \tilde{\Sigma}_{R^i R^i}^{-1} (E^i(\mathbf{R}_{t+1}^i | \phi_t^i) + (\bar{\mathbf{h}}^i + H^s \phi_t^i)) \end{aligned} \tag{8}$$

where the matrix $\tilde{\Sigma}_{R^i R^i} = \mathbf{M}^{R\phi^i} \left(\text{Var}(\phi_{t+1}^i | \phi_t^i)^{-1} + \mathbf{U}_i \right)^{-1} \mathbf{M}^{R\phi^i}$ is a transformation of the conditional covariance matrix of returns, and $\mathbf{M}^{R\phi^i}$ is such that $\mathbf{R}_t^i = \mathbf{M}^{R\phi^i} \phi_t^i$. We use this decomposition in the main text.

Solving for the optimal consumption level and the value function, we get that for given price function, optimality requires that the following constraints are met: $\tilde{\gamma} = \gamma \frac{R_f - 1}{R_f}$ and

$$\begin{aligned} \kappa &= -\log\left(\frac{R_f}{R_f - 1}\right) \\ &\quad - \frac{1}{R_f - 1} \frac{1}{2} \left(\log\left(\frac{\det \mathbf{\Omega}}{\det \mathbf{\Sigma}_{\phi\phi}}\right) + \mathbf{u}' \mathbf{M}^{\phi\varepsilon} \mathbf{\Omega} \mathbf{M}^{\phi\varepsilon'} \mathbf{u} - \bar{\mathbf{M}}^{\psi'} (\mathbf{M}^{\psi\psi})^{-1} \bar{\mathbf{M}}^{\psi} \right) \\ \mathbf{u} &= \frac{1}{R_f} \left(\bar{\mathbf{M}}^{\phi} + \mathbf{M}^{\psi\phi'} (\mathbf{M}^{\psi\psi})^{-1} \bar{\mathbf{M}}^{\psi} \right) \\ \mathbf{U} &= \frac{1}{R_f} \left(\mathbf{M}^{\phi\phi} + \mathbf{M}^{\psi\phi'} (\mathbf{M}^{\psi\psi})^{-1} \mathbf{M}^{\psi\phi} \right), \end{aligned}$$

with $\bar{\mathbf{M}}^{\phi'} = \mathbf{u}' (\mathbf{I} - \mathbf{M}^{\phi\varepsilon} \mathbf{\Omega} \mathbf{M}^{\phi\varepsilon'} \mathbf{U}) \mathbf{\Phi}$ and $\mathbf{M}^{\phi\phi} = \mathbf{\Phi}' \mathbf{U} (\mathbf{I} - \mathbf{M}^{\phi\varepsilon} \mathbf{\Omega} \mathbf{M}^{\phi\varepsilon'} \mathbf{U}) \mathbf{\Phi}$.

Finally, to solve for an equilibrium, let $\Theta_{\hat{F}^S}^S$ be the part of the first row of Ψ^S that is associated with $\hat{\mathbf{F}}_t^S$, let Θ^U be the first row of Ψ^U and let $\bar{\theta}^i$ be the mean local asset demand by investor i . The equilibrium requires

$$\begin{aligned} \Delta \bar{\theta}^U + (1 - \Delta) \bar{\theta}^S &= 1, \\ \Delta \Theta^U + (1 - \Delta) \Theta_{\hat{F}^S}^S &= \mathbf{0}. \end{aligned}$$

This is a system of nonlinear equations that can be solved for the price coefficients.

B Detrending

Data on dividends and flows exhibit trends, whereas our quantitative exercise explores a detrended economy. We now outline a consistent approach to detrending dividends and flows. To fix ideas, consider the following stylized view of the stock market. There are N firms, each with a single share, paying the same (per-share) dividend \tilde{D}_t and having the same (per-share) price \tilde{P}_t . Dividends grow at an exponential rate η . The parameter η thus captures trend firm productivity growth, which benefits owners through dividends.

An observed aggregate price index records the change in the value of the average firm, $\tilde{P}_t / \tilde{P}_{t-1}$. This change in valuation has two components: capital gains that arise from fluctuations in the firm's *stationary* price P_t / P_{t-1} and the growth in prices built in from productivity growth:

$$\tilde{P}_t / \tilde{P}_{t-1} = e^\eta P_t / P_{t-1}.$$

The observed dividend yield is $y_t = \tilde{D}_t/\tilde{P}_t = D_t/P_t$. A natural way to remove the trend from dividends is to exponentially detrend the measure $y_t\tilde{P}_t$. The observed holdings of the domestic equity index by investor i are $\tilde{P}_t\tilde{\theta}_t^i$. The observed market capitalization at the end of period t is the combined value of all plants, $\tilde{M}_t = \tilde{P}_tN$. The normalization of holdings by beginning-of-period market capitalization is thus a natural way to remove the exponential trend in holdings. The normalized holdings are

$$\theta_t^i = \frac{\tilde{\theta}_t^i\tilde{P}_t}{\tilde{M}_t} = \frac{\tilde{\theta}_t^i}{N}.$$

There is an explicit connection between dividends and equilibrium holdings before and after detrending. We can summarize an economy driven by trending exogenous variables by a tuple $\mathcal{E} = \left(\tilde{R}_f, N, \left(\tilde{D}_t, \tilde{R}_t^B\right)_{t=0}^{\infty}\right)$. Suppose that $\tilde{D}_t = e^{\eta t}D_t$ and that $\left(\tilde{P}_t, \tilde{\theta}_t, \tilde{\psi}_t^B, \tilde{c}_t\right)$ is an equilibrium of \mathcal{E} , where we suppress the indices for the different types of agents. It can be verified that the tuple

$$\left(P_t, \theta_t, e^{-\eta t}\tilde{\psi}_t^B, e^{-\eta t}\tilde{c}_t\right)$$

is an equilibrium of the detrended economy

$$\mathcal{E}_\eta = \left(\tilde{R}_f e^{-\eta}, 1, \left(D_t, e^{-\eta}\tilde{R}_t^B\right)_{t=0}^{\infty}\right).$$

In our quantitative exercise, we consider a detrended economy. We determine a stationary dividend process D_t as the residuals in a regression of average firm dividends on a time trend,

$$\log\left(y_t\tilde{P}_t\right) = E[\log D_t] + \eta t + (\log D_t - E[\log D_t]). \quad (9)$$

We then match the equilibrium flows to observed flows normalized by market capitalization. In the light of the above result, this ensures consistent detrending of dividends and flows.

We also need to select an interest rate R_f for the detrended economy. Here we use the observed average interest rate. In terms of the above notation, we are thus analyzing the economy \mathcal{E}_0 . Given our data, this is preferable to considering the economy $\mathcal{E}_{\hat{\eta}}$, where $\hat{\eta}$ is the growth rate estimate from (9). The reason is that, in a small sample such as ours, $\hat{\eta}$ is driven by medium-term developments and does not reflect the long-run average growth rate. In particular, in our sample $\hat{\eta}$ exceeds the average real riskless interest rate. We are thus not likely to learn much by considering equilibrium flows from $\mathcal{E}_{\hat{\eta}}$. At the same time, the result of the previous paragraph shows that the only role of the trend growth rate η is to shift *all* returns. This suggests that the behavior of the correlations we are interested in will be similar across all economies \mathcal{E}_η for η reasonably small.

C The Dividend Process

In this appendix we discuss the estimation of the dividend process. We derive conditions under which a general ARMA(2,2) process permits a representation of the type we assume for our dividend process:

$$\begin{aligned} F_t^D &= a_1 F_{t-1}^D + a_2 F_{t-2}^D + \varepsilon_t^{FD}, \\ D_t &= \bar{D} + F_{t-1}^D + \varepsilon_t^D, \end{aligned} \quad (10)$$

where ε_t^{FD} and ε_t^D are serially uncorrelated and independent random variables with zero mean and variances $\sigma_{\varepsilon^{FD}}^2$ and $\sigma_{\varepsilon^D}^2$, respectively. To prove our result we need to compare the correlogram of dividends under the two representations. Consider first the representation (10). The correlogram of the persistent component F_t^D is summarized by

$$\begin{aligned} \sigma^2(F_t^D) &= \left(1 - a_1^2 - a_2^2 - \frac{2a_2a_1^2}{1 - a_2}\right)^{-1} \sigma_{\varepsilon^{FD}}^2, \\ \sigma(F_t^D, F_{t-1}^D) &= \frac{a_1}{1 - a_2} \sigma^2(F_t^D), \\ \sigma(F_t^D, F_{t-2}^D) &= a_1 \sigma(F_t^D, F_{t-1}^D) + a_2 \sigma^2(F_t^D) \\ &= \left(\frac{a_1^2}{1 - a_2} + a_2\right) \sigma^2(F_t^D), \\ \sigma(F_t^D, F_{t-s}^D) &= a_1 \sigma(F_t^D, F_{t-s+1}^D) + a_2 \sigma(F_t^D, F_{t-s+2}^D); \quad s \geq 3. \end{aligned}$$

The correlogram of the dividend process is thus given by

$$\begin{aligned} \sigma^2(D_t - \bar{D}) &= \sigma^2(F_t^D) + \sigma_{\varepsilon^D}^2, \\ \sigma(D_t - \bar{D}, D_{t-1} - \bar{D}) &= \sigma(F_t^D, F_{t-1}^D) \\ &= \frac{a_1}{1 - a_2} [\sigma^2(D_t - \bar{D}) - \sigma_{\varepsilon^D}^2], \\ \sigma(D_t - \bar{D}, D_{t-2} - \bar{D}) &= \sigma(F_t^D, F_{t-2}^D) \\ &= a_1 \sigma(D_t - \bar{D}, D_{t-1} - \bar{D}) + a_2 [\sigma^2(D_t - \bar{D}) - \sigma_{\varepsilon^D}^2], \end{aligned}$$

as well as, for every $s \geq 3$,

$$\begin{aligned} \sigma(D_t - \bar{D}, D_{t-s} - \bar{D}) &= \sigma(F_t^D, F_{t-s}^D) \\ &= a_1 \sigma(D_t - \bar{D}, D_{t-s+1} - \bar{D}) + a_2 \sigma(D_t - \bar{D}, D_{t-s+2} - \bar{D}). \end{aligned}$$

Now consider a general ARMA(2,2) process

$$D_t - \bar{D} = a_1 (D_{t-1} - \bar{D}) + a_2 (D_{t-2} - \bar{D}) + u_t + \lambda_1 u_{t-1} + \lambda_2 u_{t-2}, \quad (11)$$

where u_t is serially uncorrelated with mean zero and variance σ_u^2 . Squaring both sides and taking expectations, we have

$$\begin{aligned}\sigma^2 (D_t - \bar{D}) &= a_1^2 \sigma^2 (D_{t-1} - \bar{D}) + a_2^2 \sigma^2 (D_{t-2} - \bar{D}) + \sigma_u^2 (1 + \lambda_1^2 + \lambda_2^2) \\ &\quad + 2a_1 \lambda_1 \sigma (D_{t-1} - \bar{D}, u_{t-1}) + 2a_1 \lambda_2 \sigma (D_{t-1} - \bar{D}, u_{t-2}) \\ &\quad + 2a_2 \lambda_2 \sigma (D_{t-2} - \bar{D}, u_{t-2}) + 2a_1 a_2 \sigma (D_t - \bar{D}, D_{t-1} - \bar{D}).\end{aligned}\quad (12)$$

In addition, multiplying both sides of (11) by $(D_{t-1} - \bar{D})$ and taking expectations, we have

$$\sigma (D_t - \bar{D}, D_{t-1} - \bar{D}) = \frac{a_1}{1 - a_2} \sigma^2 (D_t - \bar{D}) + \frac{\lambda_1 + \lambda_2 \lambda_1 + \lambda_2 a_1}{1 - a_2} \sigma_u^2. \quad (13)$$

Finally, multiplying both sides of (11) by $(D_{t-2} - \bar{D})$ and taking expectations, we obtain

$$\sigma (D_t - \bar{D}, D_{t-2} - \bar{D}) = a_1 \sigma (D_t - \bar{D}, D_{t-1} - \bar{D}) + a_2 \sigma^2 (D_t - \bar{D}) + \lambda_2 \sigma_u^2. \quad (14)$$

Using (12)-(14) above, the variance can be solved out in terms of parameters only:

$$\begin{aligned}\sigma^2 (D_t - \bar{D}) &= \sigma_u^2 \left(1 - a_1^2 - a_2^2 - \frac{2a_1^2 a_2}{1 - a_2} \right)^{-1} \left(1 + \lambda_1^2 + \lambda_2^2 + 2a_1 \lambda_1 \right. \\ &\quad \left. + 2a_1^2 \lambda_2 + 2a_1 \lambda_2 \lambda_1 + 2a_2 \lambda_2 + 2a_1 a_2 \frac{\lambda_1 + \lambda_2 \lambda_1 + \lambda_2 a_1}{1 - a_2} \right).\end{aligned}$$

The first and second covariances are then given by (13) and (14) and all further covariances (for $s \geq 3$) follow the recursion

$$\sigma (D_t - \bar{D}, D_{t-s} - \bar{D}) = a_1 \sigma (D_t - \bar{D}, D_{t-s+1} - \bar{D}) + a_2 \sigma (D_t - \bar{D}, D_{t-s+2} - \bar{D}).$$

It is clear that if a given ARMA(2,2) process is to have the representation (10), the autoregressive coefficients must be the same in both representations. Moreover, since the recursions for all covariances beyond lag 2 are identical, a representation of the type (10) exists if there exist $\sigma_{\varepsilon_{FD}}^2, \sigma_{\varepsilon_D}^2 > 0$ such that the variance and the first two covariances are matched, which require that

$$\begin{aligned}&\sigma_u^2 \left(1 + \lambda_1^2 + \lambda_2^2 + 2a_1 \lambda_1 + 2a_1^2 \lambda_2 + 2a_1 \lambda_2 \lambda_1 + 2a_2 \lambda_2 + 2a_1 a_2 \frac{\lambda_1 + \lambda_2 \lambda_1 + \lambda_2 a_1}{1 - a_2} \right) \\ &= \sigma_{\varepsilon_{FD}}^2 + \sigma_{\varepsilon_D}^2 \left(1 - a_1^2 - a_2^2 - \frac{2a_2 a_1^2}{1 - a_2} \right), \\ &\quad (\lambda_1 + \lambda_2 \lambda_1 + \lambda_2 a_1) \sigma_u^2 = -a_1 \sigma_{\varepsilon_D}^2, \\ &\quad \lambda_2 \sigma_u^2 = -a_2 \sigma_{\varepsilon_D}^2.\end{aligned}$$

The first and last equations can be used to calculate the implied values of $\sigma_{\varepsilon^D}^2$ and $\sigma_{\varepsilon^{FD}}^2$ and obtain two inequality constraints on the ARMA(2,2) parameters:

$$\begin{aligned}\sigma_{\varepsilon^D}^2 &= -\frac{\lambda_2}{a_2}\sigma_u^2 > 0, \\ \sigma_{\varepsilon^{FD}}^2 &= \sigma_u^2 \left[1 + \lambda_1^2 + \lambda_2^2 + 2a_1\lambda_1 + 2a_1^2\lambda_2 + 2a_1\lambda_2\lambda_1 + 2a_2\lambda_2 \right. \\ &\quad \left. + 2a_1a_2 \frac{\lambda_1 + \lambda_2\lambda_1 + \lambda_2a_1}{1 - a_2} + \frac{\lambda_2}{a_2} \left(1 - a_1^2 - a_2^2 - \frac{2a_2a_1^2}{1 - a_2} \right) \right] > 0.\end{aligned}\quad (15)$$

The second equation implies the additional constraint

$$0 = a_2\lambda_1(1 + \lambda_2) - a_1\lambda_2(1 - a_2). \quad (16)$$

TABLE A1. ESTIMATES OF ARMA(2,2) PROCESS

	a_1	a_2	λ_1	λ_2	σ_u^2	$\chi_{(1)}/p\text{-value}$
CAN	1.859	-0.896	-1.051	0.365	0.013	4.499
	6.59	-3.94	-4.23	2.91	0.78	0.033
FRA	1.369	-0.420	-0.092	0.020	0.014	5.327
	33.39	-20.60	-0.64	0.62	5.42	0.020
GER	1.734	-0.773	-0.803	0.253	0.101	2.608
	19.47	-8.99	-4.92	3.80	6.18	0.106
ITA	1.685	-0.708	-0.398	0.108	0.001	30.41
	51.41	-32.03	-4.94	4.47	8.28	0.000
JAP	1.212	-0.275	-0.002	0.0004	0.786	2.768
	4.884	-0.295	-0.815	0.272	3.141	0.096
U.K.	1.223	-0.294	-0.003	0.0005	0.575	2.089
	16.408	-5.464	-9.349	6.125	8.212	0.148
U.S.	1.679	-0.747	-0.754	0.237	2.100	3.846
	6.60	-3.18	-2.02	1.55	9.48	0.049

NOTES: For each country, the second row gives t -statistics on the corresponding estimates. $\chi_{(1)}$ and p -values are given for the non-linear constraint (16).

In a first estimation step, we impose (16) but do not impose the inequality constraint. The inequalities are not binding in all countries except for Japan and the United Kingdom. For these countries, we impose $\sigma_{\varepsilon^D}^2 = 0.001$ and reestimate the restricted ARMA(2,2) process. Setting the variance of transient shocks to dividends equal to zero implies that there are no trades based on private information, since the equilibrium is fully revealing.

Table A.1 presents the estimates for the restricted ARMA(2,2) process. These estimates are then used to produce Table 2 in the main text according to the formulas in (15). The estimated ARMA(2,2) produces statistically significant estimates of the autoregressive parameters a_1 and a_2 in all countries (except for Japan's a_2) and of the moving average parameters λ_1 and λ_2 as well (except for France and Japan). Estimates of σ_u^2 are also significant in all cases except for Canada. Finally, the constraint (16) is not rejected in three out of seven countries at the usual 5% significance level and is barely rejected in the case of the United States.

D Matching Model and Data Flows

Both flow and volume data record sums over all transactions in a given month or quarter; the TIC database does not provide guidance on which days, and hence at what prices, the transactions took place. In contrast, our discrete time model makes predictions about holdings at a point in time. To match model-implied changes in holdings to flow data, we need to normalize the latter. One convenient way to do this is to divide flows by total market capitalization at the beginning of the period. To see why this makes sense, suppose that there are n dates between t and $t + 1$ at which transactions are recorded. Let x_i denote the fraction of the net change $\Delta\theta_t^* = \theta_t^* - \theta_{t-1}^*$ in U.S. investors' holdings that takes place at date t_i (with θ_t^* measured as a fraction of outstanding shares). Then normalized net flows are given by

$$\Delta\theta_t^* = \frac{1}{\tilde{P}_t} (\theta_t^* - \theta_{t-1}^*) \sum_{i=1}^n x_i \tilde{P}_{t_i} = (\theta_t^* - \theta_{t-1}^*) \sum_{i=1}^n x_i \frac{\tilde{P}_{t_i}}{\tilde{P}_t},$$

where \tilde{P}_t is the undetrended local stock price. Appendix B above shows that this normalization is consistent with exponential detrending of dividend levels.

Normalized flows are thus equal to the change in holdings multiplied by a weighted average of within-month capital gains. In what follows, we match normalized net flows to the first term, $(\theta_t^* - \theta_{t-1}^*)$. This match is exact if all transactions take place on the first day of the month, that is, $t_1 = t$, $x_1 = 1$ and $x_i = 0$ for $i > 1$. Some evidence on the importance of the resulting bias can be obtained by comparing results to the polar opposite case, when flows are normalized by the end-of-period market capitalization (i.e., $t_n = t + 1$ and $x_n = 1$). In terms of our stylized facts, this change somewhat reduces both the contemporaneous correlation of flows and returns and the persistence of flows, but the effect is on the order of a few percentage points for all countries. We conclude that the normalization is reasonable.

E Impulse Response Functions and Variance Decompositions

The vector $\mathbf{x}_t = (F_t^D, F_{t-1}^D, \phi_t^{s'}, D_t, R_t^B)'$ has a vector AR(1) representation

$$\mathbf{x}_t = \bar{\mathbf{x}} + \mathbf{M}_{xx}\mathbf{x}_{t-1} + \mathbf{M}_{x\varepsilon}\varepsilon_t, \quad (17)$$

where ε_t is the vector of economy-wide shocks and \mathbf{M}_{xx} and $\mathbf{M}_{x\varepsilon}$ are matrices constructed using the equilibrium outcome including the matrices from the Kalman filtering problem. These matrices are built using the equations that describe the processes for F_t^D, F_{t-1}^D, D_t , and R_t^B , respectively. For $\phi_t^{s'} = (\hat{\mathbf{F}}_t^{S'}, \hat{\mathbf{F}}_t^{U'})'$ we need to convert the residuals $\hat{\varepsilon}_t^S$ into ε_t . Since $\hat{\varepsilon}_t^{FB,S} = \varepsilon_t^{FB}$, we need four equations to obtain the values of $(\hat{\varepsilon}_t^{FD,S}, \hat{\varepsilon}_{t-1}^{FD,S}, \hat{\varepsilon}_t^{D,S}, \hat{\varepsilon}_t^{B,S})$. This is done using the first two equations in

$$\hat{\varepsilon}_t^S = \mathbf{K}^S \mathbf{M}^{yS\varepsilon} \varepsilon_t,$$

obtained from (3), and the two equations that arise from

$$\begin{aligned} \mathbf{y}_t^S &= \mathbf{M}^{ySF} \mathbf{F}_{t-1} + \mathbf{M}^{yS\varepsilon} \varepsilon_t \\ &= \mathbf{M}^{ySF} \hat{\mathbf{F}}_{t-1}^S + \mathbf{M}^{yS\varepsilon} \hat{\varepsilon}_t^S. \end{aligned}$$

If \mathcal{S}_t is a vector of stock variables like $P_t, \theta_t^i, i = U, S$, we can write $\mathcal{S}_t = \bar{\mathbf{S}} + \bar{\mathbf{M}}^{Sx} \mathbf{x}_t$, and similarly for flow variables such as net purchases and returns, $\mathcal{F}_t = \bar{\mathbf{F}} + \mathbf{M}^{Fx1} \mathbf{x}_t + \mathbf{M}^{Fx} \mathbf{x}_{t-1}$.

Generating impulse response functions requires iterating on equation (17) and applying the formulas for \mathcal{S}_t and \mathcal{F}_t . Calculating unconditional first and second moments of the relevant variables is also immediate given the simple linear process in (17). For example, noting that the unconditional mean of \mathbf{x}_t is $E[\mathbf{x}_t] = [\mathbf{I} - \mathbf{M}_{xx}]^{-1} \bar{\mathbf{x}}$, we obtain the unconditional variance matrix of \mathbf{x}_t solving the Ricatti equation

$$\begin{aligned} \Sigma_{xx} &= E[(\mathbf{x}_t - E[\mathbf{x}_t])(\mathbf{x}_t - E[\mathbf{x}_t])'] \\ &= \mathbf{M}_{xx} \Sigma_{xx} \mathbf{M}_{xx}' + \mathbf{M}_{x\varepsilon} \Sigma_{\varepsilon\varepsilon} \mathbf{M}_{x\varepsilon}'. \end{aligned}$$

For more details on impulse response functions and variance decompositions see J. D. Hamilton, *Time Series Analysis* (Princeton, N. J.: Princeton University Press, 1994), pp. 318–324.