Internet Appendix for
"Asset Price Dynamics with Limited Attention"

This Internet Appendix contains the notes on MLE estimation and additional figures. References to tables and figures may correspond to those the main document. Note, to avoid confusion the numbering in this Internet Appendix starts where the numbering in the main text left off.

A Details of the Maximum Likelihood Estimation

The maximum-likelihood estimation is implemented in Python's `statsmodels` (Seabold and Perktold, 2010). More specifically, our code uses Chad Fulton's statespace model within this software module. After acceptance the code submitted with the revision will be posted online together with simulated sample to run it on. The code includes a simplified version that is easier to parse and can serve as a stating point for researchers interested in estimating related models. The paper's data falls under a nondisclosure agreement with the NYSE. Numerous researchers who were visiting economists at the NYSE have had access to the data.

In the remainder of the section we calibrate the model to find reasonable starting values for the steepest-ascent algorithm used to maximize the likelihood. A high-level summary of the way starting values are set is that we first set the correlation of dividend shocks and target portfolios to zero (i.e., $\rho = 0$) to sequentially pick parameter values. We then pick the correlation at a level that would explain potential excess negative correlation between dividend shocks and market-maker inventories in the data.

If ρ is assumed to be zero, then all remaining parameters can then be solved sequentially by matching several observed "identifying" cross-autocovariances as follows:

1. First solve for retail risk-mass (μ_{dr}, μ_{mr}, and μ_{qr}) based on retail order flows. These parameters are identified solely off of the autocovariance function for retail order flow.
2. Then, given these estimates, solve for gap-sensitivity of market-maker inventories (β_M) which is identified through a cross-autocovariance between market-maker inventories and retail order flows.
3. All these estimates are unlikely to fit the autocovariance of market-maker inventories. The differential between the implied autocovariance based on retail flows and the observed autocovariance is used to identify the risk-mass of slow institutions (μ_{di}, μ_{mi}, and μ_{qi}).
4. Now that all trade-data parameters are identified, we involve price data to solve for gap-sensitivity of price pressure (β_w) through the first-order autocovariance of market returns (which is preferred over return variance as it removes dependence on the thus far unknown σ_w).
5. Finally, these parameters imply a variance in returns. The extent to which the observed return variance exceeds this variance identifies dividend risk (σ_w).

The remainder of the section will describe all these steps in full detail. All derivations are based on the closed-form expressions for the (multivariate) autocovariance function (that includes variance) of model variables in (63)-(64) .

Starting value for retail risk-mass: μ_{dr}, μ_{mr}, and μ_{qr}. Retail-flow autocovariance does not depend on any parameter other than retail risk-mass. Picking three autocovariances should therefore identify the three retail risk-mass parameters. We pick the autocovariance at lag one, five, and 20.[49] Picking element (8,8) of the autocovariance function in (64) yields the following expression:

$$\begin{gathered}\operatorname{cov}\left(RetFlow_t, RetFlow_{t-n}\right) = \\ \mathbf{1}_{(1\times 3)}\left(I_3 - e^{-\Lambda_r}\right)e^{-(n-1)\Lambda_r}\operatorname{cov}\left(G_{rt}, RetFlow_t\right).\end{gathered} \tag{76}$$

[49] These frequencies are chosen because they are visible in Figure 1 and enable an reasonable fit to the curvatures of the lines in Figure 1 at the optimization's starting values.

The contemporaneous covariance between G_{rt} and $RetFlow_t$ is in row four through six and column eight of $\text{var}(Y_t)$ in (63). The corresponding two terms on the right-hand side of (63) are

$$e^{-\Lambda_r}\text{diag}\begin{pmatrix}\frac{\mu_{dr}^2}{2\lambda_d} & \frac{\mu_{mr}^2}{2\lambda_w} & \frac{\mu_{qr}^2}{2\lambda_m}\end{pmatrix}\left(I_3 - e^{-\Lambda_r}\right)1_{(3\times 1)} \tag{77}$$

and

$$\begin{pmatrix}0_{(1\times 3)} & -1_{(1\times 3)} & 0_{(1\times 4)} & 1_{(1\times 3)}\end{pmatrix}\text{var}(\varepsilon_t)\begin{pmatrix}0_{(3\times 3)}\\ I_3\\ 0_{(7\times 3)}\end{pmatrix}, \tag{78}$$

respectively. Therefore the model-implied autocovariance is:

$$\begin{gathered}\text{cov}(RetFlow_t, RetFlow_{t-n}) = \sum_{j\in\{d,m,q\}}\left(1-e^{-\lambda_j}\right)e^{-(n-1)\lambda_j}\times\\ \left(\frac{e^{-\lambda_j}\left(1-e^{-\lambda_j}\right)}{2\lambda_j} + \left(\frac{1-e^{-\lambda_j}}{\lambda_j} - \frac{1-e^{-2\lambda_j}}{2\lambda_j}\right)\right)\mu_{jr}^2.\end{gathered} \tag{79}$$

Note that (79) evaluated at $n \in \{1, 5, 20\}$ yields a system with three equations and three unknowns with boundary conditions because risk-mass parameters have to be non-negative. One way to find reasonable risk-mass estimates is to solve the following least-squares minimization:

$$\boxed{\begin{pmatrix}\hat{\mu}_{dr}\\ \hat{\mu}_{mr}\\ \hat{\mu}_{qr}\end{pmatrix} = \left(\text{argmin}_{x_{(3\times 1)}\geq 0}\left(A_n x - b\right)^\top\left(A_n x - b\right)\right)^{\frac{1}{2}},} \tag{80}$$

where b is a 3×1 column vector with observed retail-flow autocovariances at lags $n \in \{1, 5, 20\}$, respectively, and A_n is a matrix with as row labels the frequencies $f \in \{d, m, q\}$ and as a column labels the lags $n \in \{1, 5, 20\}$. Element (f, n) in A_n equals the coefficient of μ_{fr}^2 in (79) for lag n. As the matrix A_n is well conditioned we solve the argmin part in (80) simply by taking $x_0 = A_n^{-1}b$. Should any element of x_0 be negative we set it to zero ex-post so as to ensure risk-mass starting values are non-negative.[50]

Starting value gap-sensitivity *MMInv*: β_M. The starting values $\hat{\mu}_{dr}$, $\hat{\mu}_{mr}$, and $\hat{\mu}_{qr}$ along with the following autocovariance term identify β_M:

$$\begin{gathered}\text{cov}(RetFlow_t, MMInv_{t-1}) =\\ 1_{(1\times 3)}\left(I_3 - e^{-\Lambda}\right)e^{-(1-1)\Lambda}\text{diag}\begin{pmatrix}\frac{\hat{\mu}_{dr}^2}{2\lambda_d} & \frac{\hat{\mu}_{mr}^2}{2\lambda_w} & \frac{\hat{\mu}_{qr}^2}{2\lambda_m}\end{pmatrix}\beta_M 1_{(3\times 1)}.\end{gathered} \tag{81}$$

$\hat{\beta}_M$ therefore is:

$$\boxed{\hat{\beta}_M = \frac{\text{cov}(RetFlow_t, MMInv_{t-1})}{1_{(1\times 3)}\left(I_3 - e^{-\Lambda}\right)\text{diag}\begin{pmatrix}\frac{\hat{\mu}_{dr}^2}{2\lambda_d} & \frac{\hat{\mu}_{mr}^2}{2\lambda_w} & \frac{\hat{\mu}_{qr}^2}{2\lambda_m}\end{pmatrix}1_{(3\times 1)}}.} \tag{82}$$

Note that the beauty of picking the cross-autocovariance of market-maker inventories and retail flows is that one does not need to know the risk mass of slow *institutions* as their target portfolio changes are assumed to be orthogonal to those of retail investors.[51]

[50] In the code we use very small values instead of zero to avoid a singular prediction error covariance matrix in the Kalman filter.

[51] Note that $\hat{\beta}_M$ in (82) equals the ratio of the observed covariance and the model-implied covariance assuming β_M is one. This insight is used in the code to maximize code efficiency. This trick is used for several starting-value expressions.

Starting value for the risk mass of slow institutions: μ_{di}, μ_{mi}, and μ_{qi}. The identification of the risk masses of slow institutions is in the extent to which market-maker inventory autocovariance exceeds what the model predicts it to be solely based on $\hat{\mu}_{dr}$, $\hat{\mu}_{mr}$, $\hat{\mu}_{qr}$, and $\hat{\beta}_M$ (i.e., assuming risk mass to be zero for slow institutions). More specifically, the model-implied autocovariance of market-maker inventories at lag n is (using (63)):

$$\text{cov}\left(MMInv_t, MMInv_{t-n}\right) = \beta_M 1_{(1\times 6)} e^{-\Lambda} e^{-(n-1)\Lambda} \text{cov}\left(MMInv_t, G_t\right) \beta_M 1_{(6\times 1)} = \sum_{j\in\{d,m,q\}} \beta_M^2 e^{-n\lambda_j} \frac{\mu_{ji}^2 + \hat{\mu}_{jr}^2}{2\lambda_j}. \tag{83}$$

At this point the approach is similar to the one we used to identify the risk mass of retail investors. In other words, utilizing (83) we solve for risk masses in the same way as we did based on (80):

$$\boxed{\begin{pmatrix} \hat{\mu}_{di} \\ \hat{\mu}_{mi} \\ \hat{\mu}_{qi} \end{pmatrix} = \left(\text{argmin}_{x_{(3\times 1)}\geq 0} \left(A_n\left(x-c\right)-b\right)^\top \left(A_n\left(x-c\right)-b\right)\right)^{\frac{1}{2}},} \tag{84}$$

where b is a 3×1 column vector with observed inventory autocovariance at lags $n \in \{1, 5, 20\}$, respectively, and A_n is a matrix with as row labels the frequencies $f \in \{d, m, q\}$ and as a column labels the lags $n \in \{1, 5, 20\}$. Element (f, n) in A_n equals the coefficient of μ_{fr}^2 in (83) for lag n and

$$c = \begin{pmatrix} \frac{\hat{\mu}_{dr}^2}{2\lambda_d} & \frac{\hat{\mu}_{mr}^2}{2\lambda_m} & \frac{\hat{\mu}_{qr}^2}{2\lambda_q} \end{pmatrix}^\top. \tag{85}$$

Now that we have identified all risk-masses $\hat{\mu}_{jk}$ with $j \in \{d, m, q\}, k \in \{i, r\}$ along with $\hat{\beta}_M$, we add price data to the trade data used thus far to identify the remaining parameters β_w and σ_w.

Starting value gap-sensitivity price pressure: β_w. With all parameters calibrated thus far β_w can be calibrated by matching the first-order autocovariance in return. Note that this object does not depend on w and therefore removes dependence on σ_w which is thus far unknown. The mathematical expression for this autocovariance is in principle available but involves long mathematical expressions. Since the observed covariance is affine in the model-implied covariance assuming $\beta_w = 1$ with zero intercept and a coefficient of β_w^2. This allows us to calibrate β_w as follows (yielding the exact same result as computing the analytic expression but more easily expressed and coded, see also footnote 51):

$$\boxed{\hat{\beta}_w = \left(\frac{\text{cov}\left(Return_t, Return_{t-1}\right)}{\hat{\sigma}_{Return, Return_{t-1}}}\right)^{\frac{1}{2}}}, \tag{86}$$

where $\hat{\sigma}_{Return, Return_{t-1}}$ denotes the model-implied covariance between $Return_t$ and $Return_{t-1}$ on the assumption that β_w is one.[52]

[52] In practice, one could set β_w as the square root of the average of multiple ratios based on different lag values. In the code, we used 1, 5, and 20 to have a smoothen the calibrated value. This could also be done when calibrating the other parameters but in our application the returns variable was more noisy than others and we therefore only implemented it here.

Starting value dividend risk: σ_w. The variance of returns along identifies the remaining parameter σ_w given starting values for all other parameters:

$$\begin{gathered}\operatorname{var}\left(Return_t\right) = \\ \hat{\beta}_w 1_{(1\times 6)}\left(rI_6+\Lambda\right)^{-1}\left(I_6-e^{-\Lambda}\right)\operatorname{var}\left(G_t\right)\left(I_6-e^{-\Lambda}\right)\left(rI_6+\Lambda\right)^{-1}1_{(6\times 1)}\hat{\beta}_w + \\ \hat{\beta}_w 1_{(1\times 6)}\left(rI_6+\Lambda\right)^{-1}\operatorname{diag}\left(\tfrac{1-e^{-2\lambda_d}}{2\lambda_d}\hat{\mu}_{di}^2 \quad \ldots \quad \tfrac{1-e^{-2\lambda_m}}{2\lambda_m}\hat{\mu}_{mr}^2\right)\left(rI_6+\Lambda\right)^{-1}1_{(6\times 1)}\hat{\beta}_w + \\ \sigma_w^2\end{gathered} \tag{87}$$

and therefore

$$\boxed{\hat{\sigma}_w = \left(\operatorname{var}\left(Return_t\right) - \hat{\beta}_w^2 \sum_{j\in\{d,m,q} \frac{1-e^{-\lambda_j}}{\lambda_j\left(r+\lambda_j\right)^2}\left(\hat{\mu}_{ji}^2+\hat{\mu}_{jr}^2\right)\right)^{\frac{1}{2}}.} \tag{88}$$

Starting value correlation dividend and target portfolio innovations: ρ. Finally, ρ is estimated as:

$$\boxed{\hat{\rho} = \frac{\operatorname{cov}\left(MMInv, Return\right) - \hat{\sigma}_{MMInv,Return}}{\hat{\sigma}_{MMInv}\hat{\sigma}_w},} \tag{89}$$

where $\hat{\sigma}_{MMInv,Return}$ denotes the model-implied covariance between *MMInv* and *Return* on the assumption that ρ is zero. The same goes for $\hat{\sigma}_X$ which denotes the model-implied volatility of X. Note that the denominator is $\hat{\sigma}_w$ instead of $\hat{\sigma}_{Return}$ as ρ pertains to dividend innovations.

B Model-Implied Autocorrelations for Lower-Frequency Returns

The N-period return (used for lower frequency returns such as monthly or quarterly) is:

$$r_{m,t} = p_t - p_{t-N} \tag{90}$$

where t runs over days and it is assumed that there are N work days in a month. The price p_t consist of a martingale component m_t with daily innovations w_t plus a friction-induced pricing error s_t. The return therefore becomes:

$$r_{m,t} = m_t - m_{t-N} + s_t - s_{t-N} = (w_{t-1} + \ldots + w_{t-N}) + s_t - s_{t-N}. \tag{91}$$

The variance of monthly returns therefore is:

$$\sigma^2_{r_{m,t}} = N\sigma^2_w + 2\sigma^2_s - 2\rho_{s,N}\sigma^2_s \tag{92}$$

where $\rho_{s,N}$ is the N lag autocorrelation in the pricing error. The first-order autocovariance of monthly return is:

$$Cov(r_{m,t}, r_{m,t-N}) = -\sigma^2_s + 2\rho_{s,N}\sigma^2_s - \rho_{s,2N}\sigma^2_s. \tag{93}$$

Therefore, the first-order autocorrelation in monthly return is:

$$\frac{-\left(1 - 2\rho_{s,N} + \rho_{s,2N}\right)\sigma^2_s}{N\sigma^2_w + 2\left(1 - \rho_{s,N}\right)\sigma^2_s}. \tag{94}$$

The pricing error at any time t is defined as:

$$s_t = -\beta_w \left(rI_6 + \Lambda\right)^{-1} G_t \tag{95}$$

where the variance of G_t is in (62). The covariance of s_t and s_{t-j} with $j > 0$ is:

$$\begin{aligned} Cov(s_t, s_{t-N}) &= \beta_w \left(rI_6 + \Lambda\right)^{-1} e^{-\Lambda} e^{-(j-1)\Lambda} \Sigma_G \left(rI_6 + \Lambda\right)^{-1} \beta_w^\top \\ &= \beta_w \left(rI_6 + \Lambda\right)^{-1} e^{-j\Lambda} \Sigma_G \left(rI_6 + \Lambda\right)^{-1} \beta_w^\top \end{aligned} \tag{96}$$

and

$$Var(s_t) = \beta_w \left(rI_6 + \Lambda\right)^{-1} \Sigma_G \left(rI_6 + \Lambda\right)^{-1} \beta_w^\top. \tag{97}$$

C Daily Target Portfolio Changes

The variance of daily total target-portfolio changes of large investors are:

$$\int_0^{\Delta t} \mu^\top \mu \mathrm{d}u. \tag{98}$$

As the units of both market-maker inventories and retail flows are in million dollar and Δt is a day, this implies that the standard deviation of the total daily target portfolio changes is:

$$\left(\sum_{k \in \{d,m,q\}} \mu_{ki}^2 + \mu_{kr}^2 \right)^{\frac{1}{2}} \text{ million dollar.} \tag{99}$$

D MLE Results for Size Terciles

The following four pages show the results of our MLE for each of the three size terciles. The underlying parameter estimates are shown in Table 2 from the main paper. The figures compare with Fig 2 from the main paper.

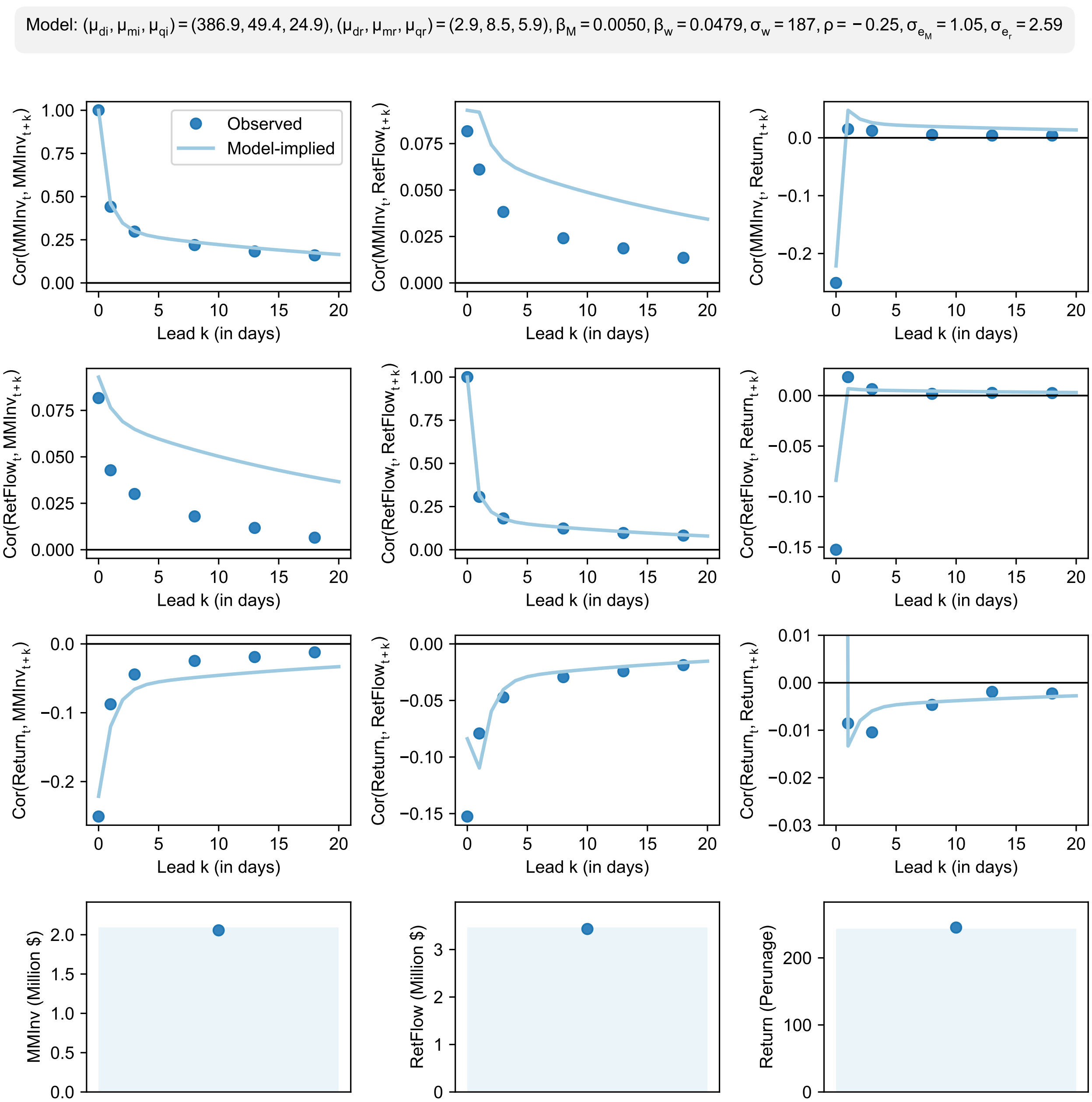

Figure 7. MLE Results for Large Stocks Only. Similar to Figure 2 except for the sample of large stocks only.

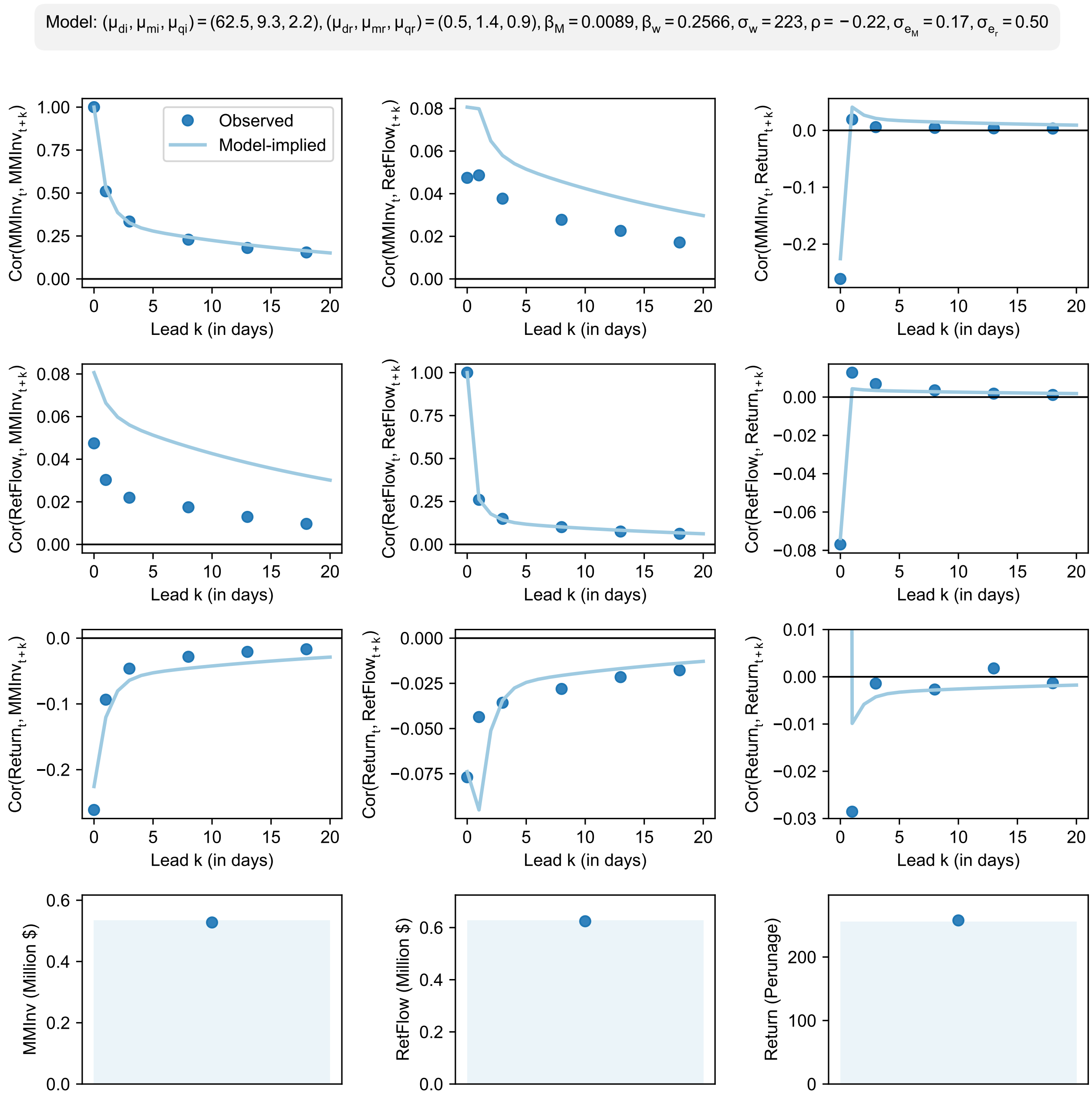

Figure 8. MLE Results for Medium Stocks Only. Similar to Figure 2 except for the sample of medium stocks only.

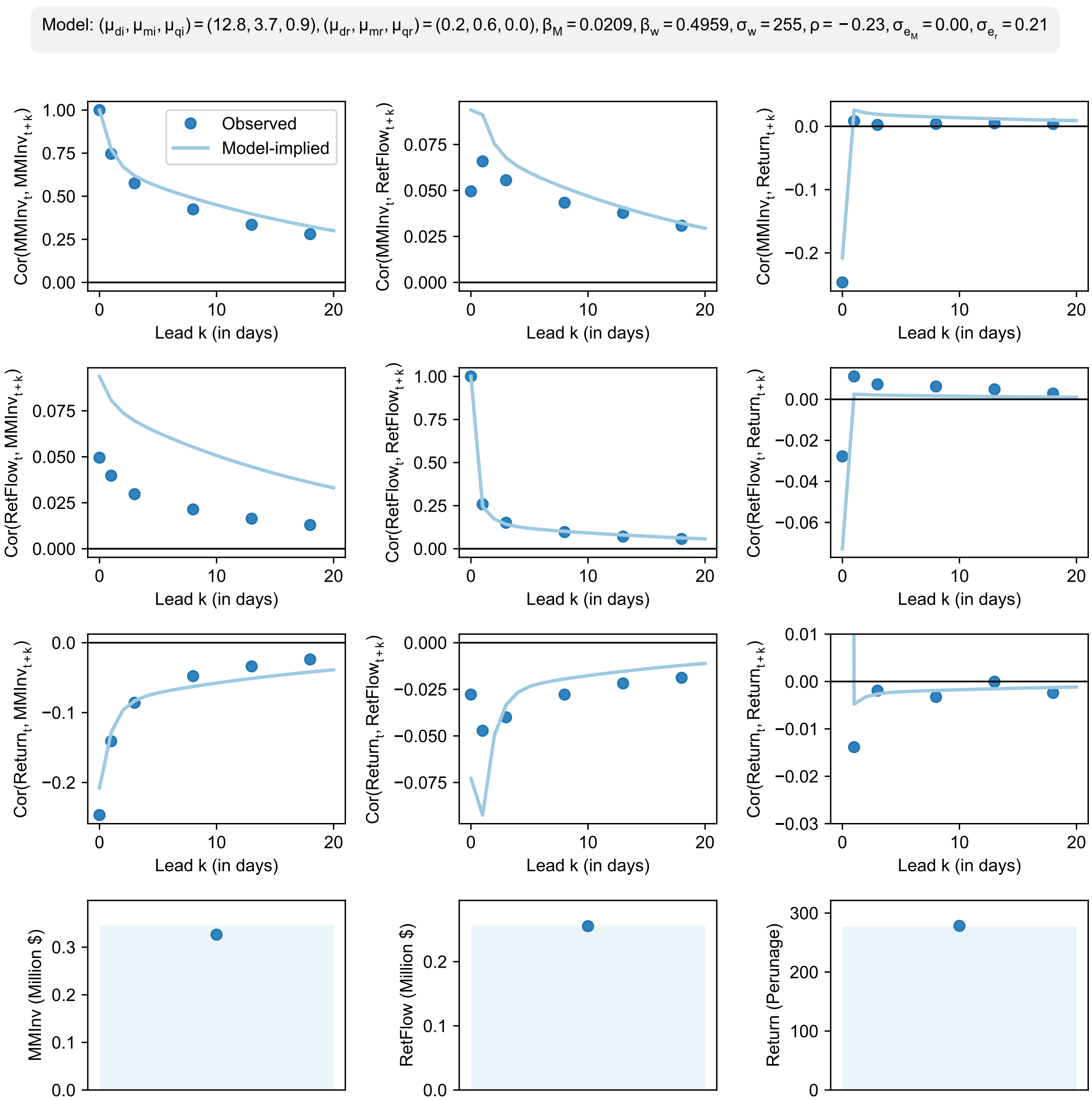

Figure 9. MLE Results for Small Stocks Only. Similar to Figure 2 except for the sample of small stocks only.

E MLE Results with Daily / Weekly / Monthly Slow Investors

Version. This table presents the maximum likelihood parameter estimates and their standard errors. We consider "All" stocks as well three size-terciles labeled "Large", "Medium", and "Small". Subscripts: "d" daily; "w" weekly; "m" monthly; "i" slow institutional investors; "r" slow retail investors. Idiosyncratic noise in dividends (σ_w); market-maker inventories (σ_{e_M}); and retail flows (σ_{e_r}); The stars (*/**/***) indicate statistical significance at a 10%, 5%, and 1% level, respectively.

	All	*Large*	*Medium*	*Small*
Panel A: Risk masses of slow institutional investors				
μ_{di}	154 *** (3.36)	381 *** (17.6)	55.0*** (2.32)	9.50*** (0.32)
μ_{wi}	0.01 (0.90)	0.01 (4.15)	0.01 (0.50)	0.17 (0.10)
μ_{mi}	28.9*** (0.65)	65.9*** (3.10)	8.97*** (0.39)	2.97*** (0.10)
Panel B: Risk masses of (slow) retail investors				
μ_{dr}	1.65 *** (0.003)	3.03 *** (0.008)	0.46 *** (0.002)	0.16 *** (0.001)
μ_{wr}	0.010 (0.03)	0.012 (0.11)	0.010 (0.02)	0.170*** (0.005)
μ_{mr}	5.23*** (0.01)	9.78*** (0.05)	1.55*** (0.01)	0.56 *** (0.005)
Panel C: Deep parameters				
β_M	0.0089*** (0.002)	0.0066*** (0.0003)	0.0100*** (0.0004)	0.0279*** (0.0009)
β_w	0.0935*** (0.0025)	0.0426*** (0.0023)	0.288 *** (0.0141)	0.720 *** (0.0333)
Panel D: Volatility related to returns, market-maker inventories, and retail flows				
σ_w	237 *** (0.33)	217 *** (0.54)	233 *** (0.58)	262 *** (0.53)
σ_{e_M}	0 (0.011)	0 (0.014)	0 *** (0.0008)	0 *** (0.0001)
σ_{e_r}	1.58 *** (0.0007)	2.53 *** (0.022)	0.50 *** (0.0004)	0.21 *** (0.0002)
Panel E: Shared component				
ρ	−0.210*** (0.0011)	−0.199*** (0.0018)	−0.208*** (0.002)	−0.245*** (0.0022)
# of stocks	689	230	229	230
# of obs	1,206,935	403,971	402,169	400,795