Online Appendix to Where Does the Predictability from Sorting on Returns of Economically Linked Firms Come From?

Aaron Burt and Christopher Hrdlicka*

July 5, 2020

ABSTRACT

This Online Appendix contains a number of additional details and results that are referenced in the body of the paper.

*Aaron Burt is at the Michael F. Price College of Business, University of Oklahoma, Adams Hall 259, Norman, Oklahoma, 73026, USA, email: aaronburt@ou.edu. Christopher Hrdlicka is at the Michael G. Foster School of Business, University of Washington, PACCAR Hall, Box 353226, Seattle, WA 98195-3226, USA, email: hrdlicka@uw.edu

Contents

O.I. Link formation summary

This section contains a summary of the formation of the economic links used in the main text. Further details can be found in the original papers.

A. Data overview: Cohen and Frazzini (2008)

Cohen and Frazzini (2008) use the economic link between customers and suppliers to measure slow information diffusion. They obtain data on these links from regulation SFAS No. 131, which requires each firm to report any customer whose sales represent more than 10% of the firm's total sales. The authors argue that shocks to a supplier's customers (as measured by a portfolio of the supplier's customer returns) will propagate through to the supplier itself.

Following their method, we successfully replicate the key results of Cohen and Frazzini (2008). Using the U.S. Customer Supplier Links data made available by the authors on Andrea Frazzini's website, we calculate the equal-weighted excess return for a portfolio of customers of each supplier in each month.[1] We then rank customer portfolio returns in month t in ascending order. We use the customer portfolio return rankings at t to assign linked supplier firms at t into quintile portfolios. We calculate the value and equal-weighted $t + 1$ portfolio excess returns of all suppliers in each quintile to generate a monthly time series of portfolio returns, which we use throughout our analysis.

B. Data overview: Cohen and Lou (2012)

Cohen and Lou (2012) demonstrate return predictability of complex conglomerate firms based on a portfolio of simple firms' returns. These simple firms are single-industry firms in the industries in which the conglomerate has operating segments. Assuming information processing constraints, the authors argue that conglomerate firms are more difficult to an-

[1] See http://www.econ.yale.edu/~af227/data_library.htm

alyze than single-industry firms. As this difficulty leads to delayed information diffusion, a portfolio of single-industry firms should predict the returns of conglomerate firms.

Following Cohen and Lou (2012), we use data on operating and business segments from the Compustat segment files to identify firms as single-industry firms or as conglomerates. We use 2 digit SIC codes to identify the segment. Using the single-industry firms only, we calculate the industry portfolio returns of these firms in each 2 digit SIC code.

For each conglomerate firm, we identify its principal operating and business segments that represent more than 10% of the firm's sales. We use operating segments of each conglomerate firm along with each segment's industry return (calculated from single-industry firms) to form a value-weighted portfolio return, which, following the original paper, we refer to as the pseudo-conglomerate returns. At each time t, we rank the pseudo-conglomerate returns in ascending order into deciles. Using these rankings, we assign the corresponding conglomerate firms at t to the (value or equal-weighted) decile portfolios (returns at $t+1$).

C. *Data overview: Cao et al. (2016)*

Cao et al. (2016) measure slow information diffusion among firms in a strategic alliance. The authors provide evidence that investor inattention, as measured by news intensity on alliance announcement days, causes information shocks to one firm to be impounded into the price of another alliance-linked firm with a delay. The data on alliances comes from the Securities Data Company (SDC) platinum database. SDC collects the data from news announcements, SEC filings and industry publications. For the sample of firms in an alliance between 1985 and 2012, the paper documents that a firm's return is predicted by the portfolio of firms linked through a strategic alliance.

We successfully replicate the study's main result by constructing the sample of alliance-linked firms as documented in Cao et al. (2016). Each month, for each firm in an alliance, we form a portfolio of all other firms in the alliance. We rank the time t portfolio returns in ascending order. Using these rankings, we assign the alliance-linked firm at t (returns at $t+1$) to one of five portfolios. We then calculate the equal- and value-weighted returns for each portfolio. We form a time series of long-short portfolio returns by subtracting the lowest quintile portfolio return from the highest portfolio return in each month.

O.II. Econometrics of lead lag sorts

In this section, we show how that the profits from trading in lagging economic firms based on sorting them into portfolios based on information about economically linked leader firms decomposes in a way similar to that from using the continous weights analogous to the framework pioneered by Lo and MacKinlay (1990) as described in Section O.IX. We show an important difference in that sorting on estimated idiosyncratic returns introduces a downward bias from estimation noise and how to correct for this bias.

Like in the main text we present the case of one-to-one economic links. This setup equivalently describes the links between two individual firms or a portfolio and a single firm, e.g., one supplier linked to one customer or to one portfolio of customers. We also focus on the case of static parameters (e.g., alphas and betas) for notational simplicity, while changing the timing convention of the correlation in parameters between the leaders and laggers to allow for a more general relationship between them that captures not only mommentum but also seasonal variation in expected returns.

A. Definitions and assumptions

Asset pricing model and factors: Given the linear asset pricing model against which the long-short alphas are measured, we have the canonical decomposition of returns of firm i into the alpha, modeled risks and idiosyncratic returns (all returns are excess returns unless otherwise noted):

$$r_{i,t} = \alpha_i + \vec{\beta}_i' \vec{f}_t + \epsilon_{i,t}. \tag{O-1}$$

Where the factors f are normal with mean $\vec{\mu}_f$ and variance-covariance matrix Σ_f. The idiosyncratic returns are mean zero by construction.

Economic links and correlated risks: Let there be two sets of firms A and B (i.e., leaders and laggards). Let the set of links L be the set of ordered pairs of firms $\{i, j\}$ where $i \in A$ and $j \in B$. We model this cross-sectional correlation between economically linked firms as

$$\alpha_j = \rho_\alpha \alpha_i + \eta_{\alpha,i,j}, \quad \{i, j\} \in L \tag{O-2}$$

$$\vec{\beta}_j = \rho_\beta \vec{\beta}_i + \vec{\eta}_{\beta,i,j}, \quad \{i, j\} \in L \tag{O-3}$$

where $\rho_\alpha > 0$, $\eta_{\alpha,i,j}$ and $\eta_{\alpha,i,j}$ are mean 0, and ρ_β is a matrix. All these terms are independent and normally distributed. Furthermore, they are independent of the firms' idiosyncratic returns and the distribution of the idiosyncratic returns. The idiosyncratic returns, themselves, are mean zero and normally distributed with variances that are a random variable.

Information delay: A link between firms means that, on average, news affects linked firms in the same direction (as measured by returns). Furthermore, the news sometimes affects firms in set A before affecting firms in set B. This (delayed) link between firm i in set A and firm j in set B can be written as

$$\epsilon_{j,t+1} = \rho_{\epsilon,i,j} \epsilon_{i,t} + \psi_{i,j,t+1}, \quad \{i, j\} \in L \tag{O-4}$$

where $\psi_{i,j,t+1}$ is a normal shock independent of both $\epsilon_{i,t}$ and the model-predicted value for the B firm, i.e., $\alpha_j + \vec{\beta}_j' f_{t+1}$. The assumption that news affects linked firms in the same direction means that a positive cross-autocorrelation exists between linked firms' returns. That is, $\rho_{\epsilon,i,j} \geq 0$ for all $\{i, j\} \in L$.

Trading strategy from sorting on returns: To learn about the predictability among these economically linked firms, the econometrician sorts the firms in set A by their excess returns at time t into N quantiles (e.g., quintiles, deciles, etc.). This sorting is used to assign the linked B firms to portfolios in the following period (their returns at $t+1$). These portfolios are either equal- or value-weighted and are rebalanced each period by repeating this procedure. This procedure produces a time-series of N portfolio returns formed from stocks in set B. Let $B(P)$ be the Pth such portfolio, where P ranges from 1, the lowest quantile, to N, the highest quantile.

The alphas of these portfolios are computed using a linear regression with *constant* betas over the sample window. The long-short alpha formed by going long portfolio N and short portfolio 1 is the statistic that the literature has used to measure the predictability present in economically linked firms.

B. *Decomposing alphas from return sorts*

PROPOSITION O-1: *For economically linked firms in set A and B, linked via the preceding assumptions but otherwise independent, if the factors are iid and there exists cross-sectional dispersion in A firms' alphas, then the expected alpha of the portfolio $B(P)$ can be decomposed into two pieces:*

$$E\left[\alpha_{B(P)}\right] = \varrho_{\alpha,P} + \varrho_{\epsilon,P}. \tag{O-5}$$

The first component, $\varrho_{\alpha,P}$, is due to correlated alphas and the second component, $\varrho_{\epsilon,P}$, is due to slow information diffusion. $\varrho_{\alpha,P}$ and $\varrho_{\epsilon,P}$ are increasing in P and can be expressed as

$$\varrho_{\alpha,P} = \rho_\alpha E[\alpha_{A(P,R)}] \quad \textit{and} \tag{O-6}$$

$$\varrho_{\epsilon,P} = \bar{\rho}_\epsilon E[\epsilon_{A(P,R)}] \tag{O-7}$$

where $\bar{\rho}_\epsilon$ is the average correlation of the idiosyncratic shocks across economically linked firms and the subscript $A(P,R)$ denotes the value conditional on being in portfolio P when *sorted by returns, R.*

The proof of Proposition O-1 and of all other propositions unless otherwise noted are in Appendix O.III. Closed-form expressions for the conditional expectations in Equations (O-6) and (O-7) are in Online Appendix O.V.

The B firm portfolio alpha has two components because the chances of an A firm falling in any quantile portfolio depends upon both on the A firm's alpha and its idiosyncratic shock. Higher alphas or higher idiosyncratic returns increase the chance a firm falls into a higher portfolio. Because the A firms' alphas and idiosyncratic returns are correlated with those of the B firms through the economic link, the B firms with higher alphas or idiosyncratic returns are, therefore, more likely to be assigned to higher quantile portfolios.

Though the betas of the A and B firms are also correlated, no term for the betas appears in the portfolio alpha because the factors are iid. When the factors are iid, however, the linear regression captures the unconditional beta of the B portfolio even though the betas of the B firm portfolio vary each period. As such, this time variation in beta does not contribute to an unconditional alpha. The variation in the B portfolio returns from this time-varying alpha shows up as additional idiosyncratic volatility in the regression residuals. See Lewellen and Nagel (2006) for further details of the conditional and unconditional alphas in the presence of time-varying betas.

i. Returns are predictable even in efficient markets

COROLLARY O-1: *The expected alpha of the long-short portfolio $B(LS)$ formed from $B(N)$ and $B(1)$ is*

$$E[\alpha_{B(LS)}] = (\varrho_{\alpha,N} - \varrho_{\alpha,1}) + (\varrho_{\epsilon,N} - \varrho_{\epsilon,1}). \tag{O-8}$$

An immediate consequence of the $\varrho_{\alpha,P}$ increasing in P is that returns are predictable–i.e., one finds a positive long-short alpha from sorting on returns–even when the market is efficient (i.e., $\rho_{\epsilon,i,j} = 0$ for all pairs of economically linked firms). More generally, the return predictability extends beyond the horizon of slow information diffusion. This occurs because the alphas of the A and B firms along with their correlation are persistent, making $(\varrho_{\alpha,N} - \varrho_{\alpha,1})$ positive beyond a single period. This persistence produces increasing cumulative long-short alphas. The one-period and cumulative long-short alphas increase in (1) the strength of the correlation in the alphas, ρ_α, and (2) in the amount of cross-sectional variation in the alphas of the A firms.

ii. Non-iid factors add a third component to portfolio alphas

If we slacken the assumption that the factors are iid, then the correlation in the modeled risk components can contribute a term to the B portfolios formed by sorting on the returns of the A firms.

PROPOSITION O-2: *With the assumptions of Proposition O-1 except now the factors may be non-iid, if there is also cross-sectional dispersion in the A firms' betas, then the expected alpha of the long-short portfolio $B(LS)$ can be decomposed into three pieces:*

$$E\left[\alpha_{B(LS)}\right] = \varrho_{\alpha,LS} + \varrho_{\beta' f,LS} + \varrho_{\epsilon,LS} \tag{O-9}$$

The first component, $\varrho_{\alpha,P}$, is due to correlated alphas and is positive. The second component, $\varrho_{\beta' f}$, is due to correlation in the modeled risk components. The third component, $\varrho_{\epsilon,LS}$, is due to slow information diffusion and is positive when $\bar{\rho}_\epsilon > 0$.

The term $\varrho_{\beta' f,LS}$ appears, because the constant beta regression cannot fully capture the expected return of the B firms. The correlation of the risk exposures, i.e., betas, of the economically linked firms does not directly determine this additional term. The component is determined by the interaction of the time-varying beta induced in the B firm portfolios and the time-series properties of the factors which are characterized by Lewellen and Nagel (2006). This additional term may contribute positively or negatively to the long-short alpha in expectation.

iii. Alphas and betas in returns act as noise in the sorting

COROLLARY O-2: *Under the assumptions of Proposition O-1 or O-2, the component of the long-short portfolio from idiosyncratic returns, $\varrho_{\epsilon,LS}$, is decreasing in the cross-sectional variance of the alpha and the model-predicted returns, $\vec{\beta}_i'\vec{f}_t$, and increasing in the cross-sectional variances of the idiosyncratic returns of the A firms.*

This decrease occurs because the non-idiosyncratic components of the returns act as noise relative to the idiosyncratic return signal: firms are assigned to portfolios based on all three dimensions rather than idiosyncratic returns only.

C. *Sorting on idiosyncratic returns measures information diffusion*

Sorting directly on the idiosyncratic return of the leading economically linked firms provides two benefits. First, it isolates the component of predictability from delayed information diffusion. Second, it provides the largest possible alpha from this effect, making it the most likely way to detect any delays in information diffusion. Proposition O-3 and its corollaries state this more precisely.

PROPOSITION O-3: *Let the portfolios of linked firms in set B formed using sorts on idiosyncratic returns as $B'(P)$. If given the assumptions of Proposition O-2, then the expected alpha of the portfolio $B'(P)$ can be decomposed into two pieces:*

$$E[\alpha_{B'(P)}] = \bar{\alpha}_B + \varrho'_{\epsilon,P}. \tag{O-10}$$

The first component, $\bar{\alpha}_B$, is the average misspecification of stocks in set B and is the same for all the quantile portfolios. The second component, $\varrho'_{\epsilon,P}$, is due to slow information diffusion, is increasing in P and can be expressed as

$$\varrho'_{\epsilon,P} = \bar{\rho}_\epsilon E[\epsilon_{A(P,\epsilon)}], \tag{O-11}$$

where the subscript $A(P, \epsilon)$ denotes the value conditional on being in portfolio P when sorted on idiosyncratic returns. Furthermore, the long-short portfolio of the extreme portfolios has expected alpha:

$$E[\alpha_{B'(LS)}] = \varrho'_{\epsilon,N} - \varrho'_{\epsilon,1}. \tag{O-12}$$

The key difference compared to sorting on returns (Propositions O-1 and O-2) is that the first term in Equation O-10 is constant across all sorted portfolios.[2] Sorting on idiosyncratic returns also eliminates the variation in the B firm betas due to the sorting. Thus, the linear regression with the constant beta is able to capture the expected return component from the modeled risks of the B firms.

The long-short portfolio, denoted $B'(LS)$, as a (first) difference, eliminates the first terms in Equation (O-10), making the long-short portfolio solely a measure of slow information diffusion. Because this long-short alpha is determined solely from the idiosyncratic returns, sorts on idiosyncratic returns are the most efficient sort to detect delays in information diffusion:

COROLLARY O-3: *Under the assumptions of Proposition O-3, the slow information diffusion component of Equation* (O-12) *is greater than or equal to that of Equation* (O-8)*:*

$$\varrho'_{\epsilon,N} - \varrho'_{\epsilon,1} \geq \varrho_{\epsilon,N} - \varrho_{\epsilon,1} \tag{O-13}$$

The inequality is strict whenever there is a cross-sectional spread in the model misspecification or model-predicted returns.

[2]Proposition O-3 can be generalized to allow for arbitrary correlation of the magnitude of alphas with the variance of idiosyncratic returns (e.g., in the case of the idiosyncratic volatility puzzle). Such a generalization means the first term is no longer the unconditional alpha of the B firms but this term remains constant across all portfolios and is removed by the first differencing that occurs when forming a long-short portfolio.

D. Sorting on other return components isolates those components

We focus on the properties of sorting on idiosyncratic returns as the long-short alphas from these sorts allow us to measure delays in information diffusion. Similar results, however, go through when one sorts instead on the alphas of economically linked stocks or on their modeled risks, i.e., $\vec{\beta}'\vec{f}$. For example, the long-short portfolio obtained from sorting on those alphas isolates the component due to correlated alphas.

Just as was the case with sorts based on idiosyncratic returns, e.g., Corollary O-3, the long-short alpha from sorting on alphas will be larger than the contribution of the alpha component in the long-short alpha from sorting on returns. Nevertheless, the magnitudes of the long-short alphas produced from these alternative sorts give an indication of the importance of the different components' contributions to the overall long-short alpha from return sorts. We, therefore, provide values of these long-short alphas in the empirical section of the paper as well as those from the idiosyncratic return sorts.

E. Sorting on estimated idiosyncratic returns

When the model parameters are unknown, we cannot observe the true idiosyncratic returns. Instead, they must be estimated with a regressio. This section shows how these estimated idiosyncratic returns can be used to measure the delay in information diffusion as well as if we knew the true idiosyncratic returns.

PROPOSITION O-4: *Let $B''(P)$ denote the portfolios of linked firms in set B formed using estimated idiosyncratic returns of the A firms $\hat{\epsilon}_{i,t}$. With the assumptions of Proposition O-3, if the estimation error in estimated idiosyncratic returns of the linked A firm, $(\hat{\epsilon}_{i,t} - \epsilon_{i,t})$, is independent of the predicted idiosyncratic returns of the B firms, $\epsilon_{j,t+\tau}$, for all $\tau \geq 1$, and the estimation error is unbiased, then the expected alpha of the portfolio $B''(P)$ can be decomposed into two pieces:*

$$E[\alpha_{B''(LS)}] = \varrho''_{\epsilon,N} - \varrho''_{\epsilon,1} \tag{O-14}$$

where $\varrho''_{\epsilon,P}$ is due to slow information diffusion and is increasing in P when $\bar{\rho}_\epsilon > 0$.

Thus, the long-short alpha from sorting on estimated idiosyncratic returns solely reflects delays in information diffusion. The conditions of Proposition O-4 are easily met. Standard linear regressions give unbiased parameter estimates. The estimation error will be uncorrelated whenever the estimation window is outside the period of time for which there is a delay in information diffusion, e.g., skipping one period between estimating and predicting when information is delayed one period.[3]

Though the long-short alpha obtained from sorting on estimated idiosyncratic returns reflects only delays in information diffusion, estimation noise makes this alpha smaller than that from sorting on known idiosyncratic returns. This attenuation is characterized completely in terms of the length of the estimation window and number of factors in the model:

PROPOSITION O-5: *Following the assumptions of Proposition O-4, if the factor returns are iid, then the ratio of the expected long-short alpha from sorting on estimated idiosyncratic returns (Equation* (O-14)*) to the expected long-short alpha from sorting on known idiosyncratic returns (Equation* (O-12)*) is*

$$\frac{\varrho''_{\epsilon,N} - \varrho''_{\epsilon,1}}{\varrho'_{\epsilon,N} - \varrho'_{\epsilon,1}} = \left(1 + \frac{1}{n}\right)^{-\frac{1}{2}} \frac{B\left[\frac{n-m+1}{2}, \frac{m}{2}\right]}{B\left[\frac{n-m}{2}, \frac{m}{2}\right]} \tag{O-15}$$

where B is the Euler beta function, n is the number of periods in the estimation window and m is the number of factors in the model.

The proof of this proposition is in Appendix O.IV.[4] This ratio can be used to to eliminate the attenuation. Applying this correction to the long-short alphas obtained from sorting on estimated idiosyncratic returns yields an unbiased estimate of the expected long-short alpha that would be obtained if we knew and sorted on the actual idiosyncratic returns.[5]

---

[3]See Online Appendix O.VI for a further discussion of this issue.

[4]We thank Andy Siegel and Raymond Kan for their help in the proof of this proposition.

[5]When estimating the model with daily data but sorting at the lower monthly frequency, the ratio is slightly different because of the correlation in estimation error induced from using the same model estimate throughout the entire sorting month. The ratio for this case is presented in Online Appendix O.VIII.

This correction is multiplicative and a constant given the estimation window length and number of factors. Therefore, it also applies to the standard errors of the long-short alphas obtained from sorting on estimated idiosyncratic returns. Thus the t-statistics of the estimates of the alphas obtainable from sorting on known idiosyncratic returns are the same as the t-statistics of the estimates of the long-short alpha obtained from sorting on estimated idiosyncratic returns.

O.III. Proofs and additional details

Proof of Proposition O-1. Portfolio $B(P)$ is formed of firms in set B linked to firms in set A that, at time t, fall into the Pth highest quantile of returns. Let $\iota_{j,t,P}$ be an indicator function of whether firm j is included in portfolio P at time t. In the case of inclusion, it has value 1. It has value 0 otherwise.

Let the portfolio weights each period be denoted $w_{j,t,P}$. Which, in the case of equal weighting,

$$w_{j,t,P} = \frac{\iota_{j,t,P}}{\sum_{k\in B}\iota_{k,t,P}}. \tag{O-16}$$

And, in the case of value weighting,

$$w_{j,t,P} = \frac{\iota_{j,t,P}P_{j,t}}{\sum_{k\in B}\iota_{k,t,P}P_{k,t}}, \tag{O-17}$$

where $P_{k,t}$ is the market value of firm k at time t.

Let the sample window of sorted B firm portfolios run from $t=1$ to T. Then the beta estimate for portfolio $B(P)$ is $\hat{\vec{\beta}}_{B(P)}$. Notice by the independence of the factors over time that this value is the time series average of the weighted average of the stocks in portfolio $B(P)$. That is

$$\hat{\vec{\beta}}_{B(P)} = \frac{1}{T}\sum_{\tau=1}^{T}\sum_{j\in B} w_{j,\tau-1,P}\beta_j. \tag{O-18}$$

Notice that the weights are shifted back by one period because the sorting done at $t-1$

determines the weights for returns realized at t.

When we take the expected value of the beta estimate we can eliminate the summation across time because the weights are iid across time from the fact that the factors are iid and the idiosyncratic returns of the A firms are iid across time. Notice however the weights are not independent of the betas because of the correlation of betas across economically linked firms. This gives

$$E[\hat{\vec{\beta}}_{B(P)}] = \sum_{j \in B} \beta_j E[w_{j,\tau-1,P}]. \tag{O-19}$$

Now we turn to calculating the expected value of the alpha of a portfolio $B(P)$.

$$E[\alpha_{B(P)}] \equiv E\left[\frac{1}{T}\sum_{\tau=1}^{T} r_{B(P),\tau} - \hat{\vec{\beta}}'_{B(P)}\frac{1}{T}\sum_{\tau=1}^{T}\vec{f}_\tau\right] \tag{O-20}$$

We now decompose the portfolio return into its constituent securities and then its alpha, beta and idiosyncratic return components.

$$= E\left[\frac{1}{T}\sum_{\tau=1}^{T}\sum_{j\in B} w_{j,\tau-1,P} r_{j,\tau} - \hat{\vec{\beta}}'_{B(P)}\frac{1}{T}\sum_{\tau=1}^{T}\vec{f}_\tau\right] \tag{O-21}$$

$$\begin{aligned} &= E\left[\frac{1}{T}\sum_{\tau=1}^{T}\sum_{j\in B} w_{j,\tau-1,P}\alpha_j + \frac{1}{T}\sum_{\tau=1}^{T}\sum_{j\in B} w_{j,\tau-1,P}\epsilon_{j,\tau}\right. \\ &\quad \left. + \frac{1}{T}\sum_{\tau=1}^{T}\sum_{j\in B} w_{j,\tau-1,P}\beta'_j\vec{f}_\tau - \hat{\vec{\beta}}'_{B(P)}\frac{1}{T}\sum_{\tau=1}^{T}\vec{f}_\tau\right] \end{aligned} \tag{O-22}$$

By the factors and idiosyncratic returns of the A firms being iid the third term becomes

$$E\left[\frac{1}{T}\sum_{\tau=1}^{T}\sum_{j\in B} w_{j,\tau-1,P}\beta'_j\vec{f}_\tau\right] = \sum_{j\in B}\beta_j E[w_{j,\tau-1,P}]E[f_\tau]. \tag{O-23}$$

Similarly by the factors being iid and the findings in Equation (O-19) the fourth term becomes

$$E\left[\hat{\vec{\beta}}'_{B(P)}\sum_{\tau=1}^{T}\vec{f}_\tau\right] = \sum_{j\in B}\beta_j E[w_{j,\tau-1,P}]E[f_\tau]. \tag{O-24}$$

Thus the last two terms cancel leaving the expected alpha only dependent upon the alpha and idiosyncratic components of the returns. We can now use the correlations of the alphas and epsilons across economically linked firms to replace these values with the values of the linked A firms.

Exploiting the economic link in alphas the first term of Equation (O-22) becomes

$$E\left[\frac{1}{T}\sum_{\tau=1}^{T}\sum_{j\in B} w_{j,\tau-1,P}\alpha_j\right] = E\left[\frac{1}{T}\sum_{\tau=1}^{T}\sum_{j\in B} w_{j,\tau-1,P}(\rho_\alpha\alpha_i + \eta_{\alpha,i,j})\right] \tag{O-25}$$

We abuse notation slightly throughout the appendix by not making the i subscript a function of j since which of the alphas of the A firms denoted with i subscripts to sum over are determined by their linkage to the B firms indexed by j. Taking advantage of the iid nature of the factors and the idiosyncratic returns of the A firms, we can move the expectation inside the time summation and remove the time summation. Additionally, the independence of the weights from the η shock makes the expectation of that term zero.

$$= \rho_\alpha \sum_{j\in B} E[w_{j,\tau-1,P}\,\alpha_i] \tag{O-26}$$

This is the term $\varrho_{\alpha,P}$ in Equation (O-9) and the decomposition specified in Equation (O-6). To see that this term is increasing in P notice that we can rewrite the summation over j in terms of the cross-sectional expected value, denoted with subscript CS and where the number of linked firms in set B is N_L.

$$= \rho_\alpha E_{CS}\Big[E[w_{j,\tau-1,P}\alpha_i]\Big] \tag{O-27}$$

$$= \rho_\alpha N_L\Big[E_{CS}\big[E[w_{j,\tau-1,P}]\big]\,E_{CS}[\alpha_i] + \mathrm{cov}_{CS}\,\big[E[w_{j,\tau-1,P}],\,\alpha_i\big]\Big]. \tag{O-28}$$

Noticing that $N_L E_{CS}\big[E[w_{j,t,P}]\big] = 1$ gives

$$= \rho_\alpha E_{CS}[\alpha_i] + \rho_\alpha N_L \ \mathrm{cov}_{CS}\left[E[w_{j,t,P}], \alpha_i\right]. \tag{O-29}$$

Because a higher alpha of an A firm increases the chances that firm falls in a higher portfolio ranking, this last covariance term is increasing in portfolio P giving the desired result.

Noticing the second term of Equation (O-22) depends entirely on the slow diffusion of information, we can make a similar argument for that used for the alpha term to obtain

$$E\left[\frac{1}{T}\sum_{\tau=1}^{T}\sum_{j\in B} w_{j,\tau-1,P}\epsilon_{j,\tau}\right] = \bar{\rho}_\epsilon \sum_{j\in B} E[w_{j,\tau-1,P}\epsilon_{i,\tau-1}] = \varrho_{\epsilon,P} \tag{O-30}$$

where $\bar{\rho}_\epsilon$ is the weighted average of the $\rho_{\epsilon,i,j}$.

$\varrho_{\epsilon,P}$ is increasing in P because a larger idiosyncratic return of an A firm increases the chances a firm is sorted into a higher portfolio, creating a positive covariance between the portfolio weights and the idiosyncratic returns. $\square$

The proof of Corollary 1 is immediate.

Proof of Proposition O-2. The proof is the same as that of Proposition O-1 except that we can no longer move the expectation inside the time series sum because the factors need not be iid, thus the weights need not be iid. Nevertheless we are able to make the same decompositions of the first and second terms of Equation (O-22) into the correlation coefficients and their respective conditional expected values. Furthermore we are able to label these ϱ_α and ϱ_ϵ which are both increasing in P because the ability to reexpress these expected values in terms of cross-sectional expected values where we notice the covariance between the alphas and idiosyncratic returns of the A firms and the portfolio weights increases in P.

The other difference from the proof of Proposition O-1 is that the estimated beta, $\hat{\vec{\beta}}_{B(P)}$, is not necessarily the time series average of the weighted average of betas of the securities that make up the B portfolios. This is because the factors no longer need be iid. Without this result, the third and forth terms of Equation (O-22) need not be equal. This gives the desired third term to the expected long short alpha which we label $\varrho_{\beta' f}$. □

Proof of Proposition O-3. Portfolio $B'(P)$ is formed of firms in set B linked to firms in set A that in the previous period fell into the Pth highest quantile of news shocks $\epsilon_{i,t}$. Let $\iota'_{j,t,p}$ be an indicator function of whether firm j is included in portfolio P at time t. In the case of inclusion, it has value 1. It has value 0 otherwise. Define the weights $w'_{j,\tau,P}$ similarly to those in the proof of Proposition O-1 except using the new indicator function. Substituting these weights in we can follow that proof until the decomposition in Equation (O-22).

Now the first term of that decomposition which includes the alphas of the B firms is equal to the unconditional average of the B firm alphas because the weights are independent of the alphas.

The second term which involves the idiosyncratic returns continues to have similar properties as in the previous two proofs because using the idiosyncratic returns of the A firms continues to induce a covariance between the portfolio weights and the idiosyncratic returns of the A firms that is increasing in P. We label this term

$$\varrho'_{\epsilon,P} = \bar{\rho}_\epsilon E[\epsilon_{A(P,\epsilon)}] \tag{O-31}$$

which has the decomposition into the correlation coefficient and the conditional expected value as desired.

Though the factors are not necessarily iid over time, the portfolio weights are independents of the betas of the B firms because the sorts occur on the idiosyncratic returns of the A firms which are independent of the B firm betas. This independence of the weights and the betas makes the betas independent of the factors which together makes the third and forth terms of Equation (O-22) equal to

$$\bar{\vec{\beta}}'_B E\left[\frac{1}{T}\sum_{\tau=1}^{T} f_\tau\right] \tag{O-32}$$

where $\bar{\vec{\beta}}'_B$ is the unconditional average of the B firm betas. Since these two terms have equal value they cancel, giving the desired result. $\square$

Proof of Corollary O-3. Recall that $\varrho_{\epsilon,N} - \varrho_{\epsilon,1}$ is increasing in the cross-section variance of the alpha and model predicted returns. Then notice that $\varrho'_{\epsilon,N} - \varrho'_{\epsilon,1}$ is the same as the degenerate case where these cross-sectional variances go to zero. $\square$

Additional details on $\varrho_{\epsilon,N} - \varrho_{\epsilon,1}$ and $\varrho'_{\epsilon,N} - \varrho'_{\epsilon,1}$ under the assumption of iid factors can be found in Online Appendix Equations (O-64) and (O-67).

Proof of Proposition O-4. The proof follows that of Proposition O-3. The unbiasedness of the estimates continues to generate the independence of the weights from the the alphas and betas of the B firms in the sorted portfolios The uncorrelated measurement error of the alphas and betas gives that the idiosyncratic returns of the B firms in the sorted portfolios depend only on the idiosyncratic return of the A firms at the sort date. $\square$

O.IV. Correction ratio for estimation error bias

In this section, we provide a proof of Proposition O-5. For the case of parameter estimation at a higher frequency than the sort frequency, the correction is slightly different. See Section O.VIII for details on that correction.

Without loss of generality, the factors are demeaned throughout this appendix. Otherwise we continue with the assumptions from Proposition O-5. Standard deviations of the idiosyncratic returns cannot be negative, so we use a generic distribution for these values where all possible values are positive. The set of all possible standard deviations for stocks in set A is $\mathbb{S}$. The probability of any $\sigma_{\epsilon,i} \in \mathbb{S}$ is $Pr(\sigma_{\epsilon,i})$. There may be infinitely many of these possibilities, but for simplicity we assume a finite number of these. Unless otherwise noted ϕ_x and Φ_x are the PDF and CDF of the normal random variable x.

We now calculate the ratio of the long-short alpha obtained from sorting on idiosyncratic news returns (Equation (O-14) and expanded in the Online Appendix to Equation (O-74)) to that obtained from sorting on actual idiosyncratic returns (Equation (O-12) and expanded in the Online Appendix to Equation (O-67)):

$$\frac{\varrho''_{\epsilon,N} - \varrho''_{\epsilon,1}}{\varrho'_{\epsilon,N} - \varrho'_{\epsilon,1}} = \frac{\bar{\rho}_\epsilon \Big[E[\epsilon_{A(N,\hat{\epsilon})}] - E[\epsilon_{A(1,\hat{\epsilon})}] \Big]}{\bar{\rho}_\epsilon \left[E[\epsilon_{A(N,\epsilon)}] - E[\epsilon_{A(1,\epsilon)}] \right]} \tag{O-33}$$

Notice that the the $\bar{\rho}_\epsilon$ cancels, meaning we only need to calculate the expected idiosyncratic returns in the A firms. Closed-form solutions for the conditional expectations in Equation (O-33) are in Online Appendix O.V.

We begin with the special case of a *single* idiosyncratic variance in the cross-section, conditional on a given estimation sample. We then generalize this to an arbitrary distribution of cross-sectional idiosyncratic variances, obtaining the same ratio as in the special case. Finally, we show how to take the expectation across all estimation samples.

A. *Special case: single idiosyncratic variance*

Consider the case of a single idiosyncratic return variance, σ_i, in the cross-section and a single history of factor draws of length n used to estimate the asset pricing model for all stocks. This history of factor draws can be put in the matrix X.

The estimated idiosyncratic returns are

$$\hat{\epsilon}_i = \epsilon_i + \epsilon_i^* \tag{O-34}$$

where ϵ_i^* is the estimation noise.

As we discuss further in Online Appendix Section O.V.D, the estimated idiosyncratic returns used for sorting have mean zero and variance

$$\sigma^2(\hat{\epsilon}_i) = \sigma_i^2 \left(1 + \underbrace{\begin{bmatrix} 1 \\ \vec{f}_t \end{bmatrix}' (X'X)^{-1} \begin{bmatrix} 1 \\ \vec{f}_t \end{bmatrix}}_{\text{contribution from } \epsilon_i^*} \right) \tag{O-35}$$

Equation (O-35) accounts for variance from the true idiosyncratic returns. It also accounts for the contribution from the error in the measured alpha and the error in the measured beta scaled by the factor draw at t used for the estimation. (See Equation (O-69) for the distribution of these measurement errors.)

Thus we see the variance from the measurement error, ϵ_i^*, scales with the variance of the idiosyncratic return. This scaling is invariant to the variance of the idiosyncratic returns. It is also independent of the distribution of the alphas, betas or correlation of the alphas or idiosyncratic returns across linked stocks.

Define this scaling factor as

$$s(\vec{f}, X'X) \equiv \begin{bmatrix} 1 \\ \vec{f}_t \end{bmatrix}' (X'X)^{-1} \begin{bmatrix} 1 \\ \vec{f}_t \end{bmatrix}. \tag{O-36}$$

Another critical feature to notice is that the breakpoints for the top (or bottom quantile) of a normal distribution can be written in the form $\bar{b}_N \sigma$ where σ is the standard deviation of the normal distribution and $\bar{b}_N$ is a function of the quantile size.[6] Thus the breakpoints for the sorts on actual versus estimated idiosyncratic returns only differ by the scale factor described.

We can now write the bias ratio (the special case of Equation (O-33)) explicitly in terms of the integrals representing the expectations. The numerator and denominator are simplified version of Equations (O-74) and (O-67). We denote the special case of single idiosyncratic

[6] These properties extend to the breakpoints of intermediate quantiles as well.

returns and specific estimation windows, etc., with additional subscripts. Explicit notation of the dependence of $\sigma(\epsilon_i^*)$ on $\vec{f}$ and X is suppressed for clarity.

$$\frac{E[\epsilon_{A(N,\hat{\epsilon},\sigma_i,X,\vec{f})}] - E[\epsilon_{A(1,\hat{\epsilon},\sigma_i,X,\vec{f})}]}{E[\epsilon_{A(N,\epsilon,\sigma_i,X,\vec{f})}] - E[\epsilon_{A(1,\epsilon,\sigma_i,X,\vec{f})}]}$$

$$= \frac{\int_{-\infty}^{\infty}\left[\int_{\bar{b}\,\sigma(\hat{\epsilon}_i)-\epsilon_i^*}^{\infty} \epsilon_i\phi_{\epsilon_i}(\epsilon_i)\,d\epsilon_i\right]\phi_{\epsilon_i^*}(\epsilon_i^*)\,d\epsilon_i^* - \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\underline{b}\,\sigma(\hat{\epsilon}_i)-\epsilon_i^*} \epsilon_i\phi_{\epsilon_i}(\epsilon_i)\,d\epsilon_i\right]\phi_{\epsilon_i^*}(\epsilon_i^*)\,d\epsilon_i^*}{\int_{\bar{b}\,\sigma(\epsilon_i)}^{\infty}\epsilon_i\phi_{\epsilon_i}(\epsilon_i)\,d\epsilon_i - \int_{-\infty}^{\underline{b}\,\sigma(\epsilon_i)}\epsilon_i\phi_{\epsilon_i}(\epsilon_i)\,d\epsilon_i} \tag{O-37}$$

This simplifies to

$$= \frac{1}{\sqrt{1 + s(\vec{f}, X'X)}}. \tag{O-38}$$

Moreover, this ratio is the same for any individual quantile.

B. *General case: arbitrary distribution of idiosyncratic variances*

To extend this ratio to a cross-section of stocks with an arbitrary distribution of idiosyncratic variances, first note that the breakpoints obtained from equations (O-66) and (O-70) scale up and down with the variance, as described in the last subsection, simply with a different constant $\bar{b}^*$.

We now calculate the ratio for the top quantile (again suppressing the dependence of the noise variance on $\vec{f}$ and X)

$$\frac{E[\epsilon_{A(N,\hat{\epsilon},X,\vec{f})}]}{E[\epsilon_{A(N,\epsilon,X,\vec{f})}]} = \frac{\sum_{\mathbb{S}}\int_{-\infty}^{\infty}\left[\int_{\bar{b}^*\,\sigma(\hat{\epsilon}_i)-\epsilon_i^*}^{\infty}\epsilon_i\phi_{\epsilon_i}(\epsilon_i)\,d\epsilon_i\right]\phi_{\epsilon_i^*}(\epsilon_i^*)\,d\epsilon_i^*\,Pr(\sigma_{\epsilon,i})}{\sum_{\mathbb{S}}\int_{\bar{b}^*\,\sigma(\epsilon_i)}^{\infty}\epsilon_i\phi_{\epsilon_i}(\epsilon_i)\,d\epsilon_i\,Pr(\sigma_{\epsilon,i})} \tag{O-39}$$

We re-write each term in the numerator's sum as a product of the equivalent term in the denominator and the ratio from the last subsection

$$= \frac{\sum_{\mathbb{S}} \int_{\bar{b}^* \sigma(\epsilon_i)}^{\infty} \epsilon_i \phi_{\epsilon_i}(\epsilon_i)\, d\epsilon_i \left(\frac{1}{\sqrt{1+s(\vec{f}, X'X)}} \right) Pr(\sigma_{\epsilon,i})}{\sum_{\mathbb{S}} \int_{\bar{b}^* \sigma(\epsilon_i)}^{\infty} \epsilon_i \phi_{\epsilon_i}(\epsilon_i)\, d\epsilon_i\, Pr(\sigma_{\epsilon,i})}. \tag{O-40}$$

This ratio is constant across all terms allowing simplification to

$$= \frac{1}{\sqrt{1 + s(\vec{f}, X'X)}}. \tag{O-41}$$

The ratio is the same for all quantiles and hence the long-short portfolio as well.

For the generalization to the value-weighted case, one merely puts the market weights of the B firms in the appropriate places in the breakpoint calculations, etc. In the value-weighted case we obtain the same ratio.

C. Expectation across estimation windows

The last subsection shows the expected bias conditional on a current factor draw $\vec{f}$ and the series of n factor draws (embodied in X) that are used to estimate the model. We now take the expectation across the current and past factor draws to obtain the unconditional bias ratio:

$$\frac{E[\epsilon_{A(N,\hat{\epsilon})}] - E[\epsilon_{A(1,\hat{\epsilon})}]}{E[\epsilon_{A(N,\epsilon)}] - E[\epsilon_{A(1,\epsilon)}]} = \frac{\varrho''_{\epsilon,N} - \varrho''_{\epsilon,1}}{\varrho'_{\epsilon,N} - \varrho'_{\epsilon,1}} = E\left[\frac{1}{\sqrt{1 + s(\vec{f}, X'X)}} \right]. \tag{O-42}$$

Recall we demeaned f to give it the distribution

$$\vec{f} \sim N(0, \Sigma_f). \tag{O-43}$$

$X'X$ is of the following form

$$X'X = \begin{bmatrix} n & n\,\bar{f}' \\ n\,\bar{f} & n\,V + n\bar{f}\bar{f}' \end{bmatrix} \tag{O-44}$$

where V and $\bar{f}$ are independent random variables and nV is Wishart with $n-1$ degrees of freedom and uses the variance covariance matrix of the factors Σ_f and

$$\bar{f} \sim N\left(0, \frac{\Sigma_f}{n}\right) \tag{O-45}$$

It can be trivially shown by decomposing Σ_F via the Cholesky decomposition and substituting out to replace it with the identity matrix of the same size that the expectation depends only on the number of factors in the asset pricing model and the estimation window length. We thank Andy Siegel for his insights on characterizing the distribution of $X'X$.

$$\frac{\varrho''_{\epsilon,N} - \varrho''_{\epsilon,1}}{\varrho'_{\epsilon,N} - \varrho'_{\epsilon,1}} = E\left[\frac{1}{\sqrt{1 + s(\vec{z}, Z'Z)}}\right] \tag{O-46}$$

where

$$s(\vec{z}, Z'Z) = \begin{bmatrix} 1 \\ \vec{z} \end{bmatrix}' (Z'Z)^{-1} \begin{bmatrix} 1 \\ \vec{z} \end{bmatrix} \tag{O-47}$$

and $Z'Z$ is of form

$$Z'Z = \begin{bmatrix} n & n\,\bar{z}' \\ n\,\bar{z} & W + n\bar{z}\bar{z}' \end{bmatrix} \tag{O-48}$$

The random variables are

$$\vec{z} \sim N(0_m, I(m)), \tag{O-49}$$

where $I(x)$ is the identity matrix of size x, m is the number of factors in the linear factor model, and n is the size of the estimation window. W, $\vec{z}$ and $\bar{z}$ are independent random variables and W is Wishart with $n-1$ degrees of freedom and uses the variance-covariance matrix $I(m)$ and

$$\bar{z} \sim N\left(0, \frac{I(m)}{n}\right). \tag{O-50}$$

This expectation can further be simplified for direct calculation in terms of the Euler beta function as follows, for which we thank Raymond Kan. Using the partitioned matrix inverse formula we obtain

$$1 + s(\vec{z}, Z'Z) = 1 + \frac{1}{n} + (\bar{z} - \vec{z})'W^{-1}(\bar{z} - \vec{z}). \tag{O-51}$$

Noting that

$$y \equiv \bar{z} - \vec{z} \sim N\left(0_m, \left(1 + \frac{1}{n}\right) I_m\right) \tag{O-52}$$

which is independent of W; defining

$$u \equiv \frac{y'y}{1 + \frac{1}{n}} \sim \chi^2_m; \tag{O-53}$$

and the fact that (Muirhead (1982), Theorem 3.2.12)

$$v \equiv \frac{y'y}{y'W^{-1}y} \sim \chi^2_{n-m} \tag{O-54}$$

which itself is independent of y (and hence u) we can write

$$1 + s(\vec{z}, Z'Z) = \left(1 + \frac{1}{n}\right)\left(1 + \frac{u}{v}\right). \tag{O-55}$$

Therefore,

$$\frac{1}{\sqrt{1 + s(\vec{z}, Z'Z)}} = \left(1 + \frac{1}{n}\right)^{-\frac{1}{2}} b^{-\frac{1}{2}}, \tag{O-56}$$

where

$$b = \frac{v}{v + u} \sim Beta\left(\frac{n-m}{2}, \frac{m}{2}\right). \tag{O-57}$$

It is straightforward to show that

$$E\left[b^{-\frac{1}{2}}\right] = \frac{B\left(\frac{n-m+1}{2}, \frac{m}{2}\right)}{B\left(\frac{n-m}{2}, \frac{m}{2}\right)}. \tag{O-58}$$

Together this gives the Equation (O-15) in Proposition O-5

O.V. Closed form solutions

A. Overview

This section details the closed form solutions for various components of the long-short portfolio alpha of stocks ($j \in B$) sorted into N-quantiles based on returns, idiosyncratic returns, estimated idiosyncratic returns or estimated alphas of economically linked stocks ($i \in A$). These components are used in the derivation of the estimation error bias ratio correction presented in Appendix O.IV.

We specifically consider the case of a one-to-one link between stocks in sets A and B. This is without loss of generality equivalent to the stock in set A being a portfolio of stocks. Also for simplicity of notation we consider only the equal weighted case of the B stocks. This trivially extends to any arbitrary weighting of the B stocks.

We provide expanded notation for distributions, etc., consistent with the assumption within the main text.

Distributional assumptions: Unless otherwise noted, independence denotes both within a given stock and across stocks within and across sets A and B. The alphas and betas are independent with set A. We allow for correlation among the beta exposures within a given stock. The idiosyncratic returns realizations and their standard deviations are independent of the alphas and betas of stocks in set A.[7]

[7]These independence assumptions can be slackened without changing the results. In particular, the alphas and beta may be jointly normal and the variance of the idiosyncratic returns can be correlated with the alphas and betas, such as in the case of the idiosyncratic volatility puzzle. We retain these independence

The cross-sectional distribution of the alphas and betas in set A are

$$\alpha_i \sim N(\mu_\alpha, \sigma_\alpha^2) \text{ and} \tag{O-59}$$

$$\vec{\beta}_i \sim N(\vec{\mu}_\beta, \Sigma_\beta). \tag{O-60}$$

Standard deviations of the idiosyncratic returns cannot be negative, so we use a generic distribution for these values where all possible values are positive. The set of all possible standard deviations for stocks in set A is $\mathbb{S}$. The probability of any $\sigma_{\epsilon,i} \in \mathbb{S}$ is $Pr(\sigma_{\epsilon,i})$. There may be infinitely many of these possibilities, but for simplicity we assume a finite number of these.

As is standard, the factor realizations are independent of the alphas, betas, idiosyncratic returns and standard deviations of the idiosyncratic returns. The factor and idiosyncratic returns are independent across time. The time series distributions of the factor and idiosyncratic draws for stocks in set A are

$$\vec{f}_t \sim N(\vec{\mu}_f, \Sigma_f) \text{ and} \tag{O-61}$$

$$\epsilon_{i,t} \sim N(0, \sigma_{\epsilon,i}^2). \tag{O-62}$$

Unless otherwise noted ϕ_x and Φ_x are the PDF and CDF of the normal random variable x.

Asset pricing model estimation: Estimated values of the factor model of stock i necessary for extracting its date t *estimated* idiosyncratic returns occur via a linear regression. The regression uses data over the period $(t-n, t-1)$. This idiosyncratic return to stocks in A is for the sort occurring at time t and is used to predicted time $t+1$ returns of the economically linked B stocks.

assumptions for simplicity of notation.

i. Key statistics

We calculate the B firms' expected long-short portfolio alphas for sorts based on the A stocks' (1) returns, (2) idiosyncratic returns, (3) estimated idiosyncratic returns, (4) alphas and (5) estimated alphas.

We decompose these long-short portfolios into the components due to the correlation in alphas across stocks in sets A and B (Equation (O-2)) and the delayed information incorporation (Equation (O-4)). This decomposition uses the expected alpha and idiosyncratic returns of stocks included in the top and bottom quantiles of the A stocks. These are symmetric so we only explicitly show the top quantile calculations.

ii. Intuition

The expected values of the alpha and idiosyncratic returns depend upon the probability that stocks with different alphas or idiosyncratic returns fall in the top quantile. The probability of falling in the top quantile at any date depends on a stock's own alpha, beta and idiosyncratic return. The cross-sectional distribution of all stocks at that date matters as well, because it determines the break points for the sorts.

For sorts based on returns or estimated idiosyncratic returns, the cross-sectional distribution changes each date with the factor realizations. The factor realization increases or decreases the spread in the cross-sectional distribution, because these factor realizations are included in the sorting variable through multiplication with either beta or estimation error in beta.

There exists more than one idiosyncratic return variance in the cross-section. Multiple idiosyncratic variances makes the cross-sectional distribution a mixture of normal distributions: one for each idiosyncratic return variance. The mixing weights for these distribution is the prevalence (or probability) of each idiosyncratic return variance.

These two facts require that we first calculate conditional on factor draws and stock specific idiosyncratic return draws. Then we integrate the against the entire cross-section and time series distributions to compute the unconditional expected values of interest.

B. *Sorting on returns*

In this subsection, we calculate the $t+1$ expected long-short alpha of the B stocks formed by sorting on A stocks' time t returns.

i. Break points

The break point for the top quantile N of returns conditional on the factor draw $\vec{f}$ is the value $\bar{c}(R, \vec{f})$ that yields enough stocks in the tops of each of the component normal distributions, weighted by the size of each distribution (the probability of the given idiosyncratic return standard deviation), to add up to the desired quantile size. That is $\bar{c}(R, \vec{f})$ satisfies

$$1 - \frac{1}{N} = \sum_{\sigma_{\epsilon,i} \in \mathbb{S}} \Phi_{R,\vec{f},\sigma_{\epsilon,i}} \left(\bar{c}(R, \vec{f}) \right) Pr(\sigma_{\epsilon,i}) \tag{O-63}$$

where $\Phi_{R,\vec{f},\sigma_{\epsilon,i}}$ is the CDF of normal distribution with mean $\mu_\alpha + \vec{\mu}'_\beta \vec{f}$ and variance $\sigma^2_\alpha + \vec{f}' \Sigma_\beta \vec{f} + \sigma^2_{\epsilon,i}$.

ii. Expected idiosyncratic return

To calculate the expected idiosyncratic return of a stock included in the top quantile, we first integrate the idiosyncratic return, for a given variance, between infinity and the break point adjusted for a given alpha, beta and factor factor realization. Then we integrate over the alphas, betas, and factor realizations and sum across the idiosyncratic return variance realizations. Finally, we divide by the total weight in the portfolio which is $\frac{1}{N}$:

$$E[\epsilon_{A(N,R)}] = \frac{1}{N}\int_{\mathbb{F}}\sum_{\mathbb{S}}\Big[\int_{\mathbb{A}}\int_{\mathbb{B}}\Big(\int_{\bar{c}(R,\vec{f})-\alpha-\vec{\beta}'\vec{f}}^{\infty}\epsilon_i\phi_{\sigma_{\epsilon,i}}(\epsilon_i)\,d\epsilon_i\Big) * \phi_{\vec{\beta}}(\vec{\beta})\phi_{\alpha}(\alpha)\,d\vec{\beta}\,d\alpha\,Pr(\sigma_{\epsilon,i})\Big]\phi_{\vec{f}}(\vec{f})\,d\vec{f} \tag{O-64}$$

where $\mathbb{F}$, $\mathbb{A}$ and $\mathbb{B}$ are the domains of all possible $\vec{f}$, α and $\vec{\beta}$ and where the subscript $A(N,R)$ denotes that this is the alpha of A stocks in quantile N when sorting on returns R.

From Equation (O-64) we see how higher idiosyncratic return values make landing in the top portfolio larger. Thus the expected value of the idiosyncratic return in the Nth quantile is larger than that in the 1st quantile.

iii. Expected alpha

To calculate the expected alpha of stock included in the top quantile, we first integrate the alpha, for a given variance, between infinity and the break point adjusted for a given idiosyncratic return, beta and factor factor realization. Then we integrate against the idiosyncratic returns, betas, and factor realizations and sum against the idiosyncratic return variance realizations. Finally, we divide by the total weight in the portfolio which is $\frac{1}{N}$:

$$E[\alpha_{A(N,R)}] = \frac{1}{N}\sum_{\mathbb{S}}\int_{\mathbb{F}}\int_{\mathbb{E}_i}\int_{\mathbb{B}}\Big[\int_{\bar{c}(R,\vec{f})-\vec{\beta}'\vec{f}-\epsilon_i}^{\infty}\alpha\,\phi_{\alpha}(\alpha)\,d\alpha\Big] * \phi_{\vec{\beta}}(\vec{\beta})\phi_{\epsilon_i}(\epsilon_i)\phi_{\vec{f}}(\vec{f})\,d\vec{\beta}\,d\epsilon_i\,d\vec{f}\,P(\sigma_{\epsilon,i}) \tag{O-65}$$

where $\mathbb{F}$, $\mathbb{E}_i$, and $\mathbb{B}$ are the domains of all possible $\vec{f}$, ϵ_i, and $\vec{\beta}$.

From Equation (O-65) we see how a higher alpha raises the probability that a stock falls in the top portfolio. Therefore, the expected alpha in the Nth quantile is larger than that in the 1st quantile.

C. Sorting on idiosyncratic returns

When we sort on the idiosyncratic returns as opposed to returns, the break points for the expectations simplify. The factor realizations and the alpha and beta distributions have no influence on the portfolios to which stocks are assigned. Stocks are assigned purely on their idiosyncratic returns. We show that the break points are thus narrower, and expected idiosyncratic returns in the extreme portfolios are larger. Most importantly, we show that there is no expected alpha variation across the A stocks assigned to the extreme quantiles.

i. Break points

The break point for the top quantile N of idiosyncratic returns, is the value $\bar{c}(\epsilon)$ that yields enough stocks in the tops of each of the constituent distributions, weighted by the sizes of these distributions (the probability of the given idiosyncratic return standard deviation), to add up to the desired quantile size. That is, $\bar{c}(\epsilon)$ satisfies this equation

$$1 - \frac{1}{N} = \sum_{\sigma_i \in \mathbb{S}} \Phi_{\epsilon_i} \left(\bar{c}(\epsilon) \right) Pr(\sigma_{\epsilon,i}) \tag{O-66}$$

where Φ_{ϵ_i} is the CDF of the idiosyncratic return distribution with mean 0 and variance $\sigma^2_{\epsilon,i}$. Notice that $\bar{c}(\epsilon)$ is no longer a function of the factor realization.

ii. Expected idiosyncratic return

To calculate the expected idiosyncratic return of a stock included in the top quantile, we first integrate the idiosyncratic return, for given variance, between infinity and the break point. Then we sum against the idiosyncratic return variance realizations. Finally, we divide by the total weight in the portfolio which is $\frac{1}{N}$:

$$E[\epsilon_{A(N,\epsilon)}] = \frac{1}{N} \sum_{\mathbb{S}} \Big[\int_{\bar{c}(\epsilon)}^{\infty} \epsilon_i \phi_(\epsilon_i) \, d\epsilon_i \Big] Pr(\sigma_{\epsilon,i}) \tag{O-67}$$

where the subscript $A(N,\epsilon)$ denotes that this idiosyncratic return is for the A stocks in quantile N when sorted on idiosyncratic returns.

Equation (O-67) shows how higher idiosyncratic returns lead to a higher likelihood of inclusion in the the top quantile. Therefore the expected idiosyncratic return in the Nth quantile is higher than that in the 1st quantile.

iii. Expected alpha

Independence of the alpha distribution from the idiosyncratic returns implies a stock's alpha is unrelated to its inclusion in the top (or bottom quantiles).[8] Thus the expected alpha in all quantiles, q, is just the unconditional expected alpha:

$$E[\alpha_{A(q,\epsilon)}] = \mu_\alpha. \tag{O-68}$$

D. Sorting on estimated idiosyncratic returns

Sorting on estimated idiosyncratic returns rather than known idiosyncratic returns, the alphas and betas are replaced with measurement error that is mean zero. Because this measurement error is uncorrelated with the true alphas of the A stocks, it is also uncorrelated with the alphas of the B stocks. Also per Proposition O-4 this measurement error is uncorrelated with the idiosyncratic returns of the B firms. Thus, the measurement error has no effect on the expected alphas of the B stocks included in the quantile portfolios.

This measurement error, however, does affect the size of the idiosyncratic returns required for an A stock to be included in a quantile portfolio. Thus it affects the expected idiosyncratic return in the portfolios.

To calculate these effects of the measurement error, the conditional distribution of the alpha and beta measurement error replaces the unconditional distributions of alpha and

[8]This independence assumption can be slackened to allow correlation between the magnitude of alpha and the variance of idiosyncratic returns. Such correlation would result in a level shift of the expected alphas in both the long and short portfolios that would difference out in the long-short portfolio.

betas. The measurement error is denoted with stars and has the following joint normal distribution:

$$\begin{bmatrix} \alpha^* \\ \beta^* \end{bmatrix} \sim N\left(0, \sigma^2_{\epsilon,i}\left(X'X\right)^{-1}\right) \tag{O-69}$$

conditional on the idiosyncratic return variance $\sigma^2_{\epsilon,i}$ and X, which is the matrix of the past n factor draws augmented by a constant.

This measurement noise induces spread in the break points relative to the sorting on known idiosyncratic returns. This spread varies depending upon the draws of the factor over the past n periods when the model is estimated. Thus we must calculate the break points and the expected values conditional not only on the current factor return but also on the past n draws of the factors. For simplicity of notation, we will just keep track of these with the X matrix.[9]

Relative to sorting on *known* idiosyncratic returns, the expected idiosyncratic returns in the extreme portfolios are lower, because stocks are assigned to these portfolios based on the sum of their idiosyncratic return and measurement error.

i. Break points

The break point for the top quantile N of estimated idiosyncratic returns conditional on the current factor draw $\vec{f}$ and the past n factor draws embodied in X is the value $\bar{c}(\hat{\epsilon}, \vec{f}, X)$ that yields enough stocks in the tops of each of the component distributions, weighted by the size each distributions (the probability of the given idiosyncratic return standard deviation), to add up to the desired quantile size. That is $\bar{c}(\hat{\epsilon}, \vec{f}, X)$ satisfies

$$1 - \frac{1}{N} = \sum_{\sigma_{\epsilon,i} \in \mathbb{S}} \Phi_{\hat{\epsilon},\vec{f},X,\sigma_{\epsilon,i}}\left(\bar{c}(\hat{\epsilon}, \vec{f}, X)\right) Pr(\sigma_{\epsilon,i}). \tag{O-70}$$

[9]We discuss the distribution of X, and $X'X$ in Appendix O.IV.

$\Phi_{\hat{\epsilon},\vec{f},X,\sigma_{\epsilon,i}}$ is the CDF of normal distribution with mean 0 and variance

$$\sigma^2_{\epsilon,i}\left(1+[1,\vec{f}'](X'X)^{-1}\begin{bmatrix}1\\ \vec{f}\end{bmatrix}\right). \tag{O-71}$$

To see where this variance comes from recall that

$$\hat{\epsilon} = \epsilon + \alpha^* + \beta^{*\prime} f \tag{O-72}$$

$$= \epsilon + [\alpha^*,\ \beta^{*\prime}]\begin{bmatrix}1\\ f\end{bmatrix}. \tag{O-73}$$

Taking the variance of Equation (O-73) gives equation (O-71).

ii. Expected idiosyncratic return

To calculate the expected idiosyncratic return of a stock included in the top quantile, we first integrate the idiosyncratic return, for given variance, between infinity and the break point adjusted for a given alpha and beta measurement error draw, (their distributions are conditioned on the past n factor draws and idiosyncratic variance draw) and the current factor realization. Then we integrate against these measure error draws, current factor draw, past n factor draws and sum against the idiosyncratic return variance realizations. Finally, we divide by the total weight in the portfolio which is $\frac{1}{N}$. This gives

$$\begin{aligned}E[\epsilon_{A(N,\hat{\epsilon})}] =& \frac{1}{N}\sum_{\mathbb{S}}\int_{\mathbb{F}}\int_{\mathbb{X}}\int_{[\mathbb{A}*,\mathbb{B}^{*\prime}]'}\Big[\int_{\bar{c}(\hat{\epsilon},\vec{f},X)-\alpha^*-\vec{\beta}^{*\prime}\vec{f}}^{\infty}\epsilon_i\phi(\epsilon_i)\,d\epsilon_i\Big] \\ &*\phi_{[\alpha^*,\vec{\beta}^{*\prime}]'}\left(\begin{bmatrix}\alpha^*\\ \vec{\beta}^*\end{bmatrix}\right)\phi_{\vec{f}}(\vec{f})\phi_X(X)\,d\begin{bmatrix}\alpha^*\\ \vec{\beta}^*\end{bmatrix}\,d\vec{f}dX\,Pr(\sigma_{\epsilon,i})\end{aligned} \tag{O-74}$$

where $\mathbb{F}$, $\mathbb{X}$, $\mathbb{A}$, and $\mathbb{B}$ are the domains of all possible $\vec{f}$, X, α^*, and $\vec{\beta}^*$ and where the subscript $A(N,\hat{\epsilon})$ denotes that this is the alpha of A stocks in quantile N when sorting on the estimated idiosyncratic returns.

iii. Expected alpha

The measurement noise is unbiased and mean zero. Thus, although the measurement noise affects the probability of inclusion of a stock in a portfolio, the *actual* alphas are uncorrelated with this measurement noise.[10] Therefore the expected alpha in quantile q is the unconditional expected alpha:

$$E[\alpha_{A(q,\hat{\epsilon})}] = \mu_\alpha. \tag{O-75}$$

E. *Sorting on alphas*

Looking at the expected long-short alphas from sorting on alphas is particularly useful because if stocks were only sorted on the alphas of the A stocks, the expected alpha in the top portfolio comes from a truncated normal distribution with break point

$$\bar{c}(\alpha) = \Phi_\alpha^{-1}\left(1 - \frac{1}{N}\right). \tag{O-76}$$

Using this break point, the expected alpha is

$$E[\alpha_{A(N,\alpha)}] = \frac{\phi_\alpha\left(\bar{c}(\alpha)\right)\sigma_\alpha}{1 - \Phi_\alpha\left(\bar{c}(\alpha)\right)} + \mu_\alpha \tag{O-77}$$

where the subscript $A(N,\alpha)$ denotes that this idiosyncratic return is for the A stocks in quantile N when sorted on alphas. A similar calculation gives the expected alpha in the bottom portfolio. From this equation and the cross-sectional correlation in alphas between leading and lagging firms, we can calculate the expected long-short alpha from sorting on leaders alphas.

[10] Allowing the alphas to be correlated with the variance of the idiosyncratic returns would then create a correlation between between the measurement noise and alphas. The results go through with this slackened assumption, but we maintain the stronger assumption for simplicity of notation.

F. *Sorting on estimated alphas*

Sorting on estimated alphas rather than actual alphas requires conditioning on the history of factor draws that are used to estimate the model. The cross-section is a mixture of normal distributions where each distribution is determined by the variance of a stock's idiosyncratic return. Conditional on a past history of factor draws embodied in X and a stock's idiosyncratic return variance, the estimated alpha distribution is

$$\hat{\alpha} \sim N\left(\mu_\alpha,\, \sigma_\alpha^2 + \sigma_{\epsilon,i}^2 (X'X)^{-1}\right). \tag{O-78}$$

Because the measurement error in alpha is uncorrelated with the true alphas of the A stocks, it is also uncorrelated with the alphas of the B stocks and the realized idiosyncratic return of the A stocks in the sorting period.[11] Also per Proposition O-4 the measurement error in the alphas is independent of the idiosyncratic returns of the B firms. The measurement error, therefore, has no effect on the expected alphas of the B stocks included in the quantile portfolios. It also does not have an effect on the idiosyncratic returns of stocks included in the portfolios.

This measurement error, however, affects the size of the alphas required for an A stock to be included in a quantile portfolio. This measurement noise induces a spread in the break points relative to the sorting on known alphas. Relative to sorting on known alphas, the expected alphas in the extreme portfolios are lower, because stocks are assigned to these portfolios based on the sum of their alpha and measurement error.

i. Break points

The break point for the top quantile N of returns conditional on the the past n factor draws embodied in X is the value $\bar{c}(\hat{\alpha}, X)$ that yields enough stocks in the tops of each of the component distributions, weighted by the size of each distribution (the probability of the

[11] This result also generalizes the case of correlation between the alphas and the variance of the idiosyncratic returns.

given idiosyncratic return standard deviation), to add up to the desired quantile size. That is $\bar{c}(\hat{\alpha}, X)$ satisfies

$$1 - \frac{1}{N} = \sum_{\sigma_{\epsilon,i} \in \mathbb{S}} \Phi_{\hat{\alpha},X,\sigma_{\epsilon,i}} \left(\bar{c}(\hat{\alpha}, X)\right) Pr(\sigma_{\epsilon,i}). \tag{O-79}$$

$\Phi_{\hat{\alpha},X,\sigma_{\epsilon,i}}$ is the CDF of normal distribution with mean μ_α and variance

$$\sigma_\alpha^2 + \sigma_{\epsilon,i}^2 (X'X)^{-1}. \tag{O-80}$$

ii. Expected idiosyncratic return

The measurement noise is unbiased and mean zero. Although the measurement noise affects the probability of inclusion of a stock in a portfolio, the expected idiosyncratic return in all quantiles, q, is the unconditional expected idiosyncratic return:

$$E[\epsilon_{A(q,\alpha)}] = 0. \tag{O-81}$$

iii. Expected alpha

To calculate the expected alpha of a stock included in the top quantile, we first integrate the alpha, between infinity and the break point adjusted for a given alpha measurement error draw (their distributions are conditioned on the past n factor draws and an idiosyncratic variance). We then integrate over these measurement error draws, the past n factor draws and sum across the idiosyncratic return variance realizations. Finally, we divide by the total weight in the portfolio which is $\frac{1}{N}$:

$$E[\epsilon_{A(N,\hat{\alpha})}] = \frac{1}{N} \sum_{\mathbb{S}} \int_{\mathbb{X}} \int_{\mathbb{A}^*} \left[\int_{\bar{c}(\hat{\alpha},X)-\alpha^*}^{\infty} \alpha \phi(\alpha)\, d\alpha \right] \phi_{\alpha^*}(\alpha^*)\, \phi_X(X)\, d\alpha^*\, dX\, Pr(\sigma_{\epsilon,i}) \tag{O-82}$$

where $\mathbb{X}$ and $\mathbb{A}^*$ are the domains of all possible X and α^*, and where the subscript $A(N, \hat{\alpha})$ denotes the alpha of A stocks in quantile N when sorting on the estimated alphas.

O.VI. Implementing sorts on estimated idiosyncratic returns: details

In this section, we show how to choose the estimation window, given a hypothesized delay in information diffusion, that will yield uncorrelated estimation error necessary to implement the attenuation correction presented in the main text. We also show that the absence of long horizon cumulative alphas from sorting on estimated idiosyncratic returns provides evidence that the measurement error is uncorrelated with the B firm's idiosyncratic returns. Thus, one can use the absence of cumulative long-short alphas beyond the skipped window from the sorts on idiosyncratic returns to verify that enough periods have been skipped between the model estimation and the prediction date to generate uncorrelated measurement error.

We continue with the assumptions of the main text and the notation of the previous Online Appendix section unless otherwise noted.

A. *Estimation window that produces uncorrelated noise*

The alphas estimated to measure the idiosyncratic returns are approximately the alpha in the estimation window plus the average of the idiosyncratic returns during the estimation window. The difference between the actual estimate and this approximation is due to random correlations of the idiosyncratic returns with the factor realizations during the estimation window. For this and the following subsection to simplify notation and exposition, we ignore those random correlations and suppress the beta terms as they have no meaningful effect. We use this fact about the form of the estimated alphas to inform us as to the optimal model estimation window to generate uncorrelated measurement error in the idiosyncratic returns.

PROPOSITION O-6: *With the assumptions of Proposition O-3, if the idiosyncratic returns of the linked A firm in the estimation window are uncorrelated with the idiosyncratic returns of the linked B firms to be predicted, then the measurement error in the estimated idiosyncratic returns used for sorting is uncorrelated with the predicted idiosyncratic returns.*

The proof of Proposition O-6 is immediate from the fact that the estimated idiosyncratic return is the realized return minus the estimated alpha. The implications of this proposition is that model estimation window must be before the period in which one hypothesizes there to be delays in information diffusion. For example if information diffusion is delayed one period, one must skip one period between the estimation window and the prediction date. This is exactly what we do in the empirical section, because the evidence supports delays in information diffusion of one month or less.

B. Testing for predictability from "old news" in estimation window

To verify a sufficient gap has been used to ensure the measurement error in the idiosyncratic returns is uncorrelated with the predicted idiosyncratic returns of the lagging firms one can look at the long run cumulative alpha from the sorts on estimated idiosyncratic returns. For example absence of cumulative alpha after the initial period indicates that information diffusion has completed within one period, and that skipping one period is sufficient to assure independence of the estimation error. We formalize this for the case with an estimation window that skips one period before the predictive period. This easily generalizes to cases with longer information delay.

To understand why the cumulative alpha provides this information, recall that to measure the estimated idiosyncratic shocks at t, we run a regression of the asset returns over the window $t-T$ to $t-1$. Estimated values from this regression have subscript $t-1$. The estimated alpha from this regression, setting aside variation from the factor realization in the estimation window, is

$$\hat{\alpha}_{i,t-1} = \alpha_i + \frac{1}{T}\sum_{\tau=1}^{T} \epsilon_{i,t-\tau}. \tag{O-83}$$

We use these parameter estimates and the contemporaneous factor returns to extract the estimated idiosyncratic return at t.

$$\hat{\epsilon}_{i,t} = r_{i,t} - \hat{\alpha}_{i,t-1} - \hat{\vec{\beta}}'_{i,t-1}\vec{f}_t \tag{O-84}$$

Expanding the original return and collecting terms gives

$$= \epsilon_{i,t} + (\alpha_i - \hat{\alpha}_{i,t-1}) + \left(\vec{\beta}_i - \hat{\vec{\beta}}_{i,t-1}\right)' \vec{f}_t \tag{O-85}$$

We now suppress the beta terms since they don't matter for the analysis. We also substitute in Equation (O-83).

$$= \epsilon_{i,t} - \frac{1}{T}\sum_{\tau=1}^{T} \epsilon_{i,t-\tau} \tag{O-86}$$

The estimated idiosyncratic shock is lowered by the average of the idiosyncratic shocks over the estimation window. This is a standard result.

The way this measurement error affects our corrected measure of slow information diffusion depends upon two things. The first is the expected value of these old idiosyncratic shocks conditional on a firm being sorted into the long or short portfolio. The second is the persistence (magnitude) of the delay in information diffusion.

The conditional expected value of the *actual* idiosyncratic return of the leading firms (A firms) in portfolio P when firms are sorted on the *estimated* idiosyncratic returns is found in Appendix Equation (D-19) of the manuscript. To measure the potential bias from the measurement error, we need to know the conditional expected value of the *actual lagged* idiosyncratic returns–for all lags in the estimation window–of the leading firms in portfolio P based on the same sort.

It can be shown that

$$\bar{\epsilon}_{P,0} = -T\bar{\epsilon}_{P,-\tau} \quad \forall \tau \in \{1, 2, \ldots, T\} \tag{O-87}$$

The t subscripts are dropped and only the lag notation is kept, because we have averaged across all dates.[12]

To see how this measurement error from old idiosyncratic shocks affects the expected idiosyncratic shock of the B firms (which is what matters for the long-short portfolio alpha), let us extend Equation (O-4) of the manuscript to account for persistent delays in information diffusion

$$\epsilon_{j,t+1} = \sum_{\tau=0}^{H} \rho_{i,j,\tau+1}\epsilon_{i,t-\tau} + \phi_{i,j,t+1}, \quad \{i,j\} \in L \tag{O-88}$$

where L is the set of links, $\rho_{i,j,\tau}$ is the diffusion delay parameter and ϕ is a mean zero normally distributed shock. Let H be the horizon at which information diffusion completes, i.e., $\rho_h = 0$ for all $h > H$.[13]

We obtain the conditional expected idiosyncratic shocks at $t+1$ of the lagging firms, B firms, by applying Equation (O-88) to the conditional expected value idiosyncratic shocks at t of the leading firms (A firms) in portfolio P. Per the derivation in the manuscript, this is the expected alpha of that portfolio of B firms.

$$\bar{\epsilon}_{B,p} = \rho_1\bar{\epsilon}_{P,0} + \sum_{\tau=1}^{T} \rho_{\tau+1}\bar{\epsilon}_{P,-\tau} \tag{O-89}$$

where ρ_τ with the i,j subscript suppressed is the unconditional average information delay

[12] The intuition for this relation is that a firm can be sorted into the long portfolio either because of a large contemporaneous idiosyncratic return or because of a large negative average idiosyncratic shock over the estimation window. Because the idiosyncratic returns are iid, the distribution of the average is the same as that of the individual idiosyncratic return. Therefore, conditional average of the contemporaneous shock must equal that of the negative of the average of the idiosyncratic shocks:

$$E\left[\epsilon_{i,t} | i \text{ in portfolio } P\right] = -E\left[\frac{1}{T}\sum_{\tau=1}^{T} \epsilon_{i,t-\tau} | i \text{ in portfolio } P\right].$$

These shocks are not independent as they are conditional on the firm being sorted into portfolio P, therefore, they cannot be separated. However, since each shock individually has the same distribution, the shocks must contribute equally to the conditional average. QED.

[13] This extended equation is compatible with Equation (O-4) by choosing the correct distribution of the shock ϕ. Exactly which distribution chosen depends upon the calibration assumption one wishes to make. These are not important for this analysis.

at lag τ (justified by the independence assumption).

Using the relation of the conditional expected values of the contemporaneous and lagged shocks in Equation (O-87) gives the expected long-short alpha of the lagging firms when sorted on the estimated idiosyncratic shocks of the leading firms, i.e., Proposition O-7.

PROPOSITION O-7: *With the assumptions of Proposition O-3, if information diffusion is delayed per Equation* (O-88) *and the estimation window used to calculate parameters necessary to estimate the idiosyncratic returns is from* $t-T$ *to* $t-1$, *then the expected long short alpha from sorting on estimated idiosyncratic returns is*

$$E\left[\alpha_{B''(LS)}\right] = c\left(\rho_1 \underbrace{-\frac{1}{T}\sum_{\tau=1}^{T}\rho_{\tau+1}}_{\textit{bias term}}\right) \tag{O-90}$$

where

$$c \equiv \bar{\epsilon}_{L,0} - \bar{\epsilon}_{S,0}. \tag{O-91}$$

and $\bar{\epsilon}_{P,0}$ *is the conditional expected value of the* actual *idiosyncratic return of the leading firms (A firms) in portfolio* P *when firms are sorted on the* estimated *idiosyncratic returns.*

Thus, we see the potential downward bias is that of the average delay in information diffusion over the estimation window. If information is delayed only one period there is no bias.

As we predict further out in time, the idiosyncratic returns being iid gives that the conditional expected idiosyncratic shocks of the A firms at any date past the sorting date t are 0. The other effect of increasing h is that the delay coefficients move further in time from the sort. This movement, combined with the zero expected value of future shocks, means we shift the subscripts in Equation (O-90). This gives Corollary O-5 which extends Equation (O-90) to longer horizons predictions.

COROLLARY O-4: *Following the assumption of Proposition O-7 the expected long short* *alpha at prediction horizon h from sorting on estimated idiosyncratic returns is*

$$\bar{\epsilon}_{B,L,h} - \bar{\epsilon}_{B,S,h} = c\left(\rho_h - \frac{1}{T}\sum_{\tau=1}^{T}\rho_{\tau+h}\right) \tag{O-92}$$

Therefore unless the delay in information diffusion is constant across the estimation window, any delay in information diffusion beyond one period will generate positive cumulative alphas beyond the first prediction period. To see this, define decay in the delay in information diffusion as

$$0 \leq \rho_{\tau+1} \leq \rho_\tau \quad \forall \tau \tag{O-93}$$

with at least one inequality between the ρ_τ strict over the estimation window (i.e., $\tau \in \{0, 1, \ldots, T\}$.

Comparing the terms we immediately see the following corollary.

COROLLARY O-5: *Following the assumptions of Proposition O-7, if there is delayed information diffusion beyond one period, i.e., $\rho_\tau > 0$ for some $\tau > 1$ and if there is decay in the rate of delay in information diffusion then there exists an $h > 1$ such that*

$$\rho_h > \frac{1}{T}\sum_{\tau=1}^{T}\rho_{\tau+h}, \tag{O-94}$$

meaning the cumulative alpha increment (Equation O-92) is positive.

Therefore if there is no cumulative alpha either, then either there is no decay in rate of delay in information diffusion which is implausible or there is no delay in information diffusion beyond one period.

O.VII. Simulation: predictability from "old news" is inconsistent with data

In this section, we present results of a simulation that confirms the results of section O.VI.B that long delays in information diffusion would result in long horizon cumulative alphas from the estimated idiosyncratic returns sorts. The simulation confirms such a delay would manifest as upward slopes as seen in Figure O-1 below. We do not see such an upward slope in Figure 2 in the main text. Table III of the main text confirms the cumulative alphas are economically and statistically insignificant.

For the simulation we draw a panel of firms the same dimension as that of the A firms for each set of economic links. We draw the alpha and beta parameters for the A firms to match the data per the calibration discussion later in this section. We bootstrap the variances of the A firms' idiosyncratic returns from the variances of the idiosyncratic returns in the data. We calibrate the data to that from portfolios of linked economic firms rather than calibrate at the firm level. In particular the portfolio level data gives more reliable measures of the idiosyncratic return variances.

With the panel we either perform sorts on the A firm returns, on the estimated idiosyncratic returns, or the known idiosyncratic returns. The value of the simulation is that we can know these idiosyncratic returns since we generated the data. With these various sorts we are able to calculate the corresponding long short alphas of the B firms. We repeat this procedure 10,000 times and report the results.

We calibrate the main parameters of interest to give the benefit of the doubt to the return sorting specification. Specifically we chose ρ_ϵ for each paper as the maximum of the value in the data or the value needed for the simulation to produce alphas from the known and estimated idiosyncratic return sorts equal to that found in the data under the assumption that information diffusion is only delayed one period.

The value of ρ_α is used directly from the data. The cross-sectional variation in alphas,

Table O-1 Simulation Parameters Panel A of this table shows the parameters used in the simulation used to produce Figure O-1. See the appendix text for a description of the calibration procedure. Case 1 matches the data assuming the delay in information diffusion lasts only one month. Cases 2 and 3 are counterfactual cases that assume information delay for months 2 through 12 is 10% and 20% that of the one-month delay. Panel B shows the bias in the long-short alpha from sorting on estimated idiosyncratic returns corrected for uncorrelated measurement error relative to the long-short alpha from sorting on the known idiosyncratic returns.

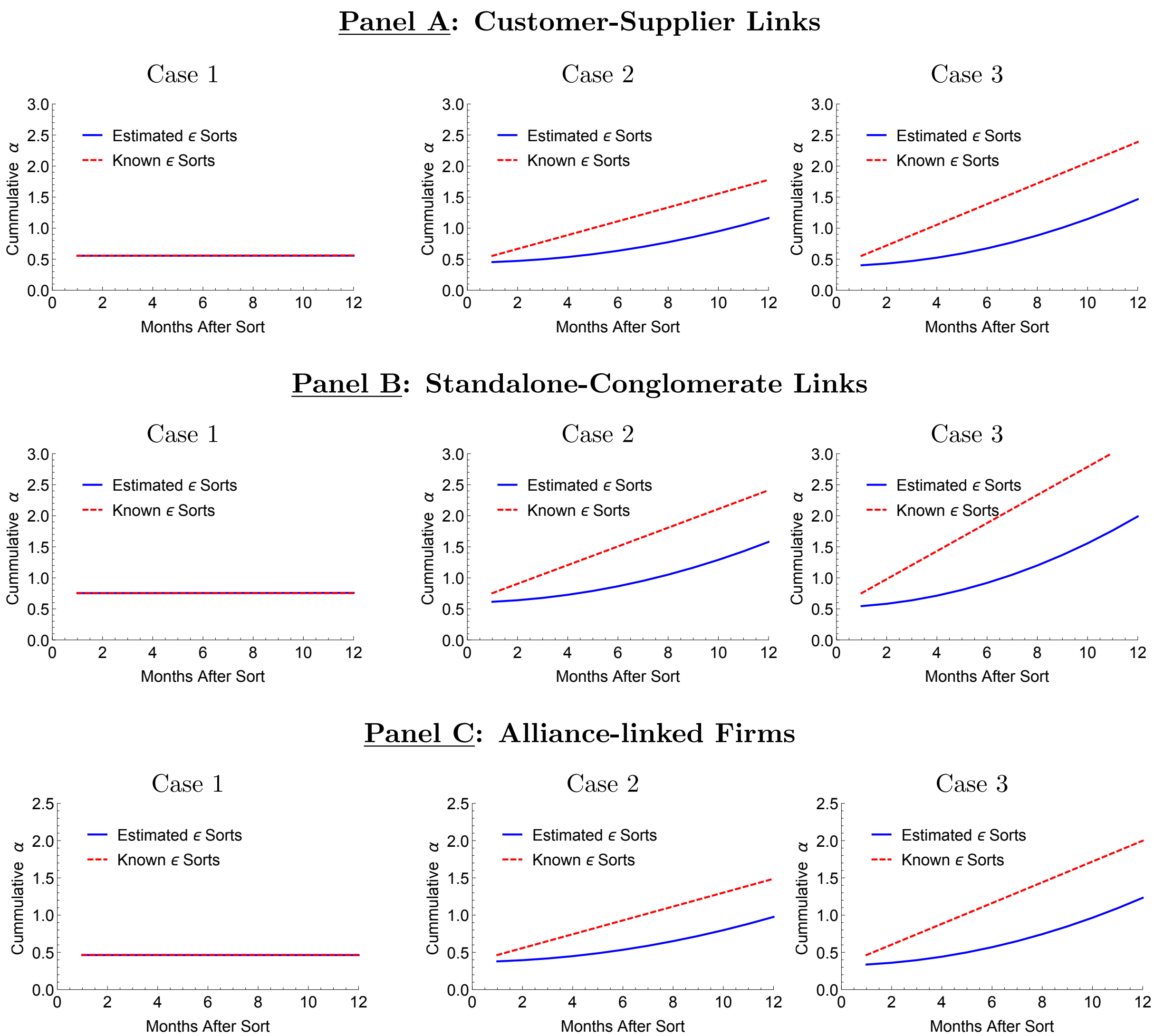
Panel A: Parameters	Customer-Supplier	Standalone-Conglomerates	Alliance-Linked Firms
$\bar{\rho}_\epsilon$	0.023	0.057	0.013
$\bar{\rho}_\alpha$	0.131	0.198	0.112
$\sigma_{CS}(\alpha)$	3.74	1.61	4.36
Panel B: One Month Bias (%)			
Case 1: No Persistent Delay	0	0	0
Case 2: 10% Persistent Delay	-10	-10	-10
Case 3: 20% Persistent Delay	-18	-18	-18

$\sigma_{CS(\alpha)}$ is the minimum of the value found in the data or that required for the return sorts to match the values found in the data. This calibration results in lower cross-sectional alpha variance than directly calculated from the data because the value directly calculated from the data is inflated due to measurement error in the alphas. These calibration values are shown in Table O-1.

For the simulation we consider three cases to illustrate the effect of persistent delays in information diffusion. In Case 1 we assume that there is no delay in information diffusion beyond one month (consistent with the data). In Case 2 and 3 we consider the counter factual cases where information diffusion extends beyond month one with the delay from months 2 to 12 being 10% or 20% that of the delay over the one month horizon.

In Panel B we see that the downward bias at the one month horizon is roughly identical to the amount of additional delay over months 2 to 12. However such additional delay results in a counter factual cumulative alpha over those month in the sorts based on the estimated idiosyncratic returns. Figure O-1 shows this, with the large upward slope for the blue lines.

Figure O-1. Simulation: Cumulative Alphas Sorting on Known and Estimated Idiosyncratic Returns This figure shows the cumulative alphas to long-short portfolios for different sets of economically linked firms under the simulation calibrated per Table O-1. At each date $t + h$, we calculate the Fama and French (1993) 3-factor model monthly alpha under sorts at t that use either known or estimated idiosyncratic returns. The cumulative alpha is the sum of alphas for each of the periods through horizon h. Long-short alphas from sorting on estimated idiosyncratic returns are corrected to eliminate attenuation from estimating the leaders' idiosyncratic returns. This correction assumes uncorrelated measurement error of the idiosyncratic returns. We consider three cases of delay in information diffusion. Case 1 is consistent with the data with one month of information delay. Cases 2 and 3 consider additional information delay between months 2 and 12 that is 10% or 20% that of the initial one month delay. Alpha is in percent per month. Panel A uses customers (leaders) to assign suppliers (laggards) into quintile portfolios. Panel B uses standalone firms (leaders) to assign conglomerates (laggards) into decile portfolios. Panel C uses alliance-linked firms (leaders) to assign another firm in the alliance (laggards) into quantile portfolios.



O.VIII. Attenuation ratio correction for daily estimation with monthly sorts

The bias ratio calculated in Appendix O.IV assumes the same data frequency (e.g., monthly) used to estimate the asset pricing model for calculating the idiosyncratic returns as that data frequency for the sorting into portfolios. The ratio changes, however, when estimating the asset pricing model at a higher data frequency (e.g., daily) than the portfolio sorting (e.g., monthly). Although the attenuation from noise decreases due to the higher frequency estimation, the same estimate of the model is used for each day of the month of the sorting period reducing the effect of this gain. This overlap in parameter estimates used throughout the lower frequency sorting period (e.g., all days of the month) must be accounted for.

We continue with the assumptions from Appendix O.IV in the main text. Recall, in particular, that the factors have been demeaned.

Under the standard assumption of i.i.d. factor returns, the variance of the factors scale linearly with frequency. Let T be the number of days in a month and let d subscripts denote daily objects.

The variance of the measurement error of the monthly idiosyncratic return estimated with daily data is

$$\sigma^2(\sigma^*_{T,i}) = \sigma^2_d(\epsilon_i) \begin{bmatrix} T \\ \sum_{\tau=1}^{T} \vec{f}_{d,\tau} \end{bmatrix}' (X_d' X_d)^{-1} \begin{bmatrix} T \\ \sum_{\tau=1}^{T} \vec{f}_{d,\tau} \end{bmatrix} \tag{O-95}$$

Thus, the scaling factor for the monthly idiosyncratic variance in this case is

$$s_{d,m}(\vec{f}, X_d' X_d) = T \begin{bmatrix} 1 \\ \frac{\vec{f}}{T} \end{bmatrix}' (X_d' X_d)^{-1} \begin{bmatrix} 1 \\ \frac{\vec{f}}{T} \end{bmatrix} \tag{O-96}$$

where

$$f \sim N(0, \Sigma_f), \tag{O-97}$$

$$X_d' X_d = \begin{bmatrix} nT & nT\bar{f} \\ nT\bar{f} & W + nT\bar{f}\bar{f}' \end{bmatrix}, \tag{O-98}$$

W is Wishart with $nT-1$ degrees of freedom and uses the variance-covariance matrix of the factors divided by T: $\frac{\Sigma_f}{T}$ and

$$\bar{f} \sim N\left(0, \frac{\Sigma_f}{nT}\right) \tag{O-99}$$

The bias ratio is the unconditional expectation of this scaling factor with these distributions:

$$E\left[\frac{1}{\sqrt{1 + s_{d,m}(\vec{f}, X_d' X_d)}}\right] \tag{O-100}$$

Similar to the case of monthly estimation, we can replace the variance-covariance matrix Σ_f with the identity matrix of the same size. Thus, only the number of factors in the asset pricing model, the length of the sample window and the difference in scale between the frequencies matter.

O.IX. Relation to Lo and MacKinlay (1990)

We can gain further intuition for the prior decomposition and related it to an earlier literature using a theoretical framework similar to that of Lo and MacKinlay (1990). The Lo and MacKinlay (1990) framework simplifies the econometrics of sorts by using continuous weights rather than discrete portfolio cutoffs. The continuous weights analogous to those in Lo and MacKinlay (1990) are based on the relative performance of the leader firms in the sorting period.[14]

The first component—the common trend—is analogous to the Lo and MacKinlay (1990) own-firm predictability components "O" and $\sigma(\mu)$. The "O" component captures firms' temporary deviations from their long-term means (i.e., time-varying alphas and betas, a.k.a. seasonalities or momentum). The $\sigma(\mu)$ term captures the permanent unconditional differences in mean returns across leaders due to both alpha and beta components. This own-firm predictability shows up in the cross-firm predictability because of the common trend in economically linked firms and the tendency of these firms to contemporaenously comove. This own-firm component of the long-short portfolio alphas obtained from sorting laggards by leaders' returns is the portion of the alpha that can be attained simply by sorting leaders on their own past returns. In essence, this component is a (noisy) repackaging of the leader's own-firm momentum (or seasonalities). The alphas from this own-firm predictability, therefore, do not represent delayed information diffusion as traditionally understood.

We are agnostic as to the source of this (predictable) common trend or momentum. We only argue it is distinct from the second component of cross-firm predictability, which is analogous to the Lo and MacKinlay (1990) cross-firm predictability component, "C". The "C" component is due to new (unpredictable) information in the leaders' returns that the laggards respond to with a delay. By tracking the response of laggards to *new* and *unpredictable* information, this component gives a measure of the direct delay in information diffusion between leaders and laggards.

Trading strategy definition Consider the leader firm i and laggard firm j with the set of all linked firms being described by the set L of ordered pairs $\{i, j\}$. Let there be N such links. Without loss of generality, for exposition we consider links between single firms (rather than portfolios and firms) and unique matching between the leaders and laggards. Within the set of linked firms, call the set of all leader firms A and the set of all such laggard firms

[14] Lo and MacKinlay (1990) examined the profits to contrarian strategies, so their weights are based on prior period returns of the firm itself.

B. The trading strategy is to invest at time t in laggard firm j according to the return of its linked leader firm i in period $t-k$ relative to the returns of all leading firms at $t-k$, where k is positive. The investment weights are

$$w_{j,t-k} = \frac{1}{N}(r_{i,t-k} - \bar{r}_{A,t-k}) \quad \forall \{i,j\} \in L, \tag{O-101}$$

where $\bar{r}_{A,t-k}$ is the average return of the A firms at time $t-k$. The raw profits (returns) to this strategy at time t are

$$\pi_t^k = \sum_{j=1}^{N} w_{j,t-k} r_{j,t}. \tag{O-102}$$

Decomposing returns using chosen asset pricing model While the raw profits (returns) of this strategy are interesting, the literature we analyze uses the risk-adjusted profits, or alpha, to measure cross-firm predictability and make inferences regarding the speed of information diffusion. Measuring alpha requires specifying an asset pricing model against which to measure the spread driven by the speed of information diffusion. Having specified this linear asset pricing model in terms of the factors $\vec{f}_t$ of interest, we can decompose the returns of the leaders and laggards:

$$r_{i,t} = \alpha_{i,t} + \vec{\beta}'_{i,t}\vec{f}_t + \epsilon_{i,t} \quad \forall i \in A \cup B. \tag{O-103}$$

The betas capture the risk exposures, the alphas capture the average difference between the model's expected returns and the firm's expected returns given the state variables $\mathbb{S}_t$ that define the asset pricing model. We will abuse notation slightly using only the t subscripts to denote these state variables. The epsilons are mean-zero normal shocks independent of the alphas given those same state variables.[15] For simplicity, assume that the alphas and betas

[15]When considering the decomposition into slow information diffusion, the prior returns or idiosyncratic returns are not part of this set of state variables. Excluding them from the state variables that define the asset pricing model allows for the epsilons of the lagging firms to have non-zero means under the expanded information set including the information contained in the leaders' returns, or more specificially, idiosyncratic returns.

are independent of each other in the cross-section of the A firms.

Predictable non-idiosyncratic vs. unpredictable idiosyncratic components We can group the terms in Equation (O-103) into the non-idiosyncratic component $\alpha_{i,t} + \vec{\beta}\,'_{i,t}\vec{f}_t$ and the idiosyncratic component $\epsilon_{i,t}$. We can say the non-idiosyncratic component is that part which is predictable when the parameters and factor realizations are known, while the idiosyncratic component is unpredictable even with that knowledge.

When the parameters are unknown, we can still form the non-idiosyncratic component up to the limit of our ability to predict (i.e., estimate with prior information) the parameters. This predicted non-idiosyncratic component is $\hat{\alpha}_{i,t-1} + \hat{\vec{\beta}}'_{i,t-1}\vec{f}_t$, where the hats and $t-1$ subscripts denote our estimates using only prior information. The unpredictable idiosyncratic component thus becomes $\hat{\epsilon}_{i,t} = r_{i,t} - \hat{\alpha}_{i,t-1} - \hat{\vec{\beta}}'_{i,t-1}\vec{f}_t$. The properties and amount of predictability depend both on the asset pricing model and the estimation method which we specify later.

Cross-sectional relation between leaders and laggards Economically linked firms are exposed to similar economic forces, news and risk. This commonality means their alpha, beta and idiosyncratic returns will comove in the cross-section contemporaneously and, in the case of delayed information diffusion, with a lag as well, i.e., directly cross-serially. Such contemporaneous comovement combined with any predictability in the time-series of firms' parameters (e.g., expected returns) leads to common trends in the economically linked firms. Modeling these relationships as follows allows us to disentangle the common momentum component of leaders and laggards from the slow information diffusion between leaders and laggards:

$$\alpha_{j,t} = \rho_\alpha \alpha_{i,t} + \phi_{i,j,t}, \tag{O-104}$$

$$\vec{\beta}_{j,t} = \rho_\beta \vec{\beta}_{i,t} + \vec{\psi}_{i,j,t} \quad \text{and} \tag{O-105}$$

$$\epsilon_{j,t} = \vec{\rho}\,'_\epsilon \vec{\epsilon}_{i,t} + \vec{\theta}_{i,j,t} \tag{O-106}$$

for all $\{i, j\} \in L$, where $\vec{\epsilon}_{i,t}$ is a vector of all past idiosyncratic returns to firm i. Using a zero-based index, the current idiosyncratic return has index 0, the previous period idiosyncratic return has index 1, etc.

ρ_α and ρ_β capture the contemporaneous commonality in the alphas and betas (e.g., model-predicted returns). $\vec{\rho}_\epsilon$ captures both the contemporaneous shocks to the leaders and laggards along with the amount of delayed information diffusion. Assume that the shocks $\phi_{i,j,t}$, $\vec{\psi}_{i,j,t}$ and $\vec{\theta}_{i,j,t}$ are independent in the cross-section at any date, independent of the other terms (e.g., $\alpha_{i,t}$ and $\vec{\beta}_{i,t}$) and have mean of 0.

A. *Profits are composed of three components*

We now show that the profits of this trading strategy can be decomposed into that due to the alphas, betas and idiosyncratic returns given the specified asset pricing model against which delayed information diffusion is being measured. We then show how these three components map into the three components "O", "C" and $\sigma^2(\mu)$ of Lo and MacKinlay (1990).

Substituting in decompositions Substituting these decompositions of the asset pricing model and the relationship between the leaders and laggards into the profit Equation (O-102) (note that we drop the j subscripts for the linked firm i subscripts) gives the time-series expectation profits of

$$E_{TS}[\pi_t^k] = E_{TS}\left[\frac{1}{N}\sum_{i=1}^{N}\begin{pmatrix} \alpha_{i,t-k} + \vec{\beta}'_{i,t-k}\vec{f} + \epsilon_{i,t-k} \\ -\bar{\alpha}_{i,t-k} - \bar{\vec{\beta}}'_{i,t-k}\vec{f} - \bar{\epsilon}_{i,t-k} \end{pmatrix}\begin{pmatrix} \rho_\alpha \alpha_{i,t} + \phi_{i,j,t} \\ +(\rho_\beta \vec{\beta}_{i,t} + \vec{\psi}_{i,j,t})'\vec{f}_t \\ +\vec{\rho}'_\epsilon \vec{\epsilon}_{i,t} + \theta_{i,j,t} \end{pmatrix}\right]. \quad \text{(O-107)}$$

The sum divided by N is the cross-sectional expectation. This fact combined with a cross-sectional independence assumption and, without loss of generality, the assumption that the average alphas and idiosyncratic shocks across the A firms are 0 at all dates, we simplify the expected profits to be

$$E_{TS}[\pi_t^k] = E_{TS}E_{CS}\begin{bmatrix} \rho_\alpha \alpha_{i,t}\alpha_{i,t-k} + \vec{\rho}'_\epsilon \vec{\epsilon}_{i,t}\epsilon_{t-k} \\ +\rho_\beta \vec{\beta}'_{i,t}\vec{f}_t(\vec{\beta}_{i,t-k} - \bar{\vec{\beta}}_{A,t-k})'\vec{f}_{t-k} \end{bmatrix} \quad \text{(O-108)}$$

Because the idiosyncratic shocks are mean-zero and uncorrelated over time, this further simplifies to

$$E_{TS}[\pi_t^k] = E_{TS}E_{CS}\begin{bmatrix} \rho_\alpha \alpha_{i,t}\alpha_{i,t-k} + \rho_{\epsilon,k}\epsilon_{t-k}^2 \\ +\rho_\beta \vec{\beta}'_{i,t}\vec{f}_t(\vec{\beta}_{i,t-k} - \bar{\vec{\beta}}'_{A,t-k})\vec{f}_{t-k} \end{bmatrix} \quad \text{(O-109)}$$

where $\rho_{\epsilon,k}$ is the relation of idiosyncratic shocks between leaders and laggards at horizon k.

Raw profit: From Equation (O-109), we see that the raw profit (returns) of this trading strategy is composed of three components. The epsilon term, $E_{TS}E_{CS}[\rho_{\epsilon,k}\epsilon_{t-k}^2]$, captures how the idiosyncratic news shocks propagate over time from leaders to laggards. It is increasing in the variance of the idiosyncratic shocks of the leaders and the amount of information delay at horizon k: $\rho_{\epsilon,k}$. In the special case of all leaders having the same idiosyncratic shock variance at time $t-k$, this simplifies to the average of these across time: $\rho_{\epsilon,t-k}\bar{\sigma}_i^2$.

The alpha term, $E_{TS}E_{CS}[\rho_\alpha \alpha_{i,t}\alpha_{i,t-k}]$, captures the contemporaneous correlation of the leaders and laggards and the leaders' own-firm predictability (momentum). The product of the leaders' alphas at different dates is increasing in the amount of own-firm predictability (i.e., predictability of the mean return given the past mean return). This term is also increasing in the amount of unconditional cross-sectional variation in alphas (mean returns adjusting for the model) and in the amount of predictable time-series variation in the alpha. Thus, this term will be positive even if there is no unconditional cross-sectional difference in leaders' alphas.[16]

The last term, $\rho_\beta \vec{\beta}'_{i,t} \vec{f}_t (\vec{\beta}_{i,t-k} - \bar{\vec{\beta}}'_{A,t-k}) \vec{f}_{t-k}$, captures the profit due to variation in the model-predicted component of returns. This term depends both upon the time-series and cross-sectional properties of the exposures $\vec{\beta}$ and the time-series properties of the risk factors.

Own-firm and cross-firm predictability The three-way decomposition overlaps with a decomposition in the spirit of Lo and MacKinlay (1990) own-firm and cross-firm predictability decomposition. Although, strictly speaking, all the predictability in our setting has a cross-firm dimension, we use Lo and MacKinlay's own-firm predictability to capture the portion of cross-firm predictability flowing through the leaders' ability to predict themselves. We reserve Lo and MacKinlay's cross-firm predictability term to denote the *direct* delayed cross-firm predictability between leaders and laggards.

The alpha and beta components from Equation (O-109) capture the profits from the own-firm predictability. As we discussed, the alpha and beta terms have transient and permanent components. Thus, the own-firm component can be decomposed into a transient component analogous to the "O" term of Lo and MacKinlay (1990) and a permanent component analogous to their $\sigma^2(\mu)$ component. Together, these pieces capture the predictable non-idiosyncratic component of the profits.

The epsilon component of Equation (O-109) captures the direct cross-firm predictability analogous to the "C" term of Lo and MacKinlay (1990). This piece is captured by the unpredictable idiosyncratic component of the profits. While the total profits remain constant

[16] This limited cross-sectional dispersion in unconditional returns is the point made in Jegadeesh and Titman (2002) and, more recently, in Keloharju, Linnainmaa, and Nyberg (2019).

across all asset pricing models and estimation procedures, the assignment of the profits to these various components changes with the choice of asset pricing model against which slow information diffusion is measured. As either the model or the estimation of the predictability becomes worse, more profit is assigned to the idiosyncratic component (slow information diffusion component).

Risk-adjusting profit does not remove own-firm predictability While Lo and MacKinlay (1990) focuses on the returns of the trading strategy, the literature we address focuses on the risk-adjusted profits of long-short portfolios. The risk-adjusted profit is the unconditional alpha from a time-series regression of the profit time-series on the factors in the specified asset pricing model. Although the raw profits remain the same under any asset pricing model, the risk-adjusted profits vary with the model chosen as a benchmark.

The risk adjustment leaves the alpha and epsilon terms of the raw profits unaffected. The adjustment only removes the beta component of the raw profits. In the case of iid factors, all of the beta component will be removed. Without iid factors, the amount of the beta component removed depends upon the interaction of the time-series dynamics of both the time-varying betas of the raw profit time series and the factors of the asset pricing model. See Lewellen and Nagel (2006) for details on a decomposition that shows how much of this beta component remains as alpha after the risk-adjustment is applied.

As the alpha component and part of the beta component remain after adjusting for risk, the risk-adjusted profits still contain the own-firm predictability component. Therefore, these risk-adjusted profits do not solely reflect delays in information diffusion relative to the chosen asset pricing model. These risk-adjusted profits reflect both the own-predictability (momentum) in the leading firms and in the common time-varying risk exposures of leading and lagging firms.

B. How sorts differ from continuous weights

In practice, the literature uses trading strategies based on sorting the laggards into portfolios, typically quintiles or deciles, and measures the risk-adjusted profits (alphas) from going long the highest quantile and short the lowest quantile. These sorts can be thought of placing a ternary operator on the weighting function used so far. This operator takes on three values: a negative $\frac{q}{N}$, where q is the number of quantiles if the weight falls in the bottom quantile of weights, a positive $\frac{q}{N}$ if the weight falls in the top quantile, and zero otherwise.

This cutoff-based weighting function complicates the econometrics because one must account for these quantile break points. We show the econometrics of these sorts in Appendix O.II. Nonetheless, the main intuition remains. The profits can be decomposed into three pieces: one due to delayed information diffusion and the two other pieces due to time-variation in the modeled betas and the (semi-)persistent difference in the leaders' returns and their expected returns specified under the model.

In the case of sorts, an additional complication of "noise" arises from (1) the large difference in weights across the break points, and (2) the fact that each component of the sorting variable plays a role in the portfolio assignment relative to the breakpoints. For example, the component of returns due to systematic risk acts as noise, lowering the potential profits to be earned from trading on slow information diffusion. The intuition is that a stock can fall into the top portfolio due to a high return from factor realizations or from a good piece of news that will be transmitted to the laggard with delay. Sorting only on idiosyncratic returns will give larger trading profits from slow information diffusion than sorting on returns.

This noise matters, in particular, when sorting on the unpredictable idiosyncratic component. Even when the estimation errors (e.g., the misallocation) are uncorrelated with the future laggard idiosyncratic returns, these errors will push down the trading profits. We derive a closed form characterization of this downward bias induced from the estimation error in Appendix O.II. With this characterization, we correct for this bias throughout the empirical section.

O.X. Supplier returns do not catch up to customer returns

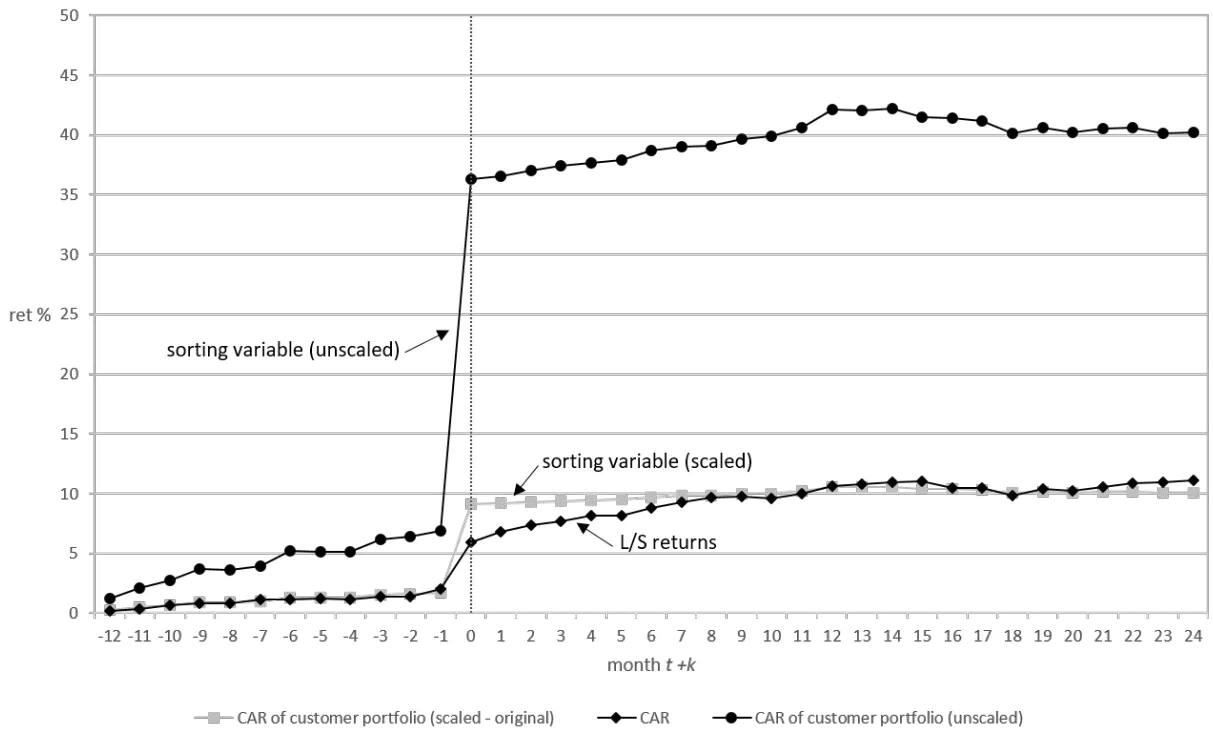
One of the most compelling pieces of evidence of slow information diffusion documented in Cohen and Frazzini (2008) is in their Figure 2. The figure reveals two important patterns in the data. First, the figure shows the long horizon predictability, with cumulative alphas lasting up to one year, under the uncorrected specification. Second, the figure shows that over time, supplier returns appear to eventually "catch up" with customer returns, signifying when information diffusion is complete. The paper suggests the two return patterns measured are due to shocks propagating between customers and suppliers though cash flows.

We argue the first pattern of long horizon predictability is due to correlation in the alphas and modeled risks across customers and suppliers, which we explore fully in the main text. This section explores the second pattern that supplier returns "catch up" with customer returns. In particular,the appearance of supplier returns catching up is a result of their unique construction of the customer cumulative abnormal returns variable. To understand the construction of the customer returns variable in the figure, we quote from the original paper.

Cohen and Frazzini (2008), pages 1989 to 1990, reads

> Figure 2 illustrates the result by reporting how customer returns predict individual stock returns at different horizons. We show the cumulative average returns in month $t + k$ on the long-short customer momentum portfolios formed on customer returns in month t. We also plot the cumulative abnormal return of the customer portfolio (the sorting variable). To allow for comparisons, we show returns of the customer portfolio *times the total fraction of supplier firm's sales accounted for by the principal customers.* [Our emphasis]

Figure O-3. Effect of Scaled and Unscaled Sorting Variables This figure replicates Figure 2 of Cohen and Frazzini (2008). It shows the cumulative abnormal returns in month $t + k$ following the sort using the uncorrected specification at t of the following; the first two of which are in the original figure. (See Table III for details of the sort). The black diamonds show the long short portfolio alphas of the suppliers. The gray squares show the long short portfolio of the sorting variable (customers) scaled down the total sales attributable to customers used for the sorting. The scale factor averages 0.25. The black circles show the long short portfolio of the sorting variable without any scaling .



We emphasize that the customer returns variable is *scaled* in the figure. This is *different* from the variable used to generate the paper's main long-short alpha results. Such a scaling may seem intuitive as shocks to customers' cash flows may not translate one-for-one to suppliers' cash flows. However, the economic basis for the direction of scaling is unclear. On the one hand, the effect of customers' poor performance on suppliers could be dampened because the measured customers make up only a fraction of the suppliers sales, justifying the downward scaling (i.e., less than one) applied in the original figure. On the other hand, a shock to customers could be magnified at the supplier level, for example, if the supplier has fixed costs of operations that must be met, justifying an upward scaling (i.e., greater than one). Ex-ante, the direction of scaling cannot be determined.

The downward scaling is also inappropriate given we show a large portion of the 1-month lagged comovement and nearly all the comovement two months and beyond is due to correlated alphas. As correlated alphas can come from common risk exposures missed by the asset pricing model, common risk would need no down-scaling. That is, returns from risk between customer and suppliers will be similar to the extent they are exposed to similar risks. For example, simply measuring a subset of the suppliers customers could give an accurate measure of the common risk exposure of the suppliers if the customers are a representative sample.[17]

To understand the effect of the downward scaling used in Cohen and Frazzini (2008), Figure O-3 plots the suppliers' cumulative alphas and the *scaled* customer returns variable (as shown in the original paper), as well as the *unscaled* customer returns variable. In the figure, the suppliers' cumulative alphas (the black diamond-ticked line) appear to converge over time to the *scaled* customers' cumulative abnormal returns variable (the gray square-ticked line). This pattern is consistent with the original paper's prediction that returns of suppliers should catch up with those of their customers as information about customers' cash flows shocks diffuses through the market. The pattern also appears to contradict our results of the long-short alphas being primarily driven by correlated alphas.

Alternatively, correlated alphas should result in the plotted line of the customers' alphas being parallel or slightly steeper (to account for imperfect correlation of alphas) than the plotted line of suppliers' cumulative alphas. In essence, the focus should be on the *slopes* of the customer and supplier alphas. The gray square ticked line is not the cumulative alpha of the customers. It is those alphas scaled down by a factor of 0.25.[18] In contrast to the original inferences, *removing the scaling* (as shown with the black circle-ticked line) reveals that the supplier returns fail to converge with the customer returns. Moreover, consistent with the

[17]Simply multiplying the average return of the long-short customer portfolio by the alpha correlation may not give the average return of the long-short supplier portfolio because of the nonlinearities documented in the closed-form calculation found in Online Appendix O.V.

[18]Cohen and Frazzini (2008) do not report their scaling. We calculate the average scaling from our replication of their data.

mechanism we document of correlated alphas, the slope of the customers' cumulative alphas beyond the first month is 0.51, which is greater than the 0.35 of the suppliers. We exclude the first month because it contains a point estimate of delayed price discovery, even under the corrected specification.[19] Removing the scaling reveals evidence consistent with our analysis: delays in information diffusion are limited to only one month or less.

O.XI. Old news has no predictive ability

See the main text Section III.C for a description of the methodology that produced Table O-2 and a description of its results.

[19]This point estimate is economically smaller than that from the uncorrected specification, and it is often statistically insignificant.

Table O-2 Sorts on old news This table shows long-short alphas obtained by sorting laggards' returns from $t+1$ to $t+12$ into portfolios based on their economically linked leaders' idiosyncratic returns (old news) from $t-12$ to $t-1$. For each leader, we estimate the alphas and factor loadings of the 5-factor model (Fama and French, 1993; Carhart, 1997; Pástor and Stambaugh, 2003) by using the 12 monthly returns from $t-24$ to $t-13$). We use these parameter estimates with the realized factor returns to compute expected returns as the predictable nonidiosyncratic component (using factor realizations). The idiosyncratic return is the realized return minus the predictable nonidiosyncratic return. At each date from $t-12$ to $t-1$, we use the idiosyncratic returns to sort the laggards into portfolios at each of the prediction dates ($t+1$ to $t+12$) and compute long-short 3-factor alphas. Alpha is reported in percent per month. Panel A uses customers (leaders) to assign suppliers (laggards) into quintile portfolios. Panel B uses standalone firms (leaders) to assign conglomerates (laggards) into decile portfolios. Panel C uses alliance-linked firms (leaders) to assign another firm in the alliance (laggards) into quantile portfolios. Statistically significant alphas at the 5% level are bolded. t-statistics are in parentheses.

Panel A: Customer-suppliers

Equal-weighted

Leader epsilon (t-k)	Laggard (t+h)											
	h = 1	2	3	4	5	6	7	8	9	10	11	h = 12
k = 1	0.261	0.183	0.034	0.124	0.342	-0.097	0.122	0.048	-0.066	0.018	0.196	-0.272
	(1.07)	(0.88)	(0.15)	(0.54)	(1.56)	(-0.37)	(0.47)	(0.17)	(-0.27)	(0.08)	(0.85)	(-1.13)
2	0.360	0.217	0.212	0.331	0.023	-0.077	0.136	-0.134	-0.191	**0.475**	-0.346	0.093
	(1.58)	(0.86)	(0.92)	(1.40)	(0.09)	(-0.28)	(0.47)	(-0.47)	(-0.83)	**(2.19)**	(-1.53)	(0.39)
3	0.158	0.030	0.204	-0.086	-0.164	0.427	0.030	0.088	0.411	-0.144	0.184	0.268
	(0.63)	(0.13)	(0.89)	(-0.35)	(-0.60)	(1.67)	(0.12)	(0.36)	(1.75)	(-0.60)	(0.79)	(1.25)
4	-0.052	0.270	-0.068	-0.044	0.364	-0.164	-0.013	**0.567**	-0.026	0.114	**0.481**	-0.234
	(-0.21)	(1.16)	(-0.27)	(-0.16)	(1.34)	(-0.67)	(-0.06)	**(2.44)**	(-0.11)	(0.48)	**(2.30)**	(-0.98)
5	0.126	0.048	-0.109	0.263	-0.149	-0.107	0.377	-0.088	-0.099	**0.481**	-0.218	0.002
	(0.54)	(0.18)	(-0.43)	(0.99)	(-0.62)	(-0.42)	(1.52)	(-0.36)	(-0.40)	**(2.20)**	(-0.94)	(0.01)
6	0.178	0.193	0.200	0.006	-0.041	**0.554**	-0.167	0.215	0.385	-0.131	-0.020	-0.022
	(0.67)	(0.73)	(0.73)	(0.02)	(-0.16)	**(2.40)**	(-0.71)	(0.88)	(1.79)	(-0.59)	(-0.09)	(-0.10)
7	0.372	0.438	-0.033	0.000	**0.863**	-0.305	0.165	0.377	-0.432	0.002	-0.080	0.463
	(1.40)	(1.57)	(-0.14)	(-0.00)	**(3.41)**	(-1.33)	(0.63)	(1.73)	(-1.91)	(0.01)	(-0.34)	(1.73)
8	0.261	-0.114	0.089	**0.658**	-0.318	0.239	0.332	-0.349	-0.134	-0.300	0.417	-0.236
	(1.02)	(-0.45)	(0.36)	**(2.63)**	(-1.28)	(0.98)	(1.45)	(-1.54)	(-0.59)	(-1.29)	(1.63)	(-0.94)
9	0.002	0.213	**0.577**	-0.271	0.044	-0.032	-0.500	-0.033	-0.381	0.299	-0.546	0.367
	(0.01)	(0.88)	**(2.07)**	(-1.06)	(0.17)	(-0.15)	(-2.08)	(-0.14)	(-1.56)	(1.14)	(-2.23)	(1.35)
10	0.288	**0.671**	-0.323	0.173	0.125	-0.320	0.095	-0.204	0.120	-0.449	0.236	-0.159
	(1.17)	**(2.54)**	(-1.27)	(0.67)	(0.57)	(-1.36)	(0.41)	(-0.90)	(0.45)	(-1.82)	(0.89)	(-0.64)
11	0.447	-0.059	0.427	0.204	-0.324	0.142	-0.086	0.274	-0.534	0.404	-0.294	-0.130
	(1.80)	(-0.24)	(1.68)	(0.87)	(-1.38)	(0.56)	(-0.36)	(1.07)	(-2.40)	(1.61)	(-1.21)	(-0.47)
k = 12	-0.230	0.265	0.148	-0.316	0.103	-0.056	0.042	-0.255	0.236	0.023	-0.170	0.234
	(-0.94)	(1.02)	(0.67)	(-1.33)	(0.42)	(-0.22)	(0.16)	(-1.12)	(0.96)	(0.10)	(-0.62)	(0.90)

Value-weighted

Leader epsilon (t-k)	Laggard (t+h)											
	h = 1	2	3	4	5	6	7	8	9	10	11	h = 12
k = 1	-0.114	0.233	0.519	-0.106	0.686	-0.833	0.501	0.777	0.539	0.032	0.408	-0.432
	(-0.31)	(0.59)	(1.24)	(-0.24)	(1.70)	(-1.86)	(1.06)	(1.68)	(1.22)	(0.07)	(0.88)	(-0.99)
2	0.560	0.615	0.008	**0.925**	-0.949	0.356	0.484	0.286	-0.524	0.538	-0.608	0.193
	(1.47)	(1.51)	(0.02)	**(2.22)**	(-2.21)	(0.78)	(0.94)	(0.65)	(-1.25)	(1.19)	(-1.32)	(0.42)
3	0.600	0.132	0.444	-0.784	-0.171	0.828	0.166	-0.183	-0.168	-0.219	0.232	0.688
	(1.47)	(0.33)	(1.15)	(-1.78)	(-0.38)	(1.83)	(0.41)	(-0.42)	(-0.35)	(-0.45)	(0.50)	(1.54)
4	0.021	0.322	-0.419	0.194	**0.988**	-0.081	-0.590	0.637	-0.655	-0.121	0.583	0.413
	(0.05)	(0.85)	(-1.16)	(0.44)	**(2.22)**	(-0.19)	(-1.33)	(1.39)	(-1.57)	(-0.28)	(1.28)	(1.20)
5	0.252	-0.607	-0.001	0.486	0.223	-0.551	0.271	-0.519	-0.209	0.714	0.334	-0.690
	(0.63)	(-1.47)	(-0.00)	(1.08)	(0.53)	(-1.28)	(0.60)	(-1.22)	(-0.51)	(1.60)	(0.92)	(-1.56)
6	-0.856	-0.002	0.441	0.551	-0.645	0.517	-0.411	0.474	0.460	**0.750**	-0.576	0.355
	(-2.09)	(-0.00)	(0.97)	(1.26)	(-1.46)	(1.19)	(-0.97)	(1.07)	(1.10)	**(2.02)**	(-1.35)	(0.94)
7	0.356	0.517	0.237	-1.008	**1.276**	-0.320	-0.061	**0.929**	0.271	-0.414	0.418	0.690
	(0.81)	(1.16)	(0.55)	(-2.14)	**(2.79)**	(-0.81)	(-0.14)	**(2.25)**	(0.72)	(-0.91)	(0.96)	(1.63)
8	0.758	0.109	-0.845	0.663	-0.419	0.185	0.501	0.146	-0.494	-0.108	0.096	0.260
	(1.62)	(0.26)	(-1.85)	(1.45)	(-1.05)	(0.46)	(1.15)	(0.37)	(-1.16)	(-0.25)	(0.22)	(0.67)
9	-0.062	-0.272	0.486	-0.430	-0.720	0.256	-0.361	-0.767	-0.536	-0.390	-0.560	0.480
	(-0.14)	(-0.64)	(1.21)	(-1.05)	(-1.69)	(0.61)	(-0.98)	(-1.90)	(-1.34)	(-0.89)	(-1.38)	(1.11)
10	0.103	**0.854**	-0.063	-0.556	0.687	-0.164	-0.795	-0.746	-0.573	-0.603	0.397	-0.093
	(0.26)	**(2.07)**	(-0.15)	(-1.27)	(1.59)	(-0.40)	(-2.04)	(-1.91)	(-1.32)	(-1.49)	(0.83)	(-0.22)
11	**1.031**	0.154	0.004	0.354	0.082	-0.375	-0.272	0.019	-0.705	0.765	-0.063	-0.363
	(2.26)	(0.35)	(0.01)	(0.87)	(0.20)	(-0.91)	(-0.62)	(0.04)	(-1.78)	(1.62)	(-0.15)	(-0.81)
k = 12	-0.449	-0.177	0.489	-0.214	-0.639	0.085	-0.035	-0.193	0.708	0.073	-0.402	-0.093
	(-1.05)	(-0.38)	(1.03)	(-0.53)	(-1.51)	(0.20)	(-0.08)	(-0.51)	(1.54)	(0.18)	(-0.90)	(-0.23)

Table O-2 continued

Panel B: Standalone-conglomerates

Equal-weighted

Leader epsilon (t-k)	Laggard (t+h) h = 1	2	3	4	5	6	7	8	9	10	11	h = 12
k = 1	0.083	**0.395**	0.247	0.081	**0.538**	-0.193	0.113	**0.400**	0.357	0.180	0.142	0.291
	(0.34)	**(2.03)**	(1.04)	(0.33)	**(2.40)**	(-0.84)	(0.64)	**(2.28)**	(1.83)	(0.85)	(0.71)	(1.62)
2	0.402	0.284	0.305	0.411	-0.050	0.342	**0.590**	0.307	-0.105	0.100	0.295	0.139
	(1.94)	(1.30)	(1.42)	(2.00)	(-0.24)	(1.54)	**(2.91)**	(1.11)	(-0.43)	(0.51)	(1.51)	(0.70)
3	0.142	0.306	0.215	-0.171	0.216	**0.511**	**0.424**	0.086	-0.117	0.391	-0.003	-0.146
	(0.65)	(1.74)	(0.94)	(-0.81)	(1.19)	**(2.77)**	**(2.21)**	(0.32)	(-0.47)	(1.84)	(-0.02)	(-0.65)
4	0.149	0.369	-0.297	-0.060	**0.512**	**0.534**	0.027	0.006	0.098	-0.001	-0.017	-0.136
	(0.80)	(1.57)	(-1.53)	(-0.31)	**(2.74)**	**(2.81)**	(0.07)	(0.02)	(0.34)	(-0.00)	(-0.09)	(-0.70)
5	**0.579**	-0.225	0.069	0.382	0.406	-0.073	-0.052	0.213	0.081	-0.274	0.115	-0.024
	(2.62)	(-1.11)	(0.28)	(1.92)	(1.96)	(-0.25)	(-0.15)	(0.87)	(0.42)	(-1.13)	(0.53)	(-0.14)
6	-0.108	-0.058	0.344	0.284	-0.155	-0.269	0.197	0.148	0.031	0.031	-0.023	-0.076
	(-0.54)	(-0.25)	(1.90)	(1.28)	(-0.69)	(-0.95)	(0.99)	(0.66)	(0.13)	(0.15)	(-0.13)	(-0.40)
7	-0.009	0.212	0.247	-0.081	0.087	0.354	0.253	0.211	0.001	0.152	-0.153	0.150
	(-0.04)	(1.15)	(1.25)	(-0.39)	(0.44)	(1.77)	(1.30)	(1.01)	(0.01)	(0.81)	(-0.72)	(0.78)
8	0.174	0.229	0.182	0.068	0.206	0.197	0.334	-0.005	-0.018	-0.050	-0.030	0.349
	(0.87)	(1.20)	(0.94)	(0.37)	(1.10)	(0.98)	(1.61)	(-0.02)	(-0.10)	(-0.26)	(-0.11)	(1.82)
9	0.054	0.091	-0.002	0.014	0.139	0.234	-0.012	0.034	-0.226	-0.037	0.055	-0.016
	(0.22)	(0.44)	(-0.01)	(0.06)	(0.72)	(0.83)	(-0.05)	(0.18)	(-1.09)	(-0.16)	(0.29)	(-0.06)
10	0.106	0.027	0.228	0.039	0.092	0.038	0.192	-0.255	-0.012	0.120	0.028	-0.044
	(0.53)	(0.14)	(1.13)	(0.21)	(0.45)	(0.19)	(1.04)	(-0.90)	(-0.03)	(0.61)	(0.16)	(-0.24)
11	0.144	0.378	-0.376	0.154	-0.002	0.265	-0.004	0.066	0.083	0.025	0.059	-0.022
	(0.75)	(1.93)	(-1.85)	(0.90)	(-0.01)	(1.38)	(-0.02)	(0.34)	(0.39)	(0.14)	(0.29)	(-0.11)
k = 12	0.483	-0.200	-0.044	-0.196	0.090	-0.074	0.215	0.188	0.056	0.018	0.018	0.018
	(2.52)	(-0.87)	(-0.22)	(-1.05)	(0.43)	(-0.42)	(1.03)	(0.94)	(0.28)	(0.09)	(0.09)	(0.09)

Value-weighted

Leader epsilon (t-k)	Laggard (t+h) h = 1	2	3	4	5	6	7	8	9	10	11	h = 12
k = 1	0.046	0.382	0.168	0.286	0.574	-0.261	0.161	0.290	0.511	-0.120	-0.282	0.197
	(0.15)	(1.37)	(0.54)	(0.93)	(1.81)	(-0.82)	(0.62)	(1.10)	(1.79)	(-0.40)	(-0.95)	(0.70)
2	0.413	0.034	0.015	0.276	0.081	0.341	0.539	0.232	-0.204	-0.380	0.228	0.042
	(1.50)	(0.11)	(0.05)	(0.94)	(0.29)	(1.14)	(1.81)	(0.68)	(-0.60)	(-1.33)	(0.80)	(0.14)
3	0.048	0.077	0.171	0.243	0.206	0.563	0.422	-0.129	-0.327	0.155	-0.064	-0.102
	(0.16)	(0.29)	(0.57)	(0.92)	(0.79)	(1.96)	(1.63)	(-0.39)	(-1.03)	(0.50)	(-0.22)	(-0.32)
4	0.026	0.505	-0.163	-0.099	0.459	0.491	-0.211	-0.113	0.078	-0.165	0.005	-0.071
	(0.10)	(1.61)	(-0.62)	(-0.40)	(1.69)	(1.88)	(-0.47)	(-0.27)	(0.20)	(-0.57)	(0.02)	(-0.26)
5	0.471	-0.164	0.292	0.238	0.305	0.014	-0.297	0.491	-0.060	-0.209	0.042	-0.060
	(1.59)	(-0.59)	(1.01)	(0.87)	(1.05)	(0.04)	(-0.74)	(1.41)	(-0.21)	(-0.70)	(0.15)	(-0.23)
6	0.264	0.165	0.548	-0.048	-0.141	-0.272	**0.564**	-0.025	-0.030	0.014	0.056	0.347
	(0.95)	(0.56)	(2.00)	(-0.17)	(-0.49)	(-0.83)	**(2.01)**	(-0.09)	(-0.09)	(0.05)	(0.20)	(1.25)
7	0.192	0.385	0.115	-0.322	-0.020	**0.630**	-0.081	0.121	-0.239	0.010	-0.031	0.320
	(0.65)	(1.33)	(0.41)	(-1.18)	(-0.07)	**(2.22)**	(-0.29)	(0.43)	(-0.91)	(0.04)	(-0.10)	(1.22)
8	0.358	0.295	0.187	-0.052	0.015	-0.234	0.230	0.004	-0.163	0.178	0.012	0.363
	(1.26)	(1.13)	(0.66)	(-0.21)	(0.05)	(-0.81)	(0.76)	(0.01)	(-0.60)	(0.62)	(0.04)	(1.27)
9	0.120	-0.131	-0.008	-0.169	-0.344	0.441	-0.159	0.057	-0.150	0.200	0.312	0.141
	(0.40)	(-0.49)	(-0.03)	(-0.49)	(-1.25)	(1.36)	(-0.57)	(0.22)	(-0.53)	(0.65)	(1.20)	(0.42)
10	-0.098	0.018	0.146	-0.042	0.053	-0.169	0.261	-0.009	0.148	0.273	0.286	0.248
	(-0.36)	(0.07)	(0.51)	(-0.16)	(0.19)	(-0.66)	(0.99)	(-0.03)	(0.37)	(0.97)	(1.14)	(0.99)
11	0.247	0.354	-0.539	0.210	-0.197	0.310	-0.101	0.144	0.054	0.188	0.117	-0.005
	(0.95)	(1.29)	(-1.73)	(0.80)	(-0.78)	(1.13)	(-0.35)	(0.47)	(0.19)	(0.67)	(0.44)	(-0.02)
k = 12	0.505	-0.507	0.183	-0.622	0.040	0.025	0.152	0.429	0.215	-0.026	-0.026	-0.026
	(1.88)	(-1.66)	(0.67)	(-2.47)	(0.14)	(0.09)	(0.53)	(1.50)	(0.73)	(-0.09)	(-0.09)	(-0.09)

Table O-2 continued

Panel C: Alliance-linked firms

Equal-weighted

Leader epsilon (t-k)	Laggard (t+h) h = 1	2	3	4	5	6	7	8	9	10	11	h = 12
k = 1	0.070	**0.454**	0.111	-0.151	**0.591**	0.066	0.158	0.190	0.216	0.052	0.127	0.075
	(0.36)	**(2.54)**	(0.54)	(-0.86)	**(3.05)**	(0.35)	(0.86)	(0.79)	(1.21)	(0.26)	(0.67)	(0.36)
2	0.324	0.128	-0.139	**0.485**	0.114	0.160	0.140	0.303	0.025	0.084	0.063	0.233
	(1.75)	(0.59)	(-0.75)	**(2.36)**	(0.61)	(0.80)	(0.57)	(1.58)	(0.13)	(0.44)	(0.31)	(1.34)
3	0.235	0.063	0.314	-0.072	0.087	0.182	0.179	0.188	-0.119	0.108	0.342	-0.116
	(1.10)	(0.37)	(1.59)	(-0.35)	(0.45)	(0.83)	(0.92)	(0.98)	(-0.62)	(0.54)	(1.80)	(-0.59)
4	-0.060	0.303	0.077	0.333	0.282	0.032	0.025	0.092	0.012	0.214	-0.123	0.040
	(-0.32)	(1.49)	(0.46)	(1.68)	(1.24)	(0.17)	(0.13)	(0.47)	(0.06)	(1.09)	(-0.67)	(0.21)
5	0.371	-0.008	0.368	0.235	0.066	0.163	0.095	0.004	0.235	-0.196	-0.101	-0.345
	(1.80)	(-0.04)	(1.96)	(1.00)	(0.36)	(0.82)	(0.50)	(0.02)	(1.15)	(-1.03)	(-0.51)	(-1.78)
6	-0.127	0.073	0.269	0.155	0.014	-0.013	0.078	0.204	-0.238	-0.089	-0.283	0.218
	(-0.66)	(0.36)	(1.21)	(0.84)	(0.07)	(-0.07)	(0.39)	(0.96)	(-1.17)	(-0.42)	(-1.43)	(1.14)
7	0.022	0.295	0.124	0.033	0.024	0.259	0.187	-0.169	-0.205	-0.244	0.100	0.346
	(0.10)	(1.37)	(0.68)	(0.16)	(0.12)	(1.31)	(1.00)	(-0.78)	(-1.04)	(-1.28)	(0.50)	(1.91)
8	0.374	-0.062	0.243	0.014	0.091	0.235	-0.161	-0.039	-0.289	0.020	0.077	0.282
	(1.64)	(-0.32)	(1.30)	(0.07)	(0.48)	(1.25)	(-0.80)	(-0.17)	(-1.42)	(0.09)	(0.43)	(1.44)
9	-0.077	0.363	-0.172	0.085	0.340	-0.164	-0.194	-0.281	0.066	0.012	0.118	0.084
	(-0.38)	(1.82)	(-0.94)	(0.41)	(1.76)	(-0.81)	(-0.88)	(-1.40)	(0.32)	(0.07)	(0.63)	(0.39)
10	0.149	0.002	0.139	0.286	-0.072	-0.076	-0.371	-0.093	-0.115	0.077	0.024	0.294
	(0.76)	(0.01)	(0.73)	(1.48)	(-0.37)	(-0.33)	(-1.81)	(-0.47)	(-0.64)	(0.42)	(0.11)	(1.46)
11	0.092	0.098	0.206	-0.219	0.101	-0.222	-0.154	-0.221	-0.106	-0.178	-0.034	-0.009
	(0.48)	(0.47)	(1.04)	(-1.20)	(0.47)	(-1.04)	(-0.74)	(-1.22)	(-0.49)	(-0.80)	(-0.18)	(-0.04)
k = 12	0.191	0.217	-0.185	-0.039	-0.226	-0.163	-0.016	0.248	-0.226	-0.232	-0.255	-0.051
	(0.93)	(0.97)	(-0.97)	(-0.18)	(-1.01)	(-0.75)	(-0.08)	(1.17)	(-1.00)	(-1.05)	(-1.11)	(-0.21)

REFERENCES

Cao, Jie, Tarun Chordia, and Chen Lin, 2016, Alliances and return predictability, *Journal of Financial and Quantitative Analysis* 51, 1689–1717.

Carhart, Mark M, 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Cohen, Lauren, and Andrea Frazzini, 2008, Economic links and predictable returns, *Journal of Finance* 63, 1977–2011.

Cohen, Lauren, and Dong Lou, 2012, Complicated firms, *Journal of Financial Economics* 104, 383–400.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Jegadeesh, Narasimhan, and Sheridan Titman, 2002, Cross-sectional and time-series determinants of momentum returns, *Review of Financial Studies* 15, 143–157.

Keloharju, Matti, Juhani T Linnainmaa, and Peter Nyberg, 2019, Long-term discount rates do not vary across firms, *Working Paper* .

Lewellen, Jonathan, and Stefan Nagel, 2006, The conditional capm does not explain asset-pricing anomalies, *Journal of Financial Economics* 82, 289–314.

Lo, Andrew W., and A. Craig MacKinlay, 1990, When are contrarian profits due to stock market overreaction?, *Review of Financial Studies* 3, 175–205.

Pástor, Ľuboš, and Robert F Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.