Intermediation Frictions in Equity Markets

Bryan Seegmiller*

March 2025

Abstract

Stocks with similar characteristics but different levels of ownership by financial institutions have returns and risk premia that comove very differently with shocks to the risk-bearing capacity of dealer banks. After observable stock characteristics are accounted for, excess returns on more intermediated stocks have higher betas on contemporaneous shocks to intermediary willingness to take risk and are more predictable by state variables that proxy for intermediary health. Intermediary risk-bearing capacity also explains a substantial and increasing fraction of the variation in conditional risk premia for portfolios sorted on intermediation. These effects are concentrated in stocks held by hedge funds or mutual fund investors who are more likely to be exposed to dealer banks. The empirical evidence supports the predictions of asset pricing models in which financial intermediaries are marginal investors but face frictions that induce changes in their risk-bearing capacity. This suggests that these models are useful in explaining price movements not only in markets for complex financial assets but also in asset classes in which households face comparatively low barriers to direct participation.

*Northwestern University, Kellogg School of Management. Nikolai Roussanov was the editor for this article. I would like to thank the editor and an anonymous reviewer for comments that improved the paper; Adrien Verdelhan, Thummim Cho (discussant), Hui Chen, Larry Schmidt, Leonid Kogan, Jonathan Parker, Tyler Muir, David Thesmar, and participants at the Midwest Finance Association and MIT Sloan Finance Lunch Seminar for their helpful comments and suggestions; and Leonid Kogan and Vikas Agarwal for generously sharing data.

1 Introduction

Recent empirical evidence has confirmed many predictions of asset pricing models featuring frictions and sophisticated financial intermediaries as the primary marginal investors. This has been particularly true for complex asset classes, which are more difficult for households to access (Eisfeldt, Lustig, and Zhang, 2023).[1] Empirical measures based on theories of frictional intermediary-based asset pricing models have connected the health of the financial sector to security broker-dealers (Adrian, Etula, and Muir, 2014) and Federal Reserve primary dealer banks (He, Kelly, and Manela, 2017), each with considerable success in pricing diverse cross sections of assets. However, as both Adrian et al. (2014) and He et al. (2017) point out, because of the relative ease of household stock market participation, the comparative importance of these models and their empirical proxies in actually *causing* stock price movements remains unclear. Despite the success of such intermediary-based empirical asset pricing models in explaining cross sections of returns on stocks and other asset classes, such tests cannot rule out that a household-based pricing kernel also holds for stocks because households also participate heavily in equity markets, both directly and indirectly, alongside financial institutions.

When households and institutional investors have differential preferences for direct holding of certain stocks for any reason unrelated to the true distributions of future cashflows, whether because of heterogeneous beliefs or differential trading costs, intermediation dispersion independent of asset fundamentals can arise naturally in the cross section of equities even when households are not prevented from trading directly. This dispersion causes endogenous segmentation *within* the equity asset class rather than *between* equities and, say, credit defaults swaps or collateralized debt obligations. Accordingly, the types of empirical tests that have been used to detect a causal role for intermediary frictions in explaining asset price movements for complex assets can also be applied to detect their importance within equity markets—the single asset class where intermediaries seem likeliest to act as a veil that simply passes on household preference. The basic prediction is that for two otherwise similar assets, the more intermediated asset exhibits a larger contemporaneous price response and variation in conditional risk premia due to shifts in intermediary risk-bearing capacity.

I find evidence strongly in support of this prediction within the equity asset class. After firm characteristics are accounted for, excess returns on stocks that are held more intensively

[1]These markets include those for credit default swaps (Siriwardane, 2019; Mitchell and Pulvino, 2012), mortgage-backed securities (Krishnamurthy, 2010; Gabaix, Krishnamurthy, and Vigneron, 2007; Diep, Eisfeldt, and Richardson, 2021), foreign exchange (Du, Tepper, and Verdelhan, 2018; Du, Hébert, and Huber, 2022), convertible bonds (Mitchell, Pedersen, and Pulvino, 2007), corporate bonds (He, Khorrami, and Song, 2022), life insurance (Koijen and Yogo, 2015), and even treasuries (Haddad and Sraer, 2018; Anderson and Liu, 2018), to name just a few examples.

by the largest and most active institutional investors in equity markets (mutual funds, hedge funds, and other investment advisors) covary more with theoretically motivated empirical proxies for shocks that shift intermediary sector risk-bearing capacity due to financial frictions. This finding holds for both portfolios and individual stocks and in terms of both the amount of contemporaneous return comovement and relative ex ante return predictability.

My main empirical proxy for financial sector frictions combines two influential empirical measures of shocks to an intermediary sector marginal utility proposed in the literature: Adrian et al. (2014) propose shocks to broker-dealer book leverage, while He et al. (2017) use shocks to the market equity capital ratio of Federal Reserve primary dealer bank holding companies. The former is a proxy for models of margin constraints that become more binding during a liquidity crisis (Brunnermeier and Pedersen, 2008; Adrian and Boyarchenko, 2012), while the latter captures the effects of models where a "skin-in-the-game" equity constraint due to a moral hazard problem leads to elevated risk premia in times of distress (He and Krishnamurthy, 2012, 2013; Brunnermeier and Sannikov, 2014). I simply standardize both of these measures and take the average of the two, similarly to Haddad and Muir (2021), to capture a powerful composite measure of shocks to the financial sector.

I find that stocks sorted on a measure of intermediation that holds stock fundamentals constant have monotonically increasing contemporaneous comovement with my measure of intermediary capital shocks: A one-standard-deviation positive shock to my combined intermediary factor increases the returns for stocks in the top-quintile portfolio of my intermediation measure by approximately 5.2% (annualized) but by only 1.1% for those in the lowest-quintile portfolio. The t-statistic on the spread in return responses between the top and bottom portfolios is 5.51, and there is a monotonic increase in stock return responses as intensity of intermediary holdings increases across all portfolios.

I further include tests with the He et al. (2017) and Adrian et al. (2014) measures separately and show that both independently yield essentially the same monotonic sorting patterns of increased comovement as intermediation rises, despite not being highly correlated with one another. An additional theoretically motivated credible proxy for shocks to financial sector risk-bearing capacity—the excess return on the financial sector—also displays the same empirical pattern of increased stock return exposure to shocks to intermediary risk-bearing capacity along the dimension of increased intermediation.

Supporting my baseline portfolio-level analysis of contemporaneous comovement, I also utilize a natural experiment from S&P 500 membership to test how stock comovement with the He et al. (2017) intermediary capital factor changes before and after index inclusion. Echoing prior work (Aghion, Van Reenen, and Zingales, 2013; Boller and Scott Morton, 2020), I confirm that institutional holdings rise significantly between the quarter before and the quarter after index inclusion. Stocks recently added to the S&P 500 experience a large and highly statistically significant increase in their betas on the He et al. (2017) intermediary capital factor relative to comparison stocks in the time period, corroborating my findings that otherwise similar stocks that institutions tend to hold more heavily also covary more with shocks to the capitalization of key intermediaries.

Tests of ex ante stock return predictability further reveal that discount rates on highly intermediated stocks respond more to changes in intermediary risk-bearing capacity, which is a fundamental feature of intermediary asset pricing models with some degree of market segmentation. This basic mechanism is inspired by numerous papers from the intermediary asset pricing literature.[2] A state variable that proxies for current financial sector constraints predicts larger ex ante risk premia for otherwise similar stocks that are more intermediated. I combine information from two state variables in my predictability tests: the squared market leverage ratio of Federal Reserve primary dealer bank holding companies and the level of the book leverage ratio of broker-dealers obtained from the flow-of-funds account. I again take the average of the standardized versions of these two ex ante state variables as my main proxy for intermediary risk aversion at the current date.[3] These measures perform in accordance with theory when included separately, and another theoretically motivated proxy for financial sector health—the financial sector stock market wealth share—also predicts returns more strongly among more intermediated stocks.

In my return predictability tests, I control for potential variation in conditional risk premia coming from sources outside of the intermediaries' health by constructing a powerful composite conditional risk premium proxy using the joint information found in a host of return predictor variables discovered in the literature. I do this following Kelly and Pruitt (2013) and Huang, Jiang, Tu, and Zhou (2014) in using partial least squares (PLS) to aggregate the information in many individual variables into the components most informative about future aggregate stock returns. Using this control to bound the share of total predictability that can be attributed directly to intermediaries, intermediary risk-bearing capacity proves to be a strong predictor of conditional risk premia in general: My proxy for the (lack of) current intermediary risk-bearing capacity predicts between 23% and 44% of

[2]Such implications are discussed in more detail in section 2, where I present a simple model in which intermediary risk tolerance can shift because of shocks to some underlying state variables (implicitly because they cause financial constraints on intermediaries to be more binding).

[3]As He et al. (2017) point out, their model implies that squared primary dealer leverage is a predictor of stock returns, while the theoretical motivation for the Adrian et al. (2014) measure implies that the level of broker-dealer leverage is a state variable that should also predict returns. The different particular theoretical motivations imply that these two state variables should predict stock market returns with opposite sign. Hence, when I take the average, I use the negative of the broker-dealer leverage ratio so that the composite measure predicts returns with a positive sign.

the variation in the conditional equity risk premium for the least intermediated portfolio, between 38% and 59% for the most intermediated portfolio, and nearly the entirety of the *relative* risk premium variation between the two portfolios. The share of risk premia variation coming from intermediary risk aversion is also monotonically increasing from low- to high-intermediation portfolios. Previous work (Haddad and Muir, 2021) finds that a large portion of the return predictability in equity markets cannot be directly attributed to either households or intermediaries; my estimates suggest that the bulk of this unexplained variation in fact comes from changes in the risk-bearing capacity of central financial institutions. My predictability tests also relate to work by Weber (2023), who finds that stocks with more institutional ownership have more time-series predictability by price–dividend ratios; I differ by connecting the relative time-series predictability of intermediated portfolios directly to a measure of time-varying financial sector risk-bearing capacity and demonstrate that this measure has considerable incremental information for the relative return predictability of highly intermediated assets over a composite of many powerful return predictor variables.

I include an additional test confirming a feature in the cross section of return predictability that is consistent with theory, though it is not explicitly laid out in my stylized static model. I find that the predictability for the return spread between high- and low-intermediation portfolios is positive but declining with the time horizon of the monthly returns being predicted and the R^2 is also decreasing with the time horizon. This suggests shocks to intermediaries induce temporary distortions in relative discount rates between more and less intermediated stocks, with such distortions reverting over time as intermediary capital recovers. This is consistent with theoretical mechanisms highlighted in Duffie (2010) and Gromb and Vayanos (2018), for example.

The proxies for intermediary risk tolerance shocks proposed by He et al. (2017) and Adrian et al. (2014) focus on a particularly influential set of levered institutions—namely, dealer banks and other broker-dealers—that have been argued in the literature to occupy a place of central importance in financial markets and as marginal investors in pricing numerous asset classes; however, they are not necessarily the set of institutions most directly active in equity markets and so are not the exact same set of institutions whose stock holdings I capture using 13F data (though there is some overlap via internal capital markets). Hence, the empirical evidence I present in section 4 implies that shocks to dealer banks and broker-dealers pass through to the risk-bearing capacity of the mutual funds, hedge funds, and other investment advisors whose stock holdings are the focus of my analysis.

In section 5.1, I show that my findings concentrate in the subset of institutional investors whose fortunes are likeliest to be interlinked with those of dealer banks and broker-dealers. In particular, these patterns are especially evident when I restrict the analysis to hedge funds using a list of 13F hedge fund managers.[4] They are also prominent for mutual funds and non–hedge fund investment advisors, but only when they are exposed indirectly via interactions with dealer-exposed hedge funds in equity markets, or when they also invest in bond markets where dealer banks play an especially prominent direct intermediation role (as established by Haddad and Muir (2021); He et al. (2022); Li and Xu (2024), for example).

Since dealer banks play a key role in the direct provision of funding liquidity, shocks to dealer banks' capitalization can lead to dry-ups in market liquidity among more intermediated stocks.[5] Consistent with this mechanism, I also show in section 5.2 that positive shocks to dealer banks and broker-dealers' risk-bearing capacity predict larger improvements in average stock-level liquidity (as measured by the Amihud (2002) illiquidity index) among highly intermediated stocks, but again only for the subset of hedge funds, mutual funds, and investment advisors whose fortunes are likeliest to be interlinked with those of dealer banks/broker-dealers.

A prior literature documents the importance of non-fundamental pressure driven by mutual fund flows (Coval and Stafford, 2007; Lou, 2012; Frazzini and Lamont, 2008; Dou, Kogan, and Wu, 2023). The intermediary mechanisms I focus on in this paper are distinct from these fund flow channels, which I demonstrate by showing in section 5.3 that common shocks to mutual fund flows from Dou et al. (2023) can say little about the relative price movements of portfolios sorted on intermediation and controlling for them has no effect on my findings. Instead, my estimates are consistent with models where central financial intermediaries face capital and funding constraints that transmit across financial markets to cause movements in asset prices. While evidence for such mechanisms has been readily demonstrated in complex asset markets (as I detail later in the literature review), a primary contribution of this paper is to demonstrate how these frictions also spill over to generate differential price movements in equity markets—perhaps the least intermediated asset class (Haddad and Muir, 2021).

I explain the theoretical backdrop to my empirical strategy in a simple economic setting introduced in section 2. The model shows that if households are relatively more willing to hold one asset for any reason unrelated to the true distribution of cash flows, assets that are less preferred by households become more intermediated and have risk premia that respond more to shocks to the intermediaries' risk tolerance. In my setting, this increased intermediary willingness to hold certain assets arises because households either have heterogeneous expectations errors or view direct investing in certain assets as relatively more or less costly.

[4] This list of hedge fund managers has been used in a series of papers studying risks and return patterns in hedge fund holdings, including Agarwal, Ruenzi, and Weigert (2024) Agarwal, Ruenzi, and Weigert (2017), Agarwal, Fos, and Jiang (2013), and Agarwal, Jiang, Tang, and Yang (2013). I thank Vikas Agarwal for generously sharing these data.

[5] See Brunnermeier and Pedersen (2008) for a discussion of such liquidity spirals.

The empirical implication is that relatively more intermediated assets that are similar on fundamentals should have prices that move more with contemporaneous intermediary shocks and possess time-varying risk premia that are relatively more predictable by state variables representing intermediary risk tolerance.

The paper proceeds as follows: In section 2, I detail my stylized model; in section 3, I describe the data and sampling criteria. In section 4, I detail my empirical strategy and present my empirical findings. Specifically, in subsection 4.1, I explain in more detail the construction of my stock-level intermediation measure and why the intermediary shocks suggested by He et al. (2017) and Adrian et al. (2014) can be considered proxies for changes in financial risk-bearing capacity; in sections 4.2 through 4.5, I present my main empirical estimates, including those from my portfolio- and stock-level analysis and robustness checks. Section 5 explores the mechanisms driving my empirical findings; finally, section 6 provides some short concluding remarks. Before proceeding with all of the above, I briefly provide more detail on the related literature and my paper's contributions.

1.1 Related Literature

Cross-sectional and time-series asset pricing tests find a role for intermediaries in explaining variation in expected returns.[6] In the most related prior work, Haddad and Muir (2021) point out that these tests are not sufficient to establish a unique causal role for moving prices, and they construct empirical tests designed to detect whether intermediaries matter for asset price movements or if they act merely as a veil in passing on household preferences. Their estimates imply that intermediation frictions do matter, especially in credit default swap, foreign exchange, commodities, and sovereign bond markets. On the other hand, they argue that equities are the asset class for which price movements coming from intermediation frictions are least likely to be detected (while the authors are clear that they also cannot rule this out). Relative to the approach of Haddad and Muir (2021), who focus on the relative time-series predictability of proxies for intermediary frictions across asset classes, my model and empirical approach emphasize the duality of contemporaneous comovement and time-series predictability in establishing a role for intermediaries for moving prices within an asset class. In doing so, I uncover a prominent role for intermediaries in equity markets and thus directly complement and add to their cross–asset class comparisons.

Others in the intermediary asset pricing literature have also expressed skepticism about the relevance of these theories in explaining price movements in equity markets. While He et al. (2017) find that their proxy for a representative intermediary stochastic discount factor

[6] See Adrian et al. (2014), He et al. (2017), Kargar (2021), and Ma (2017) for cross-sectional tests and Muir (2017) and Chen, Joslin, and Ni (2018) for evidence on time-series predictability.

performs reasonably well in describing cross sections of equity returns, they also argue that equity may be the asset class where intermediaries may act as a veil in merely passing through the preferences of households in equity markets. Similarly, in the theoretical literature, He and Krishnamurthy (2013) think of their model in the context of complex asset markets such as mortgage-backed securities as opposed to equities. While Koijen and Yogo (2019) estimate a characteristics-based demand system for heterogeneous financial intermediaries in equity markets, they do not attempt to test how their findings relate to friction-based intermediary asset pricing.[7]

In a related paper, Cho (2020) finds that anomaly portfolios with a higher arbitrage position (determined by abnormally high/low short interest in a stock) also have higher betas on shocks to the Adrian et al. (2014) leverage factor in the post-1993 period when hedge funds became more active in equity markets. I also focus on equity markets, but I consider the holdings of a much larger class of financial institutions and for multiple definitions of intermediary shocks, analyze effects at both the portfolio and individual stock levels, and include contemporaneous and predictive tests using shocks to and levels of the state variables implied by intermediary asset pricing models.

While Cho (2020) contextualizes his findings within the Kondor and Vayanos (2019) setting of minimal frictions, I prefer the friction-based interpretation since the empirical measures proposed by Adrian et al. (2014) and He et al. (2017) are constructed to proxy for mechanisms outlined in friction-based models—Brunnermeier and Pedersen (2008) and Adrian and Boyarchenko (2012) in the case of Adrian et al. (2014) and He and Krishnamurthy (2013) and Brunnermeier and Sannikov (2014) in the case of He et al. (2017). In Adrian et al. (2014), the underlying friction comes from time-varying margin constraints, while in He et al. (2017), the friction entails an equity capital constraint imposed by investors on the equity of the intermediary because of moral hazard problems in delegation to professional asset managers. These frictions naturally lead to time-varying intermediary risk-bearing capacity, which is the key mechanism I focus on in my model to derive the predictions that I test in the data.

In addition to its primary connection with the theoretical and empirical literature in intermediary asset pricing, this paper has connections with research areas such as the limits to arbitrage (Shleifer and Vishny, 1997; Duffie, 2010), the effects of institutional ownership on asset prices (Gompers and Metrick, 2001; Nagel, 2005; Basak and Pavlova, 2013), and other sources of nonfundamental price pressure such as mutual fund flows (Coval and Stafford, 2007; Frazzini and Lamont, 2008; Lou, 2012). These papers all deal with price dislocations

[7] I draw from the set of stock characteristics that they use to create my primary measure of stock-level intermediation (as detailed in section 4).

due to institutional features that allow the direction of trading or asset holding to move prices even absent fundamental information. Relative to prior work in these research areas, my finding that shocks to broker-dealers and dealer banks cause larger price movements among more intermediated stocks is a distinctly new mechanism.

2 Economic Framework for Empirical Tests

Theories linking asset price movements to intermediary health are broadly divided into equity constraint models where financial constraints bind when intermediaries' net worth is low[8] and another a class of models where constraints explicitly limit the amount of leverage or risk that intermediaries can take on.[9] To set the stage for my empirical tests, I present a simple model that takes the middle ground between these two broad classes of intermediary asset pricing models by allowing risk-bearing capacity to vary because of underlying state variables, which could be proxies for net worth shocks or changes in leverage/margin constraints. The intended interpretation is that these shifts in willingness to take on risk come from constraints that exist because of underlying agency frictions in delegation to intermediaries, which is a unifying theme in these models. The intuition and consequent empirical implications in this section draw from the models of He and Krishnamurthy (2018), Haddad and Muir (2021), and Koijen and Yogo (2019).

Here, I present a simple economic setting that leads to the empirical specification for the residual intermediation measure that I use to form portfolios and delivers clear predictions about differential asset price responses to intermediary shocks across these portfolios. As in He and Krishnamurthy (2018), Kondor and Vayanos (2019), Haddad and Muir (2021), Koijen, Richmond, and Yogo (2023), and He et al. (2022), agents have mean–variance preferences in final wealth, arising here from constant absolute risk aversion utility and normally distributed asset payoffs. There are two investors, a representative institutional investor/intermediary and a representative household, indexed by I and H, respectively. Intermediaries and households have respective risk tolerance ρ_I and ρ_H, which may be dependent on current state variables—for example, intermediary asset pricing models suggest that ρ_I should be increasing in intermediary wealth and decreasing in Lagrange multipliers on leverage, value-at-risk, or margin constraints.[10] While I follow suit by modeling ρ_I to

[8] See Bernanke and Gertler (1989) and Holmstrom and Tirole (1997) for early examples and Brunnermeier and Sannikov (2014) and He and Krishnamurthy (2013) for more recent works.

[9] Examples from this literature include Brunnermeier and Pedersen (2008), Adrian and Shin (2014), and Gârleanu, Panageas, and Yu. (2015).

[10] Additionally, as Koijen et al. (2023) and Makarov and Schornick (2010) explain, parameterizing agents' risk tolerances to be dependent on their initial wealth is particularly useful since it preserves the tractability of a setting with constant absolute risk aversion (CARA) preferences and normally distributed returns while

be dependent on intermediary wealth and other constraints, for expositional convenience, I initially suppress this dependence in my notation; I return to this idea later on.

There are N assets in net supply normalized to 1, whose cashflows are distributed multivariate normal, $D \sim N(\mu, \Sigma)$. Similarly to Koijen and Yogo (2019) , I assume that Σ can be decomposed as $\Sigma = \beta\beta' + \lambda^2 I$, where β contains asset factor loadings, λ^2 is idiosyncratic variance, and β is of dimension $N \times 1$. There is also a risk-free asset whose gross return R_f is fixed exogenously. Let X be a $N \times k$ matrix of stock characteristics. The representative household and institutional investor agree that

$$\beta = X\Pi + \pi$$

, where Π is a $k \times 1$ vector and π is a constant $N \times 1$ vector. Hence, fundamental loadings β are affine in characteristics.[11]

Now, assume that μ is linear in characteristics but that households and institutional investors may disagree on the mapping from characteristics to μ in the following manner. Households believe that the mean μ follows

$$\mu_H = X\Phi_H + \phi_H + \epsilon_H \tag{1}$$

while institutional investors' estimate of the mean μ is given by

$$\mu_I = X\Phi_I + \phi_I \tag{2}$$

Here, ϕ_H and ϕ_I are constant across assets, and ϵ_H may differ across assets. The residual ϵ_H is the component of households' beliefs about the mean of the asset payoff distribution that are uncorrelated with the asset characteristics (in the sense that $\epsilon_H{}'X = \mathbf{0}$, with $\mathbf{0}$ a k-vector of zeros) and could be a stand-in for expectations errors or real or perceived costs of trading an asset. Adding an intermediary error vector ϵ_I in (2) would yield all the same key equilibrium expressions, except with ϵ_H replaced by $\epsilon_H - \epsilon_I$; thus, my choice to place the error vector on the household side is without loss of generality.

also allowing investment in risky assets to be wealth dependent. Moreover, He et al. (2022) allow risk aversion to depend on intermediary wealth in a setting with mean–variance preferences to capture the effects of net worth constraints, while Haddad and Muir (2021) think of intermediary risk aversion as being dependent on net worth or leverage constraints in a setting with CARA preferences and normally distributed payoffs.

[11] Since any multifactor model of payoffs/returns implies a single-factor model where the stochastic discount factor (SDF) is the lone factor, this essentially assumes that loadings on the SDF are affine in characteristics.

Given constant absolute risk aversion utility, the optimal demand for agent j is

$$\theta_j = \rho_j \Sigma^{-1}(\mu_j - R_f P)$$

Imposing market clearing ($\theta_I + \theta_H = 1$) gives the following expression for prices:

$$P = \frac{\rho_I \mu_I + \rho_H \mu_H - \Sigma \mathbf{1}}{R_f(\rho_I + \rho_H)} \tag{3}$$

Substituting out price using market clearing gives the following for intermediary demand (or percentage intermediated):

$$\begin{aligned}
\theta_I &= \rho_I \Sigma^{-1} \left[\frac{\rho_H(\mu_I - \mu_H) + \Sigma \mathbf{1}}{\rho_I + \rho_H} \right] \\
&= \alpha \left(\beta\beta' + \lambda^2 I \right)^{-1} (X\Delta\Phi + \Delta\phi - \epsilon_H) + \delta \mathbf{1} \\
&= \frac{\alpha}{\lambda^2} \left(I + \frac{1}{\kappa}\beta\beta' \right) (X\Delta\Phi + \Delta\phi - \epsilon_H) + \delta \mathbf{1} \\
&= \frac{\alpha}{\lambda^2} \left(X\Delta\Phi + \Delta\phi - \epsilon_H + (X\Pi + \pi)\eta \right) + \delta \mathbf{1} \\
&= \frac{\alpha}{\lambda^2}(\Delta\phi + \pi\eta) + \delta \mathbf{1} + X\frac{\alpha}{\lambda^2}(\Delta\Phi + \Pi\eta) - \frac{\alpha}{\lambda^2}\epsilon_H \\
&\equiv a + XB + \tilde{\epsilon}
\end{aligned} \tag{4}$$

where the terms in the above are defined as follows:

$$\alpha = \frac{\rho_I \rho_H}{\rho_I + \rho_H}, \;\; \delta = \frac{\rho_I}{\rho_I + \rho_H}, \;\; \kappa = -(\lambda^2 + \beta'\beta),$$

$$\Delta\Phi = \Phi_I - \Phi_H, \;\; \Delta\phi = \phi_I - \phi_H, \;\; \eta = \frac{1}{\kappa}\beta'(X\Delta\Phi + \Delta\phi - \epsilon_H),$$

$$B = \frac{\alpha}{\lambda^2}(\Pi\eta + \Delta\Phi), \;\; a = \frac{\alpha}{\lambda^2}(\Delta\phi + \pi\eta) + \delta\mathbf{1}, \text{ and } \tilde{\epsilon} = -\frac{\alpha}{\lambda^2}\epsilon_H$$

The relation between the second and third lines follows from the Woodbury matrix identity and then simplifying. The constant η is obtained by multiplying β by X and ϵ_H, the current characteristics of all assets and the residual component of the households' estimate of the mean for all assets.

I now show that, under the assumptions above, the residual $\tilde{\epsilon}$ recovers a component of intermediary demand along which the price response to intermediary risk tolerance shocks is strictly increasing. Returning to the equation for prices,

$$\begin{aligned} P =& \frac{\rho_I \mu_I + \rho_H \mu_H - \Sigma \mathbf{1}}{R_f(\rho_I + \rho_H)} \\ =& \frac{\rho_I(\omega)(X\Phi_I + \phi_I) + \rho_H(\zeta)(X\Phi_H + \phi_H + \epsilon_H) - \Sigma \mathbf{1}}{R_f(\rho_I(\omega) + \rho_H(\zeta))} \end{aligned} \tag{5}$$

Now, I let ρ_I depend on the state variable ω and ρ_H on the state variable ζ. Here, ω proxies for an empirical measure of shocks to financial intermediaries that moves ρ_I in practice, as in the classes of intermediary asset pricing models discussed at the start of this section; I also allow some state variable ζ to move the household risk tolerance (for example, because of time-varying risk aversion). I assume that $\rho_I^{'}(\omega) > 0$, so that an increase in ω increases risk tolerance (i.e., proxies for constraints becoming less binding); similarly, I assume $\rho_H^{'}(\zeta) > 0$. We can then take the total derivative of asset price changes with respect to a local shock to these variables. This leads to the first key proposition from the model:

Proposition 1 *Suppose $\rho_I^{'}(\omega) > 0$. Then, the component of the total derivative dp due to changes in ω, β_ω, is strictly increasing in $\tilde{\epsilon}$.*

Taking the total derivative immediately yields 1:

$$\begin{aligned} \mathrm{d}P =& \frac{\rho_I^{'}(\omega)\rho_H(\zeta)\left[\Delta\Phi X + \Delta\phi - \epsilon_H\right] + \rho_I^{'}(\omega)\Sigma\mathbf{1}}{R_f(\rho_I(\omega) + \rho_H(\zeta))^2}\mathrm{d}\omega \\ &+ \frac{\rho_H^{'}(\zeta)\rho_I(\omega)\left[-\Delta\Phi X - \Delta\phi + \epsilon_H\right] + \rho_H^{'}(\zeta)\Sigma\mathbf{1}}{R_f(\rho_I(\omega) + \rho_H(\zeta))^2}\mathrm{d}\zeta \\ &\equiv \beta_\omega d\omega + \beta_\zeta d\zeta \end{aligned} \tag{6}$$

Since $\rho_I^{'}(\omega) > 0$, the expression for β_ω in (6) is strictly decreasing in ϵ_H or, equivalently, is strictly increasing in $\tilde{\epsilon} = -\frac{\alpha}{\lambda^2}\epsilon_H$.

Note that (6) resembles a regression of the local change in stock price ("stock return" for a CARA investor) on shocks to ω and ζ. In other words, proposition 1 implies that the beta on a shock that increases (decreases) the intermediaries' risk tolerance is increasing (decreasing) in the percentage intermediated. This is emphasized by He and Krishnamurthy (2018) and is the first theoretical implication that I test in the data.

This setting also delivers differential conditional return predictability for high ϵ_H assets, which is the second proposition. Define the risk premium on asset k by $E[R_{p,k}] = \mu_k - R_f P_k$, and suppose that for two assets $X_1 = X_2$ so that the asset characteristics are the same but

$\epsilon_{H,1} < \epsilon_{H,2}$ (or, equivalently, $\tilde{\epsilon}_1 > \tilde{\epsilon}_2$ so that asset 1 is more intermediated). Then,

$$E[R_{p,1} - R_{p,2}] = \frac{\rho_H(\zeta)(\epsilon_{H,2} - \epsilon_{H,1}))}{R_f(\rho_I(\omega) + \rho_H(\zeta))} \tag{7}$$

which is strictly decreasing in ω. This leads to the second proposition of the model:

Proposition 2 *Consider assets 1 and 2 such that $X_1 = X_2$ while $\tilde{\epsilon}_1 > \tilde{\epsilon}_2$. Let $E[R_{p,k}] = \mu_k - R_f P_k$ denote the risk premium on asset k. Then, the difference in the risk premium on asset 1 and asset 2 decreases with ω, i.e., $\partial E[R_{p,1} - R_{p,2}]/\partial\omega < 0$.*

Hence, after differences in characteristics are netted out, the difference in conditional expected returns for high- and low-intermediated assets decreases when intermediaries are more risk tolerant, as in proposition 2, implying that empirical proxies for current intermediary risk tolerance should negatively predict the return spread between high- and low-$\epsilon_{i,t}$ assets—or equivalently, as in my primary empirical implementation, state variables that cause intermediaries to be less risk tolerant should cause the risk premium on the high $\tilde{\epsilon}$ asset to rise.

Appendix B examines an extension of the model where household risk tolerance also responds in the same direction as intermediary risk tolerance does to shocks to the intermediary state variable(s) ω. This extension leads to a third proposition:

Proposition 3 *Suppose that household risk tolerance is also a function of the same state variable(s) ω as intermediary risk tolerance and that the partial derivative of $\rho_H(\omega, \zeta)$ with respect to ω is positive. Then, if asset prices respond more to ω shocks as $\tilde{\epsilon}$ increases, this pattern must be driven by the effect of ω on intermediary risk tolerance, not by the effect on household risk tolerance.*

Proof: See appendix B.

This proposition formalizes the intuition that one can already arrive at by examining the expression in the baseline setup in equation (6), where $\rho_H(\zeta)$ does not explicitly depend directly on ω. If, in practice, shocks to ζ and ω are positively correlated empirically and $\rho_H^{'} > 0$, then the exclusion of controls that may proxy for shocks to ζ actually work *against* my finding an effect because the coefficient on dζ is decreasing in $\tilde{\epsilon}$ while the coefficient on dω is increasing. As Haddad and Muir (2021) point out, it is likely that financial institutions' risk tolerance shocks are positively correlated with those of households, so this seems to be the relevant case empirically.

Observe that the model implies the spread in betas on contemporaneous shocks are due to discount rate effects: Price appreciation in a more intermediated stock occurs because of positive shocks to intermediaries' willingness to take risk, absent any fundamental information about stock cashflows. Though the discount rate and cashflow components of returns cannot be observed perfectly, the combined presence of return predictability on the basis of predetermined state variables and price movements induced by contemporaneous shocks to the same state variables would constitute strong evidence that the effects are driven through the discount rate component of returns. Because of this, I include both contemporaneous and predictive tests of the model's implications.

The choice to have absolute risk aversion vary as a function of underlying state variables is obviously critical to the model's predictions and deserves further attention. Since the coefficient of relative risk aversion is related to the coefficient of absolute risk aversion by $w_I\gamma_I = \alpha_I$ (where w_I is the agent's wealth, γ_I the absolute risk aversion, and α_I the relative risk aversion), parameterizing γ_I to vary as a function of wealth or wealth share—as I do here, and as is also done in Haddad and Muir (2021), He et al. (2022), and Koijen et al. (2023)—tractably captures wealth effects such as those present in intermediary asset pricing models with constant relative risk aversion of specialists. He and Krishnamurthy (2013) is one such example. In this model, wealth shocks lead to changes in risk premia, as the distribution of wealth shifts between agents with different willingness or ability to bear risk. These effects also have outsize influence in the constrained region of the model, when equity capital constraints bind and intermediaries require price concessions to bear aggregate risk. In appendix B.2, I explore an alternative version of the model with log-normal payoffs couched in constant relative risk aversion preferences, which directly allows for intermediary wealth effects. A version of propositions 1 and 2 holds in this version of the model, as well, though closed-form expressions rely on a log-linear approximation.

The presence of risk aversion is not required for intermediaries to exhibit time-varying risk-bearing capacity as long as there are binding constraints that cause intermediaries to behave as if they are risk averse. Brunnermeier and Sannikov (2014) work with risk-neutral agents and find that specialists' wealth share is a critical state variable, generating large spikes in risk premia in the constrained region just as in He and Krishnamurthy (2013). Adrian et al. (2014) point out that, in a setting resembling Brunnermeier and Pedersen (2008) with margin constraints, time variation in the margin constraint can lead to nontrivial state pricing where risk-neutral intermediaries value a dollar of wealth relatively more when the Lagrange multiplier on the margin constraint is higher and the value of relaxing the constraint is larger. Thus, when margin constraints are tighter, intermediaries invest as if they were more risk averse. Adrian et al. (2014) argue that their leverage measure (which is the reciprocal of margin) proxies for the tightness of leverage constraints and hence risk-bearing capacity. In this sense, having risk-tolerance shift because of intermediary shocks is a sort of reduced-form way of capturing the price effects of such mechanisms. Furthermore, allowing households' absolute risk aversion to vary as a function of state variables can capture features related to time variation in household risk aversion, as would be found in a habit model, for example.

In summary, the crux of the model's predictions are this: If (1) intermediary risk tolerance is time varying and we have suitable proxies for this time variance, (2) directly investing households differ from interemediaries in their assessment of stocks' cashflows (e.g., because of differential expectations errors or real or perceived trading costs), and (3) there is variation in households' expectations errors (or direct investment costs) across otherwise similar assets, then we should be able to detect the effects detailed in propositions 1 and 2. The justification behind point (1) comes from the literature on friction-based intermediary asset pricing models. Point (2) can be seen as resulting from households' limited rationality/information processing capacity relative to that of more sophisticated institutional investors; similar features are present in numerous asset pricing models. I argue point (3) by demonstrating in section 4.1 that I can construct a measure that holds fundamental stock information constant yet still generates a large spread in average intermediation.

3 Data Sources and Sample Construction

Before proceeding to the empirical implementation, I first describe the datasets used and sampling procedure followed. Individual monthly firm stock returns are from the Center for Research in Security Prices (CRSP). The sample is restricted to ordinary common shares (share codes 10 or 11) that trade on the NYSE, Amex, or Nasdaq (exchange codes 1, 2, or 3). Institutional holdings data for individual stocks come from the Thomson Reuters Institutional Holdings Database (S34 file). Because of well-documented errors in the S34 database institutional type classifications, I use the corrected type codes provided by Koijen and Yogo (2019) to classify institutions into mutual funds and other investment advisors (the category that prominently includes the largest hedge funds).[12] I download the quarterly holdings data from 1980q1 to 2017q2. When processing institutional holdings, I follow the recommendations of Ben-David, Franzoni, Moussawi, and Sedunov (2021) in using CRSP-reported shares outstanding to compute the percentage of shares held by institutions and capping any individual 13F institution's holdings of a given stock at 50% of the market cap to avoid occasional extreme outliers.[13] My primary set of stock characteristics is originally

[12] For a more detailed description of these data, see Gompers and Metrick (2001) or, for a more recent contribution, Koijen and Yogo (2019).

[13] Ben-David et al. (2021) report issues with missing holdings in the Thomson Reuters data in the post-2013 period that cause the time-series share of the market cap held by institutions to be underestimated.

derived from Compustat but is taken directly from Koijen and Yogo (2019), whose paper on characteristics-based demand of financial institutions also utilizes the Thomson Reuters database. The characteristics are derived from the Fama–French 5-factor model and include past 5-year stock capital asset pricing model (CAPM) beta, log book equity as a proxy for size, gross profitability, and asset growth. I further include the book-to-market ratio as the ratio of book equity to market cap from a year prior. As in Koijen and Yogo (2019), accounting characteristics are obtained as of at least 6 months and no more than 24 months prior to the given date to ensure the data are publicly available at the time of portfolio formation.

In addition to the Koijen and Yogo (2019) characteristics, in a robustness check, I add to the set of stock characteristics dozens of financial ratios obtained from the Wharton Research Data Services (WRDS) financial ratios suite. I also obtain the quarterly and monthly series of shocks to Federal Reserve primary dealer capital introduced in He et al. (2017) and available on Asaf Manela's website. As an additional intermediary variable, I obtain the leverage of broker-dealers introduced in Adrian et al. (2014).[14] The monthly Fama–French risk factors plus momentum factor are also downloaded from Ken French's website.

The sample construction proceeds as follows. For each quarter, I take the intersection of the entire CRSP universe of stocks meeting share code and exchange code criteria described above with the Koijen and Yogo (2019) stock characteristics data, excluding any missing matches within a quarter. To ensure my findings are not driven by very small stocks where trading frictions are likely to be larger, I further exclude microcap stocks (defined as stocks beneath the NYSE 20th percentile in market cap) from the sample each quarter and stocks with price less than $5 to focus on the set of stocks that large financial institutions are able to trade most freely. I additionally exclude financial stocks (stocks with a Standard Industrial Classification [SIC] code between 6000 and 6999) from my sample. This common restriction is even more practical in my setting because the relationship between stock price movements and intermediary risk-bearing capacity is highly endogenous for financial stocks. In terms of market cap, these restrictions lead me to drop only a small portion of the CRSP equity universe—my sampling retains on average approximately 97% of the total market capitalization of non-financial stocks on the CRSP tape. These stocks constitute the primary quarterly sample. I also convert the monthly Fama–French five factors and momentum to their respective quarterly versions. Unless otherwise noted, the sample period for regressions

My empirical strategy relies on cross-sectional variation in holdings rather than the time series, so this issue is of lesser concern. Unreported estimates (available upon request) establish that all the findings in this paper are unaffected by my restricting to the pre-2013 period.

[14]Thanks to Koijen and Yogo (2019), He et al. (2017), and Adrian et al. (2014) for making their data readily available.

spans 1980q2 to 2017q3.

Additional tests in section 5 rely on alternative institutional holdings data obtained from FactSet. I discuss the FactSet data in appendix C.3.

4 Empirical Strategy and Results

4.1 Constructing the Measure of Intermediation

The characteristics-based economic framework discussed in section 2 suggests that stocks with similar characteristics but higher intermediary holdings should have higher betas on intermediary capital shocks. In this vein, I construct a measure of intermediation intended to be unrelated to key stock characteristics that proxy for information regarding cashflow distributions. Let $X_{i,t}$ be a vector of stock characteristics that are informative about the distribution of time $t+1$ cashflows of asset i. Recall in equation (4) that the constants α and δ also depend on the current risk tolerance of the agents in the model. To the extent that these and belief differences are time varying, the empirical analogue to (4) should be estimated with time-specific coefficients. Hence, at each time t, I run the following cross-sectional regression:

$$\text{Percentage Intermediated}_{i,t} = \alpha_t + \beta_t X_{i,t} + \epsilon_{i,t} \tag{8}$$

Since the residuals $\epsilon_{i,t}$ in regression equation (8) are analogous to $\tilde{\epsilon}$ in model equation (4), this implies that sorting on $\epsilon_{i,t}$ should induce variation in betas on proxies for shocks to intermediary risk tolerance. More broadly, I implement this approach to ensure that the cross-sectional spread in asset price response to shocks to intermediary risk-bearing capacity is not driven primarily by differential fundamental exposures to other risk factors.

If the risk tolerance of the financial institutions that are active in equity markets is time varying and moves because of changes in empirical proxies of financial intermediary capital, sorting on $\epsilon_{i,t}$ should induce variation in betas on shocks to intermediary capital, and current intermediary capital should contain information about the expected returns of high-$\epsilon_{i,t}$ assets relative to those of low-$\epsilon_{i,t}$ assets. Specifically, high-$\epsilon_{i,t}$ assets should have a larger contemporaneous price response due to shocks to proxies for intermediary capitalization/intermediary risk tolerance and greater return predictability by the level of intermediary capital, as outlined in propositions 1 and 2.

Here, Percentage Intermediated$_{i,t}$ denotes the percentage of shares held by mutual funds, hedge funds, and other investment advisors. I focus on these institution types because they include the set of financial intermediaries that are the largest and most active in equity markets, though my results are unaffected by my instead including all 13F institutional investor types, as I demonstrate in section 4.5.

Regression equation (8) decomposes intermediary holdings into three components: $\beta X_{i,t}$, holdings due to firm fundamentals; α_t, holdings coming from time changes in the average-level institutional holdings[15]; and $\epsilon_{i,t}$, holdings unrelated to fundamentals (possibly reflecting unobserved expectations errors or perceived direct holding costs). In the empirical tests to follow, I show in a variety of settings that $\epsilon_{i,t}$ is strongly related to intermediary capital betas on both predictive and contemporaneous variables.

Implementing (8) requires that $X_{i,t}$ be related to asset fundamentals that are informative about the distribution of future cashflows. Following Koijen and Yogo (2019), I focus on a set of stock characteristics derived from the Fama and French (2015) empirical asset pricing model that is known to have significant explanatory power for the cross section of stock returns and, hence, presumably provides considerable fundamental information about the distributions of asset cash flows.[16] The empirical implementation of (8) includes the following set of stock characteristics in $X_{i,t}$: a second-degree polynomial in log book equity, gross profitability to book equity, annual growth in firm assets (as a proxy for investment), book-to-market ratio using one-year-lagged market cap, and 5-year rolling monthly preranking CAPM beta (requiring at least 24 observations to be included). These are derived from the sorting characteristics used to construct the risk factors in the Fama and French (2015) model. I use these characteristics as constructed by Koijen and Yogo (2019), who, in turn, construct them from Compustat to align with the procedure in Fama and French (2015).

In robustness checks, I demonstrate that the set of characteristics included in (8) is not particularly important after stock size is controlled for. My proxy for stock size is log book equity rather than market equity because market equity is an endogenous equilibrium outcome that is potentially directly affected by intermediary and household demand for the asset (though the findings hold in unreported regressions where I use market equity instead of book equity as my size proxy). The model in section 2 demonstrates why it is important to control for stock size. The empirical tests described in propositions 1 and 2 require that the means and variances of underlying cashflows be held constant. Since I normalize net supply to one, the price of a given asset has the interpretation of the stock market cap, and the means and variances are the means and variances of the total dividends paid out to all

[15] An upward trend in institutional holdings has been well documented. See Stambaugh (2014), for example.

[16] For example, Kewei, Xue, and Zhang (2015) argue that their empirical asset pricing model, which is closely related to the model of Fama and French (2015), performs particularly well in describing the cross section of returns when microcap stocks are not overweighted in portfolio formation. Thus, (8) is likely to be more relevant among the set of the larger, more liquid stocks (non-microcap stocks and stocks with a share price above $5) that I consider in my analysis.

the stock's shareholders. Hence, μ and Σ are highly dependent on the stock size in practice.

In Table 1, I estimate the Fama–Macbeth time series average coefficients from each cross-sectional regression (8). By far the strongest predictor of intermediary holdings is stock size, as proxied by log book equity and log book equity squared, although each of the other characteristics is statistically significant in explaining institutional holdings. These institutions tend to overweight large stocks, profitable stocks, and stocks with high asset growth and CAPM betas and tend to underweight value stocks. Note also that the average cross-sectional R^2 is only approximately 0.11, which still leaves a substantial portion of intermediary holdings unexplained each period. I use this unexplained portion to proxy for variation in $\tilde{\epsilon}$ from the model that is unrelated to stock fundamentals.

4.2 Empirical Results For Portfolios Sorted on Residual Intermediation

Though the theoretical propositions in section 2 required that stock fundamentals be held constant, as a practical matter, the empirical predictions in propositions 1 and 2 still hold as long as two assets look sufficiently similar but have a wide spread in ϵ_H (and, hence, in percentage intermediated). In terms of model parameters, if $\epsilon_{H,1} << \epsilon_{H,2}$, and characteristics $X_1 \approx X_2$, then the empirical implications of propositions 1 and 2 should still hold: Namely, if asset 1 is much more intermediated than asset 2, asset 1 should have a higher beta on shocks to intermediary risk-bearing capacity and should be more predictable by state variables capturing intermediary risk tolerance (or lack thereof).

I now organize around this idea by forming equal-weighted portfolios on the quintiles of the residual institutional holding measure $\epsilon_{i,t}$; later, I will support the portfolio analysis with stock-level panel regressions interacting intermediary shocks with $\epsilon_{i,t}$. The portfolios are rebalanced quarterly. I focus on equal-weighted portfolios because it is simpler to net out differences in average portfolio stock characteristics with equal weights, and this also ensures that they are well-diversified portfolios. However, I show in section 4.5 that my findings are robust to value-weighting. Moreover, since I drop microcap stocks (those below the NYSE 20th percentile of market cap) and also stocks with a start-of-period share price below \$5 from my sample, this mitigates any concerns that the equal-weighted portfolios are driven by tiny stocks facing major liquidity issues and large limits to arbitrage.

Figure 2 shows that the portfolio formation does an excellent job in holding characteristics constant between top- and bottom-quintile portfolios sorted on $\epsilon_{i,t}$ while inducing substantial variation in intermediation—the average institutional holdings quintile is just below five at each point in time for the top-$\epsilon_{i,t}$-quintile portfolio, while it is just above one for the bottom-

$\epsilon_{i,t}$-quintile portfolio. Meanwhile, the average quintiles of the rest of the characteristics all hover around three for both portfolios. Thus, these portfolios look almost exactly the same on the key stock characteristics that form the basis of the Fama and French (2015) asset pricing model, which is known to describe the cross section of stock returns well.

Table 2 shows the means and medians of stock characteristics for each of the five portfolios formed on quintiles of $\epsilon_{i,t}$, including percentage holdings by mutual funds, hedge funds, and investment advisors; log of market and book equity; book-to-market ratio; asset growth; profitability/book equity; and preranking CAPM beta estimated over the past 60 months (and a minimum of 24 months). In line with the graphical evidence in Figure 2, the means and medians of each characteristic other than percentage intermediated are extremely close for the top- and bottom-quintile portfolios formed on $\epsilon_{i,t}$ and are also fairly close for the middle three portfolios (though they tend to have slightly higher profitability and book to market in the middle). Meanwhile, there is a large spread in average percentage intermediated between the top- and bottom-quintile portfolios, with 62% intermediated at the top and only 19% intermediated at the bottom. Thus, Table 2 provides further confirmation that sorting on $\epsilon_{i,t}$ isolates variation in holdings by financial institutions while holding other stock fundamentals essentially constant, particularly when the top- and bottom-quintile portfolios are compared.

The top- and bottom-$\epsilon_{i,t}$ portfolios also have a high degree of comovement, as can be seen graphically in Figure 1. The correlation in excess returns on the two portfolios is 0.96. Table 3 shows the means, standard deviations, and Sharpe ratios of the five portfolios formed on $\epsilon_{i,t}$. Focusing on the top and bottom quintiles, we see that the annualized excess return standard deviations of 41.74% and 38.66% of the top- and bottom-quintile portfolios are also close to one another, as are their Sharpe ratios, which are, respectively, 0.22 for the top quintile and 0.26 for the bottom quintile. The top-quintile portfolio has slightly lower returns, though the spread is not very large at -0.853% per year, and carries an insignificant t-statistic of -1.439. Given that I construct the portfolios after netting out differences in characteristics that are known to predict the cross section of stock returns, it is not particularly surprising that I find little action on differences in average returns; in any case, my purpose is to examine heterogeneity in stock price responses to intermediary shocks, rather than to analyze average returns, so this fact is not of crucial importance.

The model implies that the portfolios should have monotonically increasing exposures to shocks to intermediary risk-bearing capacity. I use four proxies for this (from here, I abbreviate the references to He et al. (2017) and Adrian et al. (2014) as HKM and AEM, respectively): shocks to primary dealer market equity capital ratio from HKM, broker-dealer book leverage shocks from AEM, value-weighted excess returns on the financial sector (stocks with SIC codes between 6000 and 6999), and my primary proxy, which standardizes the AEM and HKM measures individually, takes the average of the two, and then restandardizes this average to zero mean and unit variance.

The reason I combine the AEM and HKM shocks is to take a weighted average of financial sector risk-bearing capacity using the most prominent and successful proxies for shocks to intermediaries proposed in the literature, analogously to Haddad and Muir (2021). A seeming source of tension between these two measures is that AEM find that broker-dealers have procyclical leverage while HKM find that bank holding companies have countercyclical leverage. While this fact initially seems hard to reconcile, follow-on work has demonstrated that the two results can be easily squared in a more general framework featuring heterogeneous types of key intermediaries: Both Kargar (2021) and Ma (2017) demonstrate that a heterogeneous intermediary SDF can be constructed as a function of shocks to two state variables that are closely related to the AEM and HKM measures. Such an SDF arises when different classes of intermediaries face heterogeneous financial constraints or have different risk aversion and yields leverage patterns that are simultaneously consistent with the findings of both AEM and HKM.

I take the two measures at face value: AEM capture shocks to institutions that lever up in good times and face more binding margin constraints when market conditions deteroriate and, hence, must delever in a liquidity crisis (Brunnermeier and Pedersen, 2008; Adrian and Boyarchenko, 2012); HKM capture shocks to central financial institutions that face an equity constraint that binds during crisis times when their net worth is low and, hence, leverage is high (He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014). Combining the two results yields a powerful measure of overall shocks to intermediary sector constraints. In additional tests, I examine comovement with value-weighted returns on the financial sector, as this measure is directly related to the wealth share shocks that are important in equity-constraint-based intermediary asset pricing models.

4.2.1 Portfolio Contemporaneous Comovement

I test proposition 1 by running regressions of the form

$$R^i_{t+1} - R_{f,t} = \alpha_i + \beta_{1,i} F_{t+1} + \beta_{2,i}(\text{Mkt}^{NonFin}_{t+1} - R_{f,t}) + \nu_{i,t} \tag{9}$$

individually for $F_{t+1} = \text{Intermediary Shock}_{t+1}$, $\text{Capital Shock}_{t+1}$, $\text{Leverage Shock}_{t+1}$, and $\text{Ex Ret. (Fin)}_{t+1}$ and also across the returns in excess of the risk-free rate for the five intermediation ($\epsilon_{i,t}$) quintile portfolios, $i = Q1, Q2, Q3, Q4, Q5$, plus the excess returns of the high-minus-low-$\epsilon_{i,t}$ ($Q5 - Q1$) portfolio. Here, $\text{Intermediary Shock}_{t+1}$ refers to the combined AEM and HKM measure, $\text{Capital Shock}_{t+1}$ the HKM shock, $\text{Leverage Shock}_{t+1}$ the AEM shock and, finally, Ex Ret. $(\text{Fin})_{t+1}$ the value-weighted stock market excess return on the financial sector. I control for a version of the value-weighted market risk factor that includes just the returns to non-financial stocks. I include this control for several reasons. First, the AEM and HKM models present asset pricing tests controlling for market risk. Second, as illustrated in equation (6), it is important to control for shocks that could proxy for changes in the risk aversion of households, and market returns relate to time variation in risk aversion for certain classes of models, such as habit models. The joint inclusion of shocks to intermediary risk-bearing capacity and non-financial stocks also directly relates asset price movements to financial and non-financial wealth share shocks. The non-financial market risk factor has a correlation of 0.99 with the value-weighted market risk factor from Ken French's website.

Table 4 shows the results of the contemporaneous portfolio tests using the combined Intermediary Shock_{t+1} measure. The same findings are illustrated graphically in Figure 3. Strikingly, there is a strong monotonically increasing relationship in the betas on the intermediary shock and no pattern whatsoever in the non-financial market return betas. The t-statistic of 5.51 on the $Q5$-minus-$Q1$-intermediation spread portfolio is highly significant. This monotonic pattern is directly in line with the theoretical implications presented in section 2 and also with the finding of He and Krishnamurthy (2018) that similar but more intermediated assets should have relatively higher betas in response to intermediary risk tolerance shocks than to household wealth shocks. Since the intermediary shock is scaled to have unit standard deviation and the returns are in annualized percentage form, the coefficient of 4.09 in column (6) of Table 4 means that the annualized return on the high-intermediation portfolio increases by 4.09% relative to that on the low-intermediation portfolio in response to a one-standard-deviation intermediary shock. Note also the relative magnitudes when columns (5) and (1) are compared: The beta on the top-quintile portfolio is nearly 5 times larger than that on the bottom-quintile one—a very large economic difference.

The empirical patterns illustrated in Figure 3 continue to hold when I examine each proposed intermediary shock individually while controlling for the market factor. This is demonstrated in Figure 4, where I also estimate a version of (11) for the HKM, AEM, and financial sector excess return. The loadings are increasing from bottom to top quintile, and the top-minus-bottom-quintile spread has a significant loading for each of the four intermediation risk-bearing capacity shocks. The exposures increase monotonically for all measures. Note also that combining the information in AEM leverage shocks and HKM capital shocks—as I do in Figure 3 and repeat for comparison in the top-left subfigure of Figure 4—leads to a more statistically significant coefficient on the top-minus-bottom-quintile spread relative to the significance of the coefficients on the individual measures.

As a final piece of evidence for the contemporaneous portfolio regressions, in Figure 6, I run regressions separating the HKM capital shocks and AEM leverage shocks but include them both within the same specification:

$$R^e_{i,t+1} = \alpha_i + \beta_{1,i} F_{t+1} + \beta_{2,i}(\text{Mkt}^{NonFin}_{t+1} - R_{f,t}) + \nu_{i,t} \tag{10}$$

The high-minus-low-intermediation excess return is significantly positive for both risk factors, with monotonicity in the betas for both the capital shocks and the leverage shocks. Thus, the HKM capital and AEM leverage factors continue to display patterns in line with proposition 1 when included together in the same regression.

4.2.2 Return Predictability and Risk Premium Variance Decomposition

I next turn to my predictability tests that relate to proposition 2, which is that otherwise similar but more intermediated stocks should also have excess returns that are ex ante more positively predictable by state variables that capture a lower risk-bearing capacity of financial institutions. I primarily focus on my composite measure of the inverse of intermediary risk aversion. I construct this proxy using information from two state variables in my predictability tests: the squared market leverage ratio of Federal Reserve primary dealer bank holding companies, as implied by HKM, and the level of the book leverage ratio of broker-dealers obtained from the flow-of-funds account. Similarly to what I did for my tests of contemporaneous comovement, I again take the average of the standardized versions of these two ex ante state variables as my main proxy for intermediary risk aversion at the current date. He et al. (2017) point out that the model of Adrian et al. (2014) implies that the level of broker-dealer leverage is a state variable that should *negatively* predict returns. Hence, when I take the average, I use the negative of the broker-dealer leverage ratio so that the composite measure predicts returns with a positive sign. I call my composite measure of intermediary risk aversion η_t. In addition to testing whether η_t predicts returns relatively more for intermediated portfolios, I examine what fraction of the total variation in conditional equity risk premia I can attribute to movements in η_t.

I run regressions of the following form:

$$R^e_{i,t+1} = \alpha_i + \beta_{1,i}\eta_t + \beta_{2,i} Z_t + \nu_{i,t} \tag{11}$$

Here, i indexes the quintile $i = 1, \ldots 5$ portfolios formed on residual intermediation ϵ, as before. The control Z_t is a composite predictor for the conditional equity risk premium, comprising information coming from market return predictors previously documented in the literature. To be precise, I construct Z_t by taking the 14 predictor variables from Welch and Goyal (2008)[17]; the consumption–wealth ratio (cay) from Lettau and Ludvigson (2001); the aligned investor sentiment index from Huang et al. (2014), which uses PLS to optimize the six investor sentiment proxies in Baker and Wurgler (2006) for time-series return predictability; the aggregate short-selling index of Rapach, Ringgenberg, and Zhou (2016); and the investor attention index from Chen, Tang, Yao, and Zhou (2022). The Welch and Goyal (2008) predictors are standard macroeconomic variables used in return prediction, while the aligned sentiment, short-selling, and investor attention indices are three of the most powerful time-series predictors of the equity premium documented in recent literature.

I apply the full-sample PLS approach of Kelly and Pruitt (2013) and Huang et al. (2014) for optimizing the information in these variables to predict the equity premium. For each individual predictor j, I run a time-series regression of the time-t predictor $x_{j,t}$ on the time-$t+1$ non-financial market excess return ($\text{Mkt}_{t+1}^{NonFin} - R_{f,t}$) and extract the coefficient estimate $\hat{\pi}_j$. Then, for each time period t, I run a cross-sectional regression of $x_{j,t}$ on $\hat{\pi}_j$ and extract the estimated coefficient Z_t, which is my composite risk premium predictor. Kelly and Pruitt (2013) and Huang et al. (2014) discuss in detail how this method is effective at isolating the common component across the $x_{j,t}$ that is most informative about future stock returns, and it tends to result in highly powerful time-series return predictors. The resulting Z_t series has a correlation of 0.22 with η_t.

I consider Z_t a reasonable upper bound on the drivers of risk premium variation not related to time variation in intermediary risk aversion η_t. When estimating (11), I extract the R^2 and compare this with the R^2 resulting from regressing $R_{i,t+1}^e$ on $\eta_t^\perp$, where $\eta_t^\perp$ has been orthogonalized with respect to Z_t. The ratio of the two R^2 values gives a lower bound on the share of risk premium variation driven by η_t; similarly, regressing $R_{i,t+1}^e$ on η_t (without orthogonalizing) and taking the ratio of R^2 values from the resulting regression and also from the full specification (11) gives an upper bound on the share of return predictability driven by η_t. Note that the lower bound relies on a fairly conservative assumption, as aggregate short-selling, dividend price ratios, investor attention, and various other variables in the predictor set may plausibly be directly driven at least in part by frictions originating in the financial sector (as captured by η_t) but, at the lower bound, all of the predictability coming

[17] These predictors are as follows: log dividend–price ratio, log dividend yield, log earnings–price ratio, log dividend–payout ratio, stock variance (monthly sum of squared daily returns on the S&P 500 index), book-to-market value ratio for the Dow Jones industrial average, net equity expansion (ratio of 12-month moving sum of net equity issues by NYSE-listed stocks to the total end-of-year market capitalization of NYSE stocks), three-month treasury bill yield (secondary market), long-term government bond yield, return on long-term government bonds, long-term (10-year) government bond yield minus the treasury bill yield, default yield spread (difference between BAA- and AAA-rated corporate bond yields), and default return spread (long-term corporate bond return minus the long-term government bond return).

from this variation is instead attributed to non-intermediary sources.

I report the coefficients $\beta_{1,i}$ for portfolio quintiles $i = 1, \ldots 5$ in Table 6, and the final column of the table considers a regression of the relative returns to the spread portfolio, $R_{5,t+1} - R_{1,t+1}$, on η_t and Z_t. Coefficient estimates include Newey–West standard errors with Newey and West (1994) optimal lags. As before with the contemporaneous regressions, the slope coefficients on η_t monotonically increase from portfolios 1 to 5, and the coefficient on the excess return between the top and bottom portfolios is highly statistically significant. The sixth column implies that a one-standard-deviation increase in η_t raises the relative annualized risk premium of the most intermediated portfolio by 2.72%. There is substantial predictability, with R^2 values hovering from approximately 8% to nearly 10%, which is quite a large value for the quarterly horizon.[18]

As expected, Z_t is a highly robust predictor, with p-values below 0.01 for all portfolios. However, in contrast to η_t, the predictability of Z_t is constant across portfolios. In terms of the share of variation in conditional portfolio risk premia due to η, the R^2 value obtained from my regressing the first portfolio returns on $\eta_t^{\perp}$ is 23% of the overall R^2 when both η_t and Z_t are included; the R^2 from the regression on the raw η_t is 44% of the total, so η_t must account for between 23% and 44% of the variation in the bottom portfolio. These numbers are 38% and 59% for the top portfolio, and both bounds monotonically increase from bottom to top. The average lower bound is 31.4%, while the average upper bound is 52.6%. Hence, this exercise suggests that intermediary risk aversion explains approximately 31.4% to 52.6% of the measured variation in conditional equity returns across these portfolios. Finally, column 6 shows that η_t explains nearly the entirety of the *relative* risk premia variation between the high-intermediation portfolio ($Q5$) and low-intermediation portfolio ($Q1$).

Compare this results to the finding in Haddad and Muir (2021) that households' risk aversion can explain approximately 40% of the measured variation in conditional risk premia while their bounding exercise cannot attribute any risk premium variation in equity markets, their least-intermediated asset class, to intermediaries. My estimates imply that the bulk of the remaining unexplained variation is driven by the intermediary risk aversion proxy η_t.

I include a final test to examine the presence of a theoretical mechanism in the cross section of return predictability outlined in Gromb and Vayanos (2018). In their model, when the capital of constrained arbitrageurs depletes, the expected returns increase relatively more on the assets where arbitrageurs take larger positions. This causes the increased spread to self-correct over time as intermediary capital recovers because of the increased expected

[18] By comparison, Haddad and Muir (2021) report an R^2 of 4.4% in their analysis of quarterly stock market return predictability.

returns on their positions. To test for such effects, I run regressions of the form

$$R_{5,t+k} - R_{1,t+k} = \alpha + \beta_k \eta_t + \nu_t \tag{12}$$

where t is now at the monthly horizon and k varies from 1 month ahead to 18 months ahead.[19] Figure 6 plots $\widehat{\beta_k}$, its 90% confidence interval, and the R^2 for $k = 1, ..., 18$. As implied by the theory, the coefficients $\widehat{\beta_k}$ decrease with k, as does the R^2. Thus, the quarterly horizon used in the previous predictability tests from this section contains the bulk of the overall high-minus-low-intermediation portfolio spread return predictability of η_t, which then reverts over the ensuing months. This is consistent with temporary relative asset price distortions that are corrected over time as constraints on intermediaries relax when capitalization improves, in line with Gromb and Vayanos (2018) and, more broadly, with models where intermediary capital moves slowly because of constraints that become more binding when intermediaries are poorly capitalized.[20]

In Appendix Table A.1, I examine the predictability of the intermediation spread portfolio $R_{5,t+1} - R_{1,t+1}$ in specifications where I separate the individual predictor variables used to construct η_t in columns (1) and (3); and in columns (2) and (4), I explore the predictive power of the time-t financial sector stock market wealth share. In the last two columns, I additionally control for the composite predictor Z_t. Columns (1) and (3) demonstrate that both the predictor variables have explanatory power for the spread portfolio, though broker-dealer leverage is significant only at the 10% level in column (1) and close to but not quite significant at the 10% level in column (3)—still, theory implies a clear negative sign for the coefficient, which means that, in both cases, it is still significant at better than the 10% level for a more powerful one-tailed test. Primary dealer squared leverage is a significant individual predictor in both specifications; meanwhile, the financial sector stock market wealth share also predicts the spread portfolio with the appropriate negative sign. Overall, these tests suggest that I benefit from improved power by combining information in the two main predictors to arrive at η_t, but both individual measures contain individual incremental power. Further, the financial sector wealth share's robust predictive ability also provides supporting evidence for the role of frictions in the financial sector as a driver behind the larger relative predictability among highly intermediated stocks.

[19] Because the broker-dealer leverage is available only at the quarterly horizon, solely for this analysis, I linearly interpolate the level of broker-dealer leverage for months between adjoining quarters when constructing η_t here.

[20] See, for example, Duffie (2010) for a theoretical summary and Mitchell et al. (2007) for early empirical evidence in convertible bond markets or, for more recent work, Siriwardane (2019) in credit default swap markets.

4.3 Natural Experiment Based on S&P 500 Inclusion

I next examine a natural experiment generating variation in intermediation. In particular, as noted by Basak and Pavlova (2013) and Aghion et al. (2013), institutional investors have the incentive to increase holdings of stocks added to the S&P 500, even if they are not explicitly indexed to it, because the S&P 500 is a natural performance benchmark, which leads to an increase in institutional ownership after index inclusion. More recently, Boller and Scott Morton (2020) shows that index inclusion does increase institutional ownership, and in Appendix Table A.6, I also confirm this intuition: Mutual funds and investment advisors increase holdings by 2.0% (last two columns) between the quarter before and the quarter after inclusion relative to holdings of stocks in the same size quintile that quarter, and total institutional ownership changes by 4% (first two columns), with all estimates being significant at better than the 1% level.

Accordingly, I examine changes in exposure to the HKM intermediary capital factor 30 months before and after S&P 500 inclusion. I move to the monthly horizon to increase the number of observations available in a reasonably close window around the event, which means that I cannot use my combined intermediary shock variable that also includes the AEM leverage factor. Stocks must have full coverage of returns in the 30-month windows before and after the event to be included, and I include only stocks that enter the S&P 500 once in my sample period and do not leave for at least 30 months. Comparison firms may also be or not be S&P 500 member firms, but they must never enter or leave the index in my sample period. I construct a difference-in-differences (DID) estimator by testing whether the change in betas on the HKM intermediary capital factor before and after S&P 500 inclusion is larger than its counterpart for a set of comparison firms in the same group. To be specific, I test whether

$$\Delta\widehat{\beta}_{i,t}^{HKM} - \frac{1}{N_{i,t}} \sum_{j \in \text{Comp Group}_{i,t}} \Delta\widehat{\beta}_{j,t}^{HKM}$$

is statistically different from zero on average for firms i that joined the S&P 500 during month t. Here, $\Delta\widehat{\beta}_{i,t,HKM}$ is the difference between the HKM beta for stock i estimated from months $t+1$ to $t+30$ and the HKM beta estimate taken from months $t-1$ to $t-30$, Comp Group$_{i,t}$ is a set of counterfactual comparison firms for firm i in the month it joined the index, and $N_{i,t}$ is the number of firms in the comparison group.

Table 5 explores the results. In the table, I measure the betas on the capital risk factor in the 30 months leading up to S&P 500 inclusion and the 30 months following inclusion, leaving out the event month, and control for the Fama and French (2015) 5 factors plus momentum when estimating the HKM betas.[21] The columns of Table 5 explore estimates for different sets of comparison firms Comp Group$_{i,t}$. Specifically, I compare the change in HKM intermediary capital betas with the change in the average intermediary capital betas for stocks in the same size quintile in that year (first column) or the same size quintile interacted with CAPM beta, book to market, profitability, investment, or momentum (columns (2) through (6)). The DID estimates are all statistically significant at the 1% level or better and stable across comparison groups, all pointing to an increase in HKM betas of approximately 0.10 or 0.11. This exercise corroborates the findings based on the portfolio sorts using my residual intermediation measure and suggests that quasi-exogenous increases in institutional holdings also generate increasing betas on intermediary factors, consistent with the intuition from the economic framework in section 2.

4.4 Stock-Level Panel Regressions

This section demonstrates that the portfolio-level evidence from the previous section extends to the individual stock level. My stock-level empirical tests take the following form for the contemporaneous regressions:

$$R_{i,t+1} - R_{f,t} = \alpha_0 + \beta_1 F_{t+1} \times \epsilon_{i,t} + \beta_2 W_{t+1} \times \epsilon_{i,t} + \delta \epsilon_{i,t} + \alpha_t + \alpha_i + \nu_{i,t+1} \tag{13}$$

Here, F_{t+1} is any of the contemporaneous shocks to intermediary risk tolerance, and i now indexes individual stocks instead of portfolios. A finding of $\beta_1 > 0$ implies that betas on shocks to financial institutions increase with the component of intermediary holdings that is uncorrelated with characteristics of the stock. I control for value-weighted non-financial market excess returns and also add specifications that include the Fama and French (2015) factors plus the momentum factor for W_{t+1} in (13). I also add time fixed effects to control for common shocks to the cross section as well as stock fixed effects. Replacing the time fixed effects with uninteracted risk factors yields estimates that are essentially identical.

Once again agreeing with the theory, Table 7 shows that $\beta_1 > 0$ for all intermediary shocks considered and is strongly significant for all specifications. However, note in the first row of Table 7 that the intermediary shock measure that combines the information embedded in the AEM and HKM factors again yields a stronger individual estimate than do the measures that include the HKM or AEM factors alone. The financial sector excess return provides more evidence in agreement with the theory, as it also has a positive and significant coefficient on the residual intermediation interaction term across specifications.

[21] I include these controls because membership may affect other features of a firm, such as innovation propensity as in Aghion et al. (2013), which may effect covariances with other risk factors.

The economic magnitude of these estimates is fairly large. Consider two stocks with identical characteristics except that one is owned entirely by mutual funds, hedge funds, or other investment advisors while the alternative is owned entirely by households. Looking at the coefficients in the first row of Table 7, we see that the returns to the fully intermediated stock increase by approximately 8% per year relative to those of the unintermediated stock on an annualized basis in response to a one-standard-deviation shock to the composite intermediary factor. The point estimates on these coefficients are also quite precise, with t-stats ranging from 3.7 to 5.[22]

The only included nonintermediary risk factor whose betas significantly increase with $\epsilon_{i,t}$ is the Fama–French robust-minus-weak profitability factor. This feature is also present in portfolio regressions where I control for the Fama and French (2015) factors plus momentum in section 4.5, though, for brevity, I do not explicitly report the coefficient estimates in those specifications.[23]

For a final stock-level test, I examine the relationship between residual intermediation $\epsilon_{i,t}$ and rolling stock betas on intermediary shocks. I first compute rolling betas for each stock i and each intermediary shock:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i \mathrm{F}_t + \beta_i^M \left(\mathrm{Mkt}_t^{NonFin} - \mathrm{Rf}_t\right) + \delta_{i,t}$$

individually for F = Capital Shock, Leverage Shock, Intermediary Shock, and Ex Ret (Fin). The parameter $\beta_{i,t}$ is estimated at each time t within a rolling window of plus or minus 15 quarters, including the given quarter. I then run the panel regression

$$\widehat{\beta}_{i,t-15\to t+15} = \alpha_0 + \beta_1 \epsilon_{i,t} + \beta_2 Z_{i,t} + \alpha_t + \alpha_i + \nu_{i,t} \tag{14}$$

The controls $Z_{i,t}$ include profitability, investment, CAPM beta, book to market, a second-degree polynomial in log market cap and log book equity, in addition to stock and time fixed effects. I require the estimated betas to have all observations from $t-15$ to $t+15$ to be in the sample for (14) (though the results are not sensitive to this restriction). Because of the

[22] Standard errors are clustered by date to account for cross-sectional correlation in the residuals and are also double-clustered by firm.

[23] In unreported estimates, I find that the loading on the RMW profitability factor only obtains conditional on my controlling for the intermediary shock, which is by far the strongest individual predictor of price movements for portfolios sorted on intermediation. The RMW loading is also driven entirely by the two-year period from 1999 to the end of 2000 that immediately preceded the dotcom crash. While it is possible that intermediary marginal utility loads more on the profitability factor relative to household marginal utility, such an investigation is outside the scope of this paper. It seems likelier that the loading is driven by some small relative movements in the cashflow component of returns that is unrelated to intermediary-based mechanisms and was concentrated in the 1999–2000 period, which was a unique episode for US equity markets.

persistence in the dependent variable due to overlap in the estimation windows, I double-cluster the standard errors by stock and time. Table 8 shows that the individual stock betas centered on time t on each of the intermediary shocks are each strongly increasing in intermediation measure $\epsilon_{i,t}$, with t-stats ranging from 3.19 to 3.88. Thus, the component of intermediary holdings unrelated to characteristics has strong explanatory power for the time variation in betas even at the individual stock level. The coefficient of 5.9 on $\epsilon_{i,t}$ in column (1) for the combined AEM/HKM intermediary shock is comparable (albeit slightly lower) in magnitude to the coefficients found in Table 7 and has the interpretation that, with stock characteristics held constant, the return response of a completely intermediated stock to a one-standard-deviation intermediary shock is 5.9 percentage points higher on an annualized basis than the response of a comparable but completely household-owned stock.

4.5 Additional Tests and Robustness

A natural question regarding my findings concerns whether the results depend crucially on the characteristics included in, or excluded from, regression (8) to back out the residual intermediation component $\epsilon_{i,t}$. Accordingly, I examine the empirical robustness of my findings to my including many more characteristics or, alternatively, controlling for size. To do this, I download the set of stock financial ratios provided by the WRDS financial ratios suite. This set of stock characteristics was used by Kozak, Nagel, and Santosh (2020) to construct an SDF from a large number of potential cross-sectional return predictors.

Though I obtain the full set of 73 financial ratios from WRDS, I restrict the set of characteristics to 40 out of the 73 because of data availability restrictions that I impose.[24] In terms of the categories provided by WRDS, the 40 ratios that remain comprise 6 valuation ratios, 13 profitability ratios, 4 capitalization ratios, 7 financial soundness ratios, 3 solvency ratios, 3 efficiency ratios, and 4 other ratios. I supplement the original set of characteristics included in (8), which consisted of a second-degree polynomial in log book equity, gross profitability to book equity, annual growth in firm assets, book-to-market ratio using one-year lagged market cap, and 5-year rolling monthly preranking CAPM beta (requiring at least 24 observations to be included) with these 40 financial ratios and examine whether including the additional characteristics substantially changes anything. On the other end, I also check the robustness of my results to the inclusion of only the second-degree polynomial in log book equity. Using these alternative sets of characteristics, I reestimate $\epsilon_{i,t}$ and reform the quarterly quintile portfolios.

In further robustness checks, I alternatively value weight the portfolios using one-year

[24]I outline the process I use for selecting these characteristics in detail in appendix C.

lagged market cap and value-weighted cross-sectional regressions to back out $\epsilon_{i,t}$); drop the financial crisis from the sample (defined according to the dates calculated by the NBER as beginning after the business cycle peak in the end of the fourth quarter of 2007 and ending after the business cycle trough in the second quarter of 2009); and control for the Fama–French factors plus momentum as in the stock-level panel regressions from the last section. The contemporaneous regressions are found in Table A.2, and the predictive regressions are in Table A.3. Table A.2 uses my primary proxy for contemporaneous shocks to risk-bearing capacity, the "Intermediary Shock" (the average of the standardized AEM/HKM shocks). Meanwhile, in the Table A.3 predictive regressions, I continue to focus on the state variable η, my main proxy for the time-t intermediary risk-bearing capacity, constructed as the average of the standardized primary dealer squared leverage ratio and the negative of standardized broker-dealer leverage. The tables report regression coefficients for the high-minus-low-intermediation quintile spread portfolio.

Table A.2 demonstrates that the intermediary shock significantly explains the spread in returns between portfolios with high and low residual intermediation no matter the specification or the set of characteristics included. Interestingly, without controls for the other characteristics, the non-financial market risk factor also strongly loads on the returns to the spread portfolio, but this is not the case in any of the other specifications. Value weighting changes little; nor does controlling for the Fama and French (2015) risk factors plus the momentum factor. In the last column, we do see that both dropping the financial crisis and including the additional risk factors increases estimation noise somewhat and reduces the t-statistic on the intermediary shock to 2.77 in the specification with the full set of controls.

The predictive regressions in Table A.3 have the same features. Increasing η (i.e., decreasing intermediary risk tolerance) predicts higher returns going forward on the top-intermediation portfolio relative to those on the low-intermediation portfolio. As in Table A.2, the specification that includes only log book equity features more significant coefficients on the control predictors, but this is almost entirely attenuated in the other specifications. The coefficient on η remains quite stable, strongly significant, and positive for all specifications, where value weighting the portfolios, including more stock characteristics, or dropping the crisis period hardly affects the estimates or the significance. The coefficient on η does decline in the specification that drops the financial crisis, and the R^2 also decreases, though the coefficient is still significant at better than the 5% level outside the crisis period.

Since the intermediary-based state variables are particularly volatile for the financial crisis period (see Appendix Figure A.1 for a time-series plot), it is especially important to note that key comovement and predictability patterns hold when the crisis period is excluded, but they are also quantitatively more powerful when the crisis episode is included. Since friction-based models of financial intermediaries often contain constraints that become heavily binding in a crisis episode, the financial crisis should be one of the strongest individual periods for explaining these phenomena, but they should also still continue to be borne out in other periods. Tables A.2 and A.3 clearly demonstrate that this is the case. The crisis episode thus enhanced the relative comovement and return predictability of highly intermediated stocks, consistent with the nonlinear spikes in risk premia during crises found in models of intermediary asset pricing (He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014), but comovement and predictability are not constrained to this one episode.

While one can never account for all information regarding a stock, Tables A.2 and A.3 also illustrate that the empirical patterns are robust to my conditioning on a wide range of characteristics. It should also be noted that unobserved characteristics would tend to bias against my finding an effect. This is because stocks whose returns have naturally higher covariance with shocks to intermediaries provide very poor hedges against bad times for financial institutions and so observed holdings are unlikely to be driven by some underlying institutional preference for stocks with high intermediary shock betas. Thus, unobserved stock information would tend to result in my understating rather than overstating these effects.

Another concern may be that my findings depend on my focusing on this particular subset of 13F institutions. In Tables A.4 and A.5, I show that this is not the case: I repeat the full set of robustness exercises from Tables A.2 and A.3, respectively, except that, instead of restricting to mutual funds and investment advisors, I include the holdings of all 13F institutions. The high-minus-low-intermediated portfolio excess return continues to comove more with contemporaneous shocks to intermediary capital for all the different versions, and the relative ex ante risk premium is always more predicted by my proxy for current intermediary risk aversion, all with similar magnitudes and statistical significance.

Since the percentage of stocks held by households is one minus the percentage held by institutional investors (ignoring small institutions that do not qualify as 13F filers), these estimates from sorting on total residual institutional ownership are also equivalent to sorting on (the reverse of) residual household ownership. Thus, Tables A.4 and A.5 confirm another prediction of the theory in section 2: Prices of stocks preferred by households should move less with contemporaneous intermediary shocks and should be less predictable by state variables capturing intermediary risk aversion.

4.5.1 Panel Regressions with Illiquidity and Short Interest Controls

The stock-level panel regressions in section 4.4 did not control for liquidity-based measures, nor did they account for short interest in the stock (although my portfolio-level tests include aggregate short interest as one of the predictors in the composite risk premium Z_t). I intentionally do not net out liquidity and short interest when constructing the residual intermediation measure $\epsilon_{i,t}$ because such characteristics are equilibrium outcomes determined by the set of institutions active in trading a given asset, rather than fundamental characteristics determining stock cashflows. For example, since financial institutions may be more active traders, high institutional ownership endogenously affects liquidity. Similarly, short sellers are attracted to stocks held by institutional investors because they relax short-selling constraints (Nagel, 2005).

These measures do have modest correlations with $\epsilon_{i,t}$: The average of the log of the Amihud (2002) illiquidity index over quarters $t-1$ to $t-4$ has a correlation of -0.15 with $\epsilon_{i,t}$, while average short interest as a percentage of shares outstanding over quarters $t-1$ to $t-4$ has a correlation of 0.14 with $\epsilon_{i,t}$. On the liquidity side, this could be consistent with some preference for liquidity among institutional investors but also with stocks with greater institutional holdings improving equilibrium liquidity. Given these correlations, one may wonder whether the panel regressions are robust to my including controls for these variables. I do so in Appendix Table A.8, which repeats the analysis of Table 7 but adds illiquidity and short interest controls, and Appendix Table A.9, which repeats the analysis of Table 8 with additional liquidity and short interest controls (see table footnotes for details). In both cases, the coefficients are economically and statistically similar.

5 Exploring Mechanisms and Alternative Explanations

5.1 Importance of Exposure to Dealer Banks and Broker-Dealers

As explained in proposition 3 of section 2, unless household risk tolerance shocks are negatively correlated with the risk tolerance shocks of mutual funds, hedge funds, and other investment advisors, the increasing price responses to intermediary shocks as institutional holdings increase *must* reflect that the risk-bearing capacity of these institutions is more affected by these shocks than is households'. Thus, my findings in the last section imply that the largest institutional investors in equity markets are directly affected by shocks to dealer banks and other broker-dealers. What, then, are the reasons why the risk-bearing capacity of mutual funds, hedge funds, and other investment advisors is sensitive to the health of bank holding companies of Federal Reserve primary dealers and the broker-dealer sector in general?

As Cho (2020), Adrian et al. (2014), and He et al. (2017) discuss, broker-dealers and dealer banks occupy a key role in the financial sector due to their central positions as capital providers and as prominent intermediaries in the markets for diverse asset classes. Among institutional investors that are active in equity markets, the mechanisms described in this paper should be most active for those whose operations are most clearly connected to dealer banks and broker-dealers. For example, the connection for hedge funds seems readily apparent, as hedge funds are levered institutional investors that depend heavily on capital provision by dealer banks for their ability to trade actively in equity markets and other asset classes. Aragon and Strahan (2012) list the top prime brokers to hedge funds in the years 2002–2008 leading up to the financial crisis; the vast majority of the top ten institutions and all of the top five each year were also Federal Reserve primary dealers at the time. Cho (2020) also argues that hedge fund capital depends on the AEM broker-dealer leverage. When these institutions become distressed, capital availability declines, and hedge funds, in turn, also become distressed. In line with this, Ben-David, Franzoni, and Moussawi (2012) demonstrate that hedge funds were forced to delever when their institutional capital providers withdrew capital via margin calls and redemptions.

Nonetheless, other 13F institutions classified as investment advisors or mutual funds may also invest in a manner that directly connects their well-being to shocks to dealer banks' risk-bearing capacity. This is because 13F institutions are classified at the investment company level but even investment companies classified as mutual funds may in fact have funds that engage in investment strategies exposing them to the intermediation provided by dealer banks. To the extent that this is true, we would expect the financial health of the set of 13F institutions with exposure to such strategies to be more dependent on shocks to the health of broker-dealer banks than those without such exposure.

Given the role of dealer banks as prime brokers for hedge funds, I first examine the differential price response between stocks with high and low holdings by hedge funds. I obtain an updated list of 13F hedge fund managers from Vikas Agarwal, which has been used in a series of papers examining the characteristics of hedge fund holdings and returns.[25] Using these data, I restrict my baseline measure of mutual fund and investment advisor holdings to the subset identified as hedge fund managers, and I then back out the percentage holdings by hedge funds. I then re-estimate $\epsilon_{i,t}^{Hedge\ Funds}$ in equation (8), where I replace the percentage intermediated with the percentage of the stock held just by hedge funds. I then examine the excess returns for stocks in the top-minus-bottom-quintile porfolios sorted on $\epsilon_{i,t}^{Hedge\ Funds}$.

I examine the results in Table 9. In the first two columns of panel A, I regress the top-minus bottom-quintile excess return $R_{t+1}^{Q5,HedgeFunds} - R_{t+1}^{Q1,HedgeFunds}$ on the contemporaneous

[25]These include Agarwal et al. (2024), Agarwal et al. (2017), Agarwal et al. (2013), and Agarwal et al. (2013). Thanks to Vikas Agarwal for generously sharing the data with me.

Intermediary Shock$_{t+1}$, with different sets of controls. I obtain a statistically significant positive relationship between the high-minus-low-quintile hedge fund return portfolios and the intermediary shock, consistent with hedge funds' risk-bearing capacity directly responding to shocks to dealer banks and other broker-dealers, supporting that proposition 1 holds for hedge funds specifically. In terms of magnitude, and when we focus on the full-sample regressions with controls in column (2), a one-standard-deviation intermediary shock increases the annualized returns of top-quintile hedge fund residual holdings stocks by approximately 2.7 percentage points relative to those of bottom-quintile stocks. This holds regardless of whether I control for the non-financial market risk factor or the full set of Fama and French (2015) risk factors plus the momentum factor.

In the first two columns of panel B of Table 9, I explore predictive regressions of $R_{t+1}^{Q5,HedgeFunds} - R_{t+1}^{Q1,HedgeFunds}$ on η_t, the intermediary risk aversion index. Similarly to the results of the contemporaneous regressions in panel A, and in line with proposition 2, the estimates show a significantly positive predictive relationship between the return on the hedge fund residual intermediation spread portfolio and start-of-period intermediary risk aversion η_t. For the regression with the composite risk premium in column (2), the coefficient implies that a one-standard-deviation increase in initial intermediary risk aversion generates a 2% annualized increase in the conditional expected return for stocks in the top quintile of residual holdings by hedge funds relative to that for stocks in the bottom quintile.

A casual comparison of the coefficients for just hedge funds in Table 9 and those based off a broader set of institutions in Tables 4 and 6 could make it seem that, while hedge funds' risk-bearing capacity responds to broker-dealer shocks, the degree of responsiveness is smaller. However, the two coefficients are not directly comparable for two reasons. The first simple reason is apparent in equations (6) and (7), which make clear that the coefficients should increase in the amount of $\epsilon_{i,t}$ dispersion between the top and bottom portfolios. This dispersion is naturally larger for my baseline measure than when I focus specifically on hedge funds because of the larger set of institutions in my baseline.

Second, and more importantly, the comparison group for hedge funds is not the same set of investors as when I consider the broader set of institutional holdings. When the "intermediary" holdings are restricted to just hedge funds, the other agent currently labeled "households" in the model now also contains all the other non–hedge fund institutional investors, including mutual funds and other investment advisors. To the extent that these institutions' risk-bearing capacity can also be affected by shocks to dealer banks and broker-dealers, then equation 16 from the model extension in appendix B makes clear that the coefficient instead identifies the *relative* strength of hedge funds' risk-bearing capacity response with respect to the strength of the response of the other agent—which is now an amalgamation of mutual funds, non–hedge fund investment advisors, other institutional investors, and retail investors. Thus, while the comparison group for my baseline measure of residual intermediation is dominated by retail investors, when I sort on hedge fund holdings, the comparison group is now largely dominated by the other institutional investors, and especially mutual funds. Hence, a more accurate reading of the significantly positive coefficients for hedge funds in Table 9 is that hedge funds' risk-bearing capacity is more responsive to intermediary shocks *relative* to the responsiveness of this residual group of direct equity market participants. This directly supports the hypothesis that hedge funds should be more responsive to intermediary shocks than other institutional investors because of their more direct relationship.

This all being said, my results so far do suggest that the risk-bearing capacity of mutual funds and other non-hedge fund investment advisors also responds to shocks to dealer banks and broker-dealers. What might explain the connection? I now examine two potential mechanisms: 1) transmission of exposure to dealer banks and broker-dealers via interactions with dealer-exposed hedge funds in equity markets; and 2) exposure from investing in other asset classes (particularly bond markets) that are heavily intermediated by dealer banks/broker-dealers.

I first investigate whether these shocks can be passed through to mutual funds via interaction with directly exposed hedge funds. To test this, I re-estimate $\epsilon_{i,t}$ in equation (8) for all mutual funds and non–hedge fund investment advisors. I then examine the excess returns of two versions of the high-minus-low quintile portfolios based off how much mutual funds' holdings line up with those of hedge funds. For the first group, I require a high degree of alignment: stocks in the top quintile portfolio of $\epsilon_{i,t}$ estimated for mutual funds and non-hedge fund investment advisors must also be in at least the top two quintiles when sorting on $\epsilon_{i,t}$ estimated for hedge fund-only holdings; similarly, stocks in the bottom quintile portfolio for non-hedge funds must also at least be in the bottom two quintiles based on hedge fund holdings. In the second group, I form an alternative high-minus-low portfolio based on the opposite set of stocks which have a low degree of holdings alignment with hedge funds: I take the excess returns of top-quintile stocks that are in the *bottom* 3 quintiles of hedge fund holdings, minus the bottom-quintile stocks that are also in the top 3 quintiles for hedge fund holdings. Thus the analysis of "high hedge fund alignment" excess returns compares the relative return responses to intermediary shocks of stocks with *both* high mutual fund and high hedge fund holdings, to stocks with *both* low mutual fund and low hedge fund holdings. On the other side, the analysis of "low hedge fund alignment" excess returns compares the relative return responses to intermediary shocks of stocks with high mutual fund *but low* hedge fund holdings, to stocks with low mutual fund *but high* hedge fund holdings.

The results are included in the final four columns of Table 9. Columns (3) and (4) of the table demonstrate that the top-minus-bottom-quintile excess returns of mutual funds and non–hedge fund investment advisors covary strongly with intermediary shocks when there is a high degree of alignment with hedge fund holdings. However, this is very much *not* the case in the last two columns when the holdings of the groups are misaligned. This pattern holds true for the contemporaneous (panel A) and predictive regressions (panel B), and regardless of whether or not additional controls are included. This table, thus, suggests that within-equity market interactions with dealer-exposed hedge funds are a transmission mechanism for the broker-dealer shocks to pass on to mutual funds and other non–hedge fund investment advisors. In the next subsection I show that the shock transmission is also associated with deteriorating market liquidity conditions among the same set of stocks.

Next, to explore whether direct interaction with dealer banks in other asset classes exposes mutual funds and non-hedge fund investment advisors to dealer shocks, I obtain institutional holdings data from FactSet. The advantage of the FactSet data in this regard is that while they also categorize institutional investors into different classes of institutions, they additionally allow me to link many investment managers at the financial institution level (primarily for mutual funds) with their constituent individual funds. Furthermore, in addition to security-level detail on holdings for equities, the FactSet data provide the size of aggregated positions that individual funds take in other non-equity asset classes.[26] I focus on exposure to bond markets, including corporate bond and treasury markets. This is because dealer banks and broker-dealers are the central intermediaries in corporate bond markets, whose prices are particularly sensitive to their distress (Haddad and Muir, 2021; He et al., 2022), while, for treasuries, dealer banks play a crucial role as the primary dealers and shocks to their risk-bearing capacity can explain price patterns in these markets(Li and Xu, 2024). Moreover, while mutual funds constitute the largest active investors in equity markets, they are also very active in bond markets (He et al., 2022). Thus, mutual fund companies that invest in both bond and equity markets may be exposed to the intermediation of dealer banks and broker-dealers because of their direct interactions in bond markets.

I again re-estimate $\epsilon_{i,t}$ in equation (8), first for the mutual funds and non–hedge fund investment advisors that are currently exposed to bond markets, and then for the set that are not exposed to bond markets. The analysis is in Table 11. Constrained by FactSet data availability, I can go back only to the first quarter of 2000 for this analysis. Supporting the hypothesis that bond market exposure connects mutual funds' marginal utility with dealer banks and broker-dealers, I find a highly significant loading only when sorting on the residual holdings of bond-exposed mutual funds—as can be seen when we compare columns

[26] I offer more details on the FactSet data in appendix C.3.

(1) and (2) of Table 11, which focus on bond-exposed funds, and columns (3) and (4), which include only the holdings funds with no bond market exposure. This is true whether or not I include the full set of controls and whether I examine contemporaneous covariances in panel A or conditional predictability in panel B. Overall, the results in the first two panels of Table 11 support that investors who are likelier to interact directly with dealer banks and broker-dealers are more exposed shocks to intermediary risk-bearing capacity.

5.2 Evidence of Liquidity Spirals

When broker-dealers and dealer banks suffer losses, their ability to provide funding liquidity or to act as central intermediaries in diverse asset markets is hampered. As Brunnermeier and Pedersen (2008) show, the resulting tighter conditions can cause dry-ups in market liquidity, as the loss of access to capital or the scarce presence of intermediaries able to step in and provide liquidity diminishes investors' ability to trade freely. This can in turn trigger more losses, which worsens funding conditions and causes further market liquidity dry-ups—a "liquidity spiral." Under such a mechanism, when dealer banks/broker-dealers experience hard times, we should expect to see larger dry-ups in market liquidity among equities largely held by more dealer-connected financial institutions.

I now present an empirical test to show that such a mechanism can speak to my findings. In particular, I compute the portfolio equal-weighted average growth rate in the stock-level Amihud (2002) illiquidity index and take the difference for each of the top-minus-bottom-quintile portfolios for the different versions of residual intermediation $\epsilon_{i,t}$ discussed in the previous subsection. I then regress the relative illiquidity growth on the intermediary shock and the non-financial market index and controls:[27]

$$\begin{aligned}\Delta \log \left(\overline{\text{Illiq}}\right)^{Q5,j}_{t+1} - \Delta \log \left(\overline{\text{Illiq}}\right)^{Q1,j}_{t+1} = {} & \alpha_j + \delta_{1,j}\text{Intermediary Shock}_{t+1} \\ & + \delta_{2,j}(\text{Mkt}^{NonFin}_{t+1} - R_{f,t}) + \delta_{3,j}X_{t+1,j} + \psi_{i,t+1}\end{aligned} \tag{15}$$

Here, $Q5$ and $Q1$ denote the top- and bottom-quintile portfolios, and j indexes the versions of residual intermediation for different groups of institutions. The results from my estimating (15) are in Table 10. In the first two columns, I use my baseline measure of

[27] The Amihud (2002) index for stock k over time period t (in this case, a quarter) is defined as

$$\text{Illiq}_{k,t} = \frac{1}{D_{k,t}} \sum_{\tau=1}^{D_{k,t}} \frac{|R_{k,\tau}|}{VOLD_{k,\tau}}$$

where τ indexes trading days in period t, $D_{k,t}$ is the number of trading days available for stock k in period t, $|R_{k,\tau}|$ is the absolute value return on stock k on trading day τ, and $VOLD_{k,\tau}$ is the dollar trading value for stock k on trading day τ.

$\epsilon_{i,t}$. Columns (3) and (4) focus on the holdings of hedge funds; columns (5) and (6) form portfolios based on the holdings of mutual funds/non–hedge fund investment advisors, but also require stocks to have a high amount of alignment with the holdings of hedge funds; and, finally, columns (7) and (8) form portfolios based on the holdings of mutual funds/non–hedge fund investment advisors, but instead where there is a low amount of alignment with hedge fund holdings. Negative intermediary shocks strongly predict relative liquidity dry-ups for my baseline portfolio. This is concentrated particularly among stocks held intensively by hedge funds or mutual funds trading in stocks with high degree of hedge fund alignment; there is no relationship in relative declines in liquidity conditions for stocks with low hedge fund alignment. This holds true when I control only for the non-financial market factor or when I include the full complement of risk factor controls. This is consistent with Jylhä, Rinne, and Suominen (2013), who argue that hedge funds typically supply liquidity to other institutions such as mutual funds but demand liquidity in crisis times when funding dries up. This results in worsening liquidity conditions for stocks held by hedge funds, or those held by other institutions that are likelier to encounter hedge funds as natural liquidity providers.

In terms of magnitudes, a negative intermediary shock of one standard deviation reduces liquidity for the highly intermediated portfolio of stocks by approximately 10% (in annualized terms) for my baseline sort or within the high hedge fund alignment group. This figure is 6% for hedge funds (though again these magnitudes are not strictly comparable for the reasons described in the previous section).

I do a related analysis in panel C of Table 11, which shows that relative liquidity dry-ups in response to the intermediary shock are also much more strongly concentrated among stocks held intensively by the bond–exposed mutual funds and non–hedge fund investment advisors who are more likely to be directly exposed to dealer intermediation in bond markets. Thus, a central consequence of shocks to the intermediation capacity of dealer banks and broker-dealers is the deteriorating market liquidity conditions of the stocks that dealer-exposed institutions trade in.

5.3 Ruling Out Explanations Related to Mutual Fund Flow–Driven Price Pressures

One concern with my findings could be that retail investors make withdrawals from mutual funds in crisis periods and these flows happen to correlate with my measures of intermediary risk-bearing capacity. For example, Ben-David et al. (2012) show that mutual funds also suffered redemptions in the crisis period, although they were not as severe as those of hedge funds. A literature starting with Sirri and Tufano (1998) documents a flow–performance relationship in the mutual fund sector, which could generate a spiral of price declines as outflows force liquidations that further reduce performance. Moreover, these redemptions could have been driven by households that invest in mutual funds rather than anything originating in the intermediary sector. I now demonstrate that such a mechanism cannot account for my results.

To do this, I show that different ways of controlling for mutual fund flows do not affect my findings. I explore controlling for two measures to capture a broad range flow-related prices pressures. I follow Lou (2012) in creating a stock-level measure of mutual fund flow-induced trading; I then average flow-induced trading for portfolios sorted on residual intermediation and create a high-minus-low index of relative flow-induced trading pressure, following Li (2022).[28] I also control for systematic fund flow shocks from Dou et al. (2023).[29] Dou et al. (2023) establish that mutual funds intentionally tilt their portfolios away from stocks that naturally covary with common shocks to mutual fund flows so as to hedge away flow-related risk; the result is that common movements in mutual fund flows command a risk premium in equity markets.

In Table A.7 I regress the high- minus low-intermediation portfolio excess returns on the intermediary shock, and I explore what happens to the coefficient on the intermediary shock when I add controls for these flow-related measures. All specifications in this table additionally include controls for Fama and French (2015) risk factors, the non-financial market factor, and the momentum factor. In the first column, I show the baseline full-sample coefficient on the intermediary shock without fund flow controls for ease of comparison. Since both Lou (2012) and Li (2022) show that flow-induced price pressures can mean revert, in the second column of Table A.7, I control for both contemporaneous and lagged relative flow-induced trading. While lagged relative flow-induced trading does obtain a significantly negative coefficient in column (2), controlling for it has zero effect on the intermediary shock coefficient estimate (the coefficient in column (1) is 4.74, with t-stat of 4.5, while the coefficient estimate in column (2) is 4.64 with t-stat of 4.35). Thus the price pressures from pure relative flow-induced trading are quite uncorrelated with the dealer-induced intermediary shocks that are my main focus.

Next, I control for different versions of the Dou et al. (2023) systematic fund flow shocks. Again for ease of comparison, in the third column I report my baseline estimates without

[28]See appendix C.4 for details on data construction. Consistent with these papers, in appendix C.4 I verify that fund inflows induce upward price pressure on stocks which subsequently mean reverts, so that contemporaneous flow-induced trading over the quarter covaries positively with returns, while flow-induced trading in the prior quarter predicts it negatively, and both in a statistically significant manner.

[29]I thank Winston Dou, Leonid Kogan, and Wei Wu for generously sharing their systematic fund flow shock data.

flow controls, but now for the sample period where I can observe the Dou et al. (2023) measures, which starts in the second quarter of 1991. In the columns (4) and (5), I control for two versions of the Dou et al. (2023) systematic fund flow shocks created from either CRSP or MorningStar mutual fund holdings data. I obtain a negative coefficient on contemporaneous systematic flows and positive coefficient on lagged systematic flows, with both being significant at the 10% level for the MorningStar version of the measure. However, these columns also show that differential exposure to systematic fund flow shocks cannot explain my findings, as the coefficient on the intermediary shock is essentially unaffected in terms of both magnitude and statistical significance by the inclusion of these measures. Finally, the result is also the same when I include all contemporaneous and lagged flow-related controls in column (6), including both versions of the Dou et al. (2023) measures, in addition to the relative flow-induced trading price pressures: while some of the flow controls do significantly relate to the spread portfolio excess returns, the coefficient on the intermediary shock without flow controls in this sample period is 4.26 (with t-stat of 4.27), versus a coefficient of 4.28 (with t-stat of 4.03) with the entire set of flow controls.

In summary, while these flow-based forces do have at least some explanatory power for returns for portfolios sorted on residual intermediation, this appears to be nearly orthogonal to the relative price pressures induced by my intermediary shock measure and, hence, controlling for them cannot explain my findings. Instead, the estimates in this paper are more consistent with frictions originating in the financial sector spilling over into differential price movements within equity markets.

6 Conclusion

Building on theoretical and empirical work that features constrained intermediaries as marginal investors, I show that the asset holdings of financial institutions generate higher covariances of more intermediated stocks with shocks to intermediary risk-bearing capacity, via temporary differential movements in discount rates. After stock fundamentals are accounted for, the excess returns of stocks held more by financial intermediaries covary more with shocks to intermediaries' ability to take on risk, and state variables capturing the health of financial intermediaries better predict the returns of the highly intermediated stocks than those of the less-intermediated stocks. The beta on shocks to intermediary risk-bearing capacity on a portfolio formed on the most-intermediated stocks is more than 5 times higher than the beta on the least-intermediated portfolio, despite my constructing the portfolios to net out any differences in size, book to market, investment (asset growth), profitability, and CAPM betas. In terms of return predictability, I can attribute between 23% and 44% of the time variation in equity premia in the least-intermediated portfolio, 38% and 59% in the most intermediated portfolio, and nearly all of the gap in conditional relative predictability between the two portfolios to time variation in intermediary risk-bearing capacity.

All of these findings are concentrated among institutional investors whose investment activities make them more connected with dealer banks and broker-dealers. My findings are especially prominent for hedge funds; they are also stronger when mutual funds are likely to interact directly with hedge funds in equity markets or for mutual funds that invest in dealer-intermediated bond markets.

Previous empirical papers testing frictional intermediary asset pricing theories have tended to focus on asset markets that are comparatively difficult for households to access. By contrast, I demonstrate that effects predicted by intermediary asset pricing models persist even among equities, which is perhaps the easiest asset class for households to directly access. In this sense, the findings in this paper may provide a lower bound on the quantitative role of intermediaries in affecting asset price movements in other asset classes.

References

Adrian, T. and N. Boyarchenko (2012). Intermediary leverage cycles and financial stability. *Becker Friedman Institute for Research in Economics Working Paper* (2012-010).

Adrian, T., E. Etula, and T. Muir (2014). Financial intermediaries and the cross-section of asset returns. *The Journal of Finance 69*, 2557–2596.

Adrian, T. and H. S. Shin (2014). Procyclical leverage and value-at-risk. *Review of Financial Studies 27*(2), 373–403.

Agarwal, V., V. Fos, and W. Jiang (2013). Inferring reporting-related biases in hedge fund databases from hedge fund equity holdings. *Management Science 59*(6), 1271–1289.

Agarwal, V., W. Jiang, Y. Tang, and B. Yang (2013). Uncovering hedge fund skill from the portfolio holdings they hide. *The Journal of Finance 68*(2), 739–783.

Agarwal, V., S. Ruenzi, and F. Weigert (2017). Tail risk in hedge funds: A unique view from portfolio holdings. *Journal of Financial Economics 125*(3), 610–636.

Agarwal, V., S. Ruenzi, and F. Weigert (2024). Unobserved performance of hedge funds. *The Journal of Finance 79*(5), 3203–3259.

Aghion, P., J. Van Reenen, and L. Zingales (2013). Innovation and institutional ownership. *American Economic Review 103*(1), 277–304.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets 5*, 31–56.

Anderson, C. and W. Liu (2018). The shadow price of intermediary constraints. Working paper.

Aragon, G. A. and P. E. Strahan (2012). Hedge funds as liquidity providers: Evidence from the lehman bankruptcy. *Journal of Financial Economics 103*(2012), 570–587.

Baker, M. and J. Wurgler (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance 61*, 1645–1680.

Basak, S. and A. Pavlova (2013). Asset prices and institutional investors. *American Economic Review 103*(5), 1728–1758.

Ben-David, I., F. Franzoni, and R. Moussawi (2012). Hedge fund stock trading in the financial crisis of 2007–2009. *The Review of Financial Studies 25*(1), 1–54.

Ben-David, I., F. Franzoni, R. Moussawi, and J. Sedunov (2021). The granular nature of large institutional investors. *Management Science 67*(11), 6629–6659.

Bernanke, B. and M. Gertler (1989). Agency costs, net worth, and business fluctuations. *American Economic Review 79*(1), 14–31.

Boller, L. and F. Scott Morton (2020). Testing the theory of common stock ownership. Working Paper 27515, National Bureau of Economic Research.

Brunnermeier, M. K. and L. H. Pedersen (2008). Market Liquidity and Funding Liquidity. *The Review of Financial Studies 22*(6), 2201–2238.

Brunnermeier, M. K. and Y. Sannikov (2014). A macroeconomic model with a financial sector. *American Economic Review 104*(2), 379–421.

Campbell, J. Y. and L. M. Viceira (2002). *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press.

Chen, H., S. Joslin, and S. X. Ni (2018). Demand for Crash Insurance, Intermediary Constraints, and Risk Premia in Financial Markets. *The Review of Financial Studies 32*(1), 228–265.

Chen, J., G. Tang, J. Yao, and G. Zhou (2022). Investor attention and stock returns. *Journal of Financial and Quantitative Analysis 57*(2), 455–484.

Cho, T. (2020). Turning alphas into betas: Arbitrage and endogenous risk. *Journal of Financial Economics 137*(2), 550–570.

Coval, J. and E. Stafford (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics 86*(2), 479–512.

Diep, P., A. L. Eisfeldt, and S. Richardson (2021). The cross section of mbs returns. *The Journal of Finance 76*(5), 2093–2151.

Dou, W. W., L. Kogan, and W. Wu (2023). Common fund flows: Flow hedging and factor pricing. *The Journal of Finance*, forthcoming.

Du, W., B. Hébert, and A. W. Huber (2022). Are Intermediary Constraints Priced? *The Review of Financial Studies 36*(4), 1464–1507.

Du, W., A. Tepper, and A. Verdelhan (2018). Deviations from covered interest rate parity. *The Journal of Finance 73*(3), 915–957.

Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. *The Journal of Finance 65*, 1237–1267.

Eisfeldt, A. L., H. Lustig, and L. Zhang (2023). Complex asset markets. *The Journal of Finance 78*(5), 2519–2562.

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*(1), 1–22.

Frazzini, A. and O. A. Lamont (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics 88*(2), 299–322.

Gabaix, X., A. Krishnamurthy, and O. Vigneron (2007). Limits of arbitrage: theory and evidence from the mortgage-backed securities market. *Journal of Finance 62*(2), 557–595.

Gârleanu, N., S. Panageas, and J. Yu. (2015). Financial entanglement: A theory of incomplete integration, leverage, crashes, and contagion. *American Economic Review 105*(7), 1979–2010.

Gompers, P. A. and A. Metrick (2001). Institutional investors and equity prices. *Quarterly Journal of Economics 116*(1), 229–259.

Gromb, D. and D. Vayanos (2018). The dynamics of financially constrained arbitrage. *The Journal of Finance 74*, 371–399.

Haddad, V. and T. Muir (2021). Do intermediaries matter for aggregate asset prices? *The Journal of Finance 76*(6), 2719–2761.

Haddad, V. and D. Sraer (2018). The banking view of bond risk premia. Working Paper.

He, Z., B. Kelly, and A. Manela (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics 126*, 1–35.

He, Z., P. Khorrami, and Z. Song (2022). Commonality in Credit Spread Changes: Dealer Inventory and Intermediary Distress. *The Review of Financial Studies 35*(10), 4630–4673.

He, Z. and A. Krishnamurthy (2012). A model of capital and crises. *The Review of Economic Studies 79*(2), 735–777.

He, Z. and A. Krishnamurthy (2013). Intermediary asset pricing. *American Economic Review 103*, 732–770.

He, Z. and A. Krishnamurthy (2018). *Intermediary asset pricing and the financial crisis*, Volume 10. Annual Review of Financial Economics.

Holmstrom, B. and J. Tirole (1997). Financial intermediation, loanable funds, and the real sector. *The Quarterly Journal of Economics 112*(3), 663–691.

Huang, D., F. Jiang, J. Tu, and G. Zhou (2014). Investor Sentiment Aligned: A Powerful Predictor of Stock Returns. *The Review of Financial Studies 28*(3), 791–837.

Jylhä, P., K. Rinne, and M. Suominen (2013). Do Hedge Funds Supply or Demand Liquidity?*. *Review of Finance 18*(4), 1259–1298.

Kargar, M. (2021). Heterogeneous intermediary asset pricing. *Journal of Financial Economics 141*(2), 505–532.

Kelly, B. and S. Pruitt (2013). Market expectations in the cross-section of present values. *The Journal of Finance 68*(5), 1721–1756.

Kewei, H., C. Xue, and L. Zhang (2015). Digesting anomalies: An investment based approach. *The Review of Financial Studies 28*(3), 650–705.

Koijen, R. and M. Yogo (2019). A demand system approach to asset pricing. *Journal of Political Economy 127*(4), 1475–1515.

Koijen, R. S. J., R. J. Richmond, and M. Yogo (2023). Which Investors Matter for Equity Valuations and Expected Returns? *The Review of Economic Studies 91*(4), 2387–2424.

Koijen, R. S. J. and M. Yogo (2015). The cost of financial frictions for life insurers. *American Economic Review 105*, 445–475.

Kondor, P. and D. Vayanos (2019). Liquidity risk and the dynamics of arbitrage capital. *The Journal of Finance 74*, 1139–1173.

Kozak, S., S. Nagel, and S. Santosh (2020). Shrinking the cross-section. *Journal of Financial Economics 135*(2), 271–292.

Krishnamurthy, A. . (2010). How debt markets have malfunctioned in the crisis. *Journal of Economic Perspectives 24*(1), 3–28.

Lettau, M. and S. Ludvigson (2001). Resurrecting the (c)capm: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy 109*, 1238–1287.

Li, J. (2022). What Drives the Size and Value Factors? *The Review of Asset Pricing Studies 12*(4), 845–885.

Li, K. and C. Xu (2024). Intermediary-based equity term structure. *Journal of Financial Economics 157*, 103856.

Lou, D. (2012). A Flow-Based Explanation for Return Predictability. *The Review of Financial Studies 25*(12), 3457–3489.

Ma, S. (2017). Heterogeneous intermediaries and asset prices. Working Paper.

Makarov, D. and A. V. Schornick (2010). A note on wealth effect under cara utility. *Finance Research Letters 7*(3), 170–177.

Mitchell, M., L. H. Pedersen, and T. Pulvino (2007). Slow moving capital. *American Economic Review, Papers and Proceedings 97*, 215–220.

Mitchell, M. and T. Pulvino (2012). Arbitrage crashes and the speed of capital. *Journal of Financial Economics 104*(3), 469–490.

Muir, T. (2017). Financial crises and risk premia. *Quarterly Journal of Economics 132*, 765–809.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics 78*, 277–309.

Newey, W. K. and K. D. West (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies 61*(4), 631–653.

Rapach, D. E., M. C. Ringgenberg, and G. Zhou (2016). Short interest and aggregate stock returns. *Journal of Financial Economics 121*(1), 46–65.

Shleifer, A. and R. W. Vishny (1997). The limits of arbitrage. *Journal of Finance 52*, 35–55.

Siriwardane, E. N. (2019). Limited investment capital and credit spreads. *Journal of Finance 74*(5), 2303–2347.

Sirri, E. R. and P. Tufano (1998). Costly search and mutual fund flows. *The Journal of Finance 53*(5), 1589–1622.

Stambaugh, R. F. (2014). Investment noise and trends. *Journal of Finance 69*(4), 1415–1453.

Weber, R. (2023). Institutional investors, households, and the time-variation in expected stock returns. *Journal of Financial and Quantitative Analysis 58*(1), 352–391.

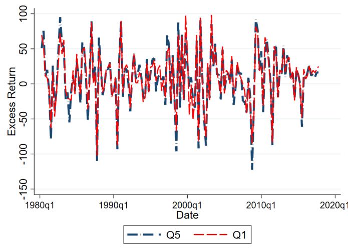
Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies 21*(4), 1455–1508.

Wermers, R. (2000). Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *The Journal of Finance 55*(4), 1655–1695.

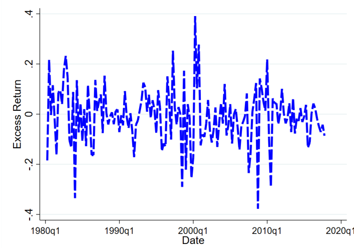
Figures

Figure 1: Annualized excess return for portfolios in the top and bottom intermediation ($\epsilon_{i,t}$) quintiles

Panel A: Top and bottom portfolios

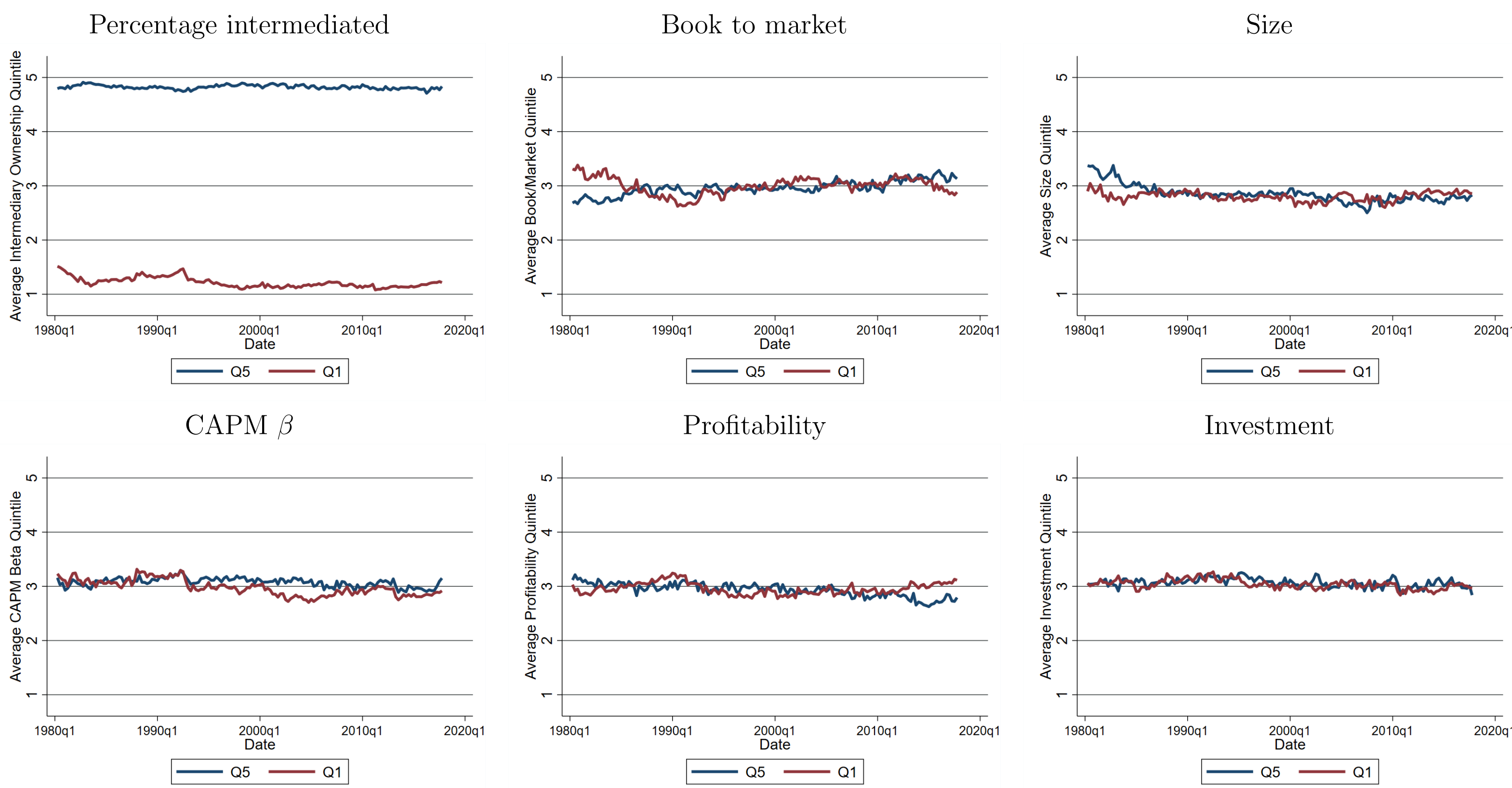


Panel B: Top-minus-bottom spread portfolio



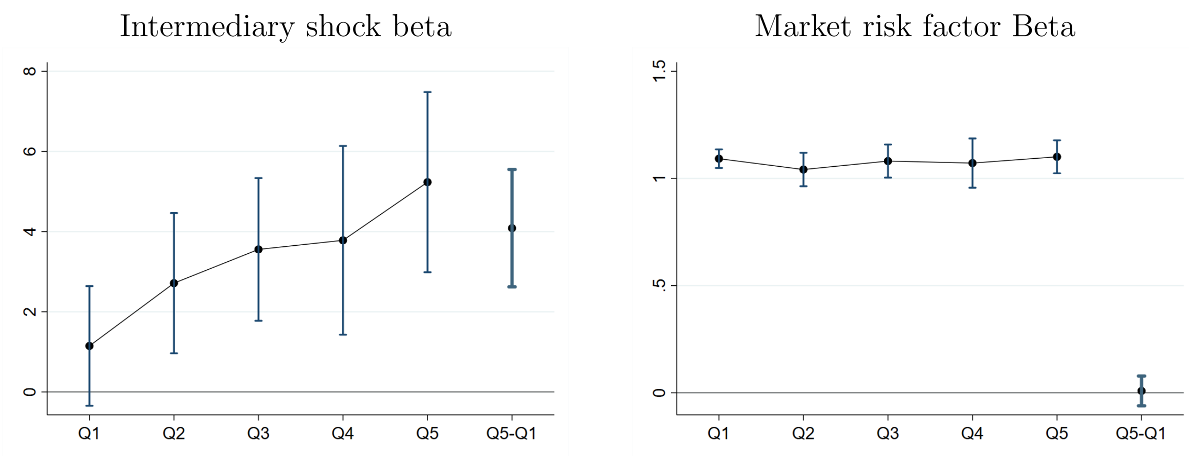
This figure shows the time series of annualized quarterly excess returns on the top- and bottom-quintile equal-weighted portfolios formed on the intermediation measure $\epsilon_{i,t}$ (panel A) and the difference between the two (panel B). Details on the construction of $\epsilon_{i,t}$ are presented in section 4.1. Sample spans 1980q2 to 2017q3.

Figure 2: Average quintile of given characteristic for top- and bottom-quintile intermediation ($\epsilon_{i,t}$) portfolios



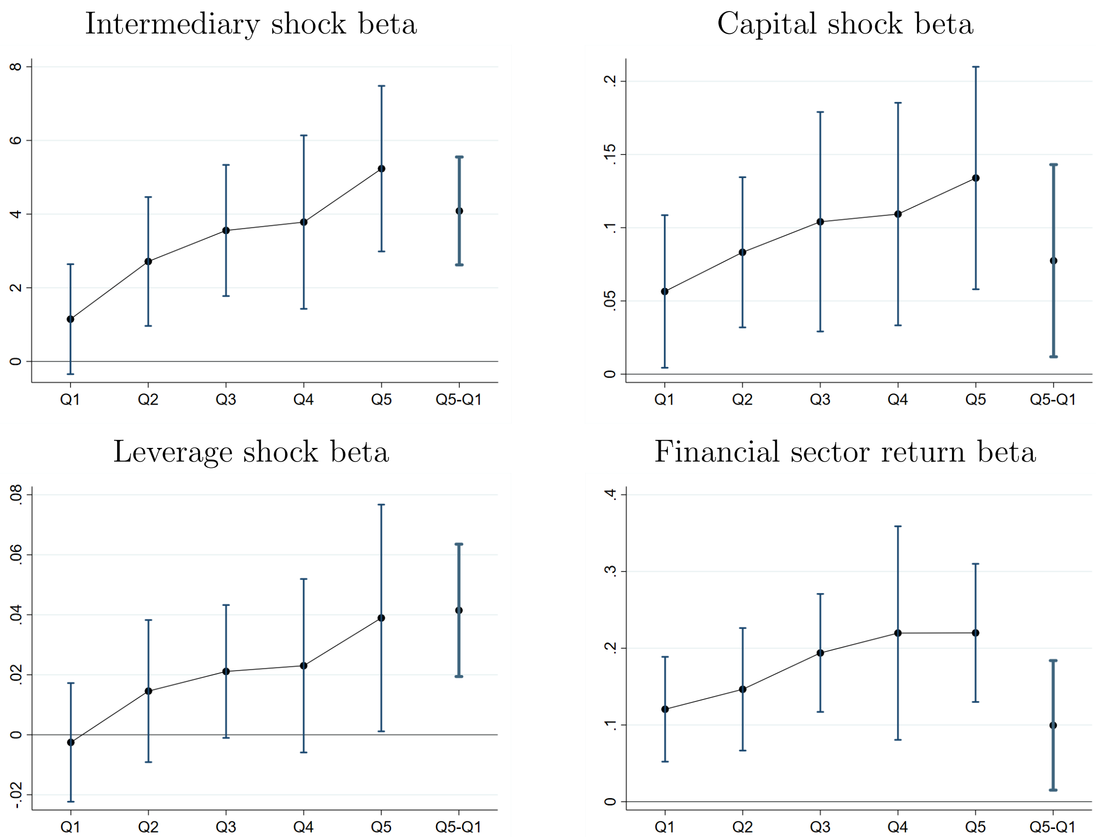
This figure shows the average quintile values over time on given characteristics for top- and bottom-quintile equal-weighted portfolios formed on the intermediation measure $\epsilon_{i,t}$. Details on the construction of $\epsilon_{i,t}$ are presented in section 4.1. Sample spans 1980q2 to 2017q3.

Figure 3: Coefficients on intermediary shocks and market risk over portfolios formed on intermediation quintile



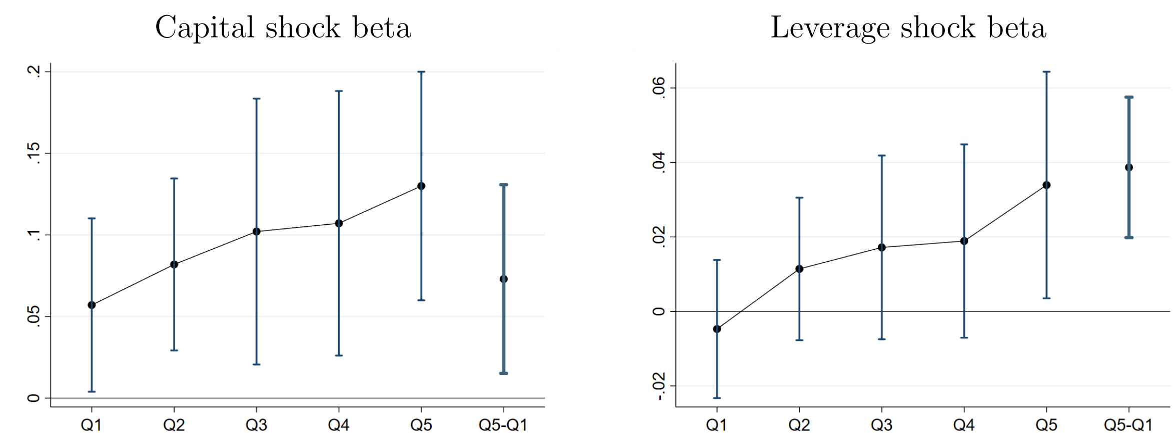
This figure plots coefficients from a regression as in (9) of the main text. The figure on the left shows the coefficient estimates on the average of the standardized Federal Reserve primary dealer equity capital ratio shocks from He et al. (2017) and the broker-dealer book leverage growth shocks from Adrian et al. (2014) for five portfolios formed on the intermediation measure $\epsilon_{i,t}$ (constructed to be uncorrelated with fundamental stock characteristics; section 4.1 discusses this in more detail), as well as the coefficient on the intermediary shocks for the top-minus-bottom-quintile spread. The figure on the right shows the corresponding betas on a version of the value-weighted market risk factor that excludes returns on financial stocks (SIC codes between 6000 and 6999). The confidence bands represent 95% confidence intervals computed from Newey–West standard errors with Newey and West (1994) optimal lags. The intermediary shock measure is standardized, and returns are in annualized percentage form. Sample spans 1980q2 to 2017q3.

Figure 4: Betas on portfolios sorted by intermediation on different shocks to intermediary risk-bearing capacity



This figure presents estimates from a regression as in (9) of the main text for each of the proposed intermediary shocks for each of the five portfolios formed on quintiles of the intermediation measure $\epsilon_{i,t}$ constructed in section 4.1 of the main text, as well as the top-minus-bottom-quintile portfolio spread. The capital shocks refer to the Federal Reserve primary dealer equity capital ratio shocks proposed in He et al. (2017), while the leverage shocks refer to the broker-dealer leverage shocks from Adrian et al. (2014). Intermediary shocks refer to the average of the standardized leverage and capital shocks. Financial sector return is the value-weighted return on the financial sector (stocks with SIC codes between 6000 and 6999). Regressions control for a version of the value-weighted market risk factor that excludes financial stocks. The intermediary shock measure is standardized, and returns are in annualized percentage form. Error bands represent 95% Newey–West confidence intervals with Newey and West (1994) optimal lags. Sample spans 1980q2 to 2017q3.

Figure 5: Betas on portfolios sorted by intermediation for capital and leverage shocks included in same specification

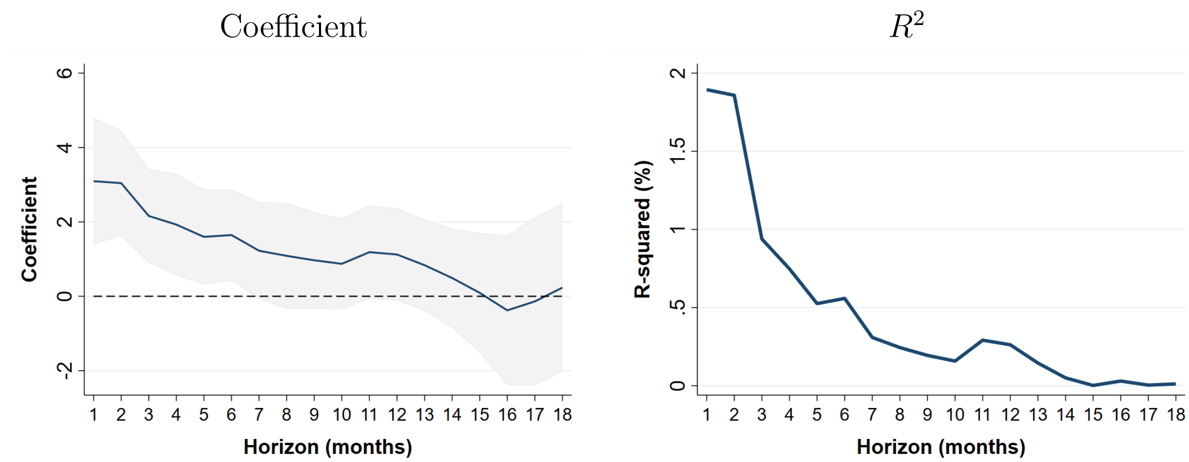


This figure presents estimates from regressions as in (10) of the main text:

$$R_t^e = \beta_{0,i} + \beta_{1,i}\text{Capital Shock}_t + \beta_{2,i}\text{Leverage Shock}_t + \beta_{3,i}\left(\text{Mkt}^{NonFin} - \text{Rf}\right)_t + \epsilon_{i,t}$$

The plots show betas on capital and leverage shocks included together in the same specification and for each of the five portfolios formed on quintiles of the intermediation measure $\epsilon_{i,t}$ constructed in section 4.1 of the main text, as well as the top-minus-bottom-quintile portfolio spread. The capital shocks refer to the Federal Reserve primary dealer equity capital ratio shocks proposed in He et al. (2017), while the leverage shocks refer to the broker-dealer leverage shocks from Adrian et al. (2014). Error bands represent 95% Newey–West confidence intervals with Newey and West (1994) optimal lags. Sample spans 1980q2 to 2017q3.

Figure 6: Predictability of lagged intermediary risk-bearing capacity for one-month high-minus-low-intermediation spread portfolio returns at different horizons



This figure shows coefficients obtained from predictive regressions of the one-month high-minus-low return spread for portfolios formed on the top and bottom quintiles of the intermediation measure $\epsilon_{i,t}$ (constructed in section 4.1 of the text) on predictor η_t at different monthly horizons. Regressions are of the form

$$R_{5,t+k} - R_{1,t+k} = \alpha + \beta_k \eta_t + \nu_t$$

as in equation (12) in the main text. The horizon k varies from 1 month to 18 months. The predictor η_t is the average of the standardized primary dealer squared leverage from He et al. (2017) and the negative of standardized broker-dealer leverage from Adrian et al. (2014). The gray shaded area corresponds to 90% Newey–West confidence intervals with Newey and West (1994) optimal lags.

Tables

Table 1: Fama–Macbeth regressions of percentage stock ownership by intermediaries on baseline stock characteristics

	Percentage Intermediated$_{i,t}$
Log Book Equity	0.061***
	(16.55)
Log Book Equity Sq.	-0.0047***
	(-12.07)
Profitability	0.027**
	(2.45)
CAPM Beta	0.046***
	(5.32)
Asset Growth	0.016***
	(2.76)
Book/Market	-0.013***
	(-6.71)
Observations	214448
Average R^2	0.11

This table shows the Fama–Macbeth time series average coefficients from the cross-sectional regression in (8):

$$\text{Percentage Intermediated}_{i,t} = \alpha_t + \beta_t X_{i,t} + \epsilon_{i,t}$$

for stocks included in the sample. At each time t, the top 1% of observations of Percentage Intermediated$_{i,t}$ are winsorized to deal with outliers in the cross section of institutional holdings. T-statistics in parentheses are computed with Fama–Macbeth standard errors, robust to 8 lags of autocorrelation. Average R^2 refers to the time-series average of the R-squared from each cross-sectional regression. The sample ranges from 1980q2 to 2017q3.

Table 2: Summary statistics of stock characteristics for portfolios sorted on quintiles of intermediation measure $\epsilon_{i,t}$

Panel A: Portfolio Characteristic Means

	% Inst	Log(ME)	Log(BE)	BE/ME	Asset Growth	Prof/BE	CAPM β
Q1	.19	6.86	6.18	.89	.14	.22	1.17
Q2	.34	7.21	6.51	.92	.12	.23	1.12
Q3	.42	7.24	6.51	.91	.12	.23	1.16
Q4	.5	7.14	6.43	.88	.13	.23	1.18
Q5	.62	6.91	6.18	.88	.14	.22	1.18

Panel B: Portfolio Characteristic Medians

	% Inst	Log(ME)	Log(BE)	BE/ME	Asset Growth	Prof/BE	CAPM β
Q1	.16	6.76	5.84	.8	.14	.22	1.16
Q2	.31	7.16	6.22	.83	.12	.23	1.1
Q3	.4	7.24	6.27	.81	.12	.23	1.13
Q4	.5	7.12	6.19	.81	.13	.23	1.17
Q5	.63	6.92	5.88	.81	.14	.22	1.18

This table shows the means and medians of percentage holdings by institutional investors (mutual funds, hedge funds, and investment advisors), log market equity, log book equity, book to market, asset growth (investment), profitability to book equity, and pre-ranking CAPM beta for the 5 portfolios formed on quintiles of the intermediation measure $\epsilon_{i,t}$. Details on the construction of $\epsilon_{i,t}$ are presented in section 4.1. Sample spans 1980q2 to 2017q3.

Table 3: Return summary stats for portfolios formed on quintiles of intermediation measure $\epsilon_{i,t}$

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
μ(Ex Ret)	9.92	11.01	10.16	11.03	9.07	-.86
t-stat	6.75	6.69	6.6	6.45	5.45	-1.53
σ(Ex Ret)	38.66	37.47	39.55	40.04	41.74	11.2
Sharpe Ratio	.26	.29	.26	.28	.22	-.08

This table reports the means, standard deviations, and Sharpe Ratios for the percentage annualized excess returns for portfolios formed on quintiles of intermediation measure $\epsilon_{i,t}$. Details on the construction of $\epsilon_{i,t}$ are presented in section 4.1. Sample spans 1980q2 to 2017q3.

Table 4: Regressions of excess returns for portfolios formed on intermediation measure $\epsilon_{i,t}$ on contemporaneous intermediary shocks

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Intermediary Shock	1.148	2.714***	3.556***	3.782***	5.234***	4.086***
	(1.52)	(3.07)	(3.95)	(3.17)	(4.60)	(5.51)
$\text{Mkt}^{NonFin} - \text{Rf}$	1.092***	1.042***	1.081***	1.072***	1.101***	0.009
	(49.74)	(26.36)	(27.68)	(18.43)	(28.28)	(0.25)
Observations	150	150	150	150	150	150
R^2	0.88	0.90	0.89	0.86	0.87	0.14

This table reports regressions of the intermediation measure $\epsilon_{i,t}$ portfolio quintile excess returns (and top minus bottom quintile spread) on risk factors as in (9) of the main text:

$$R^i_{t+1} - R_{f,t} = \alpha_i + \beta_{1,i}\text{Intermediary Shock}_{t+1} + \beta_{2,i}(\text{Mkt}^{NonFin}_{t+1} - R_{f,t}) + \nu_{i,t}$$

For the $i = Q1$ through $Q5$ quintile portfolios formed on $\epsilon_{i,t}$, as well as the high- minus- low-quintile ($Q5 - Q1$) portfolio excess returns. The first row of this table reports the coefficient estimates on the composite intermediary shock, which is the average of the standardized Federal Reserve primary dealer equity capital ratio shocks from He et al. (2017) and the broker-dealer book leverage growth shocks from Adrian et al. (2014). The intermediation measure $\epsilon_{i,t}$ is constructed to be uncorrelated with fundamental stock characteristics; section 4.1 discusses this in more detail. The second row shows the betas on a version of the value-weighted market risk factor that excludes returns on financial stocks (SIC codes between 6000 and 6999). The sample is quarterly and comprises 1980q2 to 2017q3. Newey–West t-statistics with Newey and West (1994) optimal lags are in parentheses. Quarterly excess returns are in annualized percentage form, and the intermediary shock is standardized. Sample spans 1980q2 to 2017q3.

Table 5: Changes in He et al. (2017) intermediary capital betas around S&P 500 inclusion

	(1)	(2)	(3)	(4)	(5)	(6)
Join S&P 500	0.10***	0.11***	0.10***	0.10***	0.11***	0.10***
	(3.24)	(3.13)	(2.86)	(3.07)	(3.75)	(3.00)
Observations	363	363	363	363	363	363
Comparison Group	Size	Size, Beta	Size, B/M	Size, Profit	Size, Invest	Size, Momentum

This table compares the coefficient of the He et al. (2017) intermediary capital risk factor for stocks before and after they enter the S&P 500. To be specific, it tests whether the difference-in-difference estimate

$$\Delta\widehat{\beta}^{HKM}_{i,t} - \frac{1}{N_{i,t}} \sum_{j \in \text{Comp Group}_{i,t}} \Delta\widehat{\beta}^{HKM}_{j,t}$$

for stocks i joining the S&P 500 is significantly different from zero; $N_{i,t}$ is the number of comparison firms for stock i in the month stock i joined the index, and firms j are those in the given comparison group. I measure the betas on the He et al. (2017) capital risk factor in the 30 months leading up to S&P 500 inclusion and the 30 months following and leave out the event month. I control for exposures to the Fama and French (2015) 5 factors plus momentum factor when estimating He et al. (2017) betas. I then compare the change in intermediary capital betas with the average intermediary capital betas for stocks in the same size quintile in that year (first column) or the same size quintile interacted with either CAPM beta, B/M, profitability, investment, or momentum (columns (2) through (6)): Stocks must have full coverage of returns before and after the event to be included. T-statistics in parentheses are based on Driscoll–Kraay standard errors that account for cross-sectional correlation and up to 60 lags, to allow for error correlation between observations with overlapping estimation windows for beta changes.

Table 6: Predictive regressions of excess returns for portfolios formed on intermediation measure $\epsilon_{i,t}$ on lagged intermediary risk-bearing capacity

	Q1	Q2	Q3	Q4	Q5	Q5-Q1
Intermediary Risk Aversion η_t	5.43*	6.20*	7.15**	7.34**	8.15**	2.72***
	(1.77)	(1.77)	(2.00)	(2.27)	(2.31)	(3.04)
Composite RP Z_t	8.45***	8.38***	8.46***	8.13**	8.50***	0.056
	(3.47)	(3.31)	(2.99)	(2.61)	(2.74)	(0.05)
Observations	150	150	150	150	150	150
R^2	0.081	0.093	0.095	0.091	0.097	0.056
RP η Variance Share Lower Bound	0.23	0.28	0.33	0.35	0.38	0.94
RP η Variance Share Upper Bound	0.44	0.49	0.54	0.57	0.59	0.9996

This table shows predictive regressions of the intermediation measure $\epsilon_{i,t}$ portfolio quintile excess returns on state variables η_t and Z_t:

$$R^i_{t+1} - R_{f,t} = \alpha_i + \beta_{1,i}\eta_t + \beta_{2,i}Z_t + \nu_{i,t}$$

where η_t is a measure of intermediary risk aversion defined in the main text and $i = Q1, \ldots, Q5$ denotes quintiles of portfolios sorted on residual intermediation $\epsilon_{i,t}$. The final column reports the coefficient for the regression with the high- minus low-residual intermediation portfolio excess return ($R^{Q5}_{t+1} - R^{Q1}_{t+1}$) on the left hand side. The control Z_t is a composite predictor for the conditional equity risk premium, constructed by taking the 14 predictor variables from Welch and Goyal (2008); the consumption-wealth ratio (*cay*) from Lettau and Ludvigson (2001); the aligned investor sentiment index from Huang et al. (2014); the aggregate short selling index of Rapach et al. (2016); and the investor attention index from Chen et al. (2022) and applying PLS to create an optimized single aggregate stock market return predictor from the information in each of these variables. T-statistics based on Newey-West standard errors with Newey and West (1994) optimal lags are in parentheses. The lower bound on the share of predictability coming from η is derived by regressing the given portfolio returns on a version of η that has been orthogonalized with respect to Z_t and comparing the resulting R^2 relative to the R^2 from the full bivariate regression reported in the table. The upper bound is derived in the same manner, except using the raw instead of orthogonalized version of η. Quarterly excess returns are in annualized percentage form and the independent variables are standardized. Sample spans 1980q2 to 2017q3.

Table 7: Panel regressions of stock excess returns on contemporaneous intermediary shocks interacted with intermediation measure $\epsilon_{i,t}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intermediary Shock $\times\ \epsilon_{i,t}$	8.13***			7.14***			7.71***		
	(5.01)			(4.10)			(3.72)		
Capital Shock $\times\ \epsilon_{i,t}$		0.15***			0.14**			0.17***	
		(3.24)			(2.28)			(2.77)	
Leverage Shock $\times\ \epsilon_{i,t}$		0.060**			0.060***			0.068**	
		(2.58)			(2.62)			(2.50)	
Ex Ret (Fin.) $\times\ \epsilon_{i,t}$			0.18***			0.21**			0.20*
			(3.02)			(2.10)			(1.94)
$\text{Mkt}^{NonFin} - \text{Rf} \times \epsilon_{i,t}$				0.064	0.021	-0.037	0.16*	0.11	0.074
				(0.69)	(0.18)	(-0.26)	(1.91)	(1.07)	(0.51)
SMB $\times\ \epsilon_{i,t}$							0.089	0.085	0.10
							(0.68)	(0.65)	(0.78)
HML $\times\ \epsilon_{i,t}$							-0.13	-0.17	-0.065
							(-0.92)	(-1.18)	(-0.37)
CMA $\times\ \epsilon_{i,t}$							-0.059	-0.035	-0.16
							(-0.40)	(-0.24)	(-1.05)
RMW $\times\ \epsilon_{i,t}$							0.52***	0.53***	0.50***
							(3.94)	(4.01)	(3.81)
UMD $\times\ \epsilon_{i,t}$							-0.093	-0.081	-0.070
							(-1.42)	(-1.21)	(-1.06)
$\epsilon_{i,t}$	-0.050***	-0.053***	-0.054***	-0.051***	-0.053***	-0.054***	-0.056***	-0.058***	-0.059***
	(-6.32)	(-6.56)	(-6.44)	(-6.22)	(-6.43)	(-6.37)	(-6.09)	(-6.38)	(-6.31)
Stock Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	211255	211255	211255	211255	211255	211255	211255	211255	211255
R^2	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24

This table shows estimates from panel regressions as in (13) of the main text:

$$R_{i,t+1} - R_{f,t} = \alpha_0 + \beta_1 F_{t+1} \times \epsilon_{i,t} + \beta_2 W_{t+1} \times \epsilon_{i,t} + \alpha_t + \alpha_i + \nu_{i,t+1}$$

Here, F_{t+1} denotes shocks to intermediaries, and W_{t+1} controls for other common shocks. The capital shocks refer to the Federal Reserve primary dealer equity capital ratio shocks proposed in He et al. (2017), while the leverage shocks refer to the broker-dealer leverage shocks from Adrian et al. (2014). Intermediary shocks refer to the average of the standardized leverage and capital shocks. Financial sector return is the value-weighted return on the financial sector (stocks with SIC codes between 6000 and 6999). Regressions control for a version of the value-weighted market risk factor that excludes financial stocks. Controls SMB, HML, CMA, RMW, UMD refer to the Fama and French (2015) risk factors and the up-minus-down momentum factor. In parentheses are t-statistics clustered by both firm and time to adjust for cross-sectional and time-series correlation in the residuals. The intermediary shock measure is standardized, and returns are in annualized percentage form. Sample spans 1980q2 to 2017q3.

Table 8: Panel Regressions of rolling stock-level intermediary betas on intermediation measure $\epsilon_{i,t}$

	Intermediary Shock	Capital Shock	Leverage Shock	Ex Ret (Fin.)
$\epsilon_{i,t}$	5.90***	0.14***	0.066***	0.20***
	(3.71)	(3.88)	(3.19)	(3.68)
Stock Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	130069	130069	130069	130069
R^2	0.50	0.52	0.44	0.54

This table shows the results from regressions of rolling individual stock betas on the residual intermediation measure $\epsilon_{i,t}$ (constructed to be uncorrelated with fundamental stock characteristics; section 4.1 discusses this in more detail). Stock betas are obtained from regressions of the form

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i \mathrm{F}_t + \beta_i^M \left(\mathrm{Mkt}_t^{NonFin} - \mathrm{Rf}_t\right) + \delta_{i,t}$$

in a window of plus or minus 15 quarters. I require a full window of observations for the estimated stock betas to be included in the sample. Reported coefficients are then estimated from panel regressions taking the form

$$\widehat{\beta}_{i,t-15\to t+15} = \alpha_0 + \beta_1 \epsilon_{i,t} + \beta_2 Z_{i,t} + \alpha_t + \alpha_i + \nu_{i,t}$$

Controls $Z_{i,t}$ include gross profitability, investment (asset growth), CAPM beta, book to market, second-degree polynomial in log market cap and log book equity, plus stock and time fixed effects. In parentheses are t-statistics double-clustered by stock and quarter. Returns and risk factors are expressed in annualized percentage terms, with the exception of the intermediary shock, which is standardized to have zero mean and unit variance. Sample spans 1980q2 to 2017q3.

Table 9: Regressions of high- minus low-residual intermediation portfolio returns on intermediary variables: the role of hedge funds

Panel A: Contemporaneous Regressions

	Hedge Funds Only		High HF Alignment		Low HF Alignment	
	(1)	(2)	(3)	(4)	(5)	(6)
Intermediary Shock	2.50***	2.76**	3.63***	4.50***	2.05	1.96
	(2.69)	(2.58)	(3.61)	(3.36)	(1.44)	(1.24)
$\text{Mkt}^{NonFin} - \text{Rf}$	-0.00078	0.021	0.013	0.040	-0.034	0.030
	(-0.03)	(0.79)	(0.42)	(0.63)	(-0.66)	(0.57)
Additional Risk Factor Controls	No	Yes	No	Yes	No	Yes
Observations	150	150	150	150	150	150
R^2	0.063	0.13	0.066	0.12	0.017	0.17

Panel B: Predictive Regressions

	Hedge Funds Only		High HF Alignment		Low HF Alignment	
	(1)	(2)	(3)	(4)	(5)	(6)
Intermediary Risk Aversion η_t	1.88**	1.83**	3.25***	2.99***	1.18	1.35
	(2.54)	(1.99)	(2.70)	(2.69)	(1.18)	(1.37)
Composite RP		0.20		1.22		-0.78
		(0.20)		(0.89)		(-0.75)
PLS RP Predictor Control	No	Yes	No	Yes	No	Yes
Observations	150	150	150	150	150	150
R^2	0.036	0.037	0.047	0.053	0.0075	0.011

This table shows the results from regressions of the top-minus-bottom-quintile returns for portfolios formed on residual intermediation measure $\epsilon_{i,t}$ on the intermediary shock and intermediary risk aversion proxy η_t. The contemporaneous regressions in panel A are of the form

$$R^{Q5}_{i,t+1} - R^{Q1}_{i,t+1} = \alpha + \beta_1 \text{Intermediary Shock}_{t+1} + \beta_2(\text{Mkt}^{NonFin}_{t+1} - R_{f,t}) + \beta_3 X_{t+1} + \nu_{t+1}$$

and the predictive regressions in panel B are of the form

$$R^{Q5}_{i,t+1} - R^{Q1}_{i,t+1} = \alpha + \beta_1 \eta_t + \beta_2 Z_t + \nu_{t+1}$$

Here, $\epsilon_{i,t}$ is formed on the basis of the percentage holdings for different classes of financial institutions. In the first two columns, I form portfolios on $\epsilon_{i,t}$ estimated from the holdings of hedge funds. I identify hedge funds in the 13F data using a list of hedge fund managers from Agarwal et al. (2024). In the last four columns, I form portfolios by sorting stocks into the top and bottom quintiles of residual intermediation $\epsilon_{i,t}$ from equation (8), based on the holdings of mutual funds and non–hedge fund investment advisors only, which I then split into two groups. The "High HF Alignment" group requires the stocks in the top (bottom) quintile of $\epsilon_{i,t}$ to also be in the top two (bottom two) quintiles of residual intermediation based on hedge fund holdings; the "Low HF Alignment" group compares the responses for the stocks in the top minus bottom quintiles of $\epsilon_{i,t}$, but within the opposite respective quintiles of hedge fund holdings as the "High HF Alignment". The additional risk factor controls in panel A include the Fama and French (2015) risk factors and the up-minus-down momentum factor. The predictive regressions in panel B row control for the composite risk premium measure Z_t from Table 6 obtained via PLS. Newey–West t-statistics with Newey and West (1994) optimal lags are reported in parentheses.

Table 10: Regressions of Amihud (2002) illiquidity growth for portfolios formed on intermediation measure $\epsilon_{i,t}$ on contemporaneous intermediary shocks, by institution type

	Baseline		Hedge Funds		High HF Alignment		Low HF Alignment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intermediary Shock	-9.95***	-10.3***	-5.99***	-6.20***	-10.2***	-10.3***	-3.15	-3.05
	(-4.64)	(-4.14)	(-3.16)	(-3.11)	(-3.78)	(-3.44)	(-1.43)	(-0.86)
$\text{Mkt}^{NonFin} - \text{Rf}$	0.18**	0.12	0.0074	-0.061	0.13	0.055	0.30***	0.25**
	(2.02)	(1.03)	(0.11)	(-1.14)	(1.11)	(0.39)	(3.07)	(2.00)
Additional Risk Factor Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	150	150	150	150	150	150	150	150
R^2	0.081	0.12	0.058	0.13	0.061	0.089	0.056	0.070

This table shows the results from regressions of the equal-weighted average growth rates of the Amihud (2002) illiquidity index for portfolios formed on intermediation measure $\epsilon_{i,t}$ (and the top-minus-bottom-quintile spread) on risk factors as in (15) of the main text:

$$\Delta\overline{\log(\text{Illiq})}_{5,t+1} - \Delta\overline{\log(\text{Illiq})}_{1,t+1} = \alpha + \delta_1 \text{Intermediary Shock}_{t+1} + \delta_2(\text{Mkt}_{t+1}^{NonFin} - R_{f,t}) + \delta_3 X_{t+1} + \upsilon_{i,t+1}$$

The first row of this table shows the coefficient estimates on the average of the standardized Federal Reserve primary dealer equity capital ratio shocks from He et al. (2017) and the broker-dealer book leverage growth shocks from Adrian et al. (2014) for five portfolios formed on the intermediation measure $\epsilon_{i,t}$ (constructed to be uncorrelated with fundamental stock characteristics; section 4.1 discusses this in more detail), as well as the coefficient on the intermediary shocks for the top-minus-bottom-quintile spread. The second row shows the betas on a version of the value-weighted market risk factor that excludes returns on financial stocks (SIC codes between 6000 and 6999). The sample is quarterly and comprises 1999q2 to 2017q3. Newey–West t-statistics with Newey and West (1994) optimal lags are in parentheses. Quarterly illiquidty growth rates are in annualized percentage form, and the intermediary shock is standardized.

Table 11: Regressions of high-minus-low-residual-intermediation portfolio returns and illiquidity growth on intermediary variables: Heterogeneity by mutual fund exposure to bond markets

Panel A: Contemporaneous

	Bond–Exposed MFs		Not Bond–Exposed MFs	
	(1)	(2)	(3)	(4)
Intermediary Shock	3.42***	3.82***	1.09*	-0.051
	(4.10)	(4.63)	(1.75)	(-0.04)
$\text{Mkt}^{NonFin} - \text{Rf}$	-0.00077	0.061*	-0.023	0.10*
	(-0.02)	(1.73)	(-1.21)	(1.99)
Additional Risk Factor Controls	No	Yes	No	Yes
Observations	71	71	71	71
R^2	0.11	0.22	0.011	0.13

Panel B: Predictive

	Bond–Exposed MFs		Not Bond–Exposed MFs	
	(1)	(2)	(3)	(4)
Intermediary Risk Aversion η_t	2.65***	5.28***	-1.17	-0.35
	(3.69)	(2.77)	(-1.63)	(-0.34)
Composite RP		-2.47		-0.78
		(-1.66)		(-0.84)
PLS RP Predictor Control	No	Yes	No	Yes
Observations	71	71	71	71
R^2	0.037	0.068	0.0078	0.011

Panel C: Contemporaneous (Relative Illiquidity Growth)

	Bond–Exposed MFs		Not Bond–Exposed MFs	
	(1)	(2)	(3)	(4)
Intermediary Shock	-8.69***	-7.69***	-2.42**	-0.081
	(-3.62)	(-3.78)	(-2.04)	(-0.04)
$\text{Mkt}^{NonFin} - \text{Rf}$	0.31***	0.27*	0.32***	0.25
	(2.96)	(1.75)	(3.56)	(1.60)
Additional Risk Factor Controls	No	Yes	No	Yes
Observations	71	71	71	71
R^2	0.17	0.24	0.16	0.22

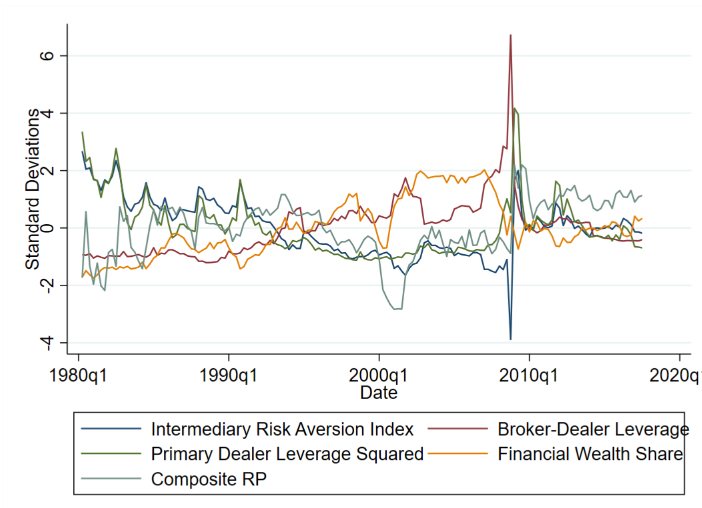
In this table I form portfolios from estimates of intermediation measure $\epsilon_{i,t}$ formed from the institutional holdings of non–hedge fund mutual funds and investment advisors, split by whether they currently report exposure to bond markets (columns (1) and (2)) or do not report such exposure (columns (3) and (4)). The holdings data and corporate bond exposure status come from FactSet. Panel A of this table reports the results from contemporaneous regressions of the average returns of the high-minus-low-intermediation portfolio formed from the top and bottom quintiles this alternative version of $\epsilon_{i,t}$ onto the intermediary shock plus controls, as in panel A of Table 9. Panel B explores the results from predictive regressions as in Panel B of Table 9. Panel C examines the contemporaneous average relative illiquidity response for the same portfolios, similarly to Table 10. See appendix C.3 for details on the FactSet data and how institutions are classified. The FactSet sample spans 2000q1 to 2017q3. Newey–West t-statistics with Newey and West (1994) optimal lags are in parentheses.

Appendices

A Appendix Figures and Tables

Appendix Figures

Figure A.1: Time series of intermediary state variables and composite risk premium used in predictability tests



This figure plots a time series of the state variables used in the return predictability tests reported in Tables 6 and A.1. Variables are standardized to have mean zero and unit standard deviation. Sample spans 1980q2 to 2017q3.

Appendix Tables

Table A.1: Predictability regressions of high- minus low-residual intermediation portfolio excess returns with alternative intermediary predictors

	(1)	(2)	(3)	(4)
PD Lev. Squared	1.996**		1.989***	
	(2.51)		(2.78)	
BD Lev.	-1.425		-1.415	
	(-1.53)		(-1.58)	
Fin. Share		-2.077***		-2.029**
		(-3.43)		(-2.37)
Composite RP Z_t			0.065	0.397
			(0.05)	(0.32)
Observations	150	150	150	150
R^2	0.057	0.032	0.057	0.034

This table shows predictive regressions of the intermediation measure $\epsilon_{i,t}$ portfolio top- minus bottom quintile portfolio excess returns on different sets of state variables X_t and Z_t:

$$R^{Q5}_{t+1} - R^{Q1}_{t+1} = \alpha + \beta_1 X_t + \beta_2 Z_t + \nu_{t+1}$$

where X_t either separately includes both the squared Federal Reserve primary dealer leverage from He et al. (2017) plus the broker leverage ratio from Adrian et al. (2014), or just the financial sector wealth share (proxied by the share of market capitalization in the aggregate stock market). The control Z_t is a composite predictor for the conditional equity risk premium (see main text or footnote to Table 6 for details). T-statistics based on Newey-West standard errors with Newey and West (1994) optimal lags are in parentheses. Quarterly excess returns are in annualized percentage form and the independent variables are standardized. Sample spans 1980q2 to 2017q3.

Table A.2: Robustness: contemporaneous portfolio regressions

	Original		Add WRDS Ratios		Just log(BE)		Value-Weighted		Drop Crisis	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intermediary Shock	4.09***	4.74***	3.14***	3.15***	2.93***	4.80***	4.73***	4.93***	3.72***	3.42***
	(5.51)	(4.51)	(4.14)	(4.23)	(2.84)	(3.84)	(4.52)	(3.70)	(3.15)	(2.77)
$\text{Mkt}^{NonFin} - \text{Rf}$	0.0087	0.024	-0.025	0.037	0.16***	0.11**	-0.0029	0.026	0.015	0.056
	(0.25)	(0.56)	(-0.88)	(1.28)	(3.05)	(2.16)	(-0.07)	(0.60)	(0.36)	(1.27)
Additional Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	150	150	150	150	150	150	150	150	144	144
R^2	0.14	0.23	0.074	0.18	0.23	0.38	0.080	0.12	0.078	0.18

This table contains the results from regressions of the excess returns of high-minus-low-residual-intermediation portfolio formed from the top and bottom quintiles of intermediation measure $\epsilon_{i,t}$ on the contemporaneous intermediary shock, with controls for other risk factors. The first two columns estimate the residual intermediation $\epsilon_{i,t}$ as done throughout the main text (and described in section 4.1); the next two add 40 financial ratios obtained from WRDS to cross-sectional regression (8) from the main text; columns (5) and (6) include only a second-degree polynomial in log book equity to estimate $\epsilon_{i,t}$. Columns (7)/(8) and (9)/(10) are, respectively, for value-weighted instead of equal-weighted portfolios and for a subsample that excludes the financial crisis (2008q1 through 2009q2). Odd columns control only for a version of the value-weighted market factor that excludes returns on financial stocks, and even columns add the Fama and French (2015) nonmarket risk factors plus the momentum factor. The intermediary shock measure is formed as an average of the standardized shocks to primary dealer equity capital from He et al. (2017) and broker-dealer leverage from Adrian et al. (2014). The intermediary shock measure is standardized, and returns are expressed in annualized percentage form. Newey–West t-statistics with Newey and West (1994) optimal lags are in parentheses. Sample spans 1980q2 to 2017q3.

Table A.3: Robustness: predictive portfolio regressions

	Original	Add WRDS Ratios	Just log(BE)	Value-Weighted	Drop Crisis
Intermediary Risk Aversion η_t	2.719***	2.687***	2.526**	3.540***	2.013***
	(3.04)	(3.63)	(2.33)	(3.13)	(3.07)
Composite RP Z_t	0.056	-0.543	1.363	-0.084	-0.095
	(0.05)	(-0.55)	(0.83)	(-0.04)	(-0.08)
Observations	150	150	150	150	144
R^2	0.056	0.065	0.043	0.053	0.027

This table shows the results from predictability regressions of the excess returns of the high-minus-low-residual-intermediation portfolio formed from the top and bottom quintiles of intermediation measure $\epsilon_{i,t}$ on η_t, a measure of intermediary risk aversion (given by the average of the standardized primary dealer squared leverage from He et al. (2017) and the negative of standardized broker-dealer leverage from Adrian et al. (2014)):

$$R^{Q5}_{t+1} - R^{Q1}_{t+1} = \alpha + \beta_1 \eta_t + \beta_2 Z_t + \nu_{t+1}$$

The control Z_t is the composite risk premium predictor used in Table 6 and constructed in section 4.2.2 of the main text. Regressions are for quarterly returns. The first column contains the original characteristics used to back out residual intermediation $\epsilon_{i,t}$ (as used throughout the main text and described in section 4.1), while the second column estimates residual intermediation $\epsilon_{i,t}$ by adding 40 financial ratios obtained from WRDS to the cross-sectional regression (8) used to obtain $\epsilon_{i,t}$; the third column includes only a second-degree polynomial in log book equity to estimate $\epsilon_{i,t}$. The last two columns are, respectively, for value-weighted instead of equal-weighted portfolios and for a subsample that excludes the financial crisis (2008q1 through 2009q2). Newey–West t-statistics with Newey and West (1994) optimal lags are in parentheses. All independent variables are standardized, and returns are in annualized percentage form. Sample spans 1980q2 to 2017q3.

Table A.4: Robustness: contemporaneous portfolio regressions (total 13F institutional holdings)

	Original		Add WRDS Ratios		Just log(BE)		Value-Weighted		Drop Crisis	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Intermediary Shock	3.74***	4.00***	2.80***	2.29***	2.77**	4.04***	4.45***	4.12***	3.50***	3.07***
	(5.17)	(4.76)	(3.50)	(4.03)	(2.17)	(3.34)	(4.17)	(4.68)	(2.75)	(3.56)
$\text{Mkt}^{NonFin} - \text{Rf}$	-0.013	0.026	-0.048	0.032	0.094	0.093*	-0.023	0.038	0.0014	0.061
	(-0.39)	(0.59)	(-1.22)	(1.25)	(1.63)	(1.68)	(-0.47)	(1.10)	(0.04)	(1.48)
Additional Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	150	150	150	150	150	150	150	150	144	144
R^2	0.10	0.25	0.061	0.23	0.13	0.34	0.088	0.18	0.060	0.23

This table contains the results from contemporaneous regressions of excess returns for the high-minus-low-residual-intermediation portfolio formed from the top and bottom quintiles of intermediation measure $\epsilon_{i,t}$ on risk factors

$$R^{Q5}_{t+1} - R^{Q1}_{t+1} = \alpha + \beta_1 \text{Intermediary Shock}_{t+1} + \beta_2(\text{Mkt}^{NonFin}_{t+1} - R_{f,t}) + \beta_3 X_{t+1} + \nu_{t+1}$$

Here, residual intermediation $\epsilon_{i,t}$ comes from total institutional holdings in the 13F data instead of only those from mutual funds and investment advisors. The first two columns estimate the residual intermediation $\epsilon_{i,t}$ as done throughout the main text (and described in section 4.1); the next two add 40 financial ratios obtained from WRDS to cross-sectional regression (8) from the main text; columns (5) and (6) include only a second-degree polynomial in log book equity to estimate $\epsilon_{i,t}$. Columns (7)/(8) and (9)/(10) are, respectively, for value-weighted instead of equal-weighted portfolios and for a subsample that excludes the financial crisis (2008q1 through 2009q2). Odd columns control only for a version of the value-weighted market factor that excludes returns on financial stocks, and even columns add the Fama and French (2015) nonmarket risk factors plus the momentum factor. The intermediary shock measure is formed as an average of the standardized shocks to primary dealer equity capital from He et al. (2017) and broker-dealer leverage from Adrian et al. (2014). The intermediary shock measure is standardized, and returns are expressed in annualized percentage form. Newey–West t-statistics with Newey and West (1994) optimal lags are in parentheses. Sample spans 1980q2 to 2017q3.

Table A.5: Robustness: Predictive portfolio regressions (total 13F institutional holdings)

	Original	Add WRDS Ratios	Just log(BE)	Value-Weighted	Drop Crisis
Intermediary Risk Aversion η_t	2.418***	2.093***	2.346**	3.285**	2.310**
	(2.90)	(2.67)	(2.29)	(2.53)	(2.52)
Composite RP Z_t	0.294	-0.191	0.789	0.944	0.325
	(0.35)	(-0.24)	(0.77)	(0.72)	(0.38)
Observations	150	150	150	150	144
R^2	0.050	0.045	0.034	0.060	0.040

This table shows the results from predictability regressions of excess returns for the high-minus-low-intermediation portfolio formed from the top and bottom quintiles of intermediation measure $\epsilon_{i,t}$ on η_t, a measure of intermediary risk aversion (given by the average of the standardized primary dealer squared leverage from He et al. (2017) and the negative of standardized broker-dealer leverage from Adrian et al. (2014)):

$$R^{Q5}_{t+1} - R^{Q1}_{t+1} = \alpha + \beta_1 \eta_t + \beta_2 Z_t + \nu_{t+1}$$

The control Z_t is the composite risk premium predictor used in Table 6 and constructed in section 4.2.2 of the main text. Here, residual intermediation $\epsilon_{i,t}$ comes from total institutional holdings in the 13F data instead of only those from mutual funds and investment advisors. Regressions are for quarterly returns. The first column contains the original characteristics used to back out residual intermediation $\epsilon_{i,t}$ (as used throughout the main text and described in section 4.1), while the second column estimates residual intermediation $\epsilon_{i,t}$ by adding 40 financial ratios obtained from WRDS to the cross-sectional regression (8) used to obtain $\epsilon_{i,t}$; the third column includes only a second-degree polynomial in log book equity to estimate $\epsilon_{i,t}$. The last two columns are, respectively, for value-weighted instead of equal-weighted portfolios and for a subsample that excludes the financial crisis (2008q1 through 2009q2). Newey–West t-statistics with Newey and West (1994) lags are in parentheses. All independent variables are standardized, and returns are in annualized percentage form. Sample spans 1980q2 to 2017q3.

Table A.6: Changes in institutional holdings around S&P 500 inclusion

	Δ Total IO		Δ Mutual/Inv. Adv. IO	
	(1)	(2)	(3)	(4)
Join S&P 500	0.040***	0.042***	0.020***	0.022***
	(4.51)	(4.81)	(2.95)	(3.29)
Observations	189354	188873	189354	188873
R^2	0.017	0.028	0.018	0.032
Size × Date FE	X	X	X	X
Stock FE		X		X

This table reports the results from regressions of the form

$$\Delta\text{Institutional Onwership}^k_{i,t} = \beta\text{Join S\&P 500}_{i,t} + \alpha_{\text{Size quintile}(i),t} + \alpha_i + \nu_{i,t}$$

where i indexes individual stocks, k denotes either all 13F institutions (columns (1) and (2)) or mutual funds and investment advisors, Join S&P $500_{i,t}$ is an indicator that stock i was added to the S&P 500 between the end of quarter $t-1$ and the start of quarter t, $\alpha_{\text{Size quintile}(i),t}$ is stock size quintile×year fixed effects, and α_i is stock fixed effects. The sample of stocks uses the same restrictions as in Table 7 plus the restriction that stocks that ever leave the S&P 500 in the sample period are excluded. T-statistics in parentheses are based on standard errors double-clustered by stock and date.

Table A.7: Regressions of high-minus-low-residual-intermediation portfolio excess returns on intermediary variables, with controls for mutual fund flow variables

	(1) Base (Full Sample)	(2) Flow Induced Trading	(3) Base (DKW Sample)	(4) CRSP Common Flow	(5) MorningStar Common Flow	(6) All
Intermediary Shock	4.74***	4.64***	4.26***	4.41***	4.32***	4.28***
	(4.51)	(4.35)	(4.27)	(4.59)	(4.79)	(4.03)
Common Fund Flow^{CRSP}				-2.57*		-1.20
				(-1.78)		(-0.45)
Common Fund Flow^{MorningStar}					-2.47*	-1.37
					(-1.88)	(-0.63)
Common Fund Flow^{CRSP}, Lagged				0.40		-0.39
				(1.28)		(-0.39)
Common Fund Flow^{MorningStar}, Lagged					0.45*	0.78
					(1.93)	(0.95)
Relative FIT		0.14				0.77***
		(0.74)				(3.52)
Relative FIT, Lagged		-0.21**				-0.61***
		(-2.48)				(-2.89)
Full Risk Factor Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	150	149	105	105	105	105
R^2	0.23	0.25	0.26	0.30	0.30	0.34

This table contains the results from regressions of excess returns for the high-minus-low-residual-intermediation portfolio formed from the top and bottom quintiles of intermediation measure $\epsilon_{i,t}$ on the contemporaneous intermediary shock, with controls for shocks related to mutual fund flows. Column (1) shows my baseline estimate for the full sample period without any additional flow-related controls. Column (2) adds contemporaneous and lagged controls for a measure of relative flow-induced trading constructed following (Lou, 2012; Li, 2022). For reference, column (3) shows the baseline estimate without flow controls, except during the Dou et al. (2023) sample period starting in 1991q3. Columns (4) and (5) control for different versions of systematic fund flow shocks from Dou et al. (2023), while column (6) additionally includes the relative flow-induced trading controls. Full risk factor controls include the value-weighted market factor excluding returns on financial stocks plus Fama and French (2015) nonmarket risk factors and the momentum factor). period, and column (4) does the same for the full sample period (1980q2 to 2017q3). See appendix C.4 for details on the construction of flow-induced trading measures. Newey–West t-statistics with Newey and West (1994) optimal lags are in parentheses.

Table A.8: Panel regressions of stock excess returns on contemporaneous intermediary shocks interacted with intermediation measure $\epsilon_{i,t}$ (with additional short interest and illiquidity controls)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intermediary Shock $\times$ $\epsilon_{i,t}$	8.25***			6.74***			5.63***		
	(4.98)			(3.81)			(2.66)		
Capital Shock $\times$ $\epsilon_{i,t}$		0.16***			0.14**			0.13**	
		(3.44)			(2.30)			(2.17)	
Leverage Shock $\times$ $\epsilon_{i,t}$		0.054**			0.054**			0.048*	
		(2.33)			(2.35)			(1.75)	
Ex Ret (Fin.) $\times$ $\epsilon_{i,t}$			0.20***			0.21**			0.15
			(3.42)			(2.27)			(1.45)
$\text{Mkt}^{NonFin} - \text{Rf} \times \epsilon_{i,t}$				0.094	0.045	-0.020	0.21**	0.15	0.13
				(1.06)	(0.42)	(-0.15)	(2.42)	(1.55)	(0.91)
SMB $\times$ $\epsilon_{i,t}$							0.32**	0.32**	0.33**
							(2.47)	(2.44)	(2.59)
HML $\times$ $\epsilon_{i,t}$							-0.076	-0.10	-0.024
							(-0.54)	(-0.72)	(-0.14)
CMA $\times$ $\epsilon_{i,t}$							-0.041	-0.031	-0.12
							(-0.29)	(-0.22)	(-0.84)
RMW $\times$ $\epsilon_{i,t}$							0.52***	0.53***	0.51***
							(4.15)	(4.17)	(4.03)
UMD $\times$ $\epsilon_{i,t}$							-0.036	-0.023	-0.017
							(-0.58)	(-0.36)	(-0.27)
$\epsilon_{i,t}$	-0.032***	-0.035***	-0.037***	-0.035***	-0.036***	-0.037***	-0.042***	-0.043***	-0.043***
	(-4.35)	(-4.61)	(-4.62)	(-4.41)	(-4.56)	(-4.60)	(-4.53)	(-4.70)	(-4.72)
Stock Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Short + Liquidity Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	207972	207972	207972	207972	207972	207972	207972	207972	207972
R^2	0.25	0.25	0.25	0.25	0.25	0.25	0.26	0.26	0.26

This table shows estimates from panel regressions as in (13) of the main text:

$$R_{i,t+1} - R_{f,t} = \alpha_0 + \beta_1 F_{t+1} \times \epsilon_{i,t} + \beta_2 W_{t+1} \times \epsilon_{i,t} + \alpha_t + \alpha_i + \nu_{i,t+1}$$

Here, F_{t+1} denotes shocks to intermediaries, and W_{t+1} controls for other common shocks. The capital shocks refer to the Federal Reserve primary dealer equity capital ratio shocks proposed in He et al. (2017), while the leverage shocks refer to the broker-dealer leverage shocks from Adrian et al. (2014). Intermediary shocks refer to the average of the standardized leverage and capital shocks. Financial sector return is the value-weighted return on the financial sector (stocks with SIC codes between 6000 and 6999). Regressions control for a version of the value-weighted market risk factor that excludes financial stocks. Controls SMB, HML, CMA, RMW, UMD refer to the Fama and French (2015) risk factors and the up-minus-down momentum factor. The additional "Short + Liquidity Controls" include the average short interest in the stock as a fraction of shares outstanding and the average of the log Amihud (2002) illiquidity index from $t-1$ to $t-4$, included individually and also interacted with the set of risk factors in the regression. In parentheses are t-statistics clustered by time and firm to adjust for cross-sectional and time-series correlation in the residuals. The intermediary shock measure is standardized, and returns are in annualized percentage form.

Table A.9: Panel regressions of rolling stock-level intermediary betas on intermediation measure $\epsilon_{i,t}$ (with additional short interest and illiquidity controls)

	Intermediary Shock	Capital Shock	Leverage Shock	Ex Ret (Fin.)
$\epsilon_{i,t}$	5.58***	0.12***	0.065***	0.18***
	(3.58)	(3.39)	(3.14)	(3.36)
Stock Fixed Effects	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Base Controls	Yes	Yes	Yes	Yes
Short + Liquidity Controls	Yes	Yes	Yes	Yes
Observations	130059	130059	130059	130059
R^2	0.50	0.53	0.44	0.54

This table shows the results from regressions of rolling individual stock betas on the residual intermediation measure $\epsilon_{i,t}$ (constructed to be uncorrelated with fundamental stock characteristics; section 4.1 discusses this in more detail). Stock betas are obtained from regressions of the form

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i \mathrm{F}_t + \beta_i^M \left(\mathrm{Mkt}_t^{NonFin} - \mathrm{Rf}_t\right) + \delta_{i,t}$$

within a window of plus or minus 15 quarters. I require a full window of observations for the estimated stock betas to be included in the sample. I then estimate the reported coefficients using panel regressions taking the form

$$\widehat{\beta}_{i,t-15\to t+15} = \alpha_0 + \beta_1 \epsilon_{i,t} + \beta_2 Z_{i,t} + \alpha_t + \alpha_i + \nu_{i,t}$$

Controls $Z_{i,t}$ include gross profitability, investment (asset growth), CAPM beta, book to market, a second-degree polynomial in log market cap and log book equity, and the average of the log Amihud (2002) illiquidity index from $t-1$ to $t-4$ plus stock and time fixed effects. Additional "Short + Liquidity Controls" include the average short interest in the stock as a fraction of shares outstanding and the average of the log Amihud (2002) illiquidity index from taken $t-1$ to $t-4$. In parentheses are t-statistics double-clustered by stock and quarter. Returns and risk factors are expressed in annualized percentage terms, with the exception of the intermediary shock, which is standardized to have zero mean and unit variance. Sample spans 1980q2 to 2017q3.

Table A.10: High– minus low–residual intermediation $\epsilon_{i,t}$ spread portfolio returns on intermediary shock and intermediary risk aversion index (alternative residualization by market cap decile)

	Contemporaneous		Predictive	
	(1)	(2)	(3)	(4)
Intermediary Shock	3.79***	3.91***		
	(5.73)	(3.62)		
Intermediary Risk Aversion η_t			2.82***	2.92***
			(2.67)	(3.18)
$\text{Mkt}^{NonFin} - \text{Rf}$	0.000062	0.029		
	(0.00)	(0.67)		
Composite RP Z_t				-0.44
				(-0.37)
Additional Risk Factor Controls	No	Yes		
Observations	150	150	150	150
R^2	0.12	0.23	0.065	0.066

This table shows the results from regressions of the excess returns from a high-minus-low-residual-intermediation portfolio formed on the top and bottom quintiles of $\epsilon_{i,t}$ on the intermediary shock or intermediary risk aversion index. The additional risk factor controls in column (2) include the Fama and French (2015) factors plus the up-minus-down momentum factor. In this table, I obtain the residual intermediation $\epsilon_{i,t}$ by sorting stocks in the sample for equation (8) into start-of-period market cap decile bins and then re-estimating (8) each quarter separately with decile bin–specific coefficients and market cap decile×year fixed effects. Newey–West t-statistics with Newey and West (1994) optimal lags are in parentheses. Sample spans 1980q2 to 2017q3.

B Model Extensions

B.1 Allowing for Correlation Between Intermediary and Household Risk Tolerance

Consider the following extension on the model from section 2: Suppose that household risk tolerance ρ_H is a function of both the state variable ζ, which does not move intermediary risk tolerance, and ω, which does induce changes in intermediary risk tolerance. Then, for local changes in ω and ζ,

$$
\begin{aligned}
\mathrm{d}P =& \frac{\left[\rho_I^{'}(\omega)\rho_H(\zeta,\omega) - \rho_I(\omega)\rho_{H,\omega}(\zeta,\omega)\right]\left[\Delta\Phi X + \Delta\phi - \epsilon_H\right] + \left[\rho_I^{'}(\omega) + \rho_{H,\omega}(\zeta,\omega)\right]\Sigma\mathbf{1}}{R_f(\rho_I(\omega)+\rho_H(\zeta,\omega))^2}\mathrm{d}\omega \\
&+ \frac{\rho_I(\omega)\rho_{H,\zeta}(\zeta,\omega)\left[-\Delta\Phi X - \Delta\phi + \epsilon_H\right] + \rho_{H,\zeta}(\zeta,\omega)\Sigma\mathbf{1}}{R_f(\rho_I(\omega)+\rho_H(\zeta,\omega))^2}\mathrm{d}\zeta
\end{aligned} \tag{16}
$$

Proposition 3 follows easily from here. I focus on the first term in (16), which is the coefficient on the intermediary risk tolerance shock $\mathrm{d}\omega$. Since in proposition 3 I assume the partial derivative $\rho_{H_\omega}(\zeta,\omega) > 0$, then because $\rho_I^{'}(\omega)$ is multiplied by ϵ_H with negative sign and $\rho_{H_\omega}(\zeta,\omega)$ is multiplied by ϵ_H with positive sign, the two effects work in opposite directions for the coefficient on $d\omega$. Moreover, as the percentage intermediated is strictly decreasing in ϵ_H, the negative sign on $\rho_I^{'}(\omega)$ causes the betas on shocks to ω to increase with residual intermediation ($\tilde{\epsilon}$ defined in equation (4)), while $\rho_{H_\omega}(\zeta,\omega)$ does the opposite. Therefore, if the betas increase with residual intermediation (all else held constant), it must be because the price is responding to changes in intermediary risk tolerance, not to correlated shocks that move the risk tolerance of other agents.

It should be noted that this does not have to be the case if $\rho_{H_\omega}(\zeta,\omega) < 0$; however, it seems highly unlikely in practice that shocks to household and intermediary risk tolerance are negatively correlated. Indeed, Haddad and Muir (2021) argue that, if anything, $\rho_{H_\omega}(\zeta,\omega) \geq 0$, as episodes where intermediaries become more risk averse are also likely to be periods of time when household risk aversion increases (the financial crisis of 2008–2009 being a particularly salient example).

B.2 Model Version With Constant Relative Risk Aversion

In the main text, I assumed CARA preferences and allowed for wealth effects indirectly by modeling wealth-dependent risk tolerance of intermediaries. I show below that the model can be couched in terms of constant relative risk aversion (CRRA) utility with log-normal payoffs, which explicitly includes equilibrium wealth effects, though closed-form expressions require some approximations. Let γ_I denote the coefficient of relative risk aversion and w_I the wealth of the intermediary.

Assume agents j believe asset payoffs $D \sim \text{LogNormal}(\mu_j - \frac{1}{2}\sigma^2, \Sigma)$, where σ^2 is the vector taken from the diagonal of Σ.[30] I make the exact same assumptions about the structure of μ_j and Σ as in the main text. Let P be a matrix with the vector of equilibrium prices on the diagonal, p the vector of equilibrium log prices, and α_j the vector of portfolio weights for agents j. As Campbell and Viceira (2002) show, the solution to the portfolio choice problem for the intermediary (and analogously for the household) can be approximated by

$$w_I\alpha_I = P\theta_I = \frac{w_I}{\gamma_I}\Sigma^{-1}(\mu_I - p - r_f\mathbf{1}) \tag{17}$$

where the approximation is exact as the length of the time interval goes to zero. Imposing $\theta_I + \theta_H = \mathbf{1}$, the log market-clearing condition is:

$$p = \log(w_I\alpha_I + w_H\alpha_H)$$

Approximate $w_I\alpha_I$ and $w_H\alpha_H$ around $w_0\alpha_0$, where $w_0 = (w_I + w_H)/2$ and $\alpha_0 = 1/N$ is the average portfolio weight. Let $a = 2w_0\alpha_0$. Then,

$$\begin{aligned} p &\approx \log(a\mathbf{1}) + \frac{w_I}{a}(\alpha_I - \alpha_0\mathbf{1}) + \frac{w_H}{a}(\alpha_H - \alpha_0\mathbf{1}) \\ &= \frac{1}{a}\left[\bar{c} + \frac{w_I}{\gamma_I}\Sigma^{-1}(\mu_I - p - r_f\mathbf{1}) + \frac{w_H}{\gamma_H}\Sigma^{-1}(\mu_H - p - r_f\mathbf{1})\right] \end{aligned}$$

where $\bar{c} = a(\log(a\mathbf{1}) - \mathbf{1})$ is a constant vector. Then,

$$p = \left(a\Sigma + (\frac{w_I}{\gamma_I} + \frac{w_H}{\gamma_H})I\right)^{-1}\left[\Sigma\bar{c} + \frac{w_I}{\gamma_I}(\mu_I - r_f\mathbf{1}) + \frac{w_H}{\gamma_H}(\mu_H - r_f\mathbf{1})\right]$$

Now define $\rho_j \equiv \frac{w_j}{\gamma_j}$. Applying the Woodbury matrix identity to $\left(a\Sigma + (\frac{w_I}{\gamma_I} + \frac{w_H}{\gamma_H})I\right)^{-1}$ and using the fact that Σ can be expressed as $\beta\beta' + \lambda^2 I$, this gives

$$p = \frac{\rho_I(\mu_I - r_f\mathbf{1}) + \rho_H(\mu_H - r_f\mathbf{1}) + \nu}{a\lambda^2 + \rho_I + \rho_H} \tag{18}$$

where $\nu = \Sigma\bar{c} - k\beta$ is a vector and k is a scalar that depends on the entire cross-sectional distribution of stock characteristics.[31] The similarity between the expressions for log price

[30] This is simply an adjustment so that the final equilibrium expressions are in terms of μ_j instead of $\mu_j + \frac{\sigma^2}{2}$ in this log-normal setting.

[31] The constant k can be expressed as $\frac{1}{c}(\frac{1}{a} + \frac{1}{c}\beta'\beta)^{-1}\beta'\left[\Sigma\bar{c} + \rho_I(\mu_I - r_f\mathbf{1}) + \rho_H(\mu_H - r_f\mathbf{1})\right]$, where $c =$

(20) under CRRA preferences and price (3) under CARA preferences in the main text is immediately evident: In both cases, $\rho_I + \rho_H$ is in the denominator, and $\rho_I \mu_I + \rho_H \mu_H$ is in the numerator. Thus, a version of proposition 1 holds in the CRRA case.

To show this, I now allow ρ_I and ρ_H to depend on state variables ω and ζ, as before. The CRRA case makes clear that this risk tolerance depends directly on the agents' wealth since $\rho_j = \frac{w_j}{\gamma_j}$ so that ω and ζ would include initial wealth. In principle, γ_j could also be affected by changes in w_j, such as to model nonlinear risk premia spikes in equity-constraint intermediary asset pricing models; as discussed in section 2, γ_j may further be affected by leverage constraints. In either case, the exact functional form for this dependence is unimportant for the results below, and I capture it broadly by the dependence of ρ_I on ω. The following is the CRRA analogy to proposition 1:

Proposition 4 *Suppose $\rho_I'(\omega) > 0$. Then, under CRRA preferences, the component of the total derivative of log price, dp, due to changes in ω (labeled by β_ω) is strictly decreasing in ϵ_H.*

Taking the total derivative of (20) immediately yields 4:

$$\begin{aligned} dp =& \frac{\rho_I'(\omega)\left[\rho_H(\zeta)(\Delta\Phi X + \Delta\phi - \epsilon_H) - \nu + a\lambda^2(\mu_I - r_f \mathbf{1})\right]}{(\rho_I(\omega) + \rho_H(\zeta) + a\lambda^2)^2} d\omega \\ &+ \frac{\rho_H'(\zeta)\left[\rho_I(\omega)(-\Delta\Phi X - \Delta\phi + \epsilon_H) - \nu + a\lambda^2(\mu_H - r_f \mathbf{1})\right]}{(\rho_I(\omega) + \rho_H(\zeta) + a\lambda^2)^2} d\zeta \\ &\equiv \beta_\omega d\omega + \beta_\zeta d\zeta \end{aligned} \tag{19}$$

As in the main text, I have used that $\mu_I = X\Phi_I + \phi_I$, $\mu_H = X\Phi_H + \phi_H + \epsilon_H$, $\Delta\Phi = \Phi_I - \Phi_H$, $\Delta\phi = \phi_I - \phi_H$ and $\Sigma = \beta\beta' + \lambda^2 I$, with $\beta = X\Pi + \pi$. Since $\rho_I'(\omega) > 0$, the expression for β_ω in (19) is strictly decreasing in ϵ_H and hence increasing in residual intermediation.

Again, note that (19) resembles a regression of the change in log price ("stock return" for a CRRA investor) on shocks to ω and ζ, a response that grows smaller as households' preference for holding the asset ϵ_H increases, in the exact same manner as in the CARA version of the model in the main text.

This setting also delivers a version of proposition 2 for CRRA preferences. Define the risk premium on asset k by $E[r_{p,k}] = \mu_k - p - r_f$, and suppose for two assets that $X_1 = X_2$ so that the asset characteristics are the same but that $\epsilon_{H,1} < \epsilon_{H,2}$. Then,

$$E[r_{p,1} - r_{p,2}] = \frac{\rho_H(\zeta)(\epsilon_{H,2} - \epsilon_{H,1}))}{a\lambda^2 + \rho_I(\omega) + \rho_H(\zeta)}$$

$a\lambda^2 + \rho_I + \rho_H$.

which is strictly decreasing in ω. This leads to the CRRA version of proposition 2:

Proposition 5 *Consider assets 1 and 2 such that $X_1 = X_2$ while $\epsilon_{H,2} > \epsilon_{H,1}$. Let $E[r_{p,k}] = \mu_k - p - r_f$ denote the risk premium on asset k. Then, the difference in the risk premium on asset 1 and asset 2 decreases with ω, i.e., $\partial E[r_{p,1} - r_{p,2}]/\partial\omega < 0$.*

Finally, I use a Taylor approximation to obtain a linear relationship between percentage intermediated, characteristics, and (a monotonic transformation of) ϵ_H—the CRRA analogy to equation (4) in the main text.

The linearization constant $a = 2w_0\alpha_0$ used above is the price that would obtain if both agents held the same wealth and set portfolio weights to $1/N$. Applying the same linearization to the intermediary's optimal portfolio (17) implies

$$\alpha_I w_I = P\theta_I = \frac{w_I}{\gamma_I}\Sigma^{-1}(\mu_I - p\mathbf{1} - r_f\mathbf{1}) \approx (a(1-\log(a))I + p)\theta_I$$

where, in an abuse of notation, I now denote by p the matrix with log prices on the diagonal and 0 elsewhere. Then,

$$\theta_I \approx ((a(1-\log(a))I + p)^{-1}\frac{w_I}{\gamma_I}\Sigma^{-1}(\mu_I - p\mathbf{1} - r_f\mathbf{1})$$

Let $\theta_{I,k}$ denote the intermediary holdings of asset k. I show that $\theta_{I,k}$ can be expressed as the ratio of two affine equations in X_k, where X_k is the vector of characteristics of asset k. I first demonstrate that the log price vector $p\mathbf{1}$ is affine in X:

$$p\mathbf{1} = \frac{\rho_I(\mu_I - r_f\mathbf{1}) + \rho_H(\mu_H - r_f\mathbf{1}) + \nu}{a\lambda^2 + \rho_I + \rho_H} \tag{20}$$

where $\nu = \Sigma\bar{c} - k\beta$. Note that $\Sigma\bar{c} = (\beta\beta' + \lambda^2 I)\bar{c} = \beta'\bar{c}(\Pi X + \pi) + \lambda^2\bar{c}$ is affine in X since $\beta'\bar{c}$ is a scalar and, similarly, $k\beta = k(X\Pi + \pi)$ is affine in X. Then, since $\mu_I = X\Phi_I + \phi_I$ and $\mu_H = X\Phi_I + \phi_H + \epsilon_H$, p is affine in X and is of the form $p\mathbf{1} = b + X\Gamma + d\epsilon_H$, where $d > 0$ so that prices are increasing in ϵ_H. The derivation of equation (4) in the main text establishes that $\Sigma^{-1}(\mu_I - p\mathbf{1} - r_f\mathbf{1})$ is also of the form $\widehat{b} + X\widehat{\Gamma} - \widehat{d}\epsilon_H$, where $\widehat{d} > 0$ and hence is decreasing in ϵ_H.

For the kth asset, we then have that $\theta_{I,k}$ can be expressed as

$$\theta_{I,k} = \frac{\widehat{b} + \widehat{\Gamma}'X_k - \widehat{d}\epsilon_{H,k}}{b + \Gamma'X_k + d\epsilon_{H,k}} \tag{21}$$

Note that (21) is strictly decreasing in ϵ_H, but in a nonlinear fashion. Without loss of generality, assume characteristics X and ϵ_H are mean 0, and approximate (21) around $X_0 = 0$

and $\epsilon_0 = 0$. The first-order Taylor approximation is

$$\begin{aligned}\theta_{I,k} &\approx \frac{\widehat{b}}{b} + \left(\frac{b\widehat{\Gamma}' - \widehat{b}\Gamma'}{b^2}\right) X_j - \left(\frac{\widehat{db} + d\widehat{b}}{b^2}\right)\epsilon_{H,j} \\ &\equiv a + B'X_k + \tilde{\epsilon}_k\end{aligned} \tag{22}$$

The local approximation (22) is centered on a constant price (i.e., market cap) at the average portfolio weight and the average characteristic. In practice, this means the approximate relationship should be more accurate when I compare stocks of a similar market capitalization. To see whether this matters, I sort stocks within my baseline sample into market-cap decile bins every quarter. I then reestimate (8) separately by market cap decile×calendar quarter to back out an alternative estimate of the residual intermediation $\epsilon_{i,t}$. The new size decile–specific high-minus-low-intermediation spread portfolio has a correlation of 0.93 with the baseline spread portfolio. In Appendix Table A.10, I sort stocks on this alternative measure of $\epsilon_{i,t}$; the findings are nearly identical quantitatively to those under the baseline for both the contemporaneous and predictive regressions.

B.3 Model Version Under Mean–Variance Preferences in Returns

The model can alternatively generate explicit wealth effects in closed form without relying on approximations if I assume that agents have mean–variance preferences in the return on initial wealth instead of the final period level of wealth.

Let P be a matrix with the vector of prices on the diagonal. Let $\tilde{\Sigma} = P^{-1}\Sigma P^{-1}$ be the variance–covariance matrix of equilibrium gross returns, where Σ is the variance–covariance matrix of asset payoffs. Under mean–variance preferences over expected returns, the solution to the intermediary's problem (in terms of portfolio shares) is

$$\alpha_I = \frac{1}{\gamma_I}\tilde{\Sigma}^{-1}(E_I[r] - R_f\mathbf{1})$$

or, in terms of dollar holdings,

$$P\theta_I = \frac{w_I}{\gamma_I}P\Sigma^{-1}P(P^{-1}\mu_I - R_f\mathbf{1})$$

$$\implies \theta_I = \frac{w_I}{\gamma_I}\Sigma^{-1}(\mu_I - R_f P\mathbf{1})$$

Everything then proceeds as before by defining $\rho_I = \frac{w_I}{\gamma_I}$. This setting is therefore isomorphic to my assuming CARA preferences and modeling wealth-dependent risk tolerance $\rho_I = \frac{w_I}{\gamma_I}$, as in Koijen et al. (2023).

C Data Appendix

C.1 Construction of AEM Leverage Factor

As noted by Cho (2020), changes to the Federal Reserve flow-of-funds data have significantly altered the implied broker-dealer leverage ratio. From the first quarter of 2014, repo assets (reverse repo) are included in assets, and only repo liabilities, rather than net repo, are included in the liabilities section. To make my leverage factor consistent with the construction in the original Adrian et al. (2014) paper, I obtain the broker-dealer leverage from Table L128 of the 2013q4 flow-of-funds release. I then compute the leverage as

$$\text{Leverage}_t = \frac{\text{Total Financial Assets}_t}{\text{Total Financial Assets}_t - \text{Total Financial Liabilities}_t} \tag{23}$$

I then seasonally adjust as described in Adrian et al. (2014). Cho (2020) suggests that the following change allows one to extend the original AEM factor:

$$\text{Leverage}_t = \frac{\text{Total Financial Assets}_t - \text{Repo Assets}_t}{\text{Total Financial Assets}_t - \text{Total Financial Liabilities}_t - \text{FDI in US}_t} \tag{24}$$

This accounts for changes to foreign direct investment reflected in liabilities in later releases of the flow of funds. However, I find that when I use the above for the most recent releases, the values I obtain under the two methods (23) and (24) coincide for quarters until the end of 2010, at which point broker-dealer leverage begins on an upward spike for (24) relative to (23); this spike becomes so extreme that leverage becomes negative toward the end of the sample. Because of this issue, I simply follow Adrian et al. (2014) and use (23) through the 2013q4 release, and then I extend the series using (24) with updated flow-of-funds data, which is also consistent with the extended leverage factor data posted on Tyler Muir's website. I further seasonally adjust the leverage growth series using expanding window regressions of leverage growth on quarterly dummies as in AEM to arrive at my final leverage factor.

C.2 Selection of WRDS Ratios for Final Sample

For my robustness checks in section 4.5, I obtain the 73 financial ratios from the WRDS financial ratios suite. I find that data availability is sparse, so I do the following:

1. When firm dividend yield and dividend–price ratios are missing, I assume they are equal to zero.

2. I replace missing values for any variables with their lags as of up to 8 quarters prior.

3. I then check the fraction of missing observations for stocks that overlap with my main sample. If this fraction is greater than 1%, I exclude the ratio from the analysis.

This leaves the following ratios:
Enterprise value multiple, price to sales, price to cash flow, dividend payout ratio, net profit margin, operating profit margin before depreciation, operating profit margin after depreciation, gross profit margin, pre-tax profit margin, cash flow margin, return on assets, return on equity, return on capital employed, after-tax return on average common equity, after-tax return on invested capital, after-tax return on total stockholders equity, gross profit to total assets, common equity to invested capital, long-term debt to invested capital, total debt to invested capital, capitalization ratio, cash balance to total liabilities, total debt to total assets, total debt to EBITDA, long-term debt to total liabilities, cash flow to total debt, total liabilities to total tangible assets, long-term debt to book equity, total debt to total assets, total debt to capital, total debt to equity, asset turnover, sales to invested capital, sales to stockholders equity, research and development to sales, advertising expenses to sales, labor expenses to sales, accruals to average assets, price to book, and dividend yield.

Though the WRDS book-to-market ratio satisfies my sampling criteria, I also exclude this variable because I already include a version of book to market in the regression. Finally, I winsorize these variables cross-sectionally at the 1% level to deal with outliers.

C.3 FactSet Data

For the estimates in section 5 of the main text, I obtain institutional holdings data from the FactSet Ownership database. In particular, I obtain the equity positions of 13F institutions from the FactSet Fund and Institutional Holdings – Equity database. These data are based on institutional 13F filings, similarly to the Thomson Reuters data used in my main sample. On the other hand, the FactSet data classify the institution types in a more granular way than the classification available based on the Thomson Reuters data. I also use these data to identify hedge funds. I follow the classification scheme outlined in Koijen et al. (2023) as closely as possible. In particular, I follow Koijen et al. (2023) in grouping all non–hedge fund investment advisors and mutual funds into one group that includes the FactSet entity subtypes “IA” (investment advisor), “IC” (investment company), “RE” (research firm), “PP” (real estate manager), “SB” (subsidiary branch), and “MF” (mutual fund). FactSet occasionally assigns a rollup entity ID for 13F-filing institutions that can be rolled into a broader institution. Because entity types are assigned at this rollup institution level, which is sometimes broader than the level of aggregation of the 13F filer FactSet entities, I use the entity subtype of the rollup entity to identify the institution types for each security-level holding position.

In addition to security-level detail for equities, FactSet provides generic asset class–level holdings for individual funds and a mapping from the FactSet fund identifiers to the 13F institution that runs each fund. This includes the size of positions in corporate bond and treasury markets. I link the information from these fund-level aggregated positions back to the 13F institutions running the funds to divide investment advisors and mutual funds into different groups. The first group is any 13F institutions that own funds that currently report positions in bond markets.[32] This classification is motivated by evidence that bond markets are heavily intermediated by dealer banks (Haddad and Muir, 2021; He et al., 2022; Li and Xu, 2024). I group the remaining mutual funds and investment advisors into their own category. If an investment advisor or mutual fund cannot be mapped to its individual funds' positions, then I automatically put it in the non-bond-exposed category. While FactSet positions are available back to the first quarter of 1999, the generic holdings data are poorly populated in 1999. Consequently, I start my FactSet-based analysis in the first quarter of 2000, similarly to Koijen et al. (2023).

I use the ratio of FactSet institutional adjusted share holdings divided by FactSet reported shares outstanding to obtain the share of each stock that is held by a given 13F entity, and then I aggregate by institution type. I then reestimate the residual intermediation in equation (8) separately considering only the holdings for the given set of institutions. I also winsorize the holdings for each institution type at the 1% level each quarter. I then group stocks into portfolios sorted on measures of residual intermediation for the different subgroups, as described in section 5.1 of the main text.

C.4 Mutual Fund Flows

I use the CRSP survivorship bias free mutual fund database to construct quarterly returns and fund flows for US equity mutual funds. I use the MFLINKS dataset available on WRDS to group the CRSP share classes at the fund level, following Wermers (2000). As in Lou (2012), the fund flow for mutual fund k at time t is given by

$$\text{Flow}_{k,t} = \frac{TNA_{k,t} - TNA_{k,t-1}(1 + \mathrm{R}_{k,t})}{TNA_{k,t-1}} \tag{25}$$

where $\mathrm{R}_{k,t}$ is the fund return over quarter t and $TNA_{k,t}$ is the end-of-quarter total net assets of fund k. Again following Lou (2012), I compute flow-induced trading pressure for stock i

[32] Because the generic positions are commonly reported biannually instead of quarterly, I consider a position current if the institution reports bond market exposure in the current quarter or the prior quarter if the current quarter is not available.

in quarter t as

$$\text{FIT}_{i,t} = \frac{\sum_k \text{SharesHeld}_{i,k,t-1}\text{Flow}_{k,t}}{\sum_k \text{SharesHeld}_{i,f,t-1}} \tag{26}$$

Where $\text{SharesHeld}_{i,k,t-1}$ is the number of shares held by fund k of stock i at the end of quarter $t-1$. I obtain this data from the Thomson Reuters S12 file. Following Li (2022), I then compute portfolio-level flow-induced trading for portfolio p at time t as

$$\text{FIT}_t^p = \frac{1}{N_{p,t}} \sum_{i \in P_t} \text{FIT}_{i,t} \tag{27}$$

where $N_{p,t}$ gives the number of stocks in portfolio p at the start of period t and P_t is the set of stocks in portfolio p at the start of the period. I verify that FIT_t^p explains variation in returns on my portfolios by estimating

$$R_{t+1}^p - R_{f,t} = \alpha + \beta_1 FIT_{t+1}^p + \beta_2 FIT_t^p + \beta_3 FIT_t^p + \nu_{t+1}^p$$

For brevity I report results for $p = Q5, Q1$, the top- and bottom-quintile portfolios sorted on residual intermediation $\epsilon_{i,t}$, respectively. Results are below:

	Dependent Variable: Excess Returns for	
	Q1	Q5
FIT_{t+1}^{Q1}	1.89***	
	(3.03)	
FIT_{t+1}^{Q5}		2.10***
		(2.78)
FIT_{t}^{Q1}	-1.21**	
	(-1.98)	
FIT_{t}^{Q5}		-1.48**
		(-2.07)
Observations	149	149
R^2	0.28	0.25

Consistent with Lou (2012); Li (2022), flow-induced trading generates contemporaneous upward price pressure, so that FIT_{t+1}^p has a positive and highly-significant coefficient, while previous flow-induced trading is associated with reversals, so that FIT_t^p and FIT_{t-1}^p have negative coefficients. Finally, relative flow-induced trading for the high- minus low-residual intermediation portfolio is given by

$$\text{Relative FIT}_{t+1} = \text{FIT}_{t+1}^{Q5} - \text{FIT}_{t+1}^{Q1} \tag{28}$$

where $Q5$ and $Q1$ denote the top- and bottom-quintile portfolios sorted on residual intermediation $\epsilon_{i,t}$, respectively.