Corporate Bond Multipliers: Substitutes Matter*

Manav Chaudhary[†]   Zhiyu Fu[‡]   Jian Li[§]

March 15, 2023

Abstract

Corporate bonds have salient characteristics, such as credit rating and maturity, that affect investors' demand. Bonds with similar characteristics are more substitutable with one another compared to other bonds. However, existing methods for estimating the multipliers (the price increase in response to demand shocks) typically assume homogeneous cross-elasticities among bonds. In this paper, we introduce rich heterogeneous substitution patterns among bonds and demonstrate that security-level multipliers are an order of magnitude smaller than previously estimated and are essentially zero. Nonetheless, aggregated portfolios exhibit substantially larger multipliers, reflecting the reduced availability of near substitutes for more aggregated portfolios. The price impact of demand shocks reverts after a quarter. Finally, we find that the multiplier is larger for high-yield bonds, longer-maturity bonds, and bonds with greater arbitrage risks.

Keywords: Corporate bonds, inelastic demand, mutual funds, demand-based asset pricing

JEL codes: G10, G12, G23

[†]Department of Economics and Booth School of Business, University of Chicago. Email: mchaudhary@chicagobooth.edu

[‡]Department of Economics and Booth School of Business, University of Chicago. Email: zhiyufu@uchicago.edu

[§]Columbia Business School. Email: jl5964@columbia.edu

1 Introduction

Our paper seeks to understand the ability of the corporate bond market to absorb demand-driven price pressures. The bond market disruption in March 2020 has raised concerns regarding whether the corporate bond market has become more vulnerable to demand shocks, particularly from mutual funds.[1] The impact of demand-driven price pressures crucially depends on the availability of substitute assets. If there is an ample supply of close substitutes, arbitrageurs can easily hedge idiosyncratic risks and aggressively trade against demand shocks. And perhaps unique to the corporate bond market, salient characteristics such as credit rating and maturity make identifying close substitutes relatively straightforward. Bonds within the same rating and maturity category are closer substitutes than bonds that do not share these similar characteristics. We show that not accounting for the heterogeneous cross-substitution patterns results in misspecification and considerably biased price impact estimates.

We measure demand-driven price pressure using multipliers—the percentage increase in the asset's price if mutual funds exogenously increased their demand for one percent of the asset's amount-outstanding (Gabaix and Koijen, 2021). In a frictionless benchmark, demand-driven price pressures will have almost no price impact; hence, the multiplier would be close to zero. In addition, we estimate the "substitute passthrough" of close substitutes, which is defined as the percentage increase in an asset's price due to its close substitutes' prices rising by one percent.

Our main innovation in estimating multipliers is to allow for heterogeneous cross-elasticities. Specifically, we allow bonds to have a different cross-elasticity for bonds with similar vs. different characteristics. Explicitly allowing for heterogeneous substitution turns out to be quantitatively important. At the security level, we find the multiplier is 0.05, and the substitute passthrough is around 1. Ignoring these heterogeneous substitution patterns leads to considerable upward bias. If we assume homogeneous substitution among all bonds, we get multipliers of around 0.4. This is an order of magnitude larger than when heterogeneous substitution patterns are taken into account, and it is also roughly the magnitude existing studies that (implicitly) impose homogeneous cross-elasticity find. Our findings suggest that bonds are highly substitutable, and the market is quite good at absorbing security-specific demand shocks. With that said, the market is not as good at absorbing demand shocks at more aggregate levels, e.g., demand shocks for a specific-rating

---

[1]See for example https://www.brookings.edu/wp-content/uploads/2021/06/financial-stability_report.pdf.

category. As figure 1 shows, the more aggregate the demand shock, the higher the multiplier and the lower the substitute passthrough. For example, at the rating level, the multiplier is 3.5, and the substitute passthrough is statistically indistinguishable from zero. The considerable difference in the price impact between security-level vs. more aggregate demand shocks highlights the double-edged nature of substitutability. While high substitutability reduces price impact to direct demand shocks, it increases price spillover from indirect demand shocks to substitutes. We also investigate the persistence of the price impact and its heterogeneity along various portfolio characteristics. We find the price impact reverts after a quarter. And that the price impact is larger for high-yield bonds, longer maturity bonds, and bonds with larger arbitrage risks.

Figure 1: Demand is more inelastic for more aggregate portfolios
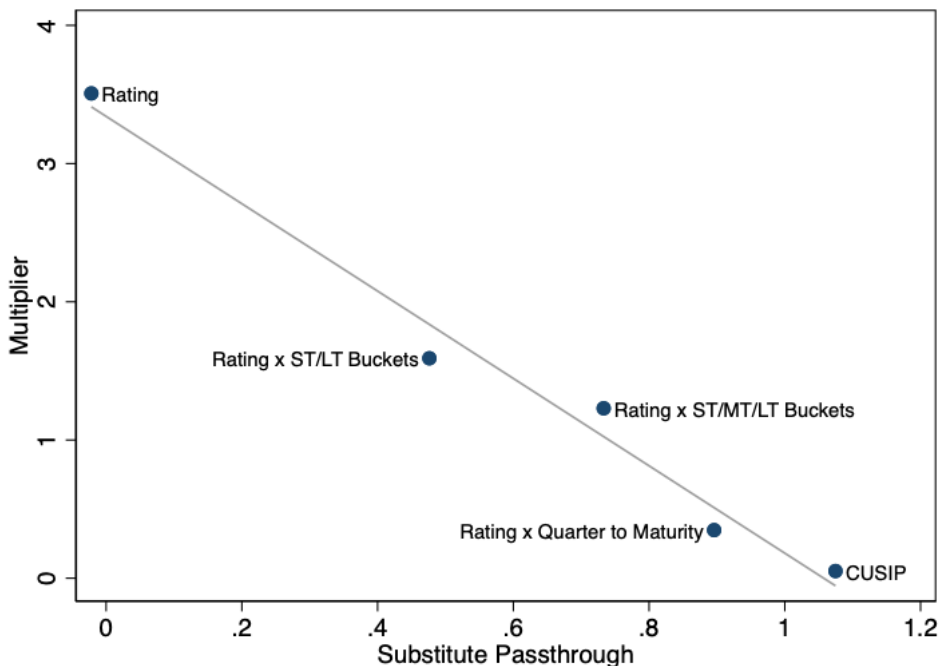
The figure shows plots the multiplier (inverse of elasticity) and substitute passthrough coefficients (link between substitute and test asset returns) for various levels of portfolio aggregation—see Table 4 for descriptions of the aggregation. The figure shows a negative relationship between multipliers and substitute passthroughs. It also shows multipliers are monotonically increasing in aggregation, whereas substitute pasthrough is monotonically decreasing in aggregation.

We construct exogenous demand shocks from flow-induced trading by mutual funds. Following

the literature, we assume mutual funds invest their flows into their existing portfolios proportionally (Lou, 2012).[2] We allow for autocorrelations in fund flows and extract the flow innovation terms. Furthermore, we adjust the demand shock to account for the expected cumulative trading from this flow innovation.[3] We then strip out common factor variations to remove potential confounding factors. Finally, we construct bond-level shocks by summing up fund-level shocks weighted by the fund's lagged holding share for each bond.

We then use these demand shocks to estimate price multipliers.[4] We show that as long as *lagged* mutual fund portfolio shares are orthogonal to unobserved demand shocks, we can use our (observed) demand shocks to identify the multipliers successfully. Our identification strategy is directly related to Bartik instrument applications where exogeneity of shares guarantees instrument validity (Goldsmith-Pinkham et al., 2020; Chaudhry, 2022). From this perspective, our identification strategy can be viewed as pooling multiple exogenous exposure research designs. Each bond's price is exposed to net retail inflows into a given mutual funds, but their degree of exposure to these flows exogenously depends on how much mutual fund's held of the bond in the previous period (conditional on controls).

To estimate multipliers for individual bonds, we regress bond returns on its demand shocks, controlling for the relevant substitute portfolio's returns. We find that allowing for heterogeneous cross-elasticity is quantitatively important for the estimate of the multiplier. Under the homogeneous cross-elasticity assumption, i.e. any given pair of bonds have the same cross-elasticity, the relevant substitute portfolio's return is simply the market return, or equivalently, a time-fixed effect. The estimated multiplier is significantly positive and around 0.4, similar to the estimate in the literature (Bretscher et al., 2022; Siani and Siani, 2022; Darmouni et al., 2023).

However, once we relax the homogeneous elasticity assumption, we find a much smaller multiplier. To allow for heterogeneous cross-elasticity, we add another control which is the returns of an additional substitute portfolio formed of bonds with the same detailed rating category as the testing asset.[5] To deal with the endogeneity of prices, we instrument the return of the substitute portfolio

---

[2]In appendix B.1 we relax this one-to-one passthrough assumption by directly looking at mutual fund rebalancing to estimate lower bounds for the demand passthrough, and use this to find upper bounds for the multiplier. This exercise suggests the security-level multiplier is at most 0.1, which is still 3-6 times smaller than what is typically found in studies that assume homogeneous cross-elasticities.

[3]This is the theoretically relevant shock when there are forward-looking investors Gabaix and Koijen (2021).

[4]We focus on the multipliers in this paper, but one can easily translates the estimates into the elasticity space, by taking the inverse of the multipliers.

[5]We refer to $AAA, AA+, AA - AA+ ...$ as detailed rating categories and $AAA, AA, A, BBB...$ as coarse rating

with the demand shocks to bonds in the substitute portfolio.[6] After allowing close substitutes to have different cross-elasticities than further away substitutes, we find the multiplier is around 0.05 and statistically indistinguishable from zero. In other words, the price impact of these demand shocks from mutual funds is close to zero. On the other hand, the substitute passthrough is close to one, i.e. the spillover from close substitute's prices to the bond's price is almost one-to-one.

Failing to account for close versus distant substitutes leads to significant over-estimation of the multiplier. Intuitively, when a bond is hit by a positive demand shock, its substitutes likely also receive similar demand shocks due to common ownership. The increase in the bond price comes from both its own demand shock and also its substitutes' prices being higher. If the latter channel is not accounted for in the right way, then the degree of price change due to its own demand shock would be mis-measured. Furthermore, the spillover effect from the asset's substitute price is larger if it is a close substitute. Assuming homogeneous substitution effectively under-weighs the impact from close substitutes, and attributes too much of the price movement to the asset's own demand shocks.

We also show that accounting for heterogeneous substitution is particularly important for the corporate bond market relative to the equity market. We repeat the same exercise for stocks to estimate the stock-level multipliers, with and without controlling for close substitutes. We define close substitutes as either stocks with similar loading on the Fama-French 3 factors, or stocks within the same industry groups. Controlling for close substitutes does reduce the multiplier estimated, but only slightly compared to the case for bonds, suggesting that the heterogeneity in cross-elasticity among stocks plays a weaker role in the equity market.

We then move on to estimate the multipliers for corporate bond portfolios. In contrast to the near-zero multiplier of individual bonds, we find that portfolios are affected more by demand shocks and have significantly positive multipliers. As the portfolio multiplier is monotonically increasing in the degree of aggregation, the substitute passthrough is monotonically decreasing. For portfolios formed by bonds in the same rating category with the same quarter-to-maturity, the multiplier is 0.35, significantly larger than the 0.05 estimate at the bond level. Furthermore, the portfolio substitute passthrough is estimated to be slightly lower than that at the bond level. At the most aggregated level, where each portfolio is formed by bonds in the same coarse rating group, the

categories.

[6]We also run a specification where we directly control substitute prices using substitute portfolio-time fixed effects rather than using instrumental variables, the multiplier estimates are essentially unchanged.

portfolio multiplier is 3.5 and the passthrough coefficient drops to 0.

Consistent with demand being more elastic in the long run, we find that the price response to demand shocks is not permanent and seems to revert fully. For our baseline level of aggregation, we find that price impact takes around two quarters to revert fully. This reversion rate is somewhat faster than the four-quarter reversion Li (2021) found for stock portfolios. Overall, our results suggest that demand is much more elastic in the long run.

Furthermore, we investigate heterogeneity in multipliers to uncover potential structural drivers of inelastic demand. We focus on portfolios formed by bonds in the same detailed rating category with the same quarter-to-maturity. First, we explore heterogeneity along credit risk and duration risk dimensions and find that portfolios consisting of bonds with lower ratings and longer maturities have significantly larger multipliers. Investment grade (IG) portfolios have near-zero multipliers, whereas high yield (HY) portfolios' have multipliers around 0.6—hence our baseline level multiplier estimate of 0.35 is predominantly driven by HY bonds. Specifically, the sharp change in multipliers around the IG/HY cutoff indicates that elasticity depends on investor clientele.

Finally, we compute the Sharpe ratio of a strategy that takes advantage of the price deviation between the testing asset and its substitute portfolio for one quarter. This is not a risk-free arbitrage strategy as price may not converge in the following quarter due to non-flow shocks and future flow shocks. We find the (annualized) Sharpe ratios are generally small, suggesting un-hedged risk may be a factor impeding arbitrage activities. Indeed, we find that in the cross-section of portfolios formed by detailed rating and maturity, portfolios with high arbitrage risks have significantly higher multipliers. A one-standard-deviation increase in arbitrage risk increases the multiplier by 0.45.

In the remainder of this section, we discuss the literature. We explain our demand structure and estimation methodology in Section 2. Section 3 describes our data and construction of demand shocks in detail. Section 4 presents the baseline estimation results and Section 5 focuses on arbitrage risks. Section 6 concludes.

1.1 Literature Review

Estimating multipliers directly contributes to the long literature in asset pricing that estimates the slope of the asset demand curve—multipliers are the inverse of demand elasticities.[7] The literature has shown asset demand is considerably less elastic than standard theories suggest. Partially due to data availability, this literature has traditionally focused on estimating the demand elasticity

---

[7]Our security-level multiplier estimate of 0.05 implies an elasticity of $1/0.05 = 20$.

of individual stocks (micro elasticities). More recently, researchers have begun estimating the micro elasticity of other asset classes, such as bonds, and more aggregate level elasticities, such as those of the entire equity market (macro elasticity). Our paper contributes to this broad body of work. Figure 2 shows how our estimates compare to a representative selection of micro elasticity estimates from the literature—our estimates confirm that bond demand is more elastic than stocks, but additionally we find that bonds are a lot more elastic than suggested by prior studies.

Figure 2: Our estimate vs. other micro stock and bond elasticity estimates
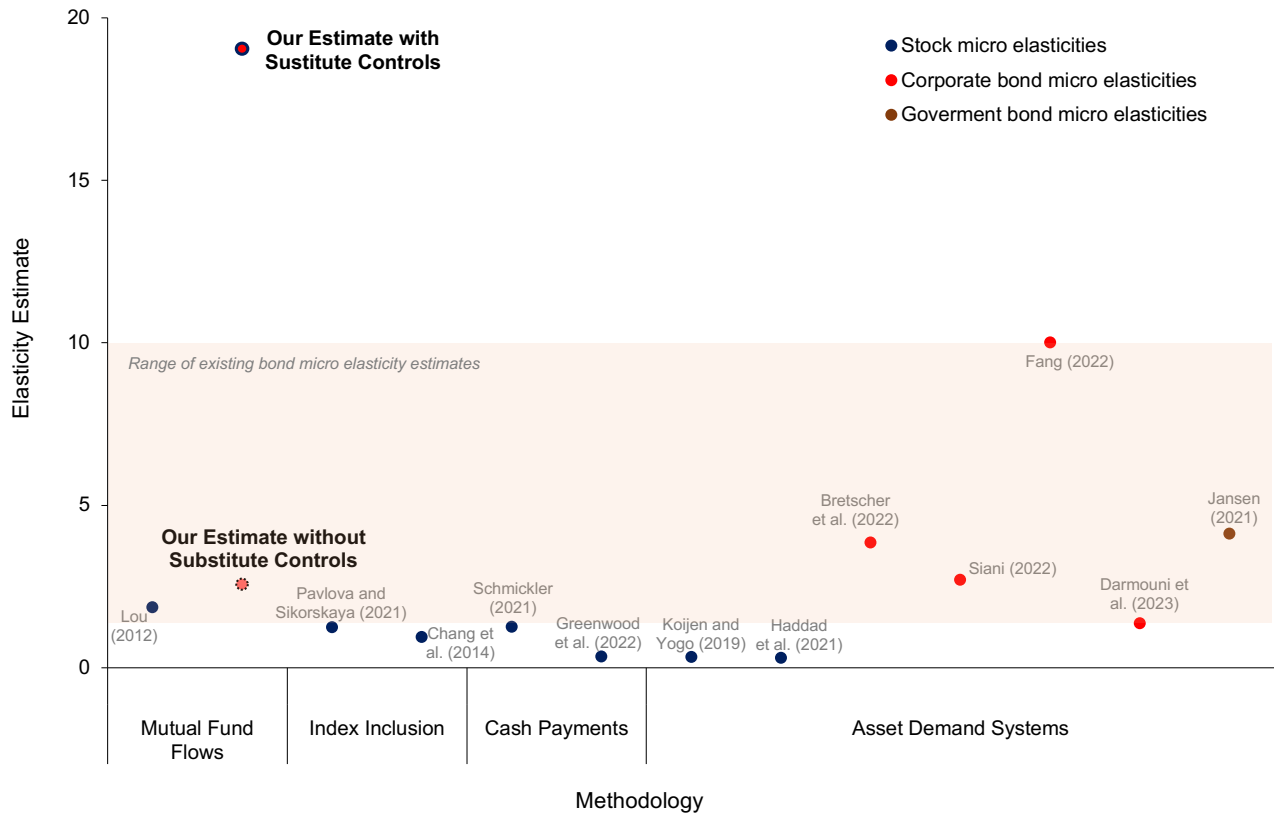
This graph presents a representative set of stock and bond micro elasticity estimates from the literature. The y-axis is the magnitude of the estimate—if authors provided a range for the estimate, the graph shows the midpoint of the range. The x-axis describes the methodologies used for estimating the elasticities.

Most directly, our estimates speak to the burgeoning literature estimating corporate bond micro elasticities. Several recent papers have adapted the asset demand system developed by Koijen

and Yogo (2019b) for the equity market to study corporate bond demand.[8] Bretscher et al. (2022) focus on the heterogeneous demand elasticities of different investors, and find the average corporate bond elasticity is around 3.8. Darmouni et al. (2023) adopt a two-layer framework to capture the interaction effect between fund flows and asset market inelasticity. Their demand elasticity estimate ranges from 0.8 to around 2, depending on the time period. Siani (2022) estimates the elasticity of the primary market demand and finds it is between 1.9 to 3.5. Fang (2022) uses a nested logit structure to allow for more flexible substitution within and across IG and HY bonds, and finds a higher elasticity of around 10. Compared to these studies, we find corporate bond markets are considerably more elastic, with a bond-level micro elasticity of around 20. Our method has several advantages over the demand system estimates. Firstly, our approach allows for greater flexibility in substitution patterns. At the investor level, standard logit demand structure imposes homogeneous cross-substitution across assets (in a nested logit system, homogeneity is within the nest). In the context of corporate bonds, this is a strong assumption. For example, it assumes AA- and CCC bonds are equally good substitutes for AA bonds. In line with findings in Fang (2022), allowing for greater flexibility in substitution results in more elastic estimates for the slope of individual bond's demand curve. Our approach is relatively model-agnostic compared to asset demand systems. Furthermore, our approach is not encumbered by not observing the holdings of investors such as households, hedge funds, private pension funds etc. Due to data-availability issues, corporate bond demand systems are typically only able to estimate the demands for insurance companies, bond mutual funds and public pension funds. Hence, the bottom-up aggregated market demand elasticity omits potentially important investors. Since our approach starts with the market clearing condition, it is not impacted by these data-availability concerns. Finally, our top-down approach makes it more suitable for studying bond specific heterogeneity in micro elasticities e.g., due to rating-based segmentation.

In terms of stock micro elasticities, an early strand of the literature shows that the inclusion of stocks in indices results in significant positive abnormal returns. For example, see Shleifer (1986); Harris and Gurel (1986); Beneish and Whaley (1996); Wurgler and Zhuravskaya (2002); Chen et al. (2004); Chang et al. (2015); Pavlova and Sikorskaya (2022) for stocks and Calomiris et al. (2022) for bonds. Generally, they find significant and permanent price impact; the estimated micro elasticities are around three orders of magnitude smaller than those implied by standard

---

[8] Jansen (2021) has also used demand systems to estimate the elasticity of European government bonds, and finds estimates of around 4.11 for maturity bucket portfolios.

models (Petajisto, 2009). Among these index-inclusion papers, Wurgler and Zhuravskaya (2002) have the closest connection with our paper. They show that demand is more elastic for stocks with closer substitutes. This finding is consistent with arbitrageurs being able to better hedge stock-specific idiosyncratic risk and trade more aggressively against demand shocks. We show this is also true for corporate bonds; those with closer substitutes exhibit more elastic demand curves. Overall, our demand shock and setting of corporate bonds have several advantages over Wurgler and Zhuravskaya (2002) for establishing the importance of close substitutes in flattening the demand curve. Firstly, stocks lack a prominent characteristic for identifying substitutes, whereas, in the case of corporate bonds, credit rating and maturity provide natural dimensions for identifying close substitutes. Secondly, index inclusion (or deletion) only impacts a handful of assets in any given year, whereas flow-induced trading shocks impact all assets to some extent. The broader exposure allows us to conduct our analysis at the bond portfolio level. At this higher level of aggregation, arbitrage risks are likely to play a more significant role in limiting arbitrage and hence is easier to detect.[9]

Another strand of the stock elasticity literature uses mechanical portfolio rebalancing as a source of demand shocks, showing that mechanical trading by mutual funds in response to flows has a significant price impact (Lou, 2012; Coval and Stafford, 2007; Edmans et al., 2012; Li, 2021). Other variants of mechanical trading-induced demand shocks include using dividend payment-induced trading (Hartzmark and Solomon, 2022), reinvestment of stimulus payments by US households (Greenwood et al., 2022), and investing refunds from unsuccessful bids in Chinese IPOs (Li et al., 2021). Overall these papers find large price impact as well. Li and Lin (2022) use equity net-flows and show stock prices respond more to demands at more aggregated levels. Our identification strategy directly builds upon the mutual flow-induced trading literature. We make three contributions compared to the standard methodology. Firstly, we take into account the close substitute portfolio's return in the estimation; this helps address the omitted variable bias generated by flow-induced demand shocks being correlated across assets. Secondly, we use the method proposed by Gabaix and Koijen (2021) to control for the omitted variable bias generated by forward-looking agents anticipating predictable fund flows. Thirdly, we formalize the identification by appealing to shares' exogeneity rather than shocks' exogeneity. This formalization of the identification strategy builds a direct parallel with exogenous share identification of Bartik instruments (Goldsmith-Pinkham et

---

[9]Inclusion in an index may bring better liquidity and better coverage by analysts, which may drive the observed price impact. Trade-induced shocks are less contaminated by such effects.

al., 2020; Chaudhry, 2022).

At more aggregate levels, Gabaix and Koijen (2021) finds that the macro elasticity of the overall equity market demand is around 0.2. Hartzmark and Solomon (2022) and Li et al. (2021) also find relatively inelastic macro elasticities. Overall, the equity market seems considerably more inelastic than individual stocks. Other papers have attempted to estimate equity portfolio level elasticities and found them to be somewhere in between equity micro and macro elasticities (Peng and Wang, 2019; Li, 2021). We show that more aggregate portfolios of individual bonds are more inelastic, and they are also less sensitive to changes in the prices of their close substitutes.

2 Estimation Methodology

We explain our estimation methodology in this section, and highlight how it allows for heterogeneous cross-substitution among bonds. The two salient characteristics of a corporate bond are its rating and time-to-maturity. These two characteristics play first order roles in investors' demand for bonds, and based on these characteristics they separate bonds into groups that are close and distant substitutes. We show that accounting for the difference in cross-elasticity for close versus distant substitutes is especially important for obtaining accurate estimates for multipliers in the corporate bond market.[10] We first present our demand framework and then discuss the identification strategy. Our specification is intentionally reduced-form to allow for greater flexibility. In Appendix D we show how our approach links to logit demand systems and nested logit demand systems as in Koijen and Yogo (2019b) and Koijen and Yogo (2019a).

2.1 Demand Framework

For a fund i in our sample (denote the set as MF), we assume its demand for the N available assets is

$$\underbrace{\mathbf{q}_{i,t}}_{N \times 1} = \underbrace{\Gamma}_{N \times N} \underbrace{\mathbf{p}_t}_{N \times 1} + \underbrace{\mathbf{u}_{i,t}}_{N \times 1} + \underbrace{\nu_{i,t}}_{N \times 1} \qquad \text{for } i \in MF \tag{1}$$

where $\mathbf{q}_{i,t}$ is a $N \times 1$ by vector, where element $q_{i,j,t}$ denotes the log quantity for bond j demanded by fund i. Similarly, $\mathbf{p}_t$ is the $N \times 1$ price vector, where element p_j denotes the log price

---

[10]Other asset classes, such as equity, may also have heterogeneous cross-elasticities among assets, but the asset characteristics that determine the degree of cross-substitution may be less salient. Indeed, in Appendix C, we conduct a similar analysis for stocks and find that accounting for heterogeneous cross-substitution matters less for the equity market.

of bond j. Γ is a $N \times N$ matrix that governs demand elasticities and substitution patterns among bonds. To simplify the notation, we assume Γ is homogeneous among investors, but this does not matter for our estimation.[11] $\frac{\partial q_j}{\partial p_j} = \Gamma_{j,j}$ is the slope of the demand curve for asset j. Furthermore, $\frac{\partial q_j}{\partial p_k} = \Gamma_{j,k}$ is the cross-elasticity of asset j to asset k's price. The observed flow-induced trading demand shock is captured by $\mathbf{u}_{i,t}$, where element $u_{i,j,t}$ denotes fund i's exogenous demand for bond j due to flows. Finally, the vector $\nu_{i,t}$ captures the unobserved demand shocks to fund i at time t.

Equation (1) can be viewed as a log-linearization of any generic demand function. For example, it nests the standard demand function derived from maximizing mean-variance utility. However, deriving the demand function from utility maximization would inevitably impose some restrictions on investors' utility. To preserve the maximum degree of flexibility, we model the demand in a reduced-form way.

For a fund that is not in our sample,[12] and we cannot observe its demand shocks, we denote their demand as

$$\mathbf{q}_{i,t} = \Gamma \mathbf{p}_t + \nu_{i,t} \qquad \text{for } i \notin MF \tag{2}$$

With the constant net supply of bonds, the market clearing condition gives:

$$\Delta \left(\sum_i \exp\left(\mathbf{q}_{i,t}\right) \right) = 0. \tag{3}$$

Log-linearize the market clearing condition around the last period values, we get for bond j

$$\Delta p_{j,t} = M u_{j,t} + \tilde{M} \Delta p_{j,t}^{sub} + \tilde{\nu}_{j,t}, \tag{4}$$

where

$$u_{j,t} \equiv \sum_i S_{i,j,t-1} u_{i,t} \equiv \sum_i \frac{\exp(q_{i,j,t-1})}{\sum_i \exp(q_{i,j,t-1})} u_{i,t}$$

$$M \equiv -\frac{1}{\Gamma_{j,j}}$$

$$\Delta p_{j,t}^{sub} \equiv \frac{\sum_{k \neq j} \Gamma_{j,k} \Delta p_{k,t}}{\sum_{k \neq j} \Gamma_{j,k}}$$

$$\tilde{M} \equiv -\frac{\sum_{k \neq j} \Gamma_{j,k}}{\Gamma_{j,j}}$$

---

[11] If Γ is heterogeneous among investors, then the final expression for market multipliers should be average of investor elasticities weighted by their AUM shares.

[12] We use the term "fund" to refer to all investors in the bond market.

and $\tilde{\nu}_{j,t} \equiv \nu_{j,t}/(-\Gamma_{j,j})$, where $\nu_{j,t}$ is the market-share weighted average of demand from the fund-level unobserved demand shocks $\nu_{i,t}$. With a slight abuse of notation, we keep using $u_{i,t}$ for shocks instead of $\Delta u_{i,t}$. If investors have downward-sloping demand with respect to the asset's own price, then the diagonal of Γ should be negative, i.e. $\Gamma_{j,j} < 0$, which implies $M > 0$. If assets are broadly speaking substitutes with each other, then the off-diagonal terms are positive, and $\tilde{M}$ should be positive. If assets are complements with each other, then the off-diagonal terms are negative, implying a negative $\tilde{M}$. We refer to M as the asset's own multiplier and $\tilde{M}$ as the substitute passthrough.

2.2 Identification Strategy

We use equation (4) to estimate the asset's multiplier M and substitute passthrough $\tilde{M}$. Under Assumption 1, the unobserved demand shock $\nu_{j,t}$ is orthogonal to the observed demand shocks $u_{k,t}$ for any k. In equilibrium, $\Delta p_{j,t}^{sub}$ is endogenous and is determined by the whole vector of demand shocks $\mathbf{u}_t$ and ν_t. We instrument Δp_j^{sub} with u_j^{sub}, defined as

$$u_{j,t}^{sub} = \frac{\sum_{k \neq j} \Gamma_{j,k} u_{k,t}}{\sum_{k \neq j} \Gamma_{j,k}} \tag{5}$$

In principle, we can use any combination of $u_{k,t}$'s to instrument for the substitution portfolio price $\Delta p_{j,t}^{sub}$. We find constructing the instrument using equation (5) to be intuitive and it also gives us high power in the first stage. Under Assumption 1, $u_{j,t}$ and $u_{j,t}^{sub}$ are both orthogonal to $\nu_{j,t}$ in the cross-section.

Assumption 1. *For any mutual fund i in each period t, the mutual fund's lagged portfolio share in bond j is orthogonal to its unobserved demand shocks,*

$$\nu_{j,t} \perp S_{i,j,t-1} \qquad \forall i, t$$

Appendix A shows that under Assumption 1, the multiplier and the passthrough coefficient are identified.[13] Our identification strategy uses the same insight as in Goldsmith-Pinkham et al. (2020) that Bartik instruments can be viewed as exogenous share instruments. As a result,

---

[13]One potential concern for Assumption 1 is that if the unobserved shock $\nu_{j,t}$ loads on a factor, such as $\nu_{j,t} = \tilde{\nu}_{j,t} + \beta_j \eta_t$, and there exist factor-focused funds whose portfolio shares are functions of bonds' factor loading, e.g., $Q_{j,t-1} = \psi_{t-1} \beta_j$, then the orthogonality condition in Assumption 1 will be violated. We address this issue in Appendix A.3. In essence, we show that as long as the fund-specific demand shock $u_{i,t}$ is orthogonal to factors, our identification is still effective. Our results are robust to different factor specifications in flows (see Appendix B).

the strategy is similar to an exposure research design, where bonds have exogenous exposures to common shocks, and the degree of exposure depends on the portfolio shares of mutual funds.

To make the pooled exposure design intuition more concrete, suppose households receive a preference shock that increases their demand for bonds (this shock may be correlated with the fundamental characteristics of the bond market). As a result of the household preference shock, bond mutual funds will receive inflows. Consider a mutual fund that holds two bonds, and before the shock held 90% of its holdings in bond A and the remaining 10% in bond B. Suppose this mutual fund receives \$10 million in inflows and mechanically invests these inflows according to its previous periods portfolio weights. In this case, bond A would experience a demand increase of \$9 million and bond B will experience an increase of \$1 million. If the previous periods portfolio weights of the mutual fund are orthogonal to the household preference shock, then the relative increase in the demand for bond A vs bond B will be as good as randomly assigned.[14] Hence, given exogeneity of a funds lagged portfolio weights and unobserved shocks, flow-induced trading by a single mutual fund amounts to an exogenous exposure research design. Our identification strategy pools many such instances of flow-induced trading, and hence it can be viewed as a pooled exposure design.

One potential threat to this assumption is that some mutual fund managers may have superior information about bond idiosyncratic returns in the future, and allocate their portfolio in advance to front-run the market. This issue is addressed in Appendix B.4 by using longer lags for shares in aggregating flow shocks.

In the special case where substitutions are homogeneous, i.e. $\Gamma_{j,k}$ is constant for all j, $\Delta p_{j,t}^{sub}$ is constant across bond j and will be absorbed by time fixed effects.[15] Existing literature that uses flow-induced trading to identify price impact with time fixed effects essentially assumes homogeneous cross-substitution (Lou, 2012). Similarly, logit-demand systems such as those in Koijen and Yogo (2019b) and Bretscher et al. (2022) are special cases of homogeneous substitution where $\Gamma_{j,k}$ is proportional to w_k (the portfolio weight of asset k).[16]

---

[14]If the fund has mandates restricting what types of bonds the fund can hold, then the previous periods portfolio weights, the part that is determined by mandates, are naturally orthogonal to demand shocks.

[15]Technically, the substitute is the market excluding the test assets. When each test asset is small relative to the whole market as in our baseline, the difference is negligible.

[16]To see this, consider the logit-demand specification $ln(\frac{w_{i,j,t}}{w_{i,0,t}}) = \gamma_i p_j + \beta^\top X_j + \epsilon_{i,j,t}$, where X_j is the relevant characteristic vector that affects demand. When asset k's price increase, $ln(\frac{w_{i,j,t}}{w_{i,0,t}})$ does not change. In other words $\frac{\partial ln(w_{i,j,t})}{\partial p_k}$ is constant for all $j \neq k$. For complete derivation, see Appendix D.

However, in the corporate bond market, the homogeneous cross-substitution assumption is likely to be violated. If the price of a BBB+ bond increases, investors will probably substitute much more into another BBB+ bond than a AAA bond (all else equal). In addition, there will probably be very little substitution into the high yield bond class. Hence, in the case of corporate bonds it is important to directly take into account the heterogeneous cross-substitution.

2.3 Demand Structure with Heterogeneous Cross-elasticity

At a high level, we want to account for the fact that cross-substitution among bonds with similar ratings and time-to-maturity may be stronger than bonds with different characteristics. Ideally, we would control for each asset i's return ($i \neq j$) separately on the right hand side of equation (4), which will allow each asset i to have a different cross-elasticity with the testing asset j. However, we do not have enough data and power to identify all the coefficients. Hence we start with allowing for two different cross-elasticities, one for bonds within the same detailed rating (close substitutes) and another for bonds outside this category (distant substitutes). Furthermore, it is intuitive to scale the cross-elasticity by the market share of substitutes, so the substitution effect is in proportion to their relative size. This implies that for each asset j, the relevant substitute portfolios are the market portfolio and the value-weighted portfolio of all bonds in the same rating category. We call the latter portfolio its close substitute. For example, for a 10-year BBB- bond, its close-substitute portfolio includes all the other BBB- bonds. In most of our analysis, we control for the time fixed effect in place of the market portfolio return.[17]

This particular demand structure maps into the following regression specification,

$$\Delta p_{j,t} = M u_{j,t} + \tilde{M} \Delta p_{j,t}^g + \text{Time fixed effect} + \tilde{\nu}_{j,t} \tag{6}$$

where $\Delta p_{j,t}^g$ is the portfolio return of assets in the same detailed rating category as asset j (excluding asset j), weighted by each asset's lagged market-value. The multiplier M measures the price response to a one-percentage point increase in demand. The passthrough coefficient $\tilde{M}$ captures the additional substitutability of within-group bonds (relative to the testing asset's own elasticity). Appendix D provides details on the exact form of Γ that leads to the specification in equation (6).

For the purpose of estimating the multiplier M alone, we can also simply absorb $\Delta p_{j,t}^g$ with group-time fixed effects, as in equation (7):

$$\Delta p_{j,t} = M u_{j,t} + \text{Fixed effects}_{g,t} + \tilde{\nu}_{j,t}. \tag{7}$$

---

[17]The time fixed effect can also partially account for substitution between corporate bond and other asset classes.

Although this group-time fixed effect specification is more straightforward, the baseline specification in equation (4) allows us to estimate the substitute passthrough coefficient $\tilde{M}$, which is informative about the spillover effect among close substitutes. In our analysis we estimate both specifications 7 and (4). We find that the choice of specification has little impact on the estimate of M, suggesting the IV estimation strategy for equation (4) successfully addresses the endogeneity issues.

Failing to control for close substitutes will lead to over-estimation of the multiplier and hence under-estimation of the elasticity. The following example showcases the source of the bias. Suppose bond-level shocks are generated as $u_{k,t} = u_{g,t} + \tilde{u}_{k,t}$, where the bond-specific shock $\tilde{u}_{k,t}$ is orthogonal to the group level shock $u_{g,t}$. Also for simplicity, we assume each bond is minuscule relative to the group so $Cov(p_k^g, \tilde{u}_k) = 0$. If the econometrician fails to control for $\Delta p_{j,t}^g$ when estimating the multiplier, the omitted variable bias can be shown as:

$$\hat{M} - M = \tilde{M} \frac{Cov(p_k^g, u_g)/Var(u_g)}{1 + \frac{Var(\tilde{u}_k)}{Var(u_g)}} \equiv \tilde{M} \frac{M_g}{1 + \frac{Var(\tilde{u}_k)}{Var(u_g)}} > 0. \tag{8}$$

The bias has three components, the passthrough coefficient $\tilde{M}$, the group-level multiplier M_g, and the variance ratio of the bond-specific shocks to group shocks. Therefore, failure to control for close substitutes will contaminate the bond-level multiplier with the group-level multiplier. As we show below, multipliers increase as we move to more aggregate portfolio levels, so the bias is positive and can be potentially large.[18]

This issue is not specific only to our empirical specification, but it arises whenever shocks across bonds are correlated—a typical scenario for most other demand shocks commonly used in the literature, such as institution mandates. When the substitution is homogeneous as commonly assumed, $\tilde{M}$ is close to zero so the bias is small. However, as we show in our estimation results, at the bond level the substitute passthrough is close to 1, so ignoring close substitutes dramatically over-estimates the multiplier.

For most of the heterogeneity analysis, we conduct the analysis at the portfolio level. In our baseline portfolio specification, we treat each asset as a market-value weighted portfolio formed by bonds with the same detailed rating category and quarter-to-maturity. Like the individual bond analysis, we assume there are two cross-elasticities, one for portfolios with the same detailed ratings (close substitutes) and another for portfolios outside this category (distant substitutes).

---

[18]We are implicitly assuming that the cross-elasticity between corporate bonds and other asset classes (such as Treasuries, mortgage-backed securities) is small. Hence the bias due to not including those asset returns explicitly is negligible.

Finally, with enough data, one can easily extend our analysis to allow for multiple layers of substitution: same detailed rating category, the same investment rating (investment-grade vs high-yield), and all the other cases. We verify that the estimation results are quantitatively very similar to the case with two different levels of cross-elasticities. So we will focus on only two levels of heterogeneity for most parts of the paper.

3 Constructing Demand Shocks

In this section, we describe the data sources, the construction of our demand shocks, and present summary statistics.

Mutual fund data: We obtain detailed data on open-end mutual funds in the U.S. from Morningstar Inc. Morningstar is one of the largest providers of investment research to the asset management industry. Detailed holdings and fund flows are collected from surveys of mutual fund managers, and cross-validated by Morningstar against publicly available sources such as regulatory filings to ensure their accuracy. Most funds report at least once per quarter. In order to keep as many funds as possible in our sample, we conduct our analysis at the quarterly frequency. The Morningstar mutual fund coverage is quite extensive—the data sets total assets under management (AUM) lines up closely with the Flow of Funds open-end mutual fund sector AUM (see Figure 10 in the Appendix).

Since our identification relies on flow induced trading in the US corporate bond market, we apply some additional filters to narrow in on domestic bond mutual funds. Firstly, we restrict our sample to funds that report their portfolio value in US dollars and also to securities denominated in US dollars. Secondly, we focus on funds that have at least $10 million dollars of bonds under management, and bonds make up 50% to 120% percent of their portfolio's asset under management (AUM). The lower bound is to filter out non-bond funds, and the upper bound is to filter out potential misreporting. Taken together this database gives us coverage of 1,151 unique funds from 2002Q1 to 2021Q3, totaling around 5.6 million bond-fund-quarter observations.

After applying our data filters, we then define mutual fund i's dollar flows F_{it} scaled by lagged $AUM_{i,t-1}$ as,

$$f_{it} := \frac{F_{it}}{AUM_{i,t-1}}.$$

This is a key building block for constructing our demand shocks. We denote the set of investors

whom we observe flow data as MF.

Bond return data: We use corporate bond returns data from WRDS Bond Returns file, which is constructed using transaction level data from FINRA's TRACE (Trade Reporting and Compliance Engine) database, and Mergent FISD data for bond issue and issuer characteristics. Specifically, our quarterly return measure is the cumulative end-of-month returns of the months that fall in a given quarter. We further merge this dataset with Morningstar's CUSIP level bond characteristics.

Additionally, we apply filters to ensure we are capturing corporate bonds with reliable return data. Firstly, we restrict our sample to securities that WRDS classifies as corporate bonds. Secondly, we restrict our attention to corporate bonds that have at least a CCC- or higher rating by S&P. Overall, we have $377,753$ bond-quarter observations.

3.1   Constructing Flow Shocks

Our shock measure builds on the flow-induced trading (FIT) measure proposed by Lou (2012). The measure exploits the fact that mutual funds tend to mechanically invest inflows/outflows according to their existing portfolio weights (Coval and Stafford, 2007). Ultimately, flow-induced demand shocks are essentially a Bartik instrument, where the key identifying assumption is the exogeneity of lagged mutual funds holding shares (Goldsmith-Pinkham et al., 2020; Chaudhry, 2022). With that said, there are several omitted variable bias concerns when directly using FIT as a demand shock measure, hence we need to apply several adjustments to get from FIT to our demand shock (Gabaix and Koijen, 2022). Below we walk through how we construct our bond-specific demand shock measure.

Step 1: Measuring flow-induced trading by mutual funds. Assuming fund i mechanically reinvests dollar flow F_{it} according to their existing portfolio weight $\theta_{ij,t-1}$. This implies that the dollar amount of flow-induced trading of bond j will be $\theta_{ij,t-1}F_{it}$. The percentage change in asset j's holdings due to flows is

$$f_{it} = \frac{\theta_{ij,t-1}F_{it}}{\theta_{ij,t-1}AUM_{i,t-1}}$$

Implicitly in this step we have assumed that the flows fully pass through to the demand shocks to each bond one-to-one. If the passthrough is less than one-to-one, then we risk underestimating the multiplier, since we overestimate the quantity of the shock. We kick-the-tyres on this assumption in appendix section B.1 by estimating the passthrough of flows to actual tradings. Notice that this passthrough is a lower-bound of the passthrough to demand, as the acutal tradings also reflect other

equilibrium objects such as prices. Overall, we find that the estimated passthroughs are sizeable (around 0.5 to 0.6). Therefore, we can use our passthrough estimates to estimate upper bounds on the multipliers, and find that these are still considerably smaller than micro multipliers typically found for stocks and bonds.[19]

Step 2: Adjusting for predictability in flows. Due to return chasing by households, mutual fund flows and flow induced trading tend to be predictable (Lou, 2012). As a result, forward-looking investors will trade today in anticipation of future flow. Hence the relevant demand shock is not just the flow induced trading today, but rather, the total amount of flow induced trading predicted by innovations in flows today (Gabaix and Koijen, 2021). To determine flow innovations, we estimate an AR(3) model with a time trend for each fund i,[20]

$$f_{it} = \rho_{i,0} + \sum_{k=1}^{3} \rho_{i,k} f_{i,t-k} + \delta_i t + \varepsilon_{it}$$

Assuming relatively little discounting by forward looking agents at the quarterly level, the relevant quantity of the demand shock is the total cumulative trading predicted by the innovation,

$$K_i \varepsilon_{it}$$

where $K_i = \frac{1}{1 - \sum_{k=1}^{3} \rho_{i,k}}$.

Step 3: Controlling for unobserved common factors. We estimate a model of the form

$$\varepsilon_{i,t} = \alpha_i + \delta_t + u_{i,t},$$

where δ_t is the time-fixed effect controlling for aggregates and $u_{i,t}$ is the fund-specific idiosyncratic flow innovations. Hence, our measure for fund-specific demand shocks is $K_i u_{it}$. We can also include additional common factors in the regression. However, we find that mutual fund flows have a weak factor structure after controlling for time-fixed effects. More generally, controlling for additional factors makes little difference to the estimates (see Appendix section B). For these reasons, in this paper's main body, we present findings for demand shocks constructed by only controlling for time-fixed effects.

Step 4: Aggregate fund-level shocks to bond-level shocks. To obtain bond-level demand shocks, we aggregate the fund-level shocks $K_i u_{it}$ to the bond level using funds' lagged holding

---

[19]At the bond level, the upper bound multiplier estimate is 0.08 (elasticity of 12.5), and for our baseline level of aggregation, it is around 0.33 (elasticity of 3).

[20]Our results are robust to changing the specification of the AR process (see appendix section B.5).

share weights. With a slight abuse of notation, we denote bond j's demand shock in quarter t as u_{jt}. Specifically, u_{jt} is defined as

$$u_{jt} = \sum_{i \in MF} S_{ij,t-1} K_i u_{it}$$

where $S_{ij,t-1}$ is the market share of bond j that is held by investor i in period $t-1$.

Figure 3 plots the 1st to 99th percentile range of the shocks for each quarter, along with its median. The magnitudes of shocks are generally quite large ranging from around 1%-3% of the total outstanding amount. The range is relatively stable in magnitude apart from two crisis episodes in 2008 and 2020. For robustness, we explore including and dropping the crisis periods from our estimation sample, and find that are results are robust to this decision.

Figure 3: Demand shocks ($N = 339,847$)
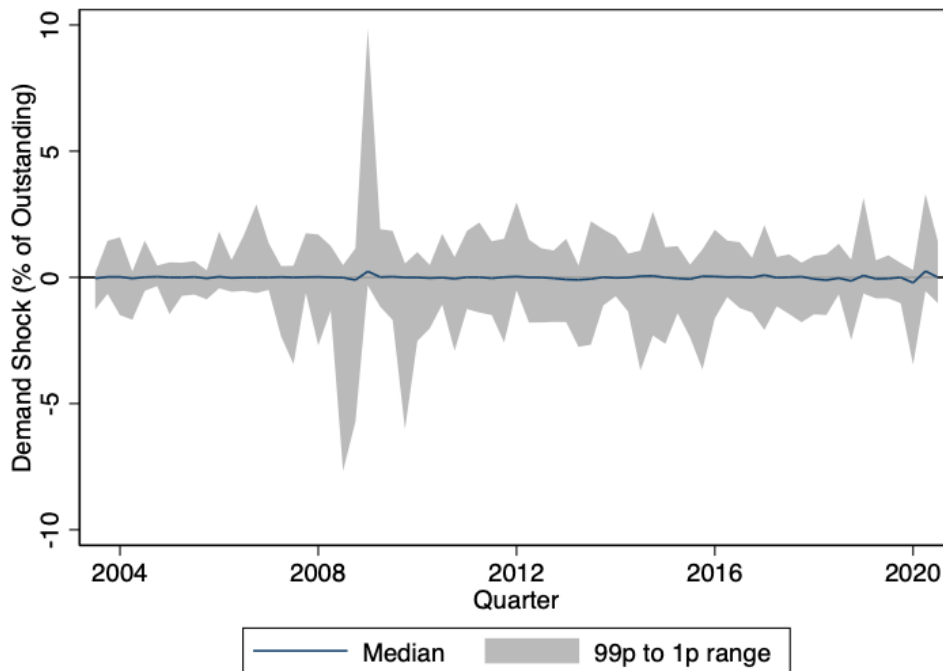
This graph plots the median value of our demand shock measure in a given quarter, and the shaded grey area denotes the range of demand shock values that fall between the 1st and 99th percentile of shocks in any given quarter. The shocks are represented as a percentage of the total outstanding amount.

4 Estimation Results

We present our main findings in this section. We first show that the multiplier at the CUSIP level drops from 0.4 to 0.05 once we account for the heterogeneity in cross-substitution. When we form bonds into our baseline portfolios, the portfolio multiplier becomes significantly larger, indicating demand is more inelastic at the portfolio level. Furthermore, the multiplier monotonically increases as we form increasingly more aggregate portfolios, while the substitute passthrough monotonically decreases. Furthermore, we investigate the dynamics of the price impact at the portfolio level and find the effect almost fully reverts after 2 quarters, suggesting demand is more elastic in the long run. Finally, we show that the multiplier is larger for bonds with higher default risks and longer time to maturity.

4.1 Baseline Estimates

CUSIP Level Table 1 presents our baseline estimates treating each CUSIP as one asset. The first column corresponds to assuming homogeneous cross-substitution. Under this assumption, every asset has the same substitute portfolio, which is simply the market portfolio. This is equivalent to including time fixed effects in the regression. Under this assumption, we get a significant CUSIP multiplier of 0.39, indicating that 1% increase in demand leads to 0.39% increase in the bond's price. This implies that the average elasticity for a bond is around 2.6. The magnitude of the average elasticity is in line with the estimate in the literature using a standard logit demand system (Bretscher et al., 2022; Siani, 2022).

Once we relax the homogeneous cross-substitution assumption and allow certain groups of bonds to be closer substitutes compared to others, the CUSIP multiplier drops to nearly zero, implying very elastic demand at the individual bond level. Column (2)-(4) include the returns of a substitute portfolio, which is defined as the market-value weighted portfolio formed by all bonds within the same detailed rating category (excluding the asset itself). We instrument the substitute portfolio's return with the market-value weighted demand shocks for all bonds in the substitute portfolio, $u_{j,t}^{sub}$. We also include a time fixed effect in all the regressions, which means we are controlling for the market portfolio return (the distant substitute) as well. Effectively, we are allowing the cross-elasticity to be different for bonds within the same rating category relative to all the other bonds.

Our result implies the demand elasticity at the CUSIP level is around 20,[21] which is much larger than the current estimates in the literature. Among the elasticity estimates in the literature, Fang (2022)'s estimate of 10 is the closest ours. Compared to other demand systems used for estimating corporate bond elasticities, Fang (2022) uses a somewhat more flexible nested logit structure, allowing for differential substitution within and across IG and HY categories. Overall, our paper emphasizes the importance of explicitly modeling the rich heterogeneous substitution patterns in the corporate bond market.

In addition to a near zero CUSIP multiplier, the substitute passthrough ($\tilde{M}$) is highly significant and is quantitatively close to 1. This implies that within the detailed rating groups, bonds are effectively perfect substitutes with each other. Finally, our point estimates are stable whether or not we exclude crisis periods, or include additional fixed effects.

Our results imply that ignoring the substitute portfolio's return in the regression leads to a significant over-estimation of the multiplier. Intuitively, when we observe a demand shock to bond j, it is likely that other similar bonds are also experiencing demand shocks. The price of bond j is higher not only because of its own demand shock, but also because all of its substitutes now have higher prices. Hence ignoring the substitute portfolio will lead to over-estimation of the price impact from its own demand shock. Furthermore, in the case when certain bonds are closer substitutes than others, we need to overweight the close substitutes in constructing the substitute portfolio, or alternatively, allow the close substitutes to have a different coefficient than the other bonds.

In Appendix C, we estimate the equity market multiplier at the stock level, following identical procedures. We define close substitutes as stocks with similar factor loading or stocks in the same industry group. We find that allowing for heterogeneous cross-substitution reduces the multiplier estimated, as in the corporate bond case. However, quantitatively, the difference is not as big as that in the bond market, suggesting that heterogeneous substitution patterns are a particularly special feature of the bond market. Perhaps due to the fact that rating and maturity play a significant role in investor's demand, accounting for heterogeneous cross-substitution is much more important in the bond market than in the equity market.

Portfolio Level Next we form portfolios of bonds based on detailed rating and quarter-to-maturity, and estimate the multiplier at the portfolio level. In addition to the market portfolio

---

[21] Under the unit-elasticity null, we estimate the 95% confidence interval for the elasticity is $(19.9, 20.1)$.

(which is taken care of by the time effect), we include the return of a substitute portfolio that is formed by all bonds in the same detailed rating category as before.

The results are presented in Table 2. Similar to the CUSIP case, we find that only including the market portfolio as the substitute significantly overestimates the portfolio multiplier. Once we include additional substitute portfolios, the multiplier estimated is considerably lower.

A priori, we expect the portfolio-level multiplier to be larger than the CUSIP multiplier. Intuitively, the substitutability among portfolios with different maturities should be smaller compared to the substitutability among CUSIPs with the same maturity. Indeed, Table 2 shows the multiplier on portfolios is significantly larger than the multiplier for individual CUSIPs—1 percent increase in demand for a given portfolio raises the price for that portfolio by 35 basis points. This estimate is stable across various specifications. Additionally, the substitute passthrough at the portfolio level is also smaller than that at the CUSIP level, consistent with the intuition that more aggregated portfolios are less substitutable.

To verify whether our demand specification accounts for important cross-substitution heterogeneity sufficiently, we further allow heterogeneity in cross-elasticities for bonds in the same investment-grade (IG) or high-yield (HY) group versus bonds that are not. We implement such specification by including an additional substitution portfolio formed by all the IG or HY bonds (excluding the own asset and close substitute's returns). The result is shown in Table 3 column (2). The own multiplier does not change quantitatively. In addition, the substitution passthrough is much larger for bonds within the same rating category compared with bonds not in the same rating category but are in the same IG/HY group. Column (3) and (4) of Table 3 further confirm that our result is not sensitive to the choice of the specific substitute portfolio. Column (3) uses all bonds in the same coarse rating group to form the substitute portfolio, i.e., using a broader definition of the set of close substitutes.[22] We can also group bonds into two maturity groups based on the partition: $\{[0, 10), [10, \infty)\}$. Column (4)'s substitute portfolio includes all bonds in the same maturity group and detailed rating category, i.e., a more narrow definition of the set of close substitutes. Across all the different alternative specifications, the estimated own multiplier is similar to that in the baseline case.

---

[22]For example, for a 11-year BBB- bonds, the substitute portfolio includes all the BBB rated bonds.

Table 1: CUSIP baseline

	Homo. OLS	OLS		First-stage	2SLS
	(1)	(2)	(3)	(4)	(5)
Shock	0.39***	0.02	0.05		0.05
	(0.06)	(0.05)	(0.05)		(0.05)
Substitute return					1.08***
					(0.05)
Group Shock				2.61***	
				(0.30)	
Quarter FE	Yes	Yes	Yes	Yes	Yes
Group x Quarter FE	No	Yes	Yes	No	No
Drop Crisis	No	No	Yes	Yes	Yes
N	333,537	333,537	314,534	314,534	314,534
R^2	0.21	0.39	0.36	0.62	
First-stage F-statistic					77.30

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table presents the CUSIP level multiplier estimates. Column 1 is the OLS estimates from regressing bond returns on the demand shock, controlling for time fixed effects—this specification corresponds to a model in which we assume homogeneous cross-elasticity with all other bonds. Columns 2 and 3 directly control for close-substitute prices using detailed rating x time fixed effects. Column 4 and 5 relate to the IV specification in which we regress bond returns on the demand shock and substitute returns, while controlling for time fixed effects (and additional controls depending on the specification). We instrument for substitute returns using demand shocks to substitute assets. The IV specification corresponds to the model that allows for heterogeneous cross-elasticities between close and distant substitutes. The parenthesis contain standard errors clustered at the substitute group x time level.

Table 2: Quarter to Maturity x Detailed Rating baseline

	Homo. OLS	OLS		First-stage	2SLS
	(1)	(2)	(3)	(4)	(5)
Shock	0.95***	0.32**	0.33***		0.35***
	(0.12)	(0.10)	(0.10)		(0.10)
Substitute return					0.90***
					(0.05)
Group Shock				2.91***	
				(0.34)	
Quarter FE	Yes	Yes	Yes	Yes	Yes
Group x Quarter FE	No	Yes	Yes	No	No
Drop Crisis	No	No	Yes	Yes	Yes
N	81,866	81,866	77,387	76,348	76,348
R^2	0.23	0.47	0.41	0.47	0.21
First-stage F-statistic					69.85

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table presents the quarter to maturity $\times$ detailed rating level multiplier estimates. Column 1 is the OLS estimates from regressing bond returns on the demand shock, controlling for time fixed effects—this specification corresponds to a model in which we assume homogeneous cross-elasticity with all other bonds. Columns 2 and 3 directly control for close-substitute prices using detailed rating x time fixed effects. Column 4 and 5 relate to the IV specification in which we regress bond returns on the demand shock and substitute returns, while controlling for time fixed effects (and additional controls depending on the specification). We instrument for substitute returns using demand shocks to substitute assets. The IV specification correspond to the model that allows for heterogeneous cross-elasticities between close and distant substitutes. The parenthesis contain standard errors clustered at the substitute group x time level.

Table 3: Alternative Substitutes

	(1)	(2)	(3)	(4)
Shock	0.35***	0.35***	0.31**	0.33**
	(0.10)	(0.10)	(0.11)	(0.11)
Detailed rating substitute return	0.90***	0.77***		
	(0.05)	(0.17)		
IG substitute return		0.18		
		(0.22)		
Coarse rating substitute return			0.79***	
			(0.08)	
Det rating x ST/LT substitute return				0.99***
				(0.08)
Quarter FE	Yes	Yes	Yes	Yes
Drop Crisis	Yes	Yes	Yes	Yes
N	76,348	76,348	76,296	76,348
First-stage F-statistic	69.85	3.37	18.19	107.94

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table shows the quarter to maturity $\times$ detailed rating level multiplier estimates are robust to the exact definition of the close substitute. Column 1 is our baseline specification in which the close substitute is defined as all other bonds in the same detailed rating category. In column 2 the substitute includes the detailed rating substitute, and all bonds with the same investment rating as an additional substitute. Column 3 includes all other bonds in the same coarse rating category as substitutes. Column 4 includes all bonds in the detailed category and similar maturity as substitutes. The parenthesis contain standard errors clustered at the substitute group x time level.

4.2 Aggregate Portfolios

Motivated by the difference in CUSIP and portfolio multiplier, we repeat our analysis for portfolios with various aggregation levels. We find that the more aggregated the portfolio is, the larger the portfolio multiplier and the smaller the substitute passthrough.

Specifically, we estimate the multiplier and substitute passthrough for all the assets/portfolios in Table 4. The second column specifies at which level we form the portfolios, and the third column specifies the substitute portfolio included in the regression in addition to the market portfolio. As we move to more aggregated portfolios, they become less substitutable with each other. As a result, demand shocks have a larger price impact and the relationship between the portfolio's price and its substitute's price becomes weaker.

Table 4: Aggregate Portfolios

	Asset	Substitute Portfolio
CUSIP	Individual bonds	Other bonds in the same detailed rating category
Rating × Quarter to Maturity	Portfolios formed by detailed rating and quarter-to-maturity	Other bonds in the same detailed rating category
Rating × ST/MT/LT Buckets	Portfolios formed by coarse rating and three maturity groups $(\{[0,4),[4,10),[10,\infty)\})$	Other bonds in the same coarse rating category
Rating × ST/LT Buckets	Portfolios formed by coarse rating and two maturity groups $(\{[0,10),[10,\infty)\})$	Other bonds in the same coarse rating category
Rating	Portfolios formed by coarse rating categories	Other bonds in the same investment grade category

This table presents all levels of portfolios for which we estimate multipliers and substitute passthroughs. The left column specifies at which level the portfolio is formed. All bonds are weighted by their market-value inside the portfolio. The right column specifies the additional substitute portfolio included in the regression in addition to the time fixed effect.

Figure 1 visualizes our findings, and Table 5 presents the exact estimates. To ensure our results are not driven by tail aggregate shocks, we exclude the crisis periods for all regressions onward. From Figure 1, we see that as we move from CUSIP-level estimate to the most aggregate portfolios formed by ratings, the asset's own multiplier increases from 0.05 to 3.5, implying the demand elasticity drops from 20 to less than 0.3. Intuitively, it is easy to find substitutes for an individual bond, whereas it is much more difficult to find substitutes for the entire portfolio of BBB bonds. Hence demand is much more elastic for individual CUSIPs compared with more aggregated portfolios. Furthermore, the substitute passthrough decreases from 1.07 to essentially 0. At the rating-portfolio level, assets are much weaker substitutes with each other compared to that at the CUSIP level.

Our aggregation results highlight the connection between micro and macro multipliers. Demand elasticities naturally depend on what we define as an asset. Micro multipliers typically refer to the multiplier on an individual security, whereas macro multipliers refer to the multiplier when we treat a whole asset class as one.[23] They are at the two ends of a spectrum, and our results show that depending on the portfolio definition, the multiplier estimated can lie anywhere in between. In fact, demand inelasticity at the portfolio level is what allows us to estimate the multipliers at more micro levels — the significance in the first stage of all the IV regressions relies on portfolio prices responding to demand shocks.

---

[23]Our identification strategy does not allow us to uncover the macro elasticity for the corporate bond market as an asset class.

Table 5: Aggregation

	CUSIP	Rat x Q to Mat	Rat x 3 Mat	Rat x 2 Mat	Rat
	(1)	(2)	(3)	(4)	(5)
Shock	0.05	0.35***	1.23*	1.59**	3.51***
	(0.05)	(0.10)	(0.48)	(0.58)	(0.87)
Substitute return	1.07***	0.90***	0.73***	0.48**	-0.02
	(0.06)	(0.05)	(0.13)	(0.15)	(0.31)
Quarter FE	Yes	Yes	Yes	Yes	Yes
Drop Crisis	Yes	Yes	Yes	Yes	Yes
N	314,534	76,348	1,407	938	469
First-stage F-statistic	72.49	69.85	33.83	16.15	17.78

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table shows the multiplier and substitute passthrough estimates for various levels of portfolio aggregation—see Table 4 for descriptions of the aggregation. All the specifications are estimated using IV, specifically we instrument the substitute's return using demand shocks to the assets in the substitute portfolio. The parenthesis contains standard errors clustered at the substitute group x time level.

4.3 Dynamic Effects

We investigate the dynamic effects of price response to demand shocks in this section. We focus on the baseline portfolio specification in Section 4.1, i.e., portfolios are formed by bonds with the same detailed rating and quarter-to-maturity. To see the dynamic effects, we regress portfolios' cumulative returns on lagged demand shocks. Specifically, we run the following regression

$$\Delta p_{j,t:t+h} = M_h u_{j,t} + \text{Fixed effects}_{g,t+h} + \varepsilon_{j,t+h} \quad \text{for} \quad h = 0, 1, \dots \tag{9}$$

Figure 4 plots the response of cumulative returns to demand shocks. Consistent with our previous result, there is a significant on-impact price increase upon positive demand shocks. The effect almost fully reverts in the following quarter, and the price impact is statistically indistinguishable from zero. In other words, the cumulative effect drops to 0 after $T + 1$. The reversal in return after 6 months is consistent with the hypothesis that long-term demand is more elastic than short-term

Figure 4: Price Impact Dynamics
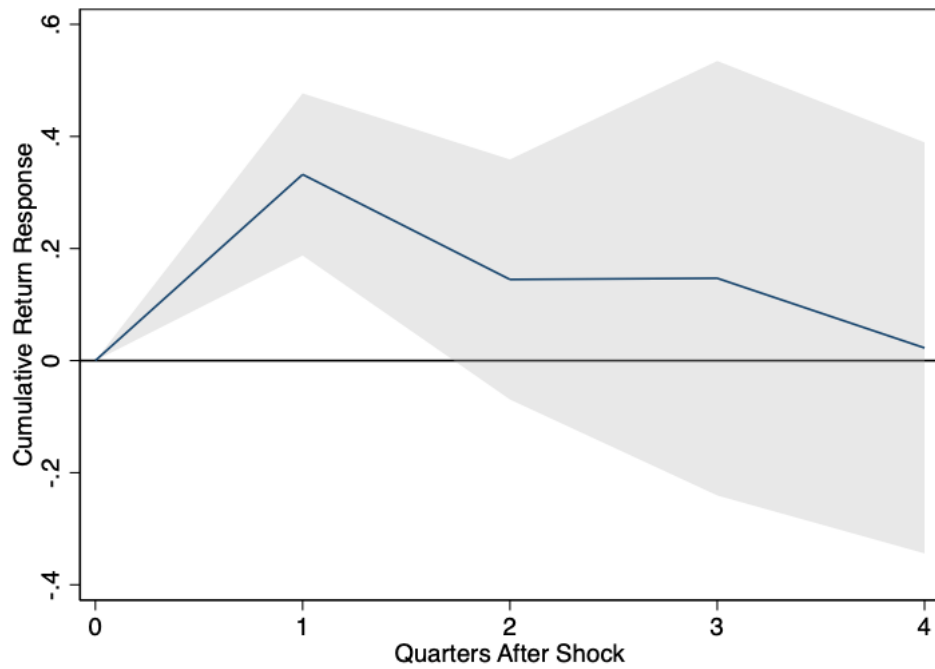
The graph plots the dynamic price impact estimates of the demand shock. The estimates are for the detailed rating × quarter to maturity level (our baseline level of aggregation). Specifically we regress various leads of the return of the portfolio in excess of its close substitute return on the demand shock. The shaded region denotes the 95% confidence interval.

demand.

Table 6: Dynamic Effects

	T	T+1	T+2	T+3
	(1)	(2)	(3)	(4)
Shock	0.33***	0.14	0.15	0.02
	(0.07)	(0.11)	(0.20)	(0.19)
Group x Quarter FE	Yes	Yes	Yes	Yes
N	77,387	55,466	45,770	39,584
R^2	0.41	0.57	0.60	0.61

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table presents the dynamic cumulative price impact estimates of the demand shock. The estimates are for the detailed rating × quarter to maturity level (our baseline level of aggregation). Specifically we regress the cumulative return on the demand shock while controlling for substitute group x time fixed effects. The parenthesis contains standard errors clustered at the substitute group x time level.

4.4 Rating and Maturity Heterogeneity

Finally, we look at how multipliers vary across different types of bonds. We focus on the heterogeneity in multipliers for bond portfolios with different maturities and ratings. As before, we use our baseline portfolio specification.[24] As shown in Section 4.1, the exact substitute definition does not matter for the estimation results.

We first estimate portfolio multipliers for IG versus HY bonds. We pool all the IG portfolios together and find the portfolio multiplier is close to 0. For the pooled HY portfolios, the estimated multiplier is around 0.6, much higher than that for IG portfolios. The overall estimate of 0.35 masks significant heterogeneity among portfolios with different ratings. We also estimate the portfolio multipliers for each rating category. The results are shown as scattered dots in Figure 5. On average, riskier portfolios have higher portfolio multipliers. HY bonds tend to be riskier with larger idiosyncratic risks. As a result, it may be more difficult to substitute to other assets, which leads to higher portfolio multipliers.

---

[24]Each asset is a portfolio formed by bonds with the same detailed rating category and quarter-to-maturity, and the substitute portfolio includes all the other bonds in the same detailed rating category.

Figure 5: IG/HY Segmentation Effect ($N = 28,192$)
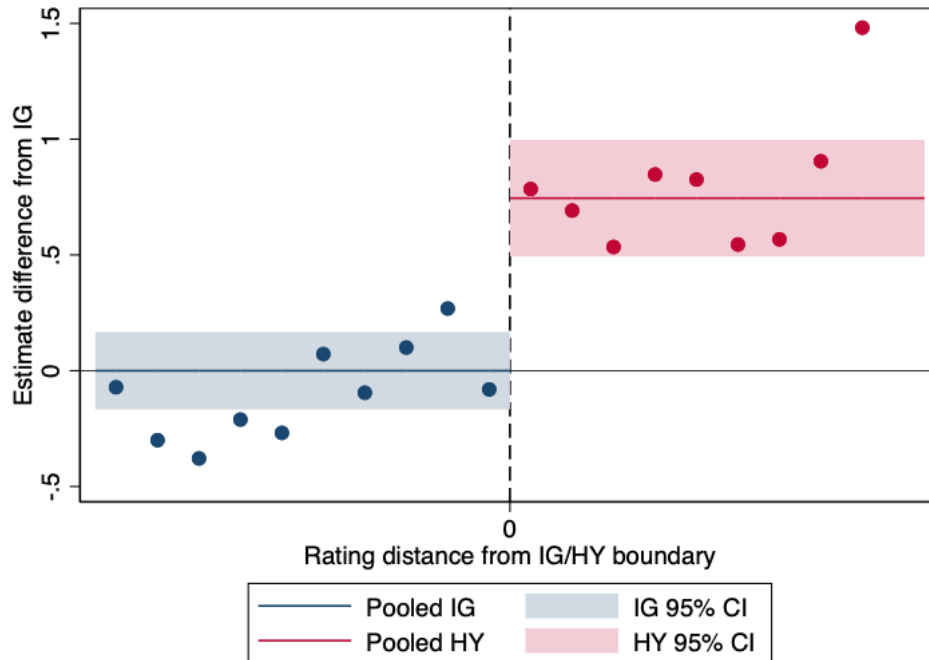
The graph shows how the detailed rating $\times$ quarter-to-maturity multiplier estimates differ across ratings. Specifically, the estimates are plotted relative to the pooled investment grade multiplier estimate. The blue and red dots are the heterogeneous estimates for IG and HY detailed rating categories, respectively. The red line represents the relative estimate for HY bonds. And the blue and red shaded areas represent the 95% confidence interval for the IG and HY pooled estimate, respectively.

Figure 6: Zooming in on IG/HY Segmentation Effect ($N = 20,372$)
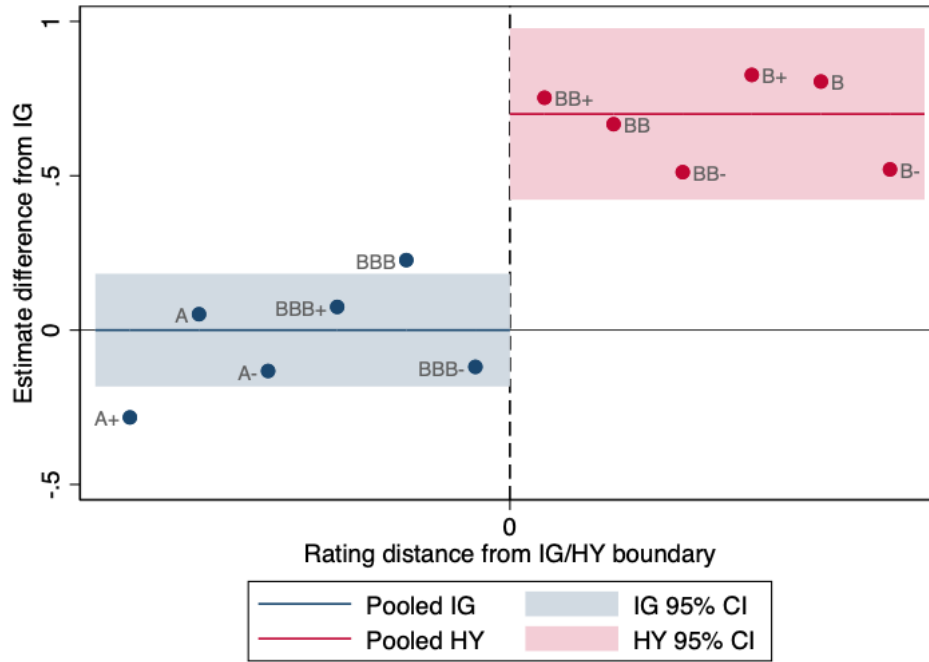
The graph shows how the detailed rating × quarter-to-maturity multiplier estimates differ across the six detailed ratings above and below the IG/HY cutoff. Specifically, the estimates are plotted relative to the multiplier estimate of the six IG categories just above the cutoff. The blue and red dots are the heterogeneous estimates for IG and HY detailed rating categories, respectively. The red line represents the relative pooled estimate for six detailed rating categories just below the cutoff. And the blue and red shaded areas represent the 95% confidence interval of the zoomed in pooled IG and HY estimates, respectively.

The discrete change in portfolio multipliers around the IG/HY cutoff suggests that investor segmentation may be a source of demand inelasticity. Figure 6 zooms in closer to the IG/HY cutoff. On the left hand side of the cutoff, the point estimates for BBB-rated portfolio and A-rated portfolios are all close to 0. Once we cross the cutoff, the estimated portfolio multipliers jump discretely upward to around 0.6. Investor segmentation is particularly strong at the IG/HY cutoff. Many investors in the corporate bond market face strong incentives to either stay within the investment grade universe or the high yield universe. For example, insurance companies face much higher capital charges when holding HY bonds — 90% of insurance companies' portfolio is in IG bonds. Mutual funds are set up as either investment-grade funds or high-yield funds. The fund mandates limit the managers from investing in the other category of bonds. As a result of these frictions, we see strong investor segmentation around the IG/HY cutoff. While mutual funds hold only about 15 percent of BBB- rated bonds, they hold about 22 percent for BB+ rated bonds. The difference in investor base likely contributes to the discrete changes in the portfolio multipliers estimated.

In addition to ratings, we also investigate heterogeneity along the maturity dimension. Specifically, we group portfolios into short, medium and long maturity groups: 0-3 years, 4-10 years and 10 years above. Figure 7 shows longer maturity portfolios have slightly higher portfolio multipliers, although the difference is not statistically significant. Short-term portfolios have a multiplier close to 0; medium-term and long-term portfolios have a higher multiplier around 0.3-0.4. The heterogeneous multipliers for portfolios with different maturity are likely linked to the clientele effects and segmentation along the yield curve (Vayanos and Vila, 2021; Kekre et al., 2022).

Figure 7: Maturity Heterogeneity $(N = 76,348)$
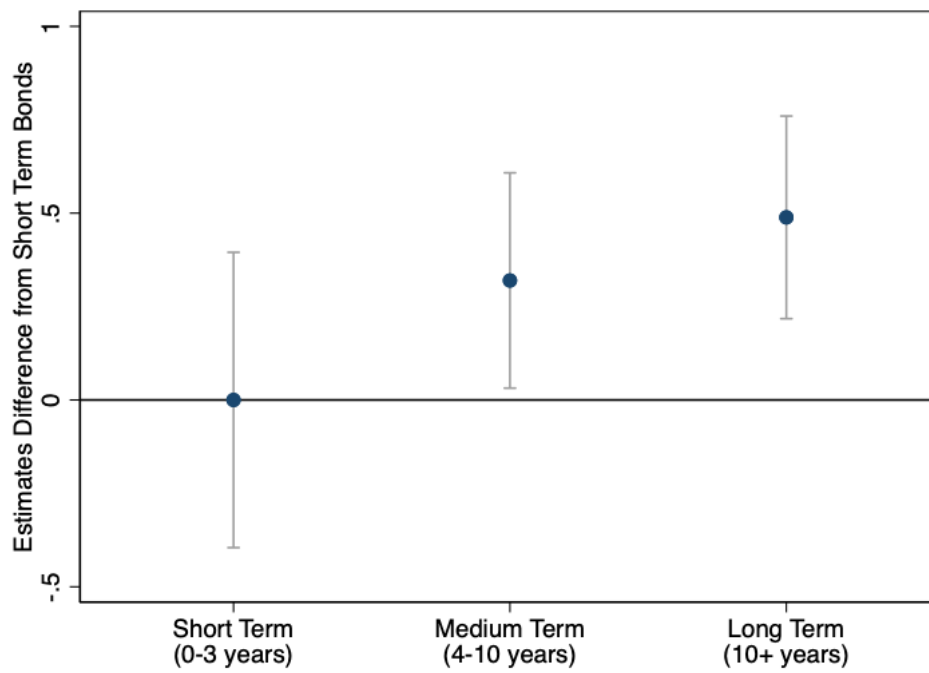
The graph shows how the detailed rating $\times$ quarter-to-maturity multiplier estimates differ across maturity. The estimates are presented relative to the short-term maturity elasticity. The bars reflect 95% confidence intervals. As the estimates suggest we can reject the null of no heterogeneity at the 95% level.

5 Arbitrage Risk

To better understand the quantitative implications of the multipliers we find in the previous section, we look at the Sharpe ratio of a strategy that takes advantage of the price deviation between the testing asset and its substitute portfolio at the one-quarter horizon. For example, if portfolio A's price is too high relative to its substitute portfolio, arbitrageurs could make money by selling portfolio A and purchasing its substitute. However, if substitute returns do not fully replicate portfolio A's returns, arbitrageurs face tracking error risk, which could result in losses. In other words, this arbitrage is not risk-free. Specifically, we consider a strategy that sells the more expensive one between asset j and its close substitute asset and buys the cheaper one. We unwind this position in the following period, given that we find the price gap disappears by the end of the next period on average.

The expected return of this strategy is $M \times Mean(|u_{j,t}|)$, where M is the multiplier estimated in Section 4 and $|u_{j,t}|$ is the size of demand shocks in absolute terms. The volatility of this return is the arbitrage risk, defined as

$$ArbRisk_j \equiv std(\tilde{\nu}_{j,t}^S + Mu_{j,t}) \tag{10}$$

The arbitrage risk comes from both non-flow shocks, as well as the next period (observable) demand shocks ($u_{j,t}$). Since we have removed all the predictable components in constructing u, the following period u is unknown to the arbitrageur when carrying out the strategy and should be considered as part of the arbitraging risk.

The average Sharpe Ratio of this arbitrage activity is then given by

$$SR = \frac{M \times Mean(|u_{j,t}|)}{Mean(ArbRisk_j)} \tag{11}$$

We find the annualized Sharpe ratio across different aggregation levels are generally small, ranging from 0.006 at the CUSIP level to 0.28 at the most aggregated rating level. The Sharpe ratio increases with the aggregation level, mainly because the multipliers are larger for more aggregated portfolios. However, even the maximum Sharpe ratio is small and it is before considering potential shorting costs. Our results reveal that the risks of engaging in these arbitrages are high relative to the average gain, which potentially explains why we observe these price deviations exist in the data.

Table 7: Arbitrage Risk and Portfolio Multipliers

	M	Arb. risk	Sharpe ratio
CUSIP	.052	.043	.006
Det. rating × Quarter to maturity	.348	.055	.026
Rating × ST/MT/LT	1.229	.037	.134
Rating × ST/LT	1.591	.048	.142
Rating	3.507	.043	.280

The table presents the estimated multiplier (M), the average arbitrage risk (as defined in Equation 10) and the implied annualized Sharpe ratio for portfolios at different aggregation levels.

5.1 Multipliers and Arbitrage Risk

Motivated by the low Sharpe ratio, we explore to what extent arbitrage risk limits investors to substitute across assets. Our arbitrage risk is the part of the risk that cannot be hedged away by constructing the substitute portfolio, and hence it acts as a limit to arbitrage Wurgler and Zhuravskaya (2002). Intuitively, the higher the arbitrage risk, the less arbitrageurs are willing to substitute to alternative portfolios and hence higher multipliers. In Appendix E, we provide a simple two-asset case illustrating this intuition.

We use our baseline portfolio (formed by bonds with the same detailed rating group and quarter-to-maturity) for analysis in this section. We first sort our portfolios into four groups based on the size of their arbitrage risk defined in Equation 10: group 1 includes portfolios with the lowest arbitrage risk, and group 4 includes portfolios with the highest arbitrage risk. We run the baseline specification in Section 4.1 including interaction terms of group indicators and demand shocks ($u_{j,t}$). In other words, we allow the portfolio multiplier M to differ by group. We plot the relative magnitudes of the estimates for each group in Figure 8.

We indeed find that portfolios with higher arbitrage risk have higher portfolio multipliers and more inelastic demand. We obtain an almost monotone pattern of multipliers across groups. Groups with higher arbitrage risk have significantly higher portfolio multipliers. For group 1, the estimate for the multiplier is close to 0, suggesting it is almost frictionless arbitrage between bonds in this group and their substitutes. However, for group 4, the estimate for the multiplier is much larger. Wurgler and Zhuravskaya (2002) find similar patterns in the stock market using index inclusion and

Table 8: Arbitrage Risk and Portfolio Multipliers

	(1)	(2)
Shock	0.35***	0.03
	(0.04)	(0.04)
ArbRisk x Shock		0.45***
		(0.03)
Substitute return	0.90***	0.91***
	(0.04)	(0.04)
Quarter FE	Yes	Yes
Drop Crisis	Yes	Yes
N	76,348	76,348
First-stage F-statistic	2544.04	2562.86

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table presents estimates for how the detailed rating $\times$ quarter-to-maturity multiplier estimates depend on the arbitrage risk of the portfolio. We standardized the arbitrage risk measurement by subtracting the mean and dividing it by the cross-section standard deviation of the arbitrage risk. Column (1) estimates our baseline specification for the post-winsorized sample. Column 2 includes an interaction term between the demand shocks and arbitrage risk. The parenthesis contains robust standard errors clustered at the substitute group $\times$ time level.

Figure 8: Arbitrage Risk and Portfolio Multipliers ($N = 76,348$)
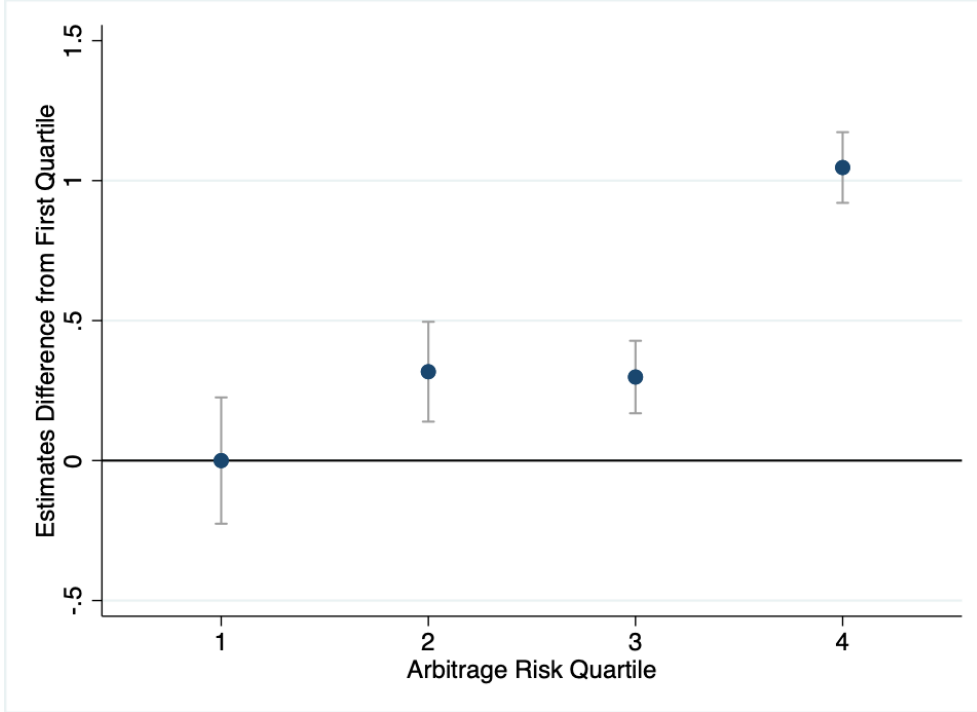
The graph shows how the detailed rating $\times$ quarter-to-maturity multiplier estimates different arbitrage risk quartiles. The estimates are presented relative to the lowest arbitrage risk quartile. The bars reflect 95% confidence intervals. As the estimates suggest we can reject the null of no heterogeneity at the 95% level.

exclusion events. We verify such pattern exists more broadly outside of the index-related events.

The relationship between arbitrage risks and portfolio multipliers is also verified in a regression specification. We add an interaction term between $ArbRisk_j$ and demand shock u in the baseline regression specification,

$$\Delta p_{j,t} = M_0 u_{j,t} + M_1 u_{j,t} \times ArbRisk_j + \tilde{M}\Delta p_{j,t}^{sub} + \epsilon_{j,t} \tag{12}$$

Table 8 presents the results. The coefficient in front of the interaction term is significantly positive. A one-standard-deviation increase in arbitrage risk increases the portfolio multiplier by 0.45. The effect is particularly strong among the high-yield bonds, which have higher overall risks and also idiosyncratic risks.

6 Conclusion

In this paper, we study a wide range of multipliers in the corporate bond market utilizing exogenous demand shocks from mutual fund flows. Studying the corporate bond market is interesting not only because it serves as an important channel for the real economy, but also because it is different from the equity market in terms of market structure, investor composition and correlation of asset payoffs. Comparing the two markets can shed light on the key drivers of market inelasticity.

We emphasize that it is important to account for the correct substitute portfolios when estimating the multipliers. Different from existing methods in the literature, we allow for close and distant substitutes in our estimation. Ignoring the heterogeneous cross-elasticity leads to underestimation of demand elasticities. At the CUSIP-level, the estimated multiplier drops from 0.3 to essentially 0 (both economically and statistically) once we allow for certain bonds to be closer substitutes than others. Relative to equities, we find that individual bonds are much more elastic.

While individual bonds are quite elastic, portfolios of bonds face more inelastic demand. The portfolios formed by rating categories and maturity have a multiplier around 0.35. Furthermore, as the portfolio becomes more aggregated, the portfolio-level multiplier increases and the substitute passthrough decreases, indicating portfolios are less substitutable at more aggregated levels. Finally, we find that riskier portfolios and long-term portfolios have higher multipliers; portfolios with larger arbitrage risks also have higher multipliers. Our results imply that both investor segmentation and arbitrage risk contribute to inelastic demand at the portfolio level. We leave the quantification of different channels to future research.

Finally, we find that the price impact at the portfolio level almost disappears in the next quarter, suggesting that long-run demand is more elastic than short-run demand. Investigating the dynamics of the price impact is interesting and can be informative about the underlying arbitrage frictions. For future work, it would be interesting to link the demand elasticities at different horizons to different arbitrage theories.

References

Beneish, Messod D. and Robert E. Whaley, "An Anatomy of the "S & P Game": The Effects of Changing the Rules," *The Journal of Finance*, 12 1996, *51* (5), 1909.

Bretscher, Lorenzo, Lukas Schmid, Ishita Sen, and Varun Sharma, "Institutional Corporate Bond Pricing," *SSRN Electronic Journal*, 1 2022.

Calomiris, Charles W., Mauricio Larrain, Sergio L. Schmukler, and Tomas Williams, "Large international corporate bonds: Investor behavior and firm responses," *Journal of International Economics*, 7 2022, *137*, 103624.

Chang, Yen-Cheng, Harrison Hong, Inessa Liskovich, Yen-Cheng Chang, Harrison Hong, and Inessa Liskovich, "Regression Discontinuity and the Price Effects of Stock Market Indexing," *Review of Financial Studies*, 1 2015, *28* (1), 212–246.

Chaudhry, Aditya, "Do Subjective Growth Expectations Matter for Asset Prices?," *SSRN Electronic Journal*, 9 2022.

Chen, Honghui, Gregory Noronha, and Vijay Singal, "The Price Response to S&P 500 Index Additions and Deletions: Evidence of Asymmetry and a New Explanation," 2004.

Coval, Joshua and Erik Stafford, "Asset fire sales (and purchases) in equity markets," *Journal of Financial Economics*, 11 2007, *86* (2), 479–512.

Darmouni, Olivier, Kerry Y Siani, and Kairong Xiao, "Nonbank Fragility in Credit Markets: Evidence from a Two-Layer Asset Demand System *," 2023.

Davis, Steven J. and John Haltiwanger, "Gross job creation, gross job destruction, and employment reallocation," *Quarterly Journal of Economics*, 1992, *107* (3), 819–863.

Edmans, Alex, Itay Goldstein, and Wei Jiang, "The Real Effects of Financial Markets: The Impact of Prices on Takeovers," *The Journal of Finance*, 6 2012, *67* (3), 933–971.

Fang, Chuck, "Monetary Policy Amplification through Bond Fund Flows," *Working paper*, 2022.

Gabaix, Xavier and Ralph S J Koijen, "In Search of the Origins of Financial Fluctuations: The Inelastic Markets Hypothesis," Technical Report 2021.

_ **and** _ , "Granular Instrumental Variables," 2022.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik Instruments: What, When, Why, and How," *American Economic Review*, 8 2020, *110* (8), 2586–2624.

Greenwood, Robin, Nber Toomas Laarits, and Jeffrey Wurgler, "Stock Market Stimulus," 2022.

Harris, Lawrence and Eitan Gurel, "Price and Volume Effects Associated with Changes in the S&P 500 List: New Evidence for the Existence of Price Pressures," *The Journal of Finance*, 9 1986, *41* (4), 815.

Hartzmark, Samuel M and David H Solomon, "Predictable Price Pressure," 2022.

Jansen, Kristy A.E., "Long-term Investors, Demand Shifts, and Yields," *SSRN Electronic Journal*, 8 2021.

Kekre, Rohan, Moritz Lenel, and Federico Mainardi, "Monetary Policy, Segmentation, and the Term Structure *," 2022.

Koijen, Ralph and Motohiro Yogo, "Exchange Rates and Asset Prices in a Global Demand System Exchange Rates and Asset Prices in a Global Demand System *," 2019.

Koijen, Ralph S.J. and Motohiro Yogo, "A Demand System Approach to Asset Pricing," *https://doi.org/10.1086/701683*, 8 2019, *127* (4), 1475–1515.

Li, Jennifer, Neil D Pearson, and Qi Zhang, "Impact of Demand Shocks on the Stock Market: Evidence from Chinese IPOS ," 2021.

Li, Jiacui, "What Drives the Size and Value Factors?," *SSRN Electronic Journal*, 6 2021.

_ and Zihan Lin, "Prices Are Less Elastic For Less Diversifiable Demand," *SSRN Electronic Journal*, 6 2022.

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

Pavlova, Anna and Taisiya Sikorskaya, "Benchmarking Intensity," *The Review of Financial Studies*, 8 2022.

Peng, Cameron and Chen Wang, "Factor Demand and Factor Returns *," 2019.

Petajisto, Antti, "Why Do Demand Curves for Stocks Slope Down?," *Source: The Journal of Financial and Quantitative Analysis*, 2009, *44* (5), 1013–1044.

Shleifer, Andrei, "Do Demand Curves for Stocks Slope Down?," *The Journal of Finance*, 7 1986, *41* (3), 579.

Siani, Kerry, "Raising Bond Capital in Segmented Markets," *SSRN Electronic Journal*, 10 2022.

Siani, Kerry Y and Kerry Yang Siani, "Raising Bond Capital in Segmented Markets *," 2022.

Vayanos, Dimitri and Jean-Luc Vila, "A Preferred-Habitat Model of the Term Structure of Interest Rates," *Econometrica*, 1 2021, *89* (1), 77–112.

Wurgler, Jeffrey and Ekaterina Zhuravskaya, "Does Arbitrage Flatten Demand Curves for Stocks?," *Journal of Business*, 10 2002, *75* (4), 583–608.

A Identification

A.1 Identification of Multipliers

We first consider the specification in Equation (7), restated here:

$$\Delta p_{j,t} = M u_{j,t} + \text{Fixed effects}_{g,t} + \tilde{\nu}_{j,t}.$$

To identify the coefficients correctly, we need the following moment conditions to be satisfied:

$$\mathbb{E}[u_{j,t}\nu_{j,t}] = 0. \tag{13}$$

We now prove that Assumption 1, i.e. $\nu_{jt} \perp S_{ij,t-1}$ for all fund i and time t, is a sufficient condition for satisfying these moment conditions, where $S_{ij,t-1}$ is the lagged market share of mutual fund i. For clarity, we use the notation $\mathbb{E}_i^{(j)}$ to denote the expectation over dimension j conditional on dimension i.

Proof.

$$
\begin{aligned}
\mathbb{E}^{(j,t)}\left[u_{j,t}\nu_{j,t}\right] &= \mathbb{E}^{(j,t)}\left[\mathbb{E}_{j,t}^{(i)}\left[S_{ij,t-1}u_{i,t}\right]\nu_{j,t}\right] && \text{(Definition of } u_{j,t}) \\
&= \mathbb{E}^{(i,j,t)}\left[S_{ij,t-1}u_{i,t}\nu_{j,t}\right] && \text{(L.I.E.)} \\
&= \mathbb{E}^{(i,t)}\left[u_{i,t}\mathbb{E}_{i,t}^{(j)}\left[S_{ij,t-1}\nu_{j,t}\right]\right] && \text{(conditional Exp.)} \\
&= \mathbb{E}^{(i)}\left[u_{i,t}0\right]. && \text{(I.D. assum.)}
\end{aligned}
$$

$\square$

The first equality is the definition of bond-level shocks as the aggregation across fund-level shocks. The second line directly follows through the law of iterated expectations. The third equality applies the law of iterated expectations again, and this time we take the expectations across bonds *within each fund*. As shocks $u_{i,t}$ are at the fund level, it can be pulled out of the expectations. This is the key property exploited in the exogenous-share design. Notice that here we do not assume orthogonal flows. The last equality uses the identification assumption.

A.2 Identification of Substitute Passthrough

We estimate the passthrough according to the following equation:

$$\Delta p_{j,t} = M u_{j,t} + \tilde{M}\Delta p_{j,t}^{sub} + \tilde{\nu}_{j,t},$$

where we instrument $\Delta p_{j,t}^{sub}$ with $u_{j,t}^{sub}$, constructed as the value-weighted group-level shocks:

$$u_{j,t}^{sub} \equiv \sum_k w_{k|g(j),t-1} u_k,$$

where $g(j)$ is the group of close substitutes for bond j, and $w_{k|g(j),t-1} = \frac{\mathbb{I}_{k \in g(j)} w_{k,t-1}}{w_{g(j),t-1}}$ is the (lagged) market share of bond k conditional on group $g(j)$. The required moment condition is:

$$\mathbb{E}[u_{j,t}^{sub} \nu_{j,t}] = 0.$$

The proof is slightly more complicated due to the weighting in u^{sub}, but essentially follows a similar strategy. We suppress the time subscript in the proof below for ease of notation.

Proof. Plug in the definitions and use the law of iterated expectations as before, we have:

$$\mathbb{E}^{(j)}\left[u_j^{sub} \nu_j\right] = \mathbb{E}^{(j)}\left[w_{k|g(j)} u_k \nu_j\right] \tag{14}$$

$$= \mathbb{E}^{(j)}\left[w_{k|g(j)} \mathbb{E}_k^{(i)}\left[u_i S_{ik}\right] \nu_j\right] \tag{15}$$

$$= \mathbb{E}^{(i)}\left[\mathbb{E}_i^{(k)}\left[\mathbb{E}_{ik}^{(j)}\left[w_{k|g(j)} \nu_j\right] S_{ik}\right] u_i\right]. \tag{16}$$

Notice that the inner expectation is virtually the group-level idiosyncratic shocks:

$$\mathbb{E}_{ik}^{(j)}\left[w_{k|g(j)} \nu_j\right] = \mathbb{E}_{ik}^{(j)}\left[\frac{\mathbb{I}_{k \in g(j)} w_k}{w_{g(j)}} \nu_j\right] = w_k \underbrace{\mathbb{E}_{ik, j \in g(k)}^{(j)} \nu_j}_{\nu_g}.$$

Plug it back into Equation 16, we have:

$$\mathbb{E}_i^{(k)}\left[\mathbb{E}_{ik}^{(j)}\left[w_{k|g(j)} \nu_j\right] S_{ik}\right] = \mathbb{E}_i^{(k)}\left[w_k S_{ik} \nu_g\right]$$

$$= \mathbb{E}_i^{(g)}\left[\underbrace{\mathbb{E}_g^{(k)}\left[w_k S_{ik}\right]}_{S_{i,g}} \nu_g\right]$$

$$= 0.$$

Here we rely on the exogenous share assumption at the group level, i.e., $S_{i,g} \perp \nu_g$. $\qquad \square$

A.3 Cases under Common Factors

A potential threat to the exogenous share assumption is that bond returns can load on common factors, and there exist factor-specialized funds whose market shares are correlated with the factor loading of bonds. More specifically, suppose $\nu_{jt} = \beta_j \eta_t + \check{\nu}_{jt}$, and there exists fund i such that $S_{ij,t-1} = \kappa_i \beta_j + \check{S}_{ij,t-1}$, then $\mathbb{E}_{i,t}[\nu_{jt} S_{ij,t-1}] = \kappa_i E[\beta_j^2] \neq 0$.

The issue is remedied by flow orthogonality to factors. That is,

$$u_{i,t} \perp \eta_t \quad \forall i \tag{17}$$

As shown in B.2, the factor structure in flows is actually very weak. In B.3, we further explicitly remove the potential common factors in flows with the parametric method to ensure condition (17) holds. We show that with condition (17) and Assumption 1, the moment condition 13 holds even when shares are not orthogonal to factor loadings.

Proof.

$$
\begin{aligned}
\mathbb{E}\left[u_{j,t}\nu_{j,t}\right] &= \mathbb{E}^{(i,j,t)}\left[u_{j,t}S_{ij,t-1}\left(\beta_j\eta_t + \check{\nu}_{j,t}\right)\right] \\
&= \mathbb{E}^{(i,j,t)}\left[u_{i,t}S_{ij,t-1}\beta_j\eta_t\right] + \mathbb{E}^{(i,j,t)}\left[u_{i,t}S_{ij,t-1}\check{\nu}_{j,t}\right] \\
&= \mathbb{E}^{(i,j)}\left[S_{ij,t-1}\beta_j\underbrace{\mathbb{E}_i^{(t)}\left[u_{i,t}\eta_t\right]}_{0}\right] + \mathbb{E}^{(i,t)}\left[u_{i,t}\underbrace{\mathbb{E}_{i,t}^{(j)}\left[S_{ij,t-1}\check{\nu}_{j,t}\right]}_{0}\right] \\
&= 0.
\end{aligned}
$$

The first term in the last equation follows from flow exogeneity with respect to common factors (17), and the second term follows from the share exogeneity in Assumption 1. $\qquad\square$

B Robustness of Shock Construction

In this section, we explore the sensitivity of our results to various deviations from our baseline assumptions used in the shock construction process.

B.1 Assessing passthrough of fund flows into trading

In constructing our shocks, we assumed that mutual funds mechanically reinvest inflows according to their existing portfolio weights. This assumption implies a one-to-one passthrough from inflows into trades. In this section, we kick the tires on this assumption and find that while the passthrough is not precisely one-to-one, it is still relatively high, around 50% at the individual bond level and 60% at our baseline level of aggregation. Overall, the less than one-to-one passthrough suggests that our multiplier estimates are potentially attenuated. However, we can find upper bounds on the multipliers using our passthrough estimates. The upper bound estimates are still considerably smaller than those typically found for stocks. At the bond level, the upper bound multiplier estimate is 0.08, and for our baseline level of aggregation, it is around 0.33.

Instead, we run the following specification ignoring the price response,

$$\Delta q_{ijt} = \alpha + \beta f_{it} + \varepsilon_{ijt}$$

where $\Delta q_{ijt} = 2\frac{Q_{ijt}-Q_{ij,t-1}}{Q_{ijt}+Q_{ij,t-1}}$ is the Davis and Haltiwanger (1992) change in mutual fund holdings, and f_{it} is the funds flows as a percentage of lagged AUM.[25] Naturally, our β estimates are biased due to omitting the prices. However, since mutual fund demand curves are presumably downward sloping, we can sign the bias as negative. As a result, our β estimate is a lower bound of the true passthrough.

Table 9 shows that passthrough is not equal to one but is still sizeable and highly significant. At the individual bond (CUSIP) level, the passthrough is around 0.51. This estimate is smaller than what Lou (2012) found for individual stocks—0.62 for inflows and 0.97 for outflows. The lower passthrough for individual bonds is, however, somewhat expected. It is presumably easier to find another GM stock, than it is to find another 34-quarter GM bond maturing in 12 months. Consistent with the difficulty in finding exact bond CUSIPs, the passthrough increases at more aggregate levels—it is easier to find another BBB bond than a specific CUSIP. At the detail rating-quarter to maturity level, the passthrough rises to around 0.58. At the rating letter level, the passthrough is closer to 0.95.

Table 9: Estimating passthrough from flows into holdings

	CUSIP		Det Rating x Q to Mat		Rating	
	(1)	(2)	(3)	(4)	(5)	(6)
Flows	0.51***	0.49***	0.58***	0.57***	0.94***	0.91***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.04)	(0.05)
Time + Fund FE	No	Yes	No	Yes	No	Yes
N	4,861,780	4,861,779	1,193,197	1,193,177	163,515	163,502

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10 separately assesses the passthrough for inflows and outflows. Two things stand out.

---

[25]Davis and Haltiwanger (1992) change is a second order approximation of the log difference, and has several desirable properties over the standard percentage change measure, for example, it is symmetric, and it does not generate outliers due to small base effects.

Firstly, we see higher passthrough at more aggregate levels. Secondly, the passthrough is higher for outflows than inflows, echoing Lou (2012) findings for equities.

Table 10: Estimating passthrough from flows into holdings

	CUSIP		Det Rating x Q to Mat		Rating	
	(1)	(2)	(3)	(4)	(5)	(6)
Flows	0.45***	0.66***	0.55***	0.79***	0.94***	1.08***
	(0.05)	(0.05)	(0.02)	(0.04)	(0.03)	(0.13)
Time + Fund FE	Yes	Yes	Yes	Yes	Yes	Yes
Flow Type	Inflow	Outflow	Inflow	Outflow	Inflow	Outflow
N	2,893,530	1,968,240	690,300	502,855	84,099	79,391

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Overall, these passthrough estimates suggest that our multiplier estimates are potentially attenuated. By assuming a passthrough of 1, we are overestimating the size of the demand shock and hence underestimating the multiplier. Since our passthrough estimates are lower bounds for the actual passthrough, we can use them to estimate the upper bounds on the true multiplier. If our lower bound passthrough estimate is $\hat{\beta} < 1$, then the true multiplier must be $M \leq \frac{1}{\hat{\beta}}\hat{M}$, where $\hat{M}$ is the multiplier we estimated in section 4.1. Accordingly, this calculation suggests that the true CUSIP multiplier could be as high as 0.08, and our detailed rating-quarter to maturity could be as high as 0.33. While these upper bounds are higher than our estimates, they are still considerably smaller than 1, the magnitude typically found for equities.

B.2 Non-parametric estimation of flow factor structure

A potential threat to the exogenous share condition is that for style-focusing funds, their shares might be correlated with the bonds' loading on common factors, which also drive bond returns. To address this concern, we exploit the shock orthogonality to common factors, as explained in Appendix A.3. In this section, we implement the factor models for flow shocks. We first show that flow shocks have very weak factor structures in this section; in the next section we explicitly remove the common factors and show our results are robust.

We first implement the non-parametric approaches to estimate common factors. Due to missing

values in our data, we estimate the factor model using alternating least squares (ALS). We find the data exhibits a relatively weak factor structure, with the first factor only explaining around 8% of the variation in mutual fund flows.

With that said, non-parametric factor estimation methods are known to have poor finite sample performance when there are a lot of missing value. In our context, due to mutual fund entry and exit, we have a highly unbalanced panel of flow data—relative to a fully balanced panel we are "missing" 49.8% of observations. Hence, there is a risk that the factor model is simply fitting noise rather than identifying actual factors.

To assess this overfitting risk, we run a five-fold cross-validation exercise. The procedure splits the sample into five subsamples, estimates the factor model on one of them, and then assess the models performance on the four subsamples left out. It then repeats this five times and calculates the average performance of the method.

Figure 9 shows the average root mean-squared errors from the five-fold cross-validation exercise. The results suggests the factors are likely being fit on noise. In fact, adding more factors seems to make the out-of-sample performance even worse. While this finding does not suggest that there is a strong factor structure in flows, it does suggest that non-parametric factor estimation may not be able to extract it in our setting with a highly unbalanced panel. Due to this concern, in the next subsection we explore a parametric approach to estimating factors—the additional structure should help reduce the risk of overfitting.

## B.3   Parametric estimation of flow factor structure

Due to the poor performance of the non-parametric method in estimating factors, in this section we follow a parametric approach below. Specifically, let the data-generating process of innovation be:

$$\epsilon_{i,t} = \delta_t + \Lambda_{i,t}\eta_t + u_{i,t}, \tag{18}$$

where δ_t is the time fixed effects, $\Lambda_i\eta_t$ is the contribution from common factors, and $u_{i,t}$ is the desired idiosyncratic demand shocks.

Following the common approach in asset pricing, we assume the factor loading are a linear function of characteristics, $\Lambda_{i,t} = C'_{i,t}\lambda$, where $C_{i,t}$ is a vector of observable characteristics of fund i, including (lagged) log AUM of the firm, the share in high-yield bonds, and the average duration in the portfolio, and λ is a constant vector. The data generating process of $\epsilon_{i,t}$ under this

Figure 9: Cross-validation of non-parametric factor estimation
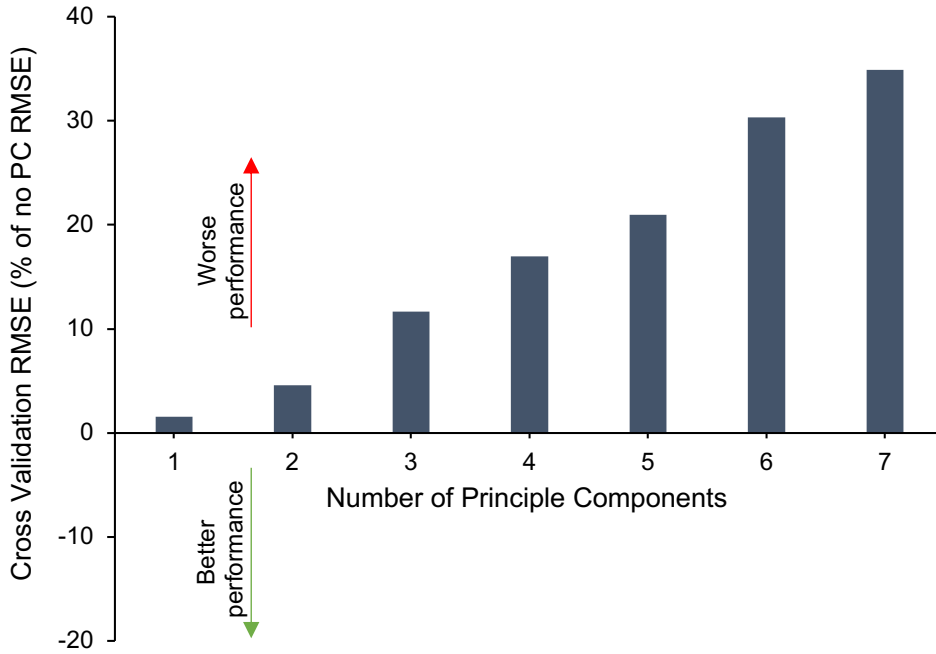
parameterization is then:

$$\epsilon_{i,t} = \delta_t + C'_{i,t}(\lambda\eta_t) + u_{i,t}. \tag{19}$$

Notice (19) can be estimated by running a panel regression of $\epsilon_{i,t}$ on $C_{i,t}$ with time-varying coefficients and time fixed effects. In the baseline where no additional factors are controlled, we simply regress $\epsilon_{i,t}$ on time fixed effects. The estimated residual is recovered as:

$$\hat{u}_{i,t} = \epsilon_{i,t} - \hat{\delta}_t - C'_{i,t}(\widehat{\lambda\eta_t}) \tag{20}$$

To minimize noise due to extreme outliers and volatile funds, we winsorize innovations $\epsilon_{i,t}$ at the 5% level before estimating (19). Importantly, the winsorized innovations are only used in the estimation of coefficients. In (20) we use the original $\epsilon_{i,t}$ to recover the demand shock $\hat{u}_{i,t}$, as outliers are also valid idiosyncratic shocks to demand.

B.4   Lagging market shares for shock construction

Another potential concern to the exogenous share condition is that some mutual fund managers have superior information and therefore can front-run the market, leading to a positive correlation between market share and bond returns. This issue can be addressed with further lagging the

market share used in aggregating fund flow shocks: it is highly unlikely for a mutual fund manager to predict *idiosyncratic* returns at the bond level in one year and allocate their portfolio in one year to benefit from it.[26]

Table 11 and 12 report our baseline results using shocks aggregated with one-year-lag market shares. The results remain largely unchanged.

Table 11: CUSIP specification lagging holding shares by one-year

	Homo. OLS	OLS		First-stage	2SLS
	(1)	(2)	(3)	(4)	(5)
Shock	0.39***	-0.03	0.01		0.02
	(0.07)	(0.06)	(0.05)		(0.06)
Substitute return					1.08***
					(0.04)
Group Shock				3.93***	
				(0.47)	
Quarter FE	Yes	Yes	Yes	Yes	Yes
Group x Quarter FE	No	Yes	Yes	No	No
Drop Crisis	No	No	Yes	Yes	Yes
N	277,336	277,336	261,144	261,144	261,144
R^2	0.21	0.40	0.37	0.62	
First-stage F-statistic					94.96

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B.5   Sensitivity of multiplier estimation to shock construction specifications

Below we explore how sensitive our multiplier estimates are to changing various choices we made in the shock construction process. Overall, our estimates do not seem to change very much if we (i) winsorize extreme values and use inverse variance weighting in calculating fixed effects, (ii)

---

[26]It might be more plausible if they are able to predict the dynamics of some systematic factors in one year and trade accordingly—the common factor issue is addressed in Section by removing the common factors from flows.

Table 12: Quarter to Maturity x Detailed Rating specification lagging holding shares by one-year

	Homo. OLS	OLS		First-stage	2SLS
	(1)	(2)	(3)	(4)	(5)
Shock	1.04***	0.31**	0.33**		0.33**
	(0.15)	(0.12)	(0.12)		(0.12)
Substitute return					0.96***
					(0.04)
Group Shock				4.36***	
				(0.51)	
Quarter FE	Yes	Yes	Yes	Yes	Yes
Group x Quarter FE	No	Yes	Yes	No	No
Drop Crisis	No	No	Yes	Yes	Yes
N	81,866	81,866	77,387	76,348	76,348
R^2	0.22	0.47	0.41	0.48	
First-stage F-statistic					69.48

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

control for additional factors as described in B.3, or (iii) use different specifications for our AR regressions. Furthermore, we also find that our substitute specification does a quite good job at control for substitutes, as the results are unchanged if we instead include detailed rating × quarter fixed effects.

Table 13: Shock robustness

|	Bench.	Robustness						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Shock	0.33***	0.29**	0.32**	0.37***	0.26**	0.28***	0.21***	0.18
	(0.10)	(0.10)	(0.11)	(0.11)	(0.08)	(0.08)	(0.06)	(0.13)
AR lags	3	3	1	2	1	2	3	0
Time Trend	Yes	Yes	Yes	Yes	No	No	No	No
Factors	No	Yes	No	No	No	No	No	No
Quarter × Sub FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Drop Crisis	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	77,387	77,387	77,387	77,387	77,387	77,387	77,387	77,387
R^2	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table summarizes the estimates for the portfolio multipliers when we use different specifications to extract demand shocks. All estimates are for portfolios formed by detailed rating and quarter-to-maturity as in Section 4.1. "Factors" refer to the parametric estimation of flow factor structure explained in Appendix B.3.

B.6   Morningstar data coverage

We find that Morningstar mutual fund dataset coverage is quite extensive and closely lines up with the coverage of the Flow of Funds estimate of mutual fund holdings.

Figure 10: Total Assets of All Mutual Funds in Flow of Funds vs. Morningstar
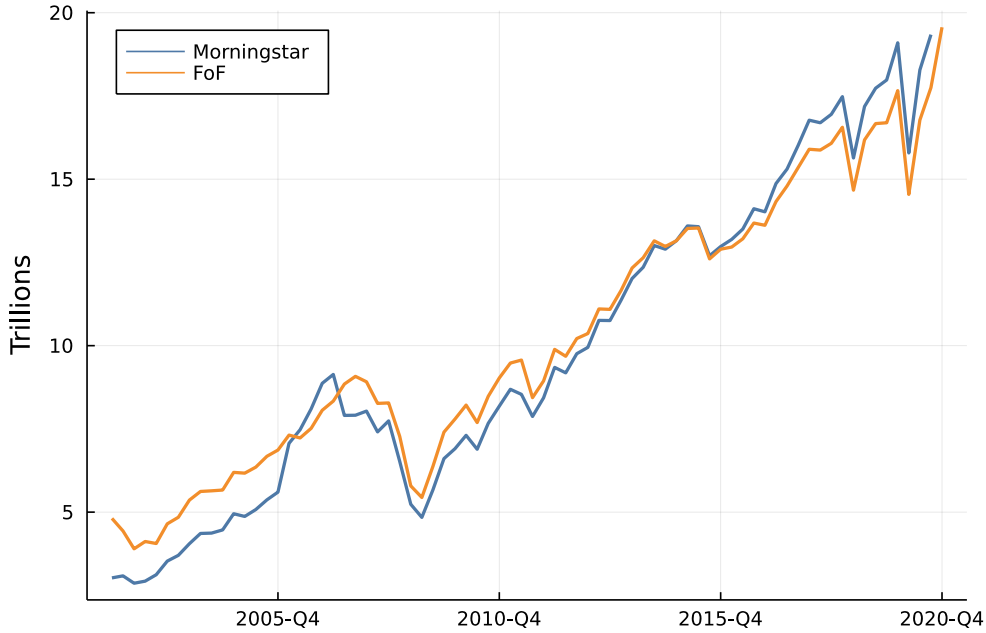
C   Substitutions in the equity market

In the main text of our paper, we focus on the corporate bond market. A priori, there are clear close substitutes in the bond market, which helps mitigate the effect of demand shocks. ignoring close substitutes leads to biased multiplier estimates. Our estimates further validate this prior. Stocks, however, do not have clear close substitutes, and therefore it is harder to accommodate demand shocks without large price impacts. Assuming homogeneous substitution when estimating the multiplier is also not far off.

In this section, we confirm this prior on the equity market. Using the same Morningstar fund data, we follow identical procedures in shock construction and identification as in our baseline estimation. We estimate the stock multiplier with and without controlling for its close substitutes— defined in terms of loading on Fama-French three factors, or its industries. We find that controlling for close substitutes does reduce the estimated multiplier, but only very slightly, suggesting that substitutions are way less influential in the equity market than in the bond market.

Table 14 reports the estimates of the equity multipliers at the stock level. The shocks are constructed in the same way as in the baseline specification for corporate bonds, and the sample covers the CRSP universe of the U.S. listed stocks from 2003Q1 to 2020Q4. In the first column, we regress stock returns on stock level shocks, controlling for the time-fixed effect only. This

Table 14: Multiplier estimates for stock markets

	Stock Return		
	(1)	(2)	(3)
Shock	0.380***	0.254***	0.252***
	(0.086)	(0.040)	(0.072)
Group x Quarter FE	None	FF3	Industry
Quarter FE	Yes	Yes	Yes
N	144,768	136,270	135,201
R^2	0.188	0.332	0.263

specification assumes homogeneous substitution patterns across stocks. The point estimate is 0.38, indicating a one-percent demand shock to a single stock leads to a 38 basis points increase in the stock price.

In Column (2), we add group-time fixed effects to control for the close substitutes. The group here is defined as the stocks with similar factor loading as the test stock. Specifically, we compute the loading of each stock on Famma-French three factors, and then triple-sort them into $3 \times 3 \times 3$ groups. The total group numbers (27) are comparable with the detailed rating groups in the bond market (21). Controlling for the group-time fixed effects does reduce the point estimate, suggesting that the equity market also exhibits heterogeneous substitution patterns. Nevertheless, the reduction in the multiplier is much less dramatic than the bond market counterpart as reported in 1. After controlling for close substitutes, the multiplier is still around 0.25, only 34% smaller than the univariate regression. To put it into perspective, in the bond market we see an almost 90% drop in the multiplier once controlled for the close substitutes. The multiplier is also much larger than the point estimate in corporate bonds, indicating much inelastic demand in the equity market. This is also consistent with the prior that it is harder to substitute away in the stock market than the bond market. In Column (3) we define the close substitutes in terms of the Global Industry Classification Standard (GICS) industry groups (27 groups in total). The estimate is very close.

In conclusion, we find that the force of heterogeneous substitution is at work in the equity market as well, but it is much weaker than the bond market, consistent with our prior.

D Interpretations based on Demand Systems

In this section, we first show the exact assumption on Γ that leads to equation (6). Formally, we let Γ be:

$$\Gamma_{j,k} = \begin{cases} \gamma^o & j = k \\ \gamma^d w_k + \gamma^c w_{k|g} & j \neq k, j, k \text{ in the same detailed rating } g \\ \gamma^d w_k & \text{otherwise} \end{cases} \quad (21)$$

where w_k is the market share of bond k, and $w_{k|g} \equiv \frac{w_k}{\sum_{k' \in g} w_g}$ is the within-group market share. This structure implies $\tilde{M} \equiv -\frac{\gamma^c}{\gamma^o}$. We scale the cross-elasticity by the market share of substitutes. This leads to market-capitalization weighted aggregation when forming the substitution portfolio.

Next, we show how our reduced form approach is linked to the logit demand system. Specifically, we derive the elasticity matrix Γ from a nested logit demand system and show that it yields the reduced-form specification as we used in the main text.

Following Koijen and Yogo (2019a), consider the demand structure for a representative investor as follows:

$$w(j \mid g) = \frac{exp(\delta(j, g))}{\sum_{j \in g} exp(\delta(j, g))}$$
$$w(g) = \frac{(\sum_{k \in g} exp(\delta(k, g)))^\lambda}{1 + \sum_{g'} (\sum_{k \in g'} exp(\delta(k, g')))^\lambda} \quad (22)$$
$$\delta(j, g) = \beta_g p_j + \beta X + u_j,$$

where $w(j \mid g)$ is the conditional share of asset j in group g, $w(g)$ is the share of group g in the whole portfolio. The unconditional share of asset j is therefore $w(j) = w(j \mid g)w(g)$. The parameter λ controls the substitution at the group level. When $\lambda = 1$, this system is reduced to the simple logit system as

$$w(j) = w(j \mid g)w(g) = \frac{exp(\delta(j, g))}{1 + \sum_{g'} (\sum_{k \in g'} exp(\delta(k, g')))}.$$

Taking the derivative of $w(j)$ with respect to p_k, we have:

$$\frac{\partial w(j)}{\partial p_k} = \begin{cases} \beta_g w(j) \left(1 - \lambda w(j) - (1 - \lambda)w(j \mid g)\right) & j = k \\ -\beta_g w(j) \left(\lambda w(k) + (1 - \lambda)w(k \mid g)\right) & j \neq k, \ j \& k \in g \\ -\beta_{g'} \lambda w(j)w(k) & j \in g, k \in g' \neq g \end{cases} \quad (23)$$

Log linearize the system (22) with the partial derivatives above, reorganize, we have:

$$\log \Delta w(j) = \beta_g \Delta p_j - \beta_g (1 - \lambda) \sum_{k \in g} w(k \mid g) \Delta p_k - \lambda \sum_{g'} \sum_{k \in g'} \beta_{g'} w(k) \Delta p_k + u_j$$

Applying market clearing condition $\log \Delta w(j) = \Delta p(j)$, we can derive the equation for estimation:

$$\Delta p_j = \underbrace{\frac{1}{(1 - \beta_g)} u_j}_{M} + \underbrace{\frac{\beta_g}{(\beta_g - 1)} (1 - \lambda) \sum_{k \in g} w(k \mid g) \Delta p_k}_{\tilde{M}} \underbrace{\vphantom{\sum_{k \in g}}}_{\Delta p_j^g} + \lambda \underbrace{\frac{1}{(\beta_g - 1)} \sum_{g'} \sum_{k \in g'} \beta_{g'} w(k) \Delta p_k}_{\text{Time FE}} + \tilde{\nu}_j \quad (24)$$

This equation maps directly to our empirical specification. When $\lambda = 1$, the nested system is reduced to standard logit demand, and $\tilde{M} = 0$. In this case, we do not need to control for price changes for close substitutes. Our estimates of $\tilde{M}$ strongly reject this assumption. Therefore, when specifying demand for corporate bonds, a nested system is preferred over the standard logit demand.

E   Arbitrage Risk

Consider a two-period economy populated with measure 1 of homogeneous investors. Investors have CARA utility, with absolute risk aversion coefficient γ. There are two risky assets with normally distributed payoff in period 2

$$\begin{pmatrix} D_A \\ D_B \end{pmatrix} \sim N(\mu, \Sigma) \quad (25)$$

$$\Sigma = \begin{pmatrix} \sigma_A^2 & \rho \sigma_A \sigma_B \\ \rho \sigma_A \sigma_B & \sigma_B^2 \end{pmatrix} \quad (26)$$

In addition, there is a risk-free asset with total return normalised to 1. Investors choose portfolio x to maximize expected utility,

$$x = \frac{\Sigma^{-1}}{\gamma} (\mu - p) \quad (27)$$

$$x_A = \frac{\gamma \sigma_B^2 (\mu_A - p_A) - \rho \gamma \sigma_A \sigma_B (\mu_B - p_B)}{|\gamma \Sigma|} \quad (28)$$

where $|\cdot|$ denotes matrix determinant. Apply market clearing, we get

$$x + u + \nu = \bar{x} \quad (29)$$

where u is the observed exogenous demand and ν is the unobserved demand shocks. Focusing on asset A, we can express its price as

$$p_A = \underbrace{\gamma\sigma_A^2(1-\rho^2)}_{=M}u_A + \frac{\rho\sigma_A}{\sigma_B}p_B + \text{constant} + \underbrace{\gamma(\sigma_A^2\nu_A + \rho\sigma_A\sigma_B\nu_B)}_{\text{error term}} \qquad (30)$$

The micro multiplier maps to the term $\gamma\sigma_A^2(1-\rho^2)$ in this case, which is determined by both the risk aversion coefficient as well as asset A's residualized risk. To understand the term $\sigma_A^2(1-\rho^2)$, regress asset A's price on its substitute, asset B's price,

$$p_A = \beta_0 + \beta_1 p_B + \epsilon \qquad (31)$$

the variance of ϵ captures the risk in asset A that cannot be hedged by holding asset B, which is equal to $\sigma_A^2(1-\rho^2)$.

Hence we have shown that the multiplier M is increasing in both the risk aversion coefficient and the residualized risk that cannot be hedged by the substitute portfolio.