

Rational Inattention and Price Underreaction

Jiacui Li*

May 16, 2019

ABSTRACT

Traditionally, researchers attribute price underreaction to liquidity or trading frictions. Using corporate bond data, I show that underreaction is better explained by investor facing information-processing (“attention”) constraints. Specifically, I test a rational inattention model under which investors optimally allocate more of their finite attention budget to more payoff-relevant risks. Consistent with model predictions, bonds with higher credit risk respond more quickly to default-relevant news, and bonds with longer duration respond more quickly to interest rate news. Investors appear to face attention constraints as larger shocks from one risk results in slower price response to the other. Because I explain differences in underreaction to different risks, this cannot be explained by liquidity which generates variation at the bond level. The amount of “money left on the table” is reasonably small so the inefficiency can be explained by actual information-processing costs. Finally, I find similar evidence in equities, suggesting my finding is general across asset classes.

Keywords: investor inattention, rational inattention, price underreaction, market efficiency

JEL classification: G14, G41

*Stanford Graduate School of Business (jiacui@stanford.edu). I am grateful for comments from Yu An, Jonathan Berk, Svetlana Bryzgalova, Darrell Duffie, Matthew Gentzkow, Steven Grenadier, Benjamin Hebert, David Hirshleifer, Robert Hodrick, Peter Koudijs, Arvind Krishnamurthy, Charles Lee, Ken Li, Yucheng Liang, Hanno Lustig, Ulrike Malmendier, Filip Matejka, Xuelang Mu, Paul Pflleiderer, Yang Song, David Sraer, Jeremy Stein, Xinyuan Tao, Laura Veldkamp, Edward Michael Watts, Joanna Yan, and David Yang, as well as seminar participants at Boston University, Chinese University of Hong Kong, Colorado, Fudan Fanghai, Hong Kong University, Indiana, Nanyang Business School, Ohio State, Peking Guanghua, Penn State, Rice, Stanford GSB, UIUC, Utah, UVA Darden, and Yale SOM. All errors are mine.

1 Introduction

Since Bernard and Thomas (1990), researchers have discovered many cases where asset prices respond slowly to public information.¹ What explains this underreaction? The predominant explanation is based on illiquidity or transaction/arbitrage costs (Mendenhall (2004), Hou and Moskowitz (2005), Sadka (2006)). Traditional economic models assume that processing public information is costless, so underreaction have to arise from these “adjustment cost” type frictions. However, in this paper, I use corporate bond data to show that underreaction may be better explained by investors facing information-processing constraints.

Using a simple rational inattention framework, I show that if investors face information constraints, then price underreaction is *risk-specific*. In the model, investors cannot immediately digest all information, but they will rationally choose to allocate more information-processing resources, or “attention”, to more payoff-relevant risks.² As a consequence, price underreaction will be smaller for risks that are more relevant for explaining payoff. This is different from illiquidity which generates variation at the security level, rather than the risk level.

I use corporate bond as the main laboratory to test the theory because it provides the necessary risk-level variation. Corporate bond value depends on both interest-rate and firm-specific default risk shocks, and there is significant variation in the payoff-relevance of these risks across bonds. Rational inattention thus predicts that underreaction to credit risk will be lower in bonds with worse credit quality, and underreaction to interest rate movements will be lower for bonds with longer duration.

The data strongly supports the rational inattention view. For each bond, I use the returns of duration-matched Treasury bonds and company stock to proxy for interest rate and firm-specific fundamental risk shocks, respectively. I measure underreaction as the fraction of the price reaction that does not happen immediately. As predicted, underreaction decreases in payoff-relevance. As an example, Figure 1 plots the response of the prices for bonds with

¹Besides the large work on post-earnings announcement drift, prior papers document that stock prices underreact to earnings news (Bernard and Thomas (1990)), new security analyst recommendations (Womack (1996)), revenue news (Jegadeesh and Livnat (2006)), credit rating downgrades (Dichev and Piotroski (2001)), and news about economically related firms (Cohen and Frazzini (2008), Menzly and Ozbas (2010)). The literature outside of stock market is smaller. In corporate bonds, researchers have documented slow reaction to earnings news (Wei, Truong, and Veeraraghavan (2012)), credit default swap movements (Blanco, Brennan, and Marsh (2005)), and equity returns (Gebhardt, Hvidkjaer, and Swaminathan (2005), Hong, Lin, and Wu (2012)). In municipal bonds, Cornaggia, Hund, and Nguyen (2018) show that prices are slow to respond to credit conditions of bond insurers.

²In this paper, “allocating attention” should be interpreted broadly as allocating resources for information-processing. This can include spending time and energy in studying news, hiring personnel with the relevant expertise, or becoming specialized in processing certain types of information.

opposite payoff-relevance profiles. Bonds with lowest quintile interest-rate relevance and highest quintile default risk relevance underreact to interest rate shocks by 75% and default risk shocks by 48% (blue lines). In contrast, bonds with the opposite risk exposure show the opposite pattern of price adjustments, underreacting to interest rate shocks only by 19% but to default risk shocks by 74% (black lines).

The data also supports the notion that investors face information-processing constraints, because there is a “distraction effect”: when one risk is more payoff-relevant or experiences larger shock realizations, consistent with it consuming more investor attention, price reaction to the other risk slows down. This is different from the alternative view that investors simply face attention costs but can flexibly expand capacity when needed, such as by working longer hours or hiring more analysts, in which case the payoff-relevance of one risk should have no effect on how prices respond to the other.

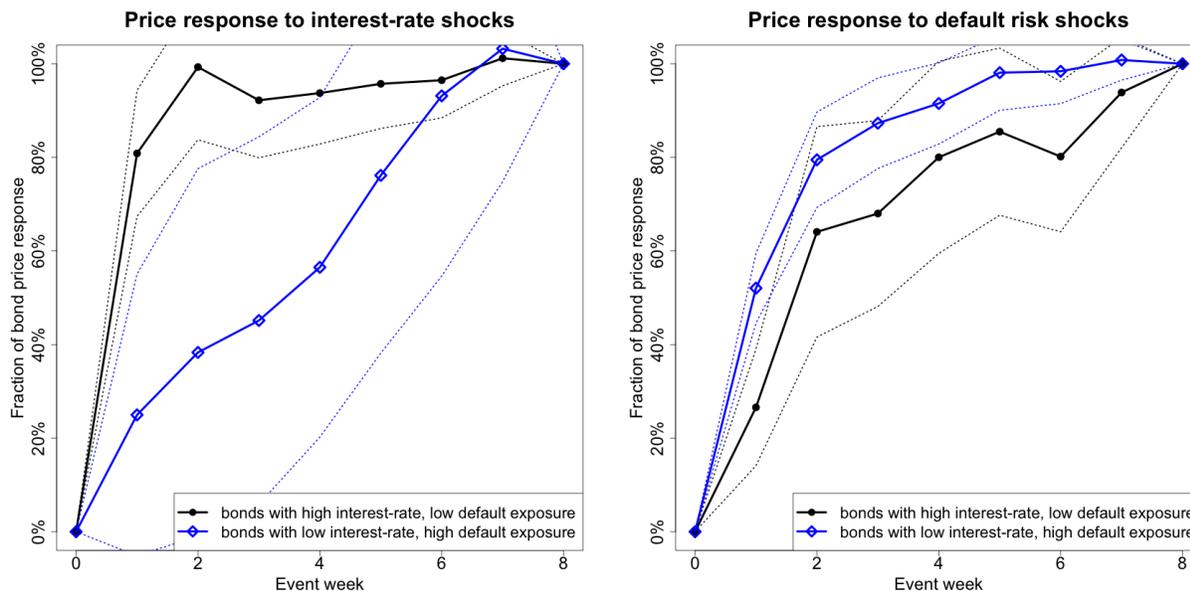


Figure 1. Price response to shocks for bonds with different risk exposures. The figures plot the fraction of price response to shocks in event week 1, assuming that the eight-week price response is 100%. Black lines represent bonds with top quintile interest rate exposure and bottom quintile default risk exposure (typically bonds with long duration and low credit risk). Blue lines represent bonds with the opposite risk exposure profile. Dotted lines are 95% confidence intervals. Interest rate and default risk shocks are proxied using returns of Treasury bonds and company stock. Details of the methodology are explained in section 4.1.

These predictions on risk-level underreaction also hold up when examining within-bond variation. That is, when the credit quality or interest rate sensitivity of a bond changes, the price underreaction patterns also change accordingly, consistent with investors adjusting

resource allocation according to circumstance changes over time. When controlling for payoff-relevance, price underreaction to interest rate shocks is 14–20% lower than to default shocks, consistent with lower costs to process interest rate information than firm fundamentals. As explained earlier, my results cannot be explained by variation in liquidity. If a bond is illiquid, it should be slow to respond to all information, but the degree of slowness cannot differ by risk. In the data, variation in bond-level liquidity proxies also explain variation in underreaction, just not as much as payoff-relevance. In a panel regression to explain the variation of price underreaction, I find that the explanatory power of payoff-relevance proxies is three times that of the liquidity measures. My results also cannot be explained by arbitrage or transaction costs because they predict the opposite of the “distraction effect” in the data.

In the final part on corporate bonds, I quantify the implied information-processing cost by measuring how much “money is left on the table.” A back-of-envelope calculation shows that the average corporate bond mutual fund, with around \$1 billion assets under management, can gain \$123,000 annually if its portfolio manager pays full attention to the shocks studied in this paper. Mutual fund managers typically have low incentives, but even if I assume a hedge-fund like contract in which the manager gets 20% of fund returns, the gain for the manager for paying more attention is only 1% (\$24,600) of her annual compensation. I argue this foregone profit is small enough that it may plausibly reflect fund managers’ information-processing costs.

To alleviate the concern that my findings are specific to corporate bonds, I also show that similar evidence exists in equity market. Specifically, I investigate the speed at which stock prices respond to market-level shocks and industry-level shocks. Similar results bear out: price underreaction is smaller to risks that are more payoff-relevant. Payoff-relevance explains roughly as much variation in underreaction as liquidity proxies. Finally, the amount of money left on the table for an average size mutual fund is roughly 2% of manager’s annual compensation.

The main message of this paper is that information-processing constraints can explain significant variation of slow price response to information. Perhaps surprisingly, this is true even in the corporate bond market which is dominated by institutional investors. A secondary contribution of this paper is providing evidence for the usefulness of rational inattention models in explaining economic behavior (Mackowiak, Matejka, and Wiederholt (2018)).

This paper is most closely related to the burgeoning literature linking investor inattention to price response to information. Making use of distractions or exogenous variation of information availability, a number of papers show that inattention causes price underreac-

tion (Hirshleifer, Lim, and Teoh (2009), Drake, Gee, and Thornock (2016), Fedyk (2018)). DellaVigna and Pollet (2009) argue that lower investor attention slows down price responses to earnings released on Fridays. The main difference of this paper is I focus on testing whether investors purposefully allocate attention according to the rational inattention framework, rather than examining how they respond to exogenous attention shocks.

There are a few recent papers applying rational inattention in other asset pricing settings. Huang, Huang, and Lin (2018) show that when Taiwan investors are distracted by large jackpot lotteries, stock return comovement increases, consistent with investors allocating more attention to market-wide information and less attention to firm-specific information. Peng, Xiong, and Bollerslev (2007) show that stock returns comove more with the market in periods of more macroeconomic uncertainty, and interpret this using endogenous investor attention allocation. Finally, Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016) show that mutual funds make more profits by timing the market in recessions and picking stocks in other periods, which can be interpreted as fund managers allocating their attention to more profitable tasks across the business cycle.

Corporate bond investors studied in this paper are likely intentionally allocating their information-processing resources. Thus, the phenomena documented in this paper is different from papers documenting retail investors' response to attention-grabbing salient information such as news, advertisement, or extreme returns (Barber and Odean (2007), Yuan (2015), Focke, Ruenzi, and Ungeheuer (2018), Kumar, Ruenzi, and Ungeheuer (2019)). The difference between my and their papers is consistent with institutional investors being more sophisticated than retail investors.

More broadly, a few papers also aim to separate inattention versus adjustment cost mechanisms in explaining other slow adjustments by economic agents. In the context of explaining slow household mortgage refinancing behavior, Andersen, Campbell, Nielsen, and Ramadorai (2018) find that inattention is the main friction for households with low socioeconomic status, and adjustment costs are the main friction for middle-aged households with high financial wealth. Mackowiak and Wiederholt (2009) argue that rational inattention, rather than menu costs, can explain why firms are slow to respond to aggregate shocks in price setting.

The remainder of the paper is organized as follows. Using a simple conceptual framework, section 2 derives predictions about price underreaction from rational inattention, and contrast them with that predicted from traditional illiquidity explanations. Section 3 explains how I measure relevant variables. Section 4 presents evidence for the rational inattention-based predictions and considers alternative hypotheses. Section 5 estimates the implied cost of investor inattention. Section 6 shows similar evidence for rational inattention in equities

and section 7 concludes.

2 Conceptual framework

I use a simple framework to show that rational inattention generates risk-specific price underreaction, and this feature distinguishes it from illiquidity or transaction/arbitrage cost explanations. To directly link to the empirical exercise, the below discussion specializes to the context of corporate bonds. The results are derived formally in Appendix A using a rational inattention model adapted from the modeling approach in Van Nieuwerburgh and Veldkamp (2010).

2.1 Rational inattention generates risk-specific underreaction

Changes in the value of a corporate bond can be decompose into two components, one due to aggregate interest-rate shocks and the other due to firm-specific default risk shocks:

$$\Delta V_t = \Delta V_t^{\text{interest}} + \Delta V_t^{\text{default}},$$

where $\Delta X_t \stackrel{\text{define}}{=} X_t - X_{t-1}$ for any variable X . I define $\sigma(\Delta V_t^{\text{interest}})$ and $\sigma(\Delta V_t^{\text{default}})$ as the payoff-relevance of interest-rate and default risks, respectively. Bonds with longer duration will have larger interest-rate relevance, and bonds with more credit risk will have higher default-risk relevance.

Attention determines price underreaction. Investors need to allocate information-processing resources (“attention”) to these two sources of risk. In practice, paying attention to interest rate entails analyzing macroeconomic news, central bank announcements, and also learning from Treasury bond price movements. Paying attention to default risk entails reading news about the company, studying its accounting statements, and also learning from its stock price movements.

I assume investors face an attention constraint and cannot process all information immediately upon release. Therefore, contemporaneous price movements will underreact to these risks:

$$\Delta P_t = (1 - \text{Underreaction}^{\text{interest}}) \cdot \Delta V_t^{\text{interest}} + (1 - \text{Underreaction}^{\text{default}}) \cdot \Delta V_t^{\text{default}} + \epsilon_t$$

I assume that, after a sufficiently long period of time (H periods), price fully incorporates the shocks. The hypothetical price path is illustrated in Figure 2. Then, we can measure

underreaction using:

$$\begin{aligned} \text{Underreaction}_{\text{risk}} &= 1 - \frac{\text{Initial price reaction}}{\text{Full price reaction}} \\ &= 1 - \frac{\text{Cov}(\Delta P_t, \Delta V_t^{\text{risk}})}{\text{Cov}(\Delta P_{t-1 \rightarrow t+H}, \Delta V_t^{\text{risk}})} \end{aligned} \quad (1)$$

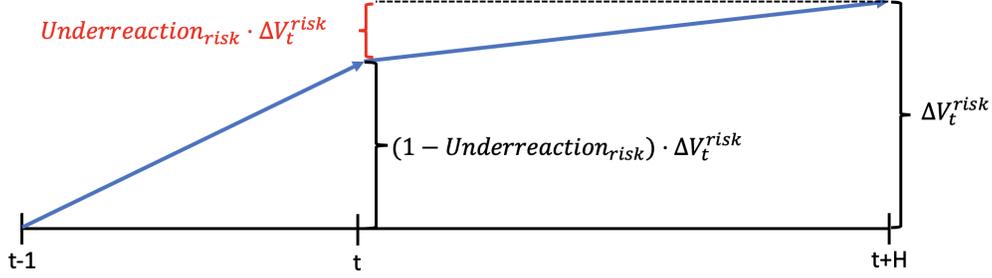


Figure 2. Illustration: how price incorporates information. In response to a value shock V_t^{risk} at time t , price partially reflects it initially, and fully reflects it H periods later.

Note that the definition of underreaction is expressed as a *fraction* of efficient price reaction to information. Intuitively, if all investors pay full attention to a risk, security price will reflect the risk fully and underreaction equals zero. If no investor pays attention, then underreaction equals one.

Prediction from rational inattention. If investors face an attention constraint, then they should equalize marginal benefit of attention across risks. In this framework, by learning about a risk ΔV_t^{risk} , investors can make profits by trading on the predictable drift from time t to $t + H$ ($\text{Drift}^{\text{risk}} = \text{Underreaction}_{\text{risk}} \cdot \Delta V_t^{\text{risk}}$). Assuming investors are risk neutral, and ignoring potential differences in the attention cost between the two risks, then the following first order condition should be satisfied:

$$\begin{aligned} \sigma(\text{Drift}_t^{\text{interest}}) &= \sigma(\text{Drift}_t^{\text{default}}) \\ \Leftrightarrow \text{Underreaction}_{\text{interest}} \cdot \sigma(\Delta V_t^{\text{interest}}) &= \text{Underreaction}_{\text{default}} \cdot \sigma(\Delta V_t^{\text{default}}) \end{aligned} \quad (2)$$

Therefore, the more payoff-relevant a risk is, the smaller its underreaction has to be. Note that this first order condition has incorporated the strategic substitutability of attention allocation across investors. Take any investor's perspective. If other investors pay more attention to interest rate, their trading behavior will lower the underreaction to interest rate, making it less profitable for me to pay attention to it, and that is taken into account in equation (2).

Of course, in reality, the attention cost to interest rate is likely lower than that to the default risk. Not only is the value-dependence of corporate bonds on interest rate more mechanical,³ it is also more systematic. A portfolio manager only needs to read interest-rate relevant news once and then deduct the impact on all bonds, but she has to read default-relevant fundamental news for all companies separately. This implies that, if the two risks have equal payoff-relevance for a bond, we should expect more underreaction to the default component. Putting these together leads to our first testable prediction.

Prediction 1 (“Main effect”): Degree of price underreaction to a risk is decreasing in the value-relevance of that risk ($\sigma(\Delta V_t^{\text{risk}})$). After controlling for payoff relevance, underreaction to risks with higher attention cost (e.g. default risk for corporate bonds) is larger.

This is the main testable prediction of rational inattention. It leads me to choose corporate bonds as the main laboratory, because there is significant variation in interest-rate and default risk payoff-relevance across bonds.

Prediction from attention constraint. Our discussion so far assumes that investors face a fixed attention budget, i.e., a constraint. This can be motivated by each portfolio manager having only 24 hours a day, so the more time she spends learning one type of information, the less she has left for the other. However, it is also reasonable to think of information-processing as a cost, rather than a constraint. That is, investors may be able to expand their information capacity whenever needed, such as hiring more personnel. These two different views lead to a second testable prediction.

Prediction 2 (“Distraction effect”): If investors face attention constraints (i.e. have fixed attention budgets), then the more payoff-relevant is one risk, the higher is the underreaction to the other risk. If investors can flexibly expand capacity when needed, then there will be no such effect.

2.2 Traditional explanations: liquidity and transaction costs

The key prediction under rational inattention is that price underreaction is risk-specific. This is not true in the traditional explanations based on illiquidity or transaction costs. We consider two types of these explanations.

³However, as shown in Appendix D, the corporate bond response to duration-matched Treasury bond returns is not completely mechanical, but depends on the information content. If the Treasury move is a pure discount rate shock, then the duration-matched corporate bonds move almost one-to-one. But if the Treasury movement is associated with market-wide change in credit environment, as proxied by changes in the Investment Grade CDS spread, then the sensitivity is much lower than one-to-one.

Type 1: Infrequent trading Infrequent trading can arise from, for instance, slowness in finding trading counterparties (Duffie, Gârleanu, and Pedersen (2005)) or infrequent portfolio rebalancing by certain investors. This is particularly relevant for the corporate bond market because many bonds trade less than once a day (Edwards, Harris, and Piwowar (2007)). This generates differences in underreaction at the bond-level: bonds that trade less frequently will underreact more to information, but the degree of underreaction should not be different depending on the source of risk.

Type 2: Transaction/arbitrage cost. Transaction cost can arise from order submission costs or bid-ask spreads (Glosten and Milgrom (1985)). To fix ideas, suppose bond investors face bid-ask spread κ , and only trade if and only if value has changed more than κ . This can be all investors, or just arbitrageurs, in which case κ can be interpreted as arbitrage costs.

Suppose price responds to information if and only if these investors trade. Then, the underreaction is a function of size of bond value change $|\Delta V_t|$. In the stark example here, we have Underreaction = $1 - \mathbf{1}_{|\Delta V_t| < \kappa}$.

This is formally shown using a simple model in Appendix A.5. If bid-ask spread κ is time-varying, then the result will not be as stark, but the general insight is robust: underreaction will be a function of realized overall value change ($|\Delta V_t|$). Once that is controlled for, the degree of underreaction does not depend on the risk source.⁴

3 Data and measures

I use Treasury returns and company-level stock returns to proxy for realization of interest-rate and default risks, respectively. This section summarizes the data, and describes how I measure the two key variables: payoff-relevance of risks and price underreaction to risks.

3.1 Data

Proxies for default and interest rate shocks For each corporate bond, I use the weekly returns of duration-matched U.S. Treasury bonds to proxy for interest rate shocks. I recompute the duration match in each week so this adjusts for the time-varying changes in

⁴To make this concrete: suppose a corporate bond's value increased by 100 basis points in a period. Under the transaction cost-based alternative hypothesis, the price reaction should not differ if interest rate movement accounts for 10 basis points and default risk accounts for 90 basis points, or the reverse (interest rate 90, default risk 10). Under the rational inattention hypothesis, if interest-rate is more payoff-relevant for the bond, then investors will pay more attention to incorporate interest-rate shocks, so underreaction will be larger in the first scenario than the second.

bond duration over time. The Treasuries data come from the Center for Research in Security Prices (CRSP) fixed-term Treasury indices.⁵

Similarly, I use company stock returns as proxies for shocks to fundamental risk. According to the Merton model, when the asset value of the company changes, the default probability of debt changes accordingly. The sensitivity of response of bond returns to stock returns is nonlinear and varies across bonds, but I will empirically estimate and adjust for that later. I use the bond-CRSP link file from Wharton Research Data Services (WRDS) to match our data to excess stock returns in CRSP. I winsorize stock returns at the (0.5%, 99.5%) levels to reduce the impact of outliers.

In using these returns as shock proxies, I am relying on Treasuries and stock returns being faster to reflect interest-rate and firm-level fundamental movements. This should not be controversial, and is also justified under the rational inattention framework: investors in Treasuries bonds should pay significantly more attention to interest-rate relevant information because it is the dominant risk. Similarly, investors in stock market should pay much more attention to firm-specific fundamental information because, being lower in the capital structure, stocks are more sensitive to firm fundamentals than corporate bonds. Empirically, Fleming and Remolona (1999) document that U.S. Treasuries react essentially immediately to macroeconomic announcements. For stocks, existing literature generally finds that equity prices reflect information faster than corporate bonds (e.g., Kwan (1996), Gebhardt et al. (2005), and Downing, Underwood, and Xing (2009)), and this is also confirmed in my subsequent analysis.⁶

Corporate bond data I use corporate bond data from enhanced Trade Reporting and Compliance Engine (TRACE) which contains all dealer-intermediated corporate bond transactions in the U.S. from July 2002 to December 2014. Because most corporate bonds do not trade every day, I conduct the analysis at weekly frequency. I first compute daily volume-weighted average prices using all institutional size trades ($\geq \$100,000$) following Bessembinder, Kahle, Maxwell, and Xu (2009), and then use the latest daily price of the week as the weekly price. After applying the code in Dick-Nielsen (2014) to delete wrong

⁵I construct duration-matched Treasury portfolios as follows. For each week t , let $\text{Dur}_{i,t-1}$ denote the modified duration of bond i at the end of the previous week. I first find the two benchmark Treasuries with nearest duration TrsyDur_{t-1}^L and TrsyDur_{t-1}^H such that $\text{TrsyDur}_{t-1}^L \leq \text{Dur}_{i,t-1} \leq \text{TrsyDur}_{t-1}^H$. I then form a Treasury portfolio with weights of $\left(w_{i,t-1}^L = \frac{\text{TrsyDur}_{t-1}^H - \text{Dur}_{i,t-1}}{\text{TrsyDur}_{t-1}^H - \text{TrsyDur}_{t-1}^L}, w_{i,t-1}^H = \frac{\text{Dur}_{i,t-1} - \text{TrsyDur}_{t-1}^L}{\text{TrsyDur}_{t-1}^H - \text{TrsyDur}_{t-1}^L} \right)$ so that the value-weighted duration of the Treasury portfolio matches the bond. I use benchmark Treasury maturities of 1, 2, 5, 7, 10, 20, and 30 years.

⁶Holden, Mao, and Nam (2018) and Addoum and Murfin (2018) find evidence that, in low credit quality firms, bond holders have private information that causes bond movements to sometimes lead equity prices. This is not inconsistent with my exercise which focuses on the incorporation of public information.

entries, I restrict attention to fixed-rate USD-denominated bonds without optionality. I also require bond issuance size to be greater than \$10 million. To reduce impact of outliers, I eliminate cases where bond price is above 150% of par value, and winsorize bond returns at (0.5%, 99.5%) levels.

I take care to not falsely find slow price reaction due to mechanical reasons. First, I only use actual transactions so stale quotes cannot impact my inference. Second, lack of trading does not mechanically translate to slow response. I only compute weekly returns if the bond traded on both the current and the previous weeks, or otherwise the return will be considered missing.⁷ Finally, I merge the data with FISD Mergent to get bond characteristics such as coupon rate, issuance size, and so forth.

After requiring stock and duration-matched Treasury returns, the sample contains 1,432,062 bond-week observations with 13,655 unique bonds issued by 1,736 companies. Summary statistics are in Table 1. Because I only calculate weekly returns when a bond trades in two successive weeks, more liquid bonds are more likely to enter into the sample. The sample captures 82% of all corporate bonds appearing in TRACE by count and 96% when value-weighting. In terms of dollar trading volume, the sample includes 74% of total trading in the U.S. corporate bond market.

Year	Bond-weeks	Bonds	Companies	Credit spread (%)	Duration (years)	Issuance size (\$mil)	Weekly volume (\$mil)	Trades per week
2002	41,094	3,586	733	6.25	4.89	541	30.5	15.4
2003	92,994	4,282	814	5.06	5.25	562	29.2	14.9
2004	96,607	4,286	862	4.83	5.31	586	28.6	12.6
2005	90,847	4,172	880	5.38	5.32	614	26.8	12.5
2006	91,365	4,197	902	6.09	5.32	632	23.7	11.0
2007	82,631	4,054	876	6.16	5.46	700	23.7	10.9
2008	82,119	3,852	805	7.09	5.34	789	22.5	13.4
2009	103,750	4,376	853	6.46	5.44	802	26.0	17.1
2010	124,783	4,802	956	4.44	5.71	816	24.0	15.2
2011	135,926	5,113	1,025	4.17	5.91	804	20.8	14.1
2012	145,270	5,605	1,067	3.40	6.13	811	26.0	16.2
2013	167,568	6,068	1,091	3.37	6.22	812	20.2	14.8
2014	177,108	6,422	1,124	3.32	6.36	840	19.6	13.8

Table 1. Summary statistics of the corporate bond sample. Corporate bond data comes from enhanced TRACE.

⁷I also verify that corporate bond-based repo transactions are not contained in the TRACE dataset. Otherwise, this would impact the subsequent analysis because the reversing repo trade will happen at predetermined prices, and thus mechanically be “slow to respond to information.” See FINRA, “Trade Reporting Frequently Asked Questions” Question 106.6, accessible on the FINRA website.

3.2 Measuring price underreaction

Determining full price response horizon. In my conceptual framework, underreaction is defined as the fraction of price response that does not happen immediately. I follow Hou and Moskowitz (2005) to use long horizon price returns to estimate full price reaction, and use that to estimate underreaction.⁸ To determine how long the horizon needs to be, I run a panel regression of weekly bond returns on current and up to 12 lagged shocks:

$$\begin{aligned} \text{BondRet}_{i,t} = & \alpha + \beta_0^{\text{stock}} \cdot \text{StockRet}_{i,t} + \dots + \beta_{12}^{\text{stock}} \cdot \text{StockRet}_{i,t-12} \\ & + \beta_0^{\text{trsy}} \cdot \text{TrsyRet}_{i,t} + \dots + \beta_{12}^{\text{trsy}} \cdot \text{TrsyRet}_{i,t-12} + \sum_i \gamma_i \cdot \mathbf{1}_{\text{bond } i} + u_{i,t} \end{aligned} \quad (3)$$

where I use bond fixed effects to absorb fixed average return differences across bonds. I cluster standard errors by week and bond.

Figure 3 plots the rolling sums of regression coefficients ($\beta_0^{\text{stock}}, \beta_0^{\text{stock}} + \beta_1^{\text{stock}}, \dots$) to trace out the price response to shocks. There is clear evidence of slow price reaction to both types of shocks. In response to one unit of stock return, the immediate responses is 0.052 in week 1, rising to 0.089 in week 2, and eventually converging to around 0.123. In response to one unit of Treasury return, the immediate price response is 0.469 and the convergent full price response is around 0.65.⁹

After inspecting Figure 3, I choose to use eight weeks to measure full responses, because responses to both shocks appear to have converged by eight weeks.¹⁰

Estimating underreaction using panel regressions. The above immediately suggests that we can estimate underreaction by taking ratios of panel regression coefficients. Specifically, I run regression (3) with seven lags, and calculate underreaction as:

$$\begin{aligned} \text{Underreaction}^{\text{stock}} &= 1 - \frac{\hat{\beta}_0^{\text{stock}}}{\hat{\beta}_0^{\text{stock}} + \hat{\beta}_1^{\text{stock}} + \dots + \hat{\beta}_7^{\text{stock}}}, \\ \text{Underreaction}^{\text{trsy}} &= 1 - \frac{\hat{\beta}_0^{\text{trsy}}}{\hat{\beta}_0^{\text{trsy}} + \hat{\beta}_1^{\text{trsy}} + \dots + \hat{\beta}_7^{\text{trsy}}}, \end{aligned}$$

⁸I do not use structural models to estimate full price reaction because they do not always generate reliable bond price sensitivity to risks. For instance, Schaefer and Strebulaev (2008) find that structural models of corporate bonds are bad at explaining how bond prices respond to interest rate shocks in practice.

⁹The fact that bond prices do not respond one-to-one with duration-matched Treasury returns is consistent with the long-standing finding that conventional bond duration measures overestimate interest rate sensitivity (Longstaff and Schwartz (1995), Schaefer and Strebulaev (2008)).

¹⁰Note that I face a bias-variance trade-off here: using a longer horizon is better for capturing complete responses but increases estimation error.

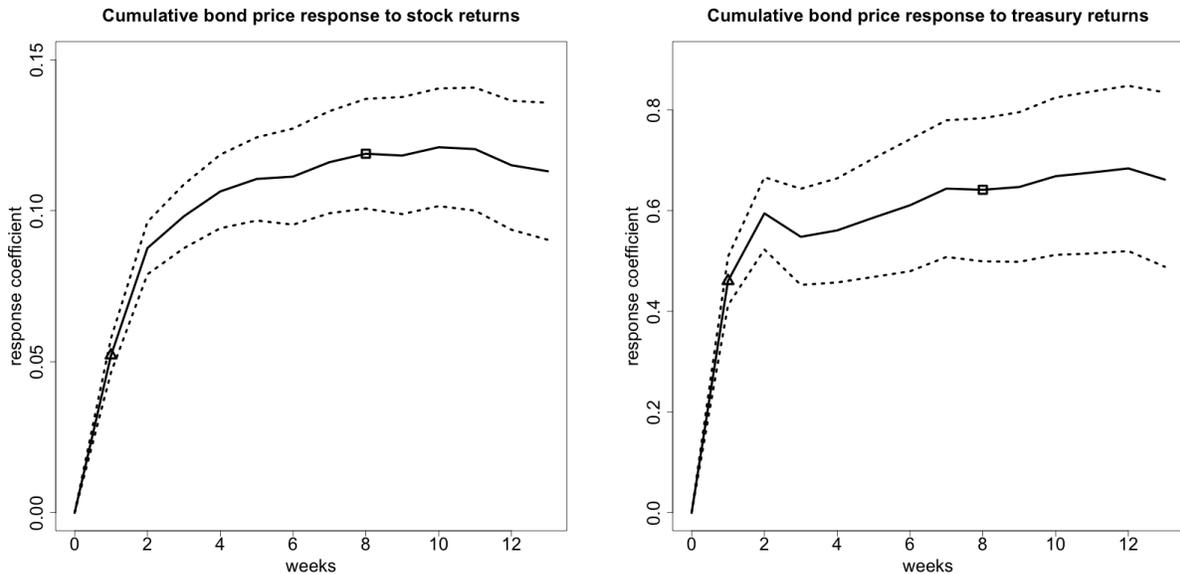


Figure 3. Cumulative response of corporate bond prices to shocks. The immediate response to shocks are marked as triangles. Because these graphs show that eight weeks is approximately enough for shocks to be fully incorporated into prices (marked by squares), I use eight weeks to measure full price responses in subsequent work. The responses are obtained by regressing weekly corporate bond returns on current and lagged stock and Treasury returns (equation (3)), and then adding up cumulative coefficients. For example, the h -week response to stock returns is obtained as $\beta_0^{\text{stock}} + \dots + \beta_{h-1}^{\text{stock}}$. The immediate reaction coefficients $\beta_0^{\text{stock}}, \beta_0^{\text{trsy}}$ are marked using triangles. Dotted lines are two standard error bounds and standard errors are clustered by week and bond.

and I calculate standard errors using the Delta method. For instance, when using the whole sample to estimate these, I get Underreaction^{stock} = 57% and Underreaction^{trsy} = 27%, with standard errors of 2.7% and 7.1%, respectively.

Testing the theory requires estimating underreaction for bonds with different payoff-relevance profiles. This panel regression method is straight-forward, but requires a lot of data, so I will not be able to subdivide the sample into many refined bins. In order to estimate underreaction using less data, I also devise a generalized method of moments (GMM) approach.

Estimating underreaction using generalized method of moments. The GMM method can estimate underreaction for much smaller subsets of data. Specifically, I estimate the fol-

lowing two equations:

$$\begin{aligned} \text{BondRet}_{i,t} &= (1 - \text{Underreaction}_i^{\text{stock}}) \cdot b_i^{\text{stock}} \cdot \text{StockRet}_{i,t} \\ &\quad + (1 - \text{Underreaction}_i^{\text{trsy}}) \cdot b_i^{\text{trsy}} \cdot \text{TrsyRet}_{i,t} + u_{i,t} \end{aligned} \quad (4)$$

$$\text{BondRet}_{i,t \rightarrow t+7} = b_i^{\text{stock}} \cdot \text{StockRet}_{i,t} + b_i^{\text{trsy}} \cdot \text{TrsyRet}_{i,t} + u_{i,t \rightarrow t+7}. \quad (5)$$

The first equation is simply the immediate price reaction to shocks, and the second is the eight-week full price reaction. The underreaction parameters are constrained to be between zero and one. I demean all left- and right-hand-side variables so the intercept term can be omitted. Note that there are four moment conditions for four parameters, the GMM is exactly identified.

This GMM estimation can be done for much smaller bins of data, allowing me to investigate how underreaction varies at a rather refined level. However, if the sample size is too small, or if the relationship between bond return and shocks are too weak ($b_i^{\text{stock}}, b_i^{\text{trsy}}$ close to zero), then the underreaction parameters can also be ill estimated. To screen out bad estimates, I apply two mild data filters. First, I require the estimated full price responses (\hat{b}_i^{stock} or \hat{b}_i^{trsy}) to be positive. Structural models of corporate bonds predict positive relationships, and previous literature predominantly find positive relationships (Collin-Dufresn, Goldstein, and Martin (2001), Schaefer and Strebulaev (2008)), so negative estimates are indications of small sample randomness. Second, I require the standard error of underreaction estimates to be below 1.¹¹

In Appendix B.2, I explain the GMM estimation and data filters in more detail. As a check on estimation precision, I also confirm that, when applied on samples large enough to use the panel-regression based estimates, the estimates from the two methods are similar.

3.3 Measuring payoff-relevance of risks

In my conceptual framework in Section 2, payoff-relevance of a risk is defined by the volatility of bond returns explained by that risk. I directly estimate this.

In the baseline specification in Section 4.1 where I do not exploit time variation in payoff-relevance, I run a time-series regression for each bond i to decompose bond returns into components explained by stock and duration-matched Treasury returns:

¹¹The reader may wonder why standard error can possibly exceed one. This is because, as is explained in more detail in Appendix B.2, I restrict these underreaction coefficients to be between 0 and 1 using a change of variable to $\eta_{i,p}^{\text{risk}}$, defined so that $\text{Underreaction}_{i,p}^{\text{risk}} = 1/(1 + e^{-\eta_{i,p}^{\text{risk}}})$ is a logistic transform. I then estimate standard errors for $\eta_{i,p}^{\text{risk}}$ and derive standard errors for $\text{Underreaction}_{i,p}^{\text{risk}}$ through the delta method, so it sometimes can exceed 1.

$$\begin{aligned} \text{BondRet}_{i,t} = & \alpha_i + \beta_{i,0}^{\text{stock}} \cdot \text{StockRet}_{i,t} + \beta_{i,1}^{\text{stock}} \cdot \text{StockRet}_{i,t-1} + \dots + \beta_{i,8}^{\text{stock}} \cdot \text{StockRet}_{i,t-8} \\ & + \beta_{i,0}^{\text{trsy}} \cdot \text{TrsyRet}_{i,t} + \beta_{i,1}^{\text{trsy}} \cdot \text{TrsyRet}_{i,t-1} + \dots + \beta_{i,8}^{\text{trsy}} \cdot \text{TrsyRet}_{i,t-8} + u_{i,t}. \end{aligned} \quad (6)$$

where I include eight lags, guided by the finding in Section 3.2 that eight lags appear to be enough to capture slow price responses.¹² I then compute the sample volatility of each estimated component and use the results as payoff-relevance proxies:

$$\sigma_i^{\text{stock}} = \hat{\sigma} \left(\hat{\beta}_{i,0}^{\text{stock}} \cdot \text{StockRet}_{i,t} + \dots + \hat{\beta}_{i,8}^{\text{stock}} \cdot \text{StockRet}_{i,t-8} \right) \quad (7)$$

$$\sigma_i^{\text{trsy}} = \hat{\sigma} \left(\hat{\beta}_{i,0}^{\text{trsy}} \cdot \text{TrsyRet}_{i,t} + \dots + \hat{\beta}_{i,8}^{\text{trsy}} \cdot \text{TrsyRet}_{i,t-8} \right) \quad (8)$$

where coefficients $\{\hat{\beta}_{i,0}^{\text{stock}}, \dots, \hat{\beta}_{i,8}^{\text{stock}}, \hat{\beta}_{i,0}^{\text{trsy}}, \dots, \hat{\beta}_{i,8}^{\text{trsy}}\}$ are the point estimates in regression (6). I then use σ_i^{stock} and σ_i^{trsy} as proxies for payoff-relevance of default and interest rate risks for bond i , respectively. In some subsequent tests where I exploit how payoff-relevance changes over time, I perform the above procedure using subsamples instead of the full sample.

In Appendix B, I verify that these payoff-relevance measures are indeed highly correlated with other commonly used payoff-relevance proxies, such as bond duration for interest rate exposure and credit spreads for default risk exposure.

I choose to use these estimated payoff-relevance measures because they are most closely aligned to the conceptual framework. Another major benefit of these two proxies is that they can be compared to each other quantitatively. For instance, 10% σ^{trsy} and 10% σ^{stock} is comparable in “payoff-relevance space”, while 5 year duration and 150 basis points credit spreads cannot be compared meaningfully. However, in the main subsequent exercises, I also verify that using alternative payoff-relevance proxies do not qualitatively alter my conclusions.

4 Testing theory predictions

I now test the two predictions of rational inattention derived in Section 2. As a reminder, they state that if one risk is more payoff-relevant, then the bond will:

1. underreact less to that risk, and (“**main effect**”)
2. underreact more to the other risk (“**distraction effect**”)

In Section 4.1, I first use panel-regression based estimates of underreaction to present

¹²The results are not sensitive to using 2, 4, or 12 lags instead of 8.

suggestive evidence at the portfolio level. In section 4.2, I examine the GMM-based underreaction estimates at the bond level, which allows me to control for bond-level liquidity proxies.

4.1 Suggestive evidence at the portfolio level

I first measure payoff-relevance ($\sigma_i^{\text{stock}}, \sigma_i^{\text{trsy}}$) for each bond, and then double sort bonds into 5×5 portfolios using common break points in σ_i^{stock} and σ_i^{trsy} .¹³ I then estimate underreaction for each of the 25 portfolios and report results in Table 2.

Panel A. Underreaction to stock returns							
σ^{stock}	σ^{trsy} quintile						
quintile	Low	2	3	4	High	High minus low	t-stat
Low	70%	71%	68%	66%	74%	4%	(0.48)
2	56%	68%	68%	66%	65%	8%	(0.88)
3	57%	60%	63%	65%	68%	11%	(1.34)
4	58%	61%	64%	60%	64%	6%	(0.81)
High	47%	53%	49%	55%	56%	8%	(1.50)
High minus low	-22%***	-19%***	-19%**	-11%	-18%**		
t-stat	(-3.00)	(-2.62)	(-2.34)	(-1.39)	(-2.51)		

Panel B. Underreaction to Treasury returns							
σ^{stock}	σ^{trsy} quintile						
quintile	Low	2	3	4	High	High minus low	t-stat
Low	35%	31%	27%	24%	18%	-16%	(-1.56)
2	56%	42%	29%	28%	18%	-38%***	(-3.25)
3	58%	38%	35%	31%	24%	-34%**	(-2.57)
4	52%	44%	37%	29%	22%	-30%*	(-1.78)
High	76%	70%	42%	34%	20%	-55%***	(-3.64)
High minus low	41%***	39%***	15%	10%	2%		
t-stat	(3.02)	(4.13)	(1.32)	(0.99)	(0.15)		

Table 2. Price underreaction of corporate bonds with different risk exposures. Bonds are double sorted into 5×5 portfolios using the payoff-relevance proxies σ^{stock} and σ^{trsy} . For each of the 25 portfolios, I estimate the bond price underreaction to stock returns (Panel A) and Treasury returns (Panel B) using panel regressions. Price underreaction is defined as the fraction of price response that happens in the delay window, i.e., from week $t + 1$ to $t + 7$, in response to information in week t . Methodological details is explained in Section 3.2. Stars *, **, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively.

There is strong evidence for the main effect. In Panel A of Table 2, for each σ^{trsy} quintile,

¹³That is, I find the quintile breakpoints of $\{\sigma_i^{\text{stock}}, \sigma_i^{\text{trsy}}\}$ and apply them in sorting both payoff-relevance measures.

bonds with higher σ^{stock} have monotonically decreasing price underreaction to stock returns. The differences between the top and bottom σ^{stock} quintiles range from 11% to 22% and are statistically significant at the 5% level in four out of five cases. Similarly, in Panel B, for each σ^{stock} quintile, bonds with higher σ^{trsy} are faster to respond to Treasury returns. The difference between the top and bottom σ^{trsy} quintiles range from 16% to 55%, and the difference is statistically significant in four out of five cases.

There is also evidence for the distraction effect. For each σ^{stock} quintile in Panel A, higher σ^{trsy} is associated with higher underreaction to stock returns. Although the differences (“high minus low” column) are not statistically significant, they all point in the predicted direction. Panel B shows that bonds with higher σ^{stock} have larger underreaction to Treasury returns, and the effect is statistically significant for the bottom two quintiles of σ^{trsy} .

It should also be pointed out that, except for some extreme cases, underreaction to Treasury returns are smaller than underreaction to stock returns. This is consistent with lower cost in processing interest rate-related information. It is worth repeating the key identification strategy in this paper: because we exploit variation in the risk-level payoff-relevance, bond-level liquidity differences cannot explain our findings.

However, this crude portfolio-level test cannot control for liquidity effects. Therefore I now switch to GMM-based bond-level estimates.

4.2 GMM-based test at the bond level

To accommodate the fact that bond credit quality and bond duration changes over time, I now separate the data into three roughly equal-length periods $p \in \{2002 - 2006, 2007 - 2010, 2011 - 2014\}$, and estimate both payoff-relevance and underreaction at the bond-period level. I require each bond-period to have at least 26 weeks of observations, and also apply the two estimation quality filters described in Section 3.2. After applying these filters, I am left with 7,480 bonds with $\text{Underreaction}_{i,p}^{\text{stock}}$ estimates and 9,014 bonds with valid $\text{Underreaction}_{i,p}^{\text{trsy}}$ estimates.

I now pool together underreaction estimates to both risks and test model predictions using a panel regression:

$$\begin{aligned} \text{Underreaction}_{i,p}^{\text{risk}} = & \underbrace{\sum_{q=1}^5 \gamma_q^{\text{risk}} \cdot \mathbf{1}_{\sigma_{i,p}^{\text{risk}} \text{ in quintile } q}}_{\text{“main effect”}} + \underbrace{\sum_{q=1}^5 \gamma_q^{\text{other risk}} \cdot \mathbf{1}_{\sigma_{i,p}^{\text{other risk}} \text{ in quintile } q}}_{\text{“distraction effect”}} \\ & + \underbrace{\beta \cdot \mathbf{1}_{\text{risk}=\text{Treasury}}}_{\text{attention cost difference}} + \text{LiquidityControls}_{i,p} + \sum_{p=1}^3 \delta_p \cdot \mathbf{1}_{\text{period } p} + \epsilon_{i,p}^{\text{risk}} \end{aligned} \quad (9)$$

where when “risk” is stock, the “other risk” refers to Treasury, and vice versa. Because the relationship can be nonlinear, similar to the portfolio level test, I sort payoff-relevance measures into quintiles and regress on the quintile indicators. To absorb the impact of liquidity, I add a number of bond liquidity controls: average bid-ask spread, log weekly trading volume, log weekly trading frequency, and log offering size. Standard errors are double clustered by bond and period.

The results are shown in Table 3, where column 1 is the main specification and columns 2 to 4 are robustness checks. We are interested in the three sets of coefficients marked with underbraces in equation (9), each capturing one aspect that the rational inattention theory has prediction over.

There are four take-aways. First, consistent with the portfolio level results, there is clear evidence for the main and distraction effects. Underreaction to a risk decreases by 31.89% when going from the bottom to top quintile of σ^{risk} (main effect), and increases by 30.75% from the bottom to top quintile of $\sigma^{\text{other risk}}$ (distraction effect). The size of these two effects are similar, consistent with most investors facing binding attention constraints.

Second, the evidence is also consistent with investors changing attention allocation over time in response to bond risk exposure changes. Column 5 adds bond fixed effects to exploit cases where a bond changes payoff-relevance quintile assignment across periods.¹⁴ In total, there are 1,231 cases of bonds changing σ^{stock} quintile assignment and 1,142 cases for σ^{trsy} .¹⁵ Both main and distraction effects are robust to adding bond fixed effects.

Third, across all specifications, I find that underreaction to Treasury is lower than underreaction to stock by 14 - 20%. This is consistent with lower attention cost in learning about interest rate risk.

Finally, while bond liquidity proxies do have explanatory power, the risk-level variables explain price underreaction much better. Although the R^2 measures are significantly downward biased because of estimation errors in the left-hand-side variables, we can still compare them R^2 across explanatory variables. The four bond level liquidity proxies jointly explain marginal R^2 of 1.98%, while the payoff-relevance indicator variables $\mathbf{1}_{\sigma^{\text{risk}}}$, $\mathbf{1}_{\sigma^{\text{other risk}}}$ explain 6.47%, which is more than 3 times as much. If I add the $\mathbf{1}_{\text{risk} = \text{Treasury}}$ indicator, the total explanatory power of risk-level variables increases to 15.31%.

¹⁴The drop in R^2 for this specification is because I use marginal R^2 which adjusts for the number of regressors. There are many bond dummies being added.

¹⁵Table ?? in Appendix ?? shows the assignment migration patterns. As expected, most of these cases see decreasing σ^{trsy} as bond duration decreases over time. There is no clear upward or downward trend for σ^{stock} changes.

Corporate bond price underreaction (percent)					
	Main	Without	Without	Value	Within
	specification	σ^{risk}	$\sigma^{\text{other risk}}$	weighted	bond
	(1)	(2)	(3)	(4)	(5)
$\mathbf{1}_{\sigma^{\text{risk}}}$ in quintile 2	-11.35*** (1.15)		7.15*** (0.37)	-13.56*** (1.61)	-13.21*** (1.61)
$\mathbf{1}_{\sigma^{\text{risk}}}$ in quintile 3	-16.91*** (1.45)		10.24*** (2.51)	-19.38*** (0.94)	-19.15*** (3.91)
$\mathbf{1}_{\sigma^{\text{risk}}}$ in quintile 4	-20.36*** (2.06)		12.79*** (1.51)	-21.92*** (1.50)	-21.47*** (2.61)
$\mathbf{1}_{\sigma^{\text{risk}}}$ in quintile 5	-31.89*** (3.80)		20.28*** (3.51)	-34.28*** (1.93)	-34.28*** (6.17)

$\mathbf{1}_{\sigma^{\text{other risk}}}$ in quintile 2	12.20*** (1.68)	-7.35*** (2.82)		12.13*** (1.00)	13.26*** (2.69)
$\mathbf{1}_{\sigma^{\text{other risk}}}$ in quintile 3	17.34*** (3.78)	-10.53*** (2.23)		17.46*** (2.99)	17.08*** (5.19)
$\mathbf{1}_{\sigma^{\text{other risk}}}$ in quintile 4	21.72*** (2.90)	-12.00*** (2.52)		21.38*** (1.21)	22.81*** (5.07)
$\mathbf{1}_{\sigma^{\text{other risk}}}$ in quintile 5	30.75*** (4.42)	-21.12*** (4.11)		30.57*** (1.74)	31.44*** (5.55)

$\mathbf{1}_{\text{risk} = \text{treasury}}$	-14.81*** (3.41)	-19.61*** (3.82)	-19.83*** (3.67)	-14.48*** (2.92)	-15.39*** (4.60)

Liquidity controls	Y	Y	Y	Y	Y
Bond fixed effect	N	N	N	N	Y
Period fixed effect	Y	Y	Y	Y	Y

Obs	16,494	16,494	16,494	16,494	16,494
Adjusted R^2	19.52%	15.08%	15.25%	22.84%	9.60%

Table 3. Explaining stock price underreaction using payoff-relevance of risks. Price underreaction to stock and Treasury returns are estimated for each bond in each period using GMM. The regression pools underreaction to both risks together. In the variables, when “risk” refers to stock, “other risk” refers to Treasury, and vice versa. The first eight regressors are payoff-relevance measures for “risk” and “other risk,” turned into quintile indicator variables to capture possible nonlinearity. The next regressor captures the average difference between underreaction to the two types of risk. The regression specification in column (4) weights data by bond issuance sizes and specification (5) includes bond fixed effects. Standard errors are clustered by bond and period. Coefficients significant at the 10%, 5%, and 1% level are noted with *, **, and ***, respectively.

4.3 Other alternative hypotheses

Transaction/arbitrage costs. Transaction/arbitrage costs, outlined in section 2.2, cannot explain my findings. This is because this mechanism generates the *opposite* of the distraction effect.

To fix ideas, suppose that investors face bid-ask spreads, and they only trade when the bond value has changed more than the bid-ask spread. Because trading activity speeds up

price responses, this can qualitatively generate our “main effect”: more payoff-relevant risks tend to have larger shocks, so the average price response would be faster.

However, transaction costs generates the opposite of the “distraction effect”. Intuitively, if default risk receives a larger shock, this increases trading and thus speeds up the incorporation of the contemporaneous interest rate shock. This is made precise using a simple model in Appendix A.5. In the inattention explanation, the opposite is true: larger default risk shocks consume investor attention, slowing down price response to interest rates.

In all earlier tests where I sort on *ex-ante* payoff-relevance of bonds, there is clear evidence for the distraction effect. This is related to, but different from, the current discussion because the transaction cost explanation prediction is based on *ex-post* realized shock sizes. However, in Appendix C.1, I find that the distraction effect is also robust when sorting on realized shock sizes. Therefore, I conclude that transaction or arbitrage costs cannot explain my findings.

Time-varying risk premium. One may wonder if the documented price underreaction patterns can be explained by frictionless asset pricing models with time-varying risk premium.¹⁶ However, Appendix C.2 shows that a simple strategy that trades on slow corporate bond price response to stock returns generates an annual Sharpe ratio of 2.61 before transaction costs. Using the logic in Hansen and Jagannathan (1991), this implies an extremely volatile stochastic discount factor. I view this as strong evidence against frictionless equilibrium pricing explanations.

5 Quantifying the implied cost of attention

The existence of price underreaction implies that investors are leaving money on the table. If investors are rationally inattentive, then the money of foregone profits equals the cost of attention.

In section 5.1, I show that the degree of return predictability is not large enough for arbitrageurs to exploit them profitably after transaction costs. More precisely, if a market participant has to initiate new trades to take advantage of the documented underreaction, it is not profitable. However, regular market participants that already trade for other reasons can still make more money by better timing their existing trades better. Therefore, in section 5.2, I take the perspective of other regular market participants and quantify their foregone

¹⁶For instance, take the fact that corporate bond prices tend to continue drifting up (or down) after periods of high (or low) stock returns. This can potentially be explained if corporate bond risk premium increases (decreases) after good (bad) news.

profits due to inattention.

5.1 Out-of-sample return predictability

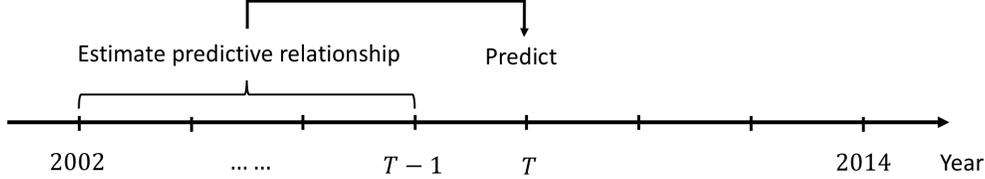


Figure 4. Illustration of the walk-forward return prediction method.

To gauge how much return predictability can actually be captured in practice, I examine a strategy that trades on out-of-sample return predictions. As illustrated in Figure 4, in each year $T = 2004, \dots, 2014$, I use data up to year $T - 1$ to estimate the predictive relationship between lagged shocks and bond returns:¹⁷

$$\text{BondRet}_{i,t} = \alpha^{\text{stock}} + \beta_1^{\text{stock}} \cdot \text{StockRet}_{i,t-1} + \dots + \beta_8^{\text{stock}} \cdot \text{StockRet}_{i,t-8} + \epsilon_{i,t} \quad (10)$$

$$\text{BondRet}_{i,t} = \alpha^{\text{trsy}} + \beta_1^{\text{trsy}} \cdot \text{TrsyRet}_{i,t-1} + \dots + \beta_8^{\text{trsy}} \cdot \text{TrsyRet}_{i,t-8} + \epsilon_{i,t}. \quad (11)$$

Because bonds with different risk exposures can depend on lagged shocks in different ways, I estimate these regressions separately for each of the 5×5 portfolios sorted by σ^{stock} and σ^{trsy} . In addition to estimating using ordinary least squares, I also do an estimation using elastic net regularization where the penalty parameter is chosen using 10-fold cross-validation (Zou and Hastie (2005)). Elastic net is a widely used method in machine learning, and the cross-validation reduces the risk of over-fitting.¹⁸ I then apply the estimated relationships to predict returns in year T .

To quantify the profits from trading on these return predictions, I compute pre-transaction cost returns to a strategy that goes long the top $\frac{x}{2}\%$ predicted returns and short the bottom $\frac{x}{2}\%$ in each of the 5×5 portfolios. When $x = 100$, this strategy trades on all predictions, and when x is small, this strategy only trades on extreme predictions. The portfolio weights are adjusted so that the gross exposure in each calendar week is always 100%.¹⁹

¹⁷Recall that $\text{BondRet}_{i,t}$ is measured using value-weighted prices from the last trading day in week $t - 1$ and week t for corporate bond i . As long as the data used in computing the week $t - 1$ price includes trades that happen before the stock and Treasury bond market closes in that week, this generates mechanical look-ahead bias due to overlap in the measurement period between $\text{BondRet}_{i,t}$ and $(\text{StockRet}_{i,t-1}, \text{TrsyRet}_{i,t-1})$. To fix this, I adjust the time interval of these two regressors. For instance, if bond i last traded on Thursday in week $t - 1$, then I modify $\text{StockRet}_{i,t-1}$ and $\text{TrsyRet}_{i,t-1}$ to end in Wednesday, and so on.

¹⁸Elastic net is a hybrid case between LASSO and ridge regression. I have tried both LASSO and ridge and the results are very similar.

¹⁹For example, if there are N longs and N shorts, then weight $\frac{1}{2N}$ is applied to each position.

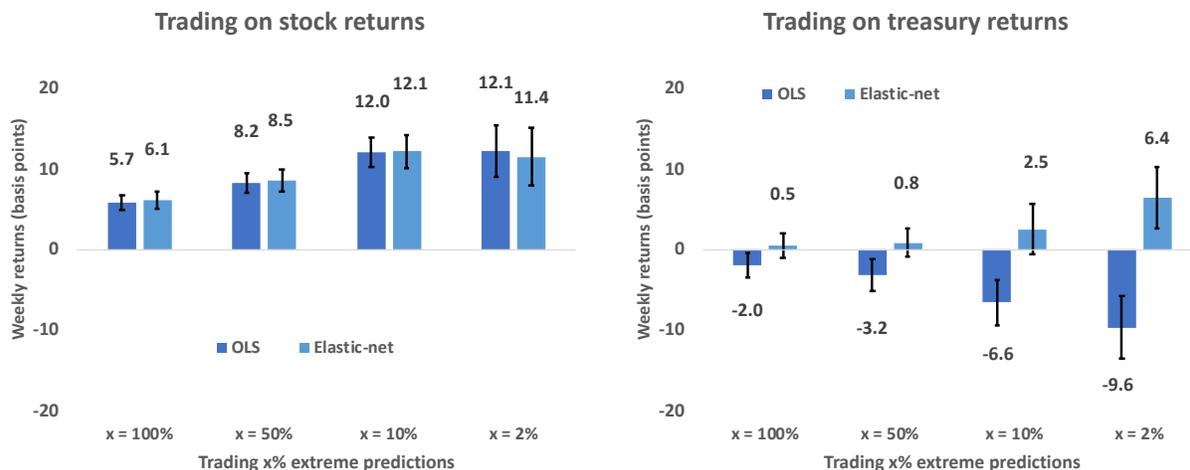


Figure 5. Weekly pre-transaction cost profits when trading on lagged shocks. I construct weekly bond return forecasts using lagged stock returns (left panel) or lagged Treasury returns (right panel). The predictive relationships are estimated using past data through either ordinary least squares or elastic net regularization. I then compute returns to hypothetical trading strategies that go long top $\frac{x}{2}\%$ predicted bond returns and short bottom $\frac{x}{2}\%$. The bar charts plot average weekly returns to these strategies along with standard errors.

Figure 5 plots the profits to these strategies. First, there is only out-of-sample return predictability when using lagged stock information. Despite some in-sample underreaction to Treasuries, there is little out-of-sample predictability, and further analysis reveal that this is because the forecasting relationship changes significantly over time. In other words, bond prices do underreact to Treasury information, but the degree of underreaction varies so much over time such that, if one has to estimate the relationship in real time, lagged Treasury returns cease to be useful.²⁰ The lack of out-of-sample predictability using Treasury information is consistent with lower costs in learning interest rate news.

This predictability is not enough for an arbitrageur to profit after transaction cost. Even when only trading on $x = 2\%$ of the stock return-based predictions, the strategy only generates 11-12 basis points weekly. In contrast, following Hong and Warga (2000), I estimate that the corporate bond average bid-ask spread in my sample to be 41 basis points.²¹ In the corporate bond market, transaction cost is lower for larger trades (Edwards et al. (2007)),

²⁰Welch and Goyal (2007) have a similar insight when examining aggregate equity return predictors. When the forecasting relationship is very unstable, it is no longer useful.

²¹The TRACE dataset reports whether a trade is a customer selling to a dealer at the bid price, a dealer selling to a customer at the ask price, or an interdealer trade. For each bond-day, I calculate trade-weighted bid and ask prices, and take their differences to get a bid-ask spread estimate when both sides are available. This is a crude estimator but Schestag, Schuster, and Uhrig-Homburg (2016) show that it is comparable to more sophisticated measures of transaction costs.

but even if I restrict to trades above \$5 million, the average bid-ask spread is still 21 basis points. Trades this large happens less than 1 out of 6 trading weeks for the median bond in my sample. Considering all the other over-the-counter trading frictions in the corporate bond market, I conclude that the return predictability is not large enough to overcome transaction costs.²² Thus, arbitrageurs are not leaving money on the table.

5.2 Quantifying the dollar value of attention

The previous section's results imply that arbitrageurs, such as hedge funds, are not the marginal players determining the documented underreactions. They are not leaving any money on the table.

Therefore, the marginal players are regular market participants such as mutual funds and pension funds that trade regularly as a part of their business. If they commit the necessary resources to estimate these predictable price drifts, they can improve their investment performance by delaying certain existing trades on which they are already paying transaction costs. Note that this approach to quantify foregone profits is different from the typical arbitrageur perspective in most papers documenting return predictability. Arbitrageurs can initiate many new positions and profit on all instances of return predictability. In my case, profits are limited by the amount of trading that market participants already do.

The way market participants can profit on predictable drifts is illustrated in Figure 6. Suppose a fund initially intends to buy a bond in week t . If the fund pays attention, it can predict bond return from week t to $t + 1$. If the predicted return is negative, which happens roughly half of the time, then the fund can delay the trade by a week and buy at a lower price.

Forgone profits for a fund. I now estimate foregone profits for an average corporate bond mutual fund. I obtain CRSP corporate bond mutual fund data over the same sample period of 2002 - 2014 and restrict attention to funds with assets under management above \$10 million. The average assets under management is \$1,002 million. According to Cici and Gibson (2012), the reported turnover in CRSP is significantly upward biased due to mistaking maturing bonds as trading, so I choose to use the annual turnover estimate of 43.2% in their paper. Because trading on lagged Treasury returns is not profitable, I suppose the fund trades only on lagged stock information, which gives an average predictable return of 5.7 basis points per trade (Figure 5). Therefore, a back-of-envelope estimate of the annual

²²The most significant friction is, perhaps, the fact that trading is non-anonymous in corporate bonds. If a hedge fund requests a quote for the trade of this size, the dealer may suspect that the hedge fund is informed and thus widen the bid-ask spread of the quote.

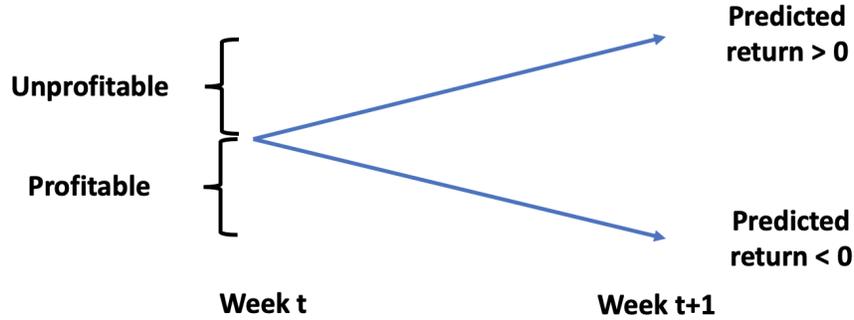


Figure 6. Illustration: profit by timing trades. Consider a fund that plans to, for exogenous reasons, buy a bond in week t . By paying attention past shocks, the fund can form predictions over future bond returns. The fund can profitably delay the trade if and only if predicted return is negative. The case for a fund planning to sell is symmetric.

foregone profit due to inattention is:

$$\underbrace{\$1,002 \text{ million AUM} \times 43.2\% \text{ Turnover}}_{\text{Annual trading volume}} \times \underbrace{\frac{5.7 \text{ bp return predictability}}{2}}_{\text{Profits from better timing trades}} \approx \$123,000$$

This is, admittedly, a crude estimate. There are a number of factors that can increase or decrease this number. On the one hand, I am conservative to assume that managers can only profit in half of the cases.²³ On the other hand, I am lenient to assume that it is possible to delay trades at will. Many corporate bonds trade infrequently, so the manager may be concerned that once she passes up a trading opportunity, she may not be able to trade the same bond again later.

Forgone profits for the portfolio manager. The above estimate is the foregone profit for the fund, but the relevant decision makers are the portfolio managers who bear the attention costs. I thus estimate how much profit accrues to managers. Because I lack data on the strength of manager incentive, I estimate an upper bound by taking the median hedge fund contract where managers get 20% of additional fund returns (Ackermann, McEnally, and Ravenscraft (1999)). With this assumption, I estimate profits to managers for paying more attention to be $\$123,000 \times 20\% \approx \$24,600$. This is only approximately 1% of the

²³Suppose there is usually a gap between the time when portfolio managers want to execute a trade to the time when the trade actually get executed. Then it is possible that managers can choose to “speed up” trades as well, thus taking advantage of the other half of the cases.

annual compensation of an average mutual fund manager (\$2.7 million).²⁴ Because mutual fund manager incentive is likely to be much weaker than that in hedge funds, this is already an overestimation.

In summary, I conclude that inattention implies sizable cost at the fund level. However, because fund managers only get a small share of the profit, the existing level of attention observed in equilibrium prices may be already optimal choices for fund managers: it is not worth their effort to pay more attention.

6 Rational inattention in U.S. stock market

To address the concern that my earlier findings may be specific to corporate bonds or the 2002 - 2014 sample, I conduct the same tests using U.S. stock data from 1931 to 2018. I find similar evidence for rational inattention.

6.1 Data and measures

I use daily returns from CRSP from 1931 to 2018. To avoid very illiquid stocks, I restrict attention to stocks with price no lower than \$5 and market capitalization above \$1 million.

Shock proxies. I examine underreaction of stock prices to market-level and industry-level shocks. The former (MarketRet_t) is proxied using returns of the CRSP total market index. The latter ($\text{IndustryRet}_{i,t}$) is proxied using returns of the Fama-French 49 industry returns, but with each stock’s own influence taken out.²⁵ In addition, because the two shock series are highly correlated, I use the residuals of the industry return series after regressing on market returns by decade. My results are not qualitatively impacted if I use the Fama-French 12, 17, or 30 industries instead.

²⁴To estimate mutual fund manager compensation, I use Table IA.VIII in the online Appendix of Ma, Tang, and Gómez (2018) which estimates that the average advisory fee per fund is around \$8 million. Using corporate bond fund data from Morningstar, I find that funds with above \$10 million in assets under management (AUM) on average have 2.96 portfolio managers. Thus I estimate that an average mutual fund manager makes \$8 million/2.96 \approx \$2.7 million annually.

²⁵Concretely, if a stock i is $w_{i,t}$ fraction of the industry it is in by market capitalization in day t , and let $\widehat{\text{IndustryRet}}_{i,t}$ be the Fama-French industry return. Then, I compute

$$X_{i,t} = \frac{\widehat{\text{IndustryRet}}_{i,t} - w_{i,t} \cdot \text{StockRet}_{i,t}}{1 - w_{i,t}}$$

where $\text{StockRet}_{i,t}$ is the return of stock i in day t . Then, for every decade, I regress $X_{i,t}$ on MarketRet_t and use the residual as $\widehat{\text{IndustryRet}}_{i,t}$.

Determining full price response horizon. As in Section 3.2, I use a full-sample panel regression to determine the appropriate horizon to measure full price response:

$$\text{StockRet}_{i,t} = b_0^{\text{market}} \cdot \text{MarketRet}_t + \dots + b_{14}^{\text{market}} \cdot \text{MarketRet}_{t-14} + b_0^{\text{industry}} \cdot \text{IndustryRet}_{i,t} + \dots + b_{14}^{\text{industry}} \cdot \text{IndustryRet}_{i,t-14} + \sum_i \gamma_i \mathbf{1}_{\text{stock } i} + \epsilon_{i,t}. \quad (12)$$

As shown in Figure 7, 15 days appear sufficient for most of the adjustment process to converge, so I use 15 days to measure full price response subsequently.

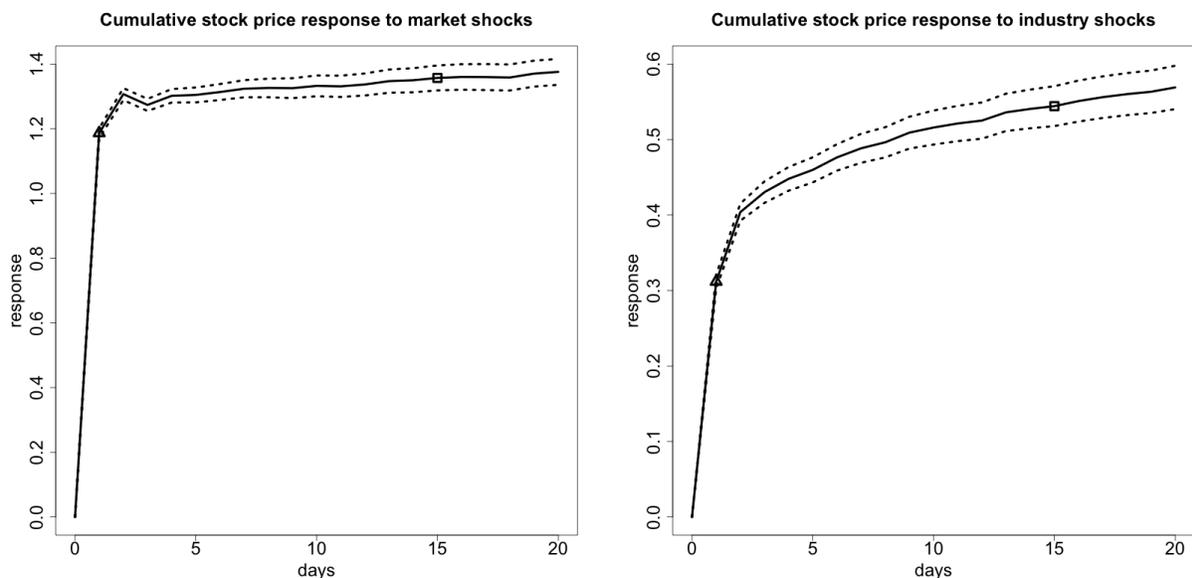


Figure 7. Cumulative response of stock prices to shocks. The responses are obtained by regressing daily stock returns on current and lagged market returns and industry returns (equation (12)). The immediate price response is marked by a triangle. These graphs reveal that 15 days (marked by square) appear roughly sufficient for most of the price response to converge.

Payoff-relevance. For each stock in each decade, I use a time series regression of its daily return on 15 lags of market and industry shocks. This decomposes its return into these two components and a residual, and I compute the volatility of the first two components. I use those volatilities, σ_i^{market} and $\sigma_i^{\text{industry}}$ as payoff-relevance proxies.

6.2 Portfolio-level evidence

Like in section 4.1, I double sort stocks by their payoff-relevance proxies into 5×5 portfolios. I then estimate underreaction using panel regressions in each portfolio. The results are shown in Table 4.

There is clear evidence for the “main effect”. Stocks in the lowest quintile of market-risk relevance underreact to market-level shocks by 49 - 57%, and that decreases monotonically to a range of 7 - 30% for the highest quintile. Similarly, stocks in the lowest quintile of industry risk exposure underreact to industry-level shocks by 65 - 81%, and that decreases monotonically to 11 - 31% for the highest quintile. The “distraction effect” evidence is unclear, but later GMM-based exercise suggests that there is such an effect, but contaminated by not properly controlling for liquidity in this portfolio-level exercise.

Underreaction to market-level shocks								
Market risk relevance (σ^{market}) quintile	Industry risk relevance (σ^{industry}) quintile						High-low	t-stat
	Low	2	3	4	High			
Low	55%	57%	55%	54%	49%	-6%	(-0.84)	
2	35%	43%	49%	49%	44%	10%***	(3.56)	
3	19%	28%	37%	41%	44%	25%***	(14.06)	
4	13%	21%	27%	31%	37%	24%***	(15.28)	
High	7%	12%	16%	22%	30%	23%***	(2.97)	
High-low	-47%***	-44%***	-39%***	-32%***	-19%***			
t-stat	(-6.17)	(-17.78)	(-21.37)	(-13.15)	(-2.85)			

Underreaction to industry-level shocks								
Market risk relevance (σ^{market}) quintile	Industry risk relevance (σ^{industry}) quintile						High-low	t-stat
	Low	2	3	4	High			
Low	81%	74%	64%	60%	11%	-70%***	(-16.04)	
2	66%	62%	57%	50%	19%	-47%***	(-13.27)	
3	65%	50%	47%	47%	28%	-37%***	(-7.25)	
4	70%	47%	44%	44%	34%	-36%***	(-7.63)	
High	69%	59%	44%	36%	31%	-37%***	(-2.71)	
High-low	-12%	-15%***	-20%***	-24%***	20%***			
t-stat	(-0.89)	(-2.99)	(-5.47)	(-4.45)	(4.91)			

Table 4. Price underreaction of stocks with different risk exposures. Stocks are double sorted into 5×5 portfolios using the payoff-relevance proxies σ^{market} and σ^{industry} . For each of the 25 portfolios, I estimate the price underreaction to market returns (Panel A) and industry returns (Panel B) using panel regressions. Price underreaction is defined as the fraction of price response that happens in the delay window, i.e., from days $t + 1$ to $t + 15$, in response to shocks on day t . Stars *, **, and *** represent statistical significance at 10%, 5%, and 1% levels, respectively.

6.3 GMM-based evidence at the stock level

I use GMM to estimate price underreaction to market and industry-level shocks for each stock i in each decade p . After applying data quality filters described in section 3.2, I end up with 31,659 stock-decade estimates of underreaction to market shocks and 22,277 estimates of underreaction to industry shocks.

Like in Section 4.2, I sort payoff-relevance proxies into quintiles, and estimate the panel regression below. Standard errors are clustered by stock and decade. For liquidity controls, I include each stock’s log market capitalization, log turnover, fraction of zero-trading days, and average bid-ask spread. The results are reported in Table 5.

$$\begin{aligned}
 \text{Underreaction}_{i,p}^{\text{risk}} = & \underbrace{\sum_{q=1}^5 \gamma_q^{\text{risk}} \cdot \mathbf{1}_{\sigma_{i,p}^{\text{risk}} \text{ in quintile } q}}_{\text{“main effect”}} + \underbrace{\sum_{q=1}^5 \gamma_q^{\text{other risk}} \cdot \mathbf{1}_{\sigma_{i,p}^{\text{other risk}} \text{ in quintile } q}}_{\text{“distraction effect”}} \\
 & + \underbrace{\beta \cdot \mathbf{1}_{\text{risk=Market}}}_{\text{attention cost difference}} + \text{LiquidityControls}_{i,p} + \sum_{p=1}^3 \delta_p \cdot \mathbf{1}_{\text{decade } p} + \epsilon_{i,p}^{\text{risk}}. \quad (13)
 \end{aligned}$$

Similar to the finding in corporate bonds, there is clear evidence for the “main effect”: going from the bottom to the top quintile in payoff-relevance reduces underreaction by 23 - 27%. The “distraction effect” is present in the equal-weighted specification but not so in the market capitalization-weighted specification, suggesting that investors in large cap stocks do not face binding information processing constraints. These results are barely affected by the inclusion of stock fixed effects. After controlling for payoff-relevance, underreaction to market-level shocks is lower by almost 30%, consistent with it being easier to process market-level shocks. Finally, in terms of marginal explanatory power, the payoff-relevance proxies are roughly half as important as the liquidity controls in the equal-weighted specification, and roughly as important as the market cap-weighted specification.

6.4 Quantifying attention costs

Finally, I also use the method in section 5 to estimate the degree of return predictability. For stocks in each of the 5×5 payoff-relevance bins, I use the most recent decade of data to estimate predictive relationships:

$$\begin{aligned}
 \text{StockRet}_{i,t} = & b_1^{\text{market}} \cdot \text{MarketRet}_{t-1} + \dots + b_{15}^{\text{market}} \cdot \text{MarketRet}_{t-15} + \\
 & b_1^{\text{industry}} \cdot \text{IndustryRet}_{i,t-1} + \dots + b_{15}^{\text{industry}} \cdot \text{IndustryRet}_{i,t-15} + \epsilon_{i,t},
 \end{aligned}$$

Stock price underreaction to risk (percent)				
	Equal-weighted		Market cap-weighted	
	Pooled regression	Within stock	Pooled regression	Within stock
	(1)	(2)	(3)	(4)
$\mathbf{1}_{\sigma^{\text{risk}}}$ in quintile 2	-6.05*** (1.16)	-6.59*** (0.90)	-8.31*** (1.52)	-6.54*** (1.58)
$\mathbf{1}_{\sigma^{\text{risk}}}$ in quintile 3	-11.73*** (1.87)	-12.97*** (1.49)	-15.09*** (1.23)	-14.02*** (1.24)
$\mathbf{1}_{\sigma^{\text{risk}}}$ in quintile 4	-16.35*** (2.07)	-17.78*** (1.71)	-20.94*** (1.78)	-19.99*** (1.41)
$\mathbf{1}_{\sigma^{\text{risk}}}$ in quintile 5	-22.96*** (2.24)	-24.60*** (1.69)	-26.67*** (1.99)	-25.59*** (2.22)

$\mathbf{1}_{\sigma^{\text{other risk}}}$ in quintile 2	1.57* (0.81)	0.00 (0.39)	-0.86 (0.85)	0.63 (1.45)
$\mathbf{1}_{\sigma^{\text{other risk}}}$ in quintile 3	3.46*** (0.78)	0.90 (0.61)	-2.18* (1.14)	-1.42 (2.11)
$\mathbf{1}_{\sigma^{\text{other risk}}}$ in quintile 4	5.54*** (1.18)	2.94*** (0.73)	-0.59 (1.55)	0.32 (2.81)
$\mathbf{1}_{\sigma^{\text{other risk}}}$ in quintile 5	9.97*** (1.86)	7.39*** (1.24)	0.94 (1.64)	2.36 (2.56)

$\mathbf{1}_{\text{risk} = \text{market}}$	-28.81*** (1.90)	-29.13*** (2.03)	-28.71*** (2.21)	-29.31*** (2.56)

Liquidity controls	Y	Y	Y	Y
Stock fixed effect	N	Y	N	Y
Decade fixed effect	Y	Y	Y	Y
Obs	53,832	53,832	53,832	53,832
Adjusted R^2	48.81%	55.54%	57.64%	54.82%
Marginal R^2 , $\mathbf{1}_{\sigma_{i,P}^{\text{risk}}}$	3.74%	3.29%	7.14%	5.95%
Marginal R^2 , $\mathbf{1}_{\sigma_{i,P}^{\text{other risk}}}$	0.79%	0.41%	0.12%	0.15%
Marginal R^2 , liq controls	7.68%	1.12%	6.04%	0.83%

Table 5. Explaining stock price underreaction. Price underreaction to market-level and industry-level shocks are estimated for each stock in each decade using GMM. The regression pools underreaction to both risks together. In the variables, when “risk” refers to market, “other risk” refers to industry, and vice versa. The first eight regressors are payoff-relevance measures for “risk” and “other risk,” turned into quintile indicator variables to capture possible nonlinearity. The next regressor captures the average difference between underreaction to two type of shocks. Standard errors are clustered by stock and decade. Coefficients significant at the 10%, 5%, and 1% level are noted with *, **, and ***, respectively.

and then apply the predictive models forward. To make sure the estimated predictability is actionable in a trading strategy, I use open-to-close prices to compute stock returns. That is, I consider hypothetical strategies that initiate positions at the opening price of the day and close positions at the end of the trading day.

Same as the corporate bond case, the implied return predictability is too large for arbitrageurs to make money after paying transaction cost. Following section 5, I compute

pre-transaction cost profits to trading strategies that long the top $\frac{x}{2}\%$ predictions and short the bottom $\frac{x}{2}\%$ predictions. The strategy returns are reported in Figure 8 by decade. Even when only trading on $x = 2\%$ extreme predictions, the predictable open-to-close return only averages 5-22 basis points a day, which is clearly below the prevailing bid-ask spread plotted using a line (right axis). Because quoted bid-ask spreads are not available throughout the sample, I use the estimated bid-ask spread in Corwin and Schultz (2012).

Implied dollar value of attention costs. To gauge how much money regular institutional market participants are “leaving on the table”, following section 5, I conduct a back-of-envelope exercise using the average U.S. mutual fund. Since 2002, the average fund has \$852 million assets under management.²⁶ The average annual turnover is 116%. Using the average predictable return of 3.1 basis points during the recent two decades (Figure 5), I estimate that the average fund loses the following amount due to insufficient processing of information:

$$\underbrace{\$852 \text{ million AUM} \times 116\% \text{ Turnover}}_{\text{Annual trading volume}} \times \underbrace{\frac{3.1 \text{ bp return predictability}}{2}}_{\text{Profits from better timing trades}} \approx \$306,379$$

Assuming 20% of profits are paid to managers, this amounts to approximately \$61,300 of compensation loss, which is roughly 2.3% of the annual compensation for mutual fund managers.

7 Conclusion

When studying price underreaction to public information, existing papers usually focus on the traditional explanations of illiquidity and trading frictions. In this paper, I propose and test the rational inattention view: underreaction arises from investors’ information-processing constraints. Using the corporate bond market as a laboratory, I find that rational inattention explains much more variation in underreaction than the traditional mechanisms.

The key distinguishing feature of rational inattention is that underreaction is risk-specific. When facing attention constraints, investors choose to pay more attention to more payoff-relevant risks, speeding up the price incorporation of those risks. In corporate bonds, I

²⁶I use domestic equity mutual funds in CRSP and restrict attention to those with equity comprising between 75% and 125% of its portfolio, and have at least \$10 million assets under management. To be able to aggregate at the fund level, I use the CRSP portfolio number, and as a consequence, my sample starts in 2002. CRSP mutual fund data is at the share-class level, not the fund level. The mapping from share-class ID to fund ID (CRSP portfolio number) starts in 2002. This sample period is similar to the corporate bond exercise where I use 2002 - 2014.

find supporting evidence that bonds with higher credit risk respond more quickly to default risk shocks, and bonds with longer duration respond more quickly to interest-rate shocks. This is not only true when comparing across bonds, but also within bond: when a bond changes credit quality, then the speeds at which it responds to shocks changes accordingly, consistent with investors reallocating attention based on changing circumstances. Because these findings are risk-specific, they cannot be explained by illiquidity or transaction costs. I also estimate that the “money left on the table” in the documented underreactions are small enough to be reasonably explained as reflecting the true cost of information-processing. Finally, I show parallel evidence in the U.S. equity market since 1930, suggesting that my earlier finding is not specific to corporate bonds or the recent sample period.

Understanding causes of market inefficiency is a first order question in asset pricing. This paper argues that information-processing frictions, rather than liquidity frictions, may be more important and should be given serious consideration.

pts)

40

32

50



References

- Ackermann, C., R. McEnally, and D. Ravenscraft. 1999. The performance of hedge funds: Risk, return, and incentives. *The journal of Finance* 54:833–874.
- Addoum, J. M., and J. Murfin. 2018. Equity price discovery with informed private debt. *Available at SSRN 2869452* .
- Andersen, S., J. Y. Campbell, K. M. Nielsen, and T. Ramadorai. 2018. Sources of inaction in household finance: Evidence from the Danish mortgage market. *Available at SSRN 2463575* .
- Barber, B. M., and T. Odean. 2007. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies* 21:785–818.
- Bernard, V. L., and J. K. Thomas. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13:305–340.
- Bessembinder, H., K. M. Kahle, W. F. Maxwell, and D. Xu. 2009. Measuring abnormal bond performance. *Review of Financial Studies* 22:4219–4258.
- Blanco, R., S. Brennan, and I. W. Marsh. 2005. An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. *The journal of Finance* 60:2255–2281.
- Cici, G., and S. Gibson. 2012. The performance of corporate bond mutual funds: Evidence based on security-level holdings. *Journal of Financial and Quantitative Analysis* 47:159–178.
- Cohen, L., and A. Frazzini. 2008. Economic links and predictable returns. *Journal of Finance* 63:1977–2011.

- Collin-Dufresn, P., R. S. Goldstein, and J. S. Martin. 2001. The determinants of credit spread changes. *The Journal of Finance* 56:2177–2207.
- Cornaggia, K. R., J. Hund, and G. Nguyen. 2018. Investor Attention and Municipal Bond Returns. *Available at SSRN 3039098* .
- Corwin, S. A., and P. Schultz. 2012. A simple way to estimate bid-ask spreads from daily high and low prices. *The Journal of Finance* 67:719–760.
- DellaVigna, S., and J. M. Pollet. 2009. Investor inattention and Friday earnings announcements. *Journal of Finance* 64:709–749.
- Dichev, I. D., and J. D. Piotroski. 2001. The long-run stock returns following bond ratings changes. *The Journal of Finance* 56:173–203.
- Dick-Nielsen, J. 2014. How to clean enhanced TRACE data. <https://ssrn.com/abstract=2337908> .
- Downing, C., S. Underwood, and Y. Xing. 2009. The relative informational efficiency of stocks and bonds: An intraday analysis. *Journal of Financial and Quantitative Analysis* 44:1081–1102.
- Drake, M. S., K. H. Gee, and J. R. Thornock. 2016. March market madness: the impact of value-irrelevant events on the market pricing of earnings news. *Contemporary Accounting Research* 33:172–203.
- Duffie, D., N. Gârleanu, and L. H. Pedersen. 2005. Over-the-counter markets. *Econometrica* 73:1815–1847.
- Duffie, D., and Y. Sun. 2012. The exact law of large numbers for independent random matching. *Journal of Economic Theory* 147:1105–1139.
- Edwards, A. K., L. E. Harris, and M. S. Piwowar. 2007. Corporate bond market transaction costs and transparency. *Journal of Finance* 62:1421–1451.

- Fama, E. F., and K. R. French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* .
- Fedyk, A. 2018. Front Page News: The Effect of News Positioning on Financial Markets .
- Fleming, M. J., and E. M. Remolona. 1999. Price formation and liquidity in the US Treasury market: The response to public information. *Journal of Finance* 54:1901–1915.
- Focke, F., S. Ruenzi, and M. Ungeheuer. 2018. Advertising, attention, and financial markets. *SSRN Working Paper Series* .
- Gebhardt, W. R., S. Hvidkjaer, and B. Swaminathan. 2005. Stock and bond market interaction: Does momentum spill over? *Journal of Financial Economics* 75:651–690.
- Glosten, L. R., and P. R. Milgrom. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics* 14:71–100.
- Hansen, L. P., and R. Jagannathan. 1991. Implications of security market data for models of dynamic economies. *Journal of Political Economy* 99:225–262.
- Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance* 64:2289–2325.
- Holden, C. W., Y. Mao, and J. Nam. 2018. Price Discovery in the Stock, OTC Corporate Bond, and NYSE Corporate Bond Markets .
- Hong, G., and A. Warga. 2000. An empirical study of bond market transactions. *Financial Analysts Journal* 56:32–46.
- Hong, Y., H. Lin, and C. Wu. 2012. Are corporate bond market returns predictable? *Journal of Banking & Finance* 36:2216–2232.
- Hou, K., and T. J. Moskowitz. 2005. Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies* 18:981–1020.

- Huang, S., Y. Huang, and T.-C. Lin. 2018. Attention allocation and return co-movement: Evidence from repeated natural experiments. *Journal of Financial Economics* .
- Jegadeesh, N., and J. Livnat. 2006. Revenue surprises and stock returns. *Journal of Accounting and Economics* 41:147–171.
- Kacperczyk, M., S. Van Nieuwerburgh, and L. Veldkamp. 2016. A rational theory of mutual funds' attention allocation. *Econometrica* 84:571–626.
- Kumar, A., S. Ruenzi, and M. Ungeheuer. 2019. Daily winners and losers. *Available at SSRN 2931545* .
- Kwan, S. H. 1996. Firm-specific information and the correlation between individual stocks and bonds. *Journal of Financial Economics* 40:63–80.
- Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica* pp. 1315–1335.
- Longstaff, F. A., and E. S. Schwartz. 1995. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance* 50:789–819.
- Ma, L., Y. Tang, and J.-P. Gómez. 2018. Portfolio manager compensation in the US mutual fund industry. *Journal of Finance (Forthcoming)* .
- Mackowiak, B., F. Matejka, and M. Wiederholt. 2018. Rational inattention: a disciplined behavioral model. *Working paper* .
- Mackowiak, B., and M. Wiederholt. 2009. Optimal sticky prices under rational inattention. *American Economic Review* 99:769–803.
- Mendenhall, R. R. 2004. Arbitrage risk and post-earnings-announcement drift. *The Journal of Business* 77:875–894.
- Menzly, L., and O. Ozbas. 2010. Market segmentation and cross-predictability of returns. *Journal of Finance* 65:1555–1580.

- Milgrom, P. R. 2004. *Putting auction theory to work*. Cambridge University Press.
- Peng, L., W. Xiong, and T. Bollerslev. 2007. Investor attention and time-varying comovements. *European Financial Management* 13:394–422.
- Sadka, R. 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. *Journal of Financial Economics* 80:309–349.
- Schaefer, S. M., and I. A. Strebulaev. 2008. Structural models of credit risk are useful: Evidence from hedge ratios on corporate bonds. *Journal of Financial Economics* 90:1–19.
- Schestag, R., P. Schuster, and M. Uhrig-Homburg. 2016. Measuring liquidity in bond markets. *Review of Financial Studies* 29:1170–1219.
- Van Nieuwerburgh, S., and L. Veldkamp. 2010. Information acquisition and underdiversification. *Review of Economic Studies* 77:779–805.
- Wei, X., C. Truong, and M. Veeraraghavan. 2012. Post-earnings Announcement Drift: Evidence from the Corporate Bond Market. *SSRN* .
- Welch, I., and A. Goyal. 2007. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21:1455–1508.
- Womack, K. L. 1996. Do brokerage analysts' recommendations have investment value? *Journal of Finance* 51:137–167.
- Yuan, Y. 2015. Market-wide attention, trading, and stock returns. *Journal of Financial Economics* 116:548–564.
- Zeileis, A. 2004. Econometric computing with HC and HAC covariance matrix estimators. *WU Vienna University of Economics and Business* .
- Zou, H., and T. Hastie. 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67:301–320.

A Model and Proofs

A.1 A rational inattention model of price underreaction

I develop a model in which price underreaction is generated from rational inattention. In the model, investors allocate their information-processing resources (“attention”) across risks. The modeling of attention choice builds upon Van Nieuwerburgh and Veldkamp (2010). The model timeline is shown in Figure 9.

Asset There is a single risky asset with zero net supply. The fundamental asset value is the sum of N independent risks,

$$V = \sum_{i=1}^N X_i, \quad X_i \sim \mathcal{N}(0, \sigma_i^2) \quad \forall i \quad (14)$$

where each risk X_i is normally distributed with mean zero and variance σ_i^2 . I assume the risks are independent for analytical tractability. The different variances $\{\sigma_i^2\}_{i=1}^N$ represent the *payoff-relevance* of risks. For example, take the asset to be a corporate bond, with X_1 representing the interest rate risk component, X_2 representing the default risk component, and X_3, \dots, X_N representing other components such as the liquidity risk components, and so forth. Bonds with longer duration will have higher σ_1 and those with more credit risk will have higher σ_2 .

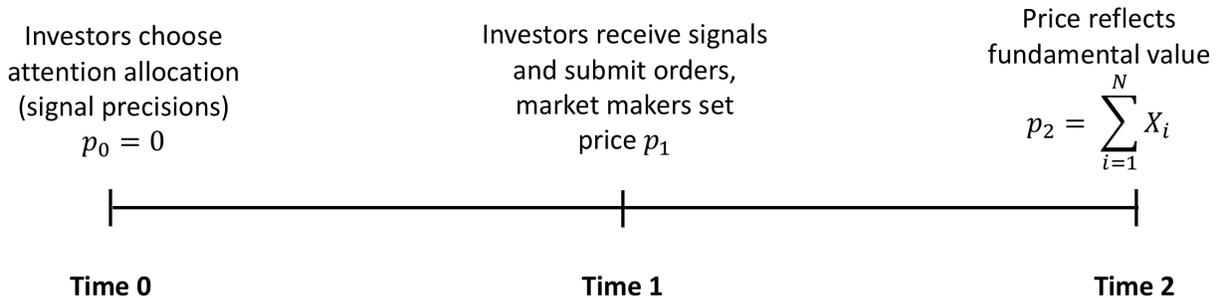


Figure 9. Sequence of events in the model.

Time 2 is taken to be sufficiently far into the future such that, for unmodeled reasons, price p_2 fully reflects fundamental asset value V . We focus on studying how much of each risk is incorporated into price after trading at time 1.

Trading At time 1, there is a continuous unit mass of competitive, identical risk-neutral investors who submit market orders. Each investor $j \in [0, 1]$ incurs a quadratic trading cost

$\frac{\psi}{2}q_j^2$ when submitting her order of q_j shares, where parameter $\psi > 0$ governs how costly trading is.²⁷ The aggregate order flow $Y = \int_0^1 q_j dj$ is then absorbed by market makers who set price equal to $p_1 = \lambda \cdot Y$.²⁸

Investor attention allocation Before trading at time 1, each investor j receives noisy signals on the N risks:

$$s_{ij} = X_i + \sigma_i \cdot u_{ij}, \quad u_{ij} \sim N(0, K_{ij}^{-1}) \quad (15)$$

where the signal errors u_{ij} are independent across risks and investors. The precisions of the normalized signals $\{s_{ij}/\sigma_i\}_{i=1}^N$ are $\{K_{1j}, \dots, K_{Nj}\}$, and they are investor choice variables.

I model investor attention using these signal precisions. If the investor pays zero attention to risk i ($K_{ij} = 0$), then her signal has infinite variance and she knows nothing about X_i . If she chooses a very large K_{ij} , then she knows X_i very precisely. I think of this precision choice as capturing investor effort to processing information. For corporate bond investors, if they very closely follow news about the company and work diligently to infer the value impact of those news, that will show up as them choosing a high precision on the default-risk component.

To capture the idea that attention is costly, I assume that investors can reduce the amount of noise in signals by paying attention resources. I closely follow the modeling technique in Van Nieuwerburgh and Veldkamp (2010). Specifically, at time 0, each investor j choose signal precisions $\{K_{ij} \geq 0\}_{i=1}^N$ subject to a cost of

$$\sum_{i=1}^N c_i \cdot K_{ij} \quad (16)$$

where $c_i > 0$ are risk-specific attention costs. I allow for different costs across risks because some risks are arguably less costly to process. In the corporate bond market, it is arguably the case that interest rate relevant news – macroeconomics news, central bank announcements, etc – are easier to process than company-level default-relevant news. For a typical corporate

²⁷This quadratic cost can be thought of as a reduced-form way to capture risk aversion of agents with constant absolute risk aversion. I can think of ψ as equal to $A \cdot \text{Var}(R|\{s_{ij}\}_{i=1}^N)$ where A is the absolute risk aversion coefficient and R is the return on the asset from time 1 to time 2. Our assumption of a constant ψ thus amounts to ignoring changes in the conditional payoff variance in $\text{Var}(R|\{s_{ij}\}_{i=1}^N)$ when investors choose different signal structures (to be explained later). This assumption should be innocuous if the amount of variance reduction from learning is small relative to the overall payoff risk, which is arguably the case in our empirical setting of corporate bonds.

²⁸While I take the price-impact parameter λ to be fixed in this section, I show in Appendix A.4.2 that it can be endogenized using the Kyle (1985) set-up. That is, λ can arise from competitive and uninformed market makers learning from order flow and setting break-even price $p_1 = E(V|Y)$.

bond portfolio manager, she only needs to learn about interest rate risk once and can apply it to all her bond holdings. Learning about default news requires reading news about all companies in her portfolio. Therefore, after dividing by the number of bonds held, the cost of learning about interest rates should be considerably lower.

A.2 Equilibrium

I characterize the equilibrium and provide intuition about the underlying mechanisms. The solution concept is Perfect Bayesian Equilibria, with each investor j choosing optimal attention allocation $(K_{1j}, K_{2j}, \dots, K_{Nj})$ and optimal order submission rules $q_j(s_{1j}, s_{2j}, \dots, s_{Nj})$. This baseline version of our model can be solved analytically.

PROPOSITION 1. (Equilibrium) *There is a unique equilibrium where time 1 price is*

$$p_1 = \sum_{i=1}^N (1 - D_i) \cdot X_i \quad (17)$$

For each risk i , if payoff-relevance is high enough ($\sigma_i \geq \sqrt{2\psi c_i}$), then the price underreaction and all investors' attention choices are:

$$D_i = \frac{\psi + \frac{\lambda\sqrt{2\psi c_i}}{\sigma_i}}{\psi + \lambda} \quad (18)$$

$$K_i = \frac{\lambda + \sigma_i \sqrt{\frac{\psi}{2c_i}}}{\lambda + \psi} \quad (19)$$

Otherwise, payoff-relevance isn't high enough, then all investors pay no attention ($K_i = 0$) and price completely underreacts ($D_i = 1$).

Finally, each investor j submits the following market order:

$$q_j(s_{1j}, \dots, s_{Nj}) = \frac{1}{\psi} \cdot \sum_{i=1}^N D_i \cdot \frac{K_i}{K_i + 1} \cdot s_{ij}. \quad (20)$$

The price path is illustrated in Figure 10. The time 1 price will partially react to risks, and D_i is the *fraction* of price underreaction to risk i . I make a few remarks about the model:

1. Investor attention determines price underreaction. If investors pay more attention to a risk by choosing higher K_i , then they will trade more heavily on that risk (equation (20)), speeding up price incorporation of information.

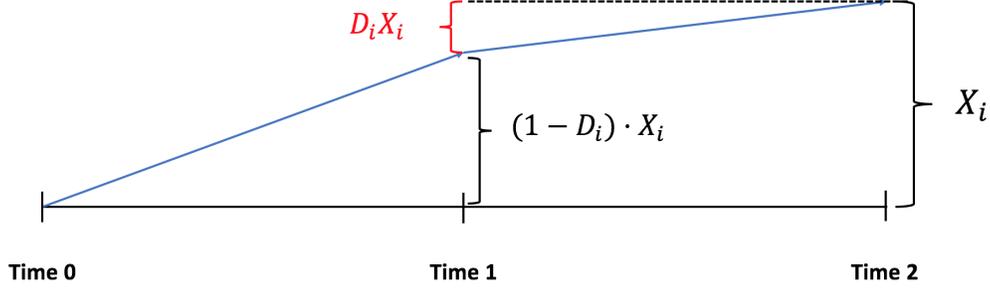


Figure 10. Illustration of price response to risk i in the model. Price incorporates $1 - D_i$ fraction of the realization X_i at time 1, so D_i (red) is the fraction of underreaction.

2. Investors profit from more precise signals by trading on price drifts. For each risk X_i , the time 1 to time 2 price drift is $D_i \cdot X_i$.
3. Investor attention allocation to each risk i is decreasing in the cost of attention c_i and increasing in the payoff-relevance σ_i (equation (19)).
4. The equilibrium uniqueness comes from two forces. First, investors are competitive and take D_i as exogenously given. Second, the attention allocation of different investors are strategic substitutes. That is, if other investors pay more attention to risk i , D_i will decrease, then some investors will react by decreasing attention to risk i . In contrast, strategic complementarity tends to lead to multiple equilibria.

Note that λ is assumed fixed in this baseline model. In Appendix A.4.2, I show that the equilibrium structure is the same if I endogenize λ as in Kyle (1985).

A.3 Testable predictions

The key feature of our model is that price underreaction will be risk-specific. I derive two predictions to bring to data.

1. **Main effect:** *for each risk i , price underreaction D_i is weakly decreasing in payoff-relevance σ_i and weakly increasing in attention cost c_i .*

This is a direct corollary from the equilibrium solution in Proposition 1. The “weak” part in the statement is only for extreme cases where investors pay zero attention. Otherwise, the comparative statics is strict.

This prediction can be best understood from a first order condition intuition. Because investors trade on drifts $D_i \cdot X_i$, the degree of profitability from learning risk X_i is related to $D_i \cdot \sigma_i$. Obviously, higher attention cost c_i means investors will optimally

pay less attention and thus price underreaction increases. Fixing attention cost c_i , the higher the payoff-relevance σ_i , the smaller D_i has to be for agents to be optimizing.

2. **“Distraction effect”**: *If investors face attention constraints, i.e. they have a finite attention budget ($\sum_{i=1}^N c_i K_{ij} \leq \bar{K}_j$), then price underreaction to each risk i (D_i) is weakly increasing in the payoff-relevance σ_r and weakly decreasing in the attention cost c_r of other risks ($r \neq i$).*

This is a test of whether investors have attention constraints. If they don't, as in the baseline model, then the payoff-relevance and attention costs of other risks are irrelevant for D_i .

Mapping to the empirical setting, I would expect investors to not face constraints if they can flexibly add attention capacity by working harder or hiring more analysts. This may be a reasonable assumption for institutional investors. At the same time, if there are frictions in hiring and delegating tasks to others, then investment managers may indeed face constraints.

In our empirical tests, the main variation will come from payoff-relevance. Corporate bond value depends on interest rate risks and default risks, but depending on the duration and credit quality of the bond, there is significant variation in payoff-relevance. In terms of the attention cost prediction, the only testable variation is comparing between the two type of risks. I think attention cost to interest rate should be lower, so after controlling for payoff-relevance, underreaction to interest rate should be smaller.

A.4 Proofs

A.4.1 Proof for Proposition 1 (equilibrium)

I start with conjecturing the pricing functional form in (17):

$$p_1 = \sum_{i=1}^N (1 - D_i) \cdot X_i + \lambda \cdot Z. \quad (21)$$

Under this conjecture, asset return from time 1 to time 2 reflects continued drifts in the

N components plus reversion of the noise-trading component:

$$R = V - p_1 = \sum_{i=1}^N D_i \cdot X_i - \lambda \cdot Z. \quad (22)$$

Time 1 trading with fixed attention allocation Investors with more accurate signals can make more profits from trading on the drifts. I first derive investor portfolio choice with fixed attention choices. After receiving signals $\{s_{ij}\}_{i=1}^N$, investor $j \in [0, 1]$ updates belief about return:

$$\begin{aligned} E(R|\{s_{ij}\}_{i=1}^N) &= \sum_{i=1}^N D_i \cdot E(X_i|s_{ij}) \\ &= \sum_{i=1}^N D_i \cdot \theta_{ij} \cdot s_{ij} \end{aligned}$$

where the projection coefficient θ_{ij} is an alternative way to parametrize the attention choices:

$$\begin{aligned} \theta_{ij} &= \frac{\text{Cov}(X_i, s_{ij})}{\text{Var}(s_{ij})} = \frac{\text{Cov}(X_i, X_i + \sigma_i \cdot u_{ij})}{\text{Var}(X_i + \sigma_i \cdot u_{ij})} \\ &= \frac{\sigma_i^2}{\sigma_i^2(1 + K_{ij}^{-1})} = \frac{K_{ij}}{K_{ij} + 1}. \end{aligned}$$

Investor submits optimal market order q_j , given by

$$\begin{aligned} q_j &= \arg \max_q q \cdot E(R|\{s_{ij}\}_{i=1}^N) - \frac{\psi}{2} q^2 \\ &= \frac{E(R|\{s_{ij}\}_{i=1}^N)}{\psi} = \frac{\sum_{i=1}^N D_i \cdot \theta_{ij} \cdot s_{ij}}{\psi}. \end{aligned}$$

The aggregate order flow is:

$$\begin{aligned} Y &= \underbrace{\int_0^1 q_j dj}_{\text{Investors}} + \underbrace{Z}_{\text{noise traders}} \\ &= \frac{1}{\psi} \sum_{i=1}^N D_i \cdot \left(\int_0^1 \theta_{ij} \cdot s_{ij} dj \right) + Z \\ &= \frac{1}{\psi} \sum_{i=1}^N D_i \cdot \bar{\theta}_i \cdot X_i + Z \end{aligned} \quad (23)$$

where $\bar{\theta}_i = \int_0^1 \theta_{ij} dj$ and the last step uses the exact law of large numbers (Duffie and Sun

(2012)). Thus, the aggregate order flow will more heavily trade on risks with more underreaction (higher D_i) and if investors paid more attention to that risk (higher $\bar{\theta}_i$). Now, under the exogenous λ set up, I have verified the pricing conjecture (21):

$$\begin{aligned} p_1 &= \lambda \cdot Y \\ &= \sum_{i=1}^N \underbrace{\frac{\lambda}{\psi} \cdot D_i \cdot \bar{\theta}_i \cdot X_i}_{=1-D_i} + \lambda_Z \end{aligned}$$

Solve for attention allocation at time 0 I now solve for the investor attention choice at time 0. After receiving signals $\{s_{ij}\}_{i=1}^N$, investor j 's expected utility is:

$$\begin{aligned} U_j(\{s_{ij}\}_{i=1}^N) &= \underbrace{q_j(\{s_{ij}\}_{i=1}^N) \cdot E(R|\{s_{ij}\}_{i=1}^N)}_{\text{expected profit}} - \underbrace{\frac{\psi}{2} q_j^2(\{s_{ij}\}_{i=1}^N)}_{\text{trading cost}} \\ \text{optimal } q_j(\{s_{ij}\}_{i=1}^N) &= \frac{E(R|\{s_{ij}\}_{i=1}^N)}{\psi} \frac{E(R|\{s_{ij}\}_{i=1}^N)^2}{2\psi}. \end{aligned}$$

To get ex ante expected utility, I simply integrate over signal realizations:²⁹

$$\begin{aligned} U_j &= E(U_j(\{s_{ij}\}_{i=1}^N)) \\ &= \frac{1}{2\psi} \cdot \text{Var}(R|\{s_{ij}\}_{i=1}^N) \end{aligned} \tag{24}$$

where the last step uses the law of total variance and the fact that $\text{Var}(R) = 0$. Equation (24) is intuitive: investor utility only depends on how informative her signals are in predicting returns. I further expand out (24) as a function of signal precisions:

$$\begin{aligned} U_j &= \frac{1}{2\psi} \cdot \text{Var} \left(\sum_{i=1}^N D_i \cdot X_i \middle| \{s_{ij}\}_{i=1}^N \right) \\ &= \frac{1}{2\psi} \cdot \sum_{i=1}^N D_i^2 \cdot \text{Var}(X_i|s_{ij}) \\ &= \frac{1}{2\psi} \cdot \sum_{i=1}^N \underbrace{\frac{K_{ij}}{K_{ij} + 1}}_{=\theta_{ij}} \cdot \underbrace{D_i^2 \sigma_i^2}_{\text{utility from knowing } X_i} \end{aligned} \tag{25}$$

Equation (25) has a natural interpretation: an investor who knows X_i completely (taking

²⁹Due to the normal-normal updating structure, $\text{Var}(R|\{s_{ij}\}_{i=1}^N)$ does not depend on signal realizations.

signal precision K_{ij} to infinity) will increase her utility by a factor proportional to $D_i^2 \sigma_i^2$. The more payoff-relevant is risk i (higher σ_i), the higher the degree of underreaction (D_i), the more profitable it is to learn about and trade on. The $\frac{K_{ij}}{K_{ij}+1}$ factor is increasing in the signal precision chosen, and it is concave in K_{ij} due to decreasing returns to scale in precision acquisition: the more precise the signals, the less valuable is another unit of precision in reducing variance.

I now solve for optimal precision choices for the two attention specifications. Because investors face exactly the same optimization problem, they will make identical attention choices.

1. If I use the attention cost specification (without an upper bound), then, notice that the objective (25) is separable and concave in precisions while the cost $\sum_{i=1}^N c_i K_{ij}$ is linear. Thus, as long as attention costs are not too large, the optimal solution comes from taking first order conditions separately for each risk:

$$\begin{aligned} \frac{dU_j}{dK_{ij}} &= \frac{D_i^2 \sigma_i^2}{2\psi} \cdot \frac{1}{(K_{ij} + 1)^2} \leq c_i \\ \Rightarrow K_{ij}^* &= \max\left(0, \frac{D_i \sigma_i}{\sqrt{2\psi c_i}} - 1\right) \end{aligned} \quad (26)$$

2. If the total investor attention has an upper bound, then I take first order conditions of the Lagrangian:

$$\mathcal{L} = \frac{1}{2\psi} \sum_{i=1}^N \frac{K_{ij}}{K_{ij} + 1} \cdot D_i^2 \sigma_i^2 + \Phi \cdot \left(\bar{K} - \sum_{i=1}^N c_i K_{ij} \right) \quad (27)$$

$$\Rightarrow K_{ij}^* = \max\left(0, \frac{D_i \sigma_i}{\sqrt{2\psi \Phi c_i}} - 1\right) \quad (28)$$

where Φ is the Lagrange multiplier.

Solving for underreaction parameters So far I have solved the model taking $\{D_1, \dots, D_N\}$ as given. Now I solve for them. Note that price is given by:

$$\begin{aligned} p_1 &= \lambda \cdot Y \\ &= \lambda \cdot \left(\frac{1}{\psi} \sum_{i=1}^N D_i \cdot \frac{K_i^*}{K_i^* + 1} \cdot X_i + Z \right) \end{aligned} \quad (29)$$

where K_i^* are the optimal attention choices in (26) (or (28)) and I omit the agent index j for simplicity. Notice that these attention choices are functions of $\{D_1, \dots, D_N\}$. Equating

coefficients in (29) and (21) gives equations for D 's:

$$1 - D_i = \frac{\lambda}{\psi} \cdot D_i \cdot \frac{K_i^*}{K_i^* + 1} \quad (30)$$

1. In the attention cost specification, if $\sigma_i \leq \sqrt{2\psi c_i}$, then $K_i^* = 0$ and $D_i = 1$ (full underreaction). Otherwise, substitute in $K_i^* = \frac{D_i \sigma_i}{\sqrt{2\psi c_i}} - 1$:

$$\begin{aligned} 1 - D_i &= \frac{\lambda}{\psi} D_i \cdot \frac{D_i \sigma_i - \sqrt{2\psi c_i}}{D_i \sigma_i} \\ \Rightarrow D_i &= \frac{\psi + \frac{\lambda \sqrt{2\psi c_i}}{\sigma_i}}{\psi + \lambda} \quad \text{as long as } \sigma_i \geq \sqrt{2\psi c_i} \end{aligned}$$

Thus, underreaction D_i increases in attention cost c_i and decreases in payoff-relevance σ_i . Note that it does not depend on the parameters of the other risks.

2. In the attention constraint specification, the expression becomes:

$$D_i = \frac{\psi + \frac{\lambda \sqrt{2\psi \Phi c_i}}{\sigma_i}}{\psi + \lambda} \quad \text{as long as } \sigma_i \geq \sqrt{2\psi \Phi c_i}$$

where Φ is the Lagrange multiplier that makes the attention constraint bind.

A.4.2 Proving equilibrium under endogenous λ

I now show that the equilibrium structure and uniqueness do not depend on an exogenous λ . Following Kyle (1985), I assume market makers are competitive and uninformed, and they learn from aggregate order flow and set prices equal to $p_1 = E(V|Y)$ where Y is the aggregate order flow.

I again start by conjecturing the pricing equation (21) and solving through to get aggregate order flow expression (23). Because Y is linear in $\{X_1, \dots, X_N, Z\}$, p_1 will also be linear in these variables through standard Bayesian updating. Competitive market makers will set

time 1 price equal to expected value conditional on their information:

$$\begin{aligned}
p_1 &= E(V|Y) = \sum_{i=1}^N E(X_i|Y) \\
&= \left(\sum_{i=1}^N \frac{\text{Cov}(X_i, Y)}{\text{Var}(Y)} \right) \cdot Y \\
&= \left(\sum_{i=1}^N \frac{\text{Cov}(X_i, \frac{1}{\psi} \cdot D_i \bar{\theta}_i X_i)}{\text{Var}(\frac{1}{\psi} \sum_{i=1}^N D_i \bar{\theta}_i X_i + Z)} \right) \cdot Y \\
&= \frac{\frac{1}{\psi} \cdot \sum_{i=1}^N D_i \bar{\theta}_i \sigma_i^2}{\underbrace{\frac{1}{\psi^2} \cdot \sum_{i=1}^N D_i^2 \bar{\theta}_i^2 \sigma_i^2 + \sigma_Z^2}_{\equiv \lambda(\{D_k\}_{k=1}^N)}} \cdot \left(\frac{1}{\psi} \sum_{i=1}^N D_i \bar{\theta}_i X_i + Z \right) \tag{31}
\end{aligned}$$

Therefore the pricing conjecture is verified. Later I will show that investors choose unique identical attention choices $\{K_i^*\}_{i=1}^N$ in equilibrium, therefore $\bar{\theta}_i = \frac{K_i^*}{K_i^*+1}$, so the λ parameter is given by:

$$\lambda = \frac{\psi \sum_{i=1}^N D_i^2 \frac{K_i^*}{K_i^*+1} \sigma_i^2}{\sum_{i=1}^N D_i^2 \left(\frac{K_i^*}{K_i^*+1} \right)^2 \sigma_i^2 + \psi^2 \sigma_Z^2} \tag{32}$$

Solving for underreaction parameters (D_1, \dots, D_N) I have solved the model taking $\{D_1, \dots, D_N\}$ as given. I now solve for them explicitly. For notational convenience, express attention choices in the θ -space as functions of underreaction parameters:

$$\begin{aligned}
\theta_{ij}^*(D_i) &= \frac{K_{ij}^*}{K_{ij}^* + 1} \\
&= \max \left(0, 1 - \frac{\sqrt{2\psi c_i}}{D_i \sigma_i} \right) \tag{33}
\end{aligned}$$

Note that all investors will make identical choices so $\bar{\theta}_i(D_i) = \theta_{ij}^*(D_i)$ for any j . Plugging (33) into (31) and matching coefficients with (21), I get a system of N equations in $\{D_i\}_{i=1}^N$:

$$1 - D_i = \frac{\lambda(\{D_k\}_{k=1}^N)}{\psi} \cdot \bar{\theta}_i(D_i) \cdot D_i \tag{34}$$

where $\lambda(\{D_i\}_{i=1}^N)$ is defined in (31). I can have corner solutions: if c_i is very large or σ_i is very small, then I should expect $\bar{\theta}_i = 0$ and $D_i = 1$.

Existence

I now treat the case with interior solutions as it is easy to adapt the proof to treat corner solutions. Rewrite (34) as:

$$D_i = \frac{1}{1 + \frac{\lambda(D_1, \dots, D_N)}{\psi} \bar{\theta}_i(D_i)} \quad (35)$$

Define the mapping $f : [0, 1]^N \rightarrow [0, 1]^N$ using the right-hand side of (35):

$$f_i(D_1, \dots, D_N) = \frac{1}{1 + \frac{\lambda(D_1, \dots, D_N)}{\psi} \bar{\theta}_i(D_i)}$$

Clearly, for the interior case, f is a continuous function so there exists at least one fixed point in D_1, \dots, D_N by Brouwer's fixed point theorem.

Uniqueness

Now I just need to show that the solution is unique. I note that it is sufficient to show that λ cannot take on two values. For a given λ , I can solve for θ_i and D_i from (33) and (35):

$$D_i(\lambda) = \frac{1 + \frac{\lambda}{\psi} \frac{\sqrt{2\psi c_i}}{\sigma_i}}{1 + \frac{\lambda}{\psi}} \quad (36)$$

$$\theta_i(\lambda) = \frac{\sigma_i - \sqrt{2\psi c_i}}{\sigma_i + \frac{\lambda}{\psi} \sqrt{2\psi c_i}} \quad (37)$$

So a unique value of λ will yield unique solutions of D_1, \dots, D_N . I now plug (36) and (37) into the expression of λ in (31) to get a mapping from λ to itself:

$$\begin{aligned} g(\lambda) &\equiv \frac{\frac{1}{\psi} \cdot \sum_{i=1}^N D_i(\lambda) \theta_i(\lambda) \sigma_i^2}{\frac{1}{\psi^2} \cdot \sum_{i=1}^N D_i^2(\lambda) \theta_i^2(\lambda) \sigma_i^2 + \sigma_Z^2} \\ &\stackrel{\text{plug in (36) and (37), define } x_i \equiv \frac{\sqrt{2\psi c_i}}{\sigma_i}}{=} \frac{\frac{1}{\psi} \sum_{i=1}^N \frac{(1-x_i) \sigma_i^2}{1+\lambda/\psi}}{\frac{1}{\psi^2} \sum_{i=1}^N \frac{(1-x_i)^2 \sigma_i^2}{(1+\lambda/\psi)^2} + \sigma_Z^2} \end{aligned} \quad (38)$$

Let $u(\lambda)$ and $v(\lambda)$ denote the numerator and denominator in (38). By the ranking lemma in Milgrom (2004), to prove uniqueness it suffices to show that, for all fixed point $\bar{\lambda}$, I have

$g'(\bar{\lambda}) < 1$. I can check this via brute force differentiation:

$$\begin{aligned}
g'(\bar{\lambda}) &= \frac{u'(\bar{\lambda})}{v(\bar{\lambda})} - \frac{u(\bar{\lambda}) \cdot v'(\bar{\lambda})}{v^2(\bar{\lambda})} \\
&\stackrel{\frac{u(\bar{\lambda})}{v(\bar{\lambda})} = \bar{\lambda}}{=} \frac{u'(\bar{\lambda}) - \bar{\lambda} \cdot v'(\bar{\lambda})}{v(\bar{\lambda})} \\
&= \frac{\frac{2\lambda}{\psi^3} \sum_{i=1}^N \frac{(1-x_i)^2 \sigma_i^2}{(1+\lambda/\psi)^3} - \frac{1}{\psi^2} \sum_{i=1}^N \frac{(1-x_i) \sigma_i^2}{(1+\lambda/\psi)^2}}{\frac{1}{\psi^2} \sum_{i=1}^N \frac{(1-x_i)^2 \sigma_i^2}{(1+\lambda/\psi)^2} + \sigma_Z^2} \\
&< \frac{\frac{2\lambda}{\psi^3} \sum_{i=1}^N \frac{(1-x_i)^2 \sigma_i^2}{(1+\lambda/\psi)^3} - \frac{1}{\psi^2} \sum_{i=1}^N \frac{(1-x_i)^2 \sigma_i^2}{(1+\lambda/\psi)^2}}{\frac{1}{\psi^2} \sum_{i=1}^N \frac{(1-x_i)^2 \sigma_i^2}{(1+\lambda/\psi)^2} + \sigma_Z^2} \\
&= \frac{\frac{\lambda-\psi}{\lambda+\psi} \cdot \left(\frac{1}{\psi^2} \sum_{i=1}^N \frac{(1-x_i)^2 \sigma_i^2}{(1+\lambda/\psi)^2} \right)}{\left(\frac{1}{\psi^2} \sum_{i=1}^N \frac{(1-x_i)^2 \sigma_i^2}{(1+\lambda/\psi)^2} \right) + \sigma_Z^2} < 1
\end{aligned}$$

Thus the equilibrium has to be unique.

A.5 Alternative model: transaction cost determines underreaction

In this section, I write a simple model in which price underreaction is driven by transaction cost, rather than inattention. Importantly, this model generate (qualitatively) the “main effect” in the inattention model, but generates the opposite prediction to the “distraction effect”.

In this alternative model, investors observe value innovation V perfectly, but needs to pay transaction cost $c > 0$ to submit an order. I think of this cost as capturing various explicit and implicit costs, such as bid-ask spreads, book-keeping, etc, associated with trading. Investors submit market orders, and market makers set price to $p = \lambda \cdot Q$ where Q is the aggregate order submitted.

This alternative model produces an intuitive equilibrium where the degree of price delays depends on the total size of the asset value change $|V|$, as illustrated in Figure 11. When $|V|$ is small relative to the transaction cost, investors choose not to trade and price does not reflect information. When $|V|$ is large, all investors trade and price fully reflects information. For intermediate values of $|V|$, a fraction of investors trade, and the degree of price inefficiency is such that investors are indifferent between trading or not.

PROPOSITION 2. (*Equilibrium*) *There is a unique equilibrium in which price after trading at time 1 is given by $p(V) = (1 - D(V)) \cdot V$. Let $k(V)$ denote the fraction of investors*

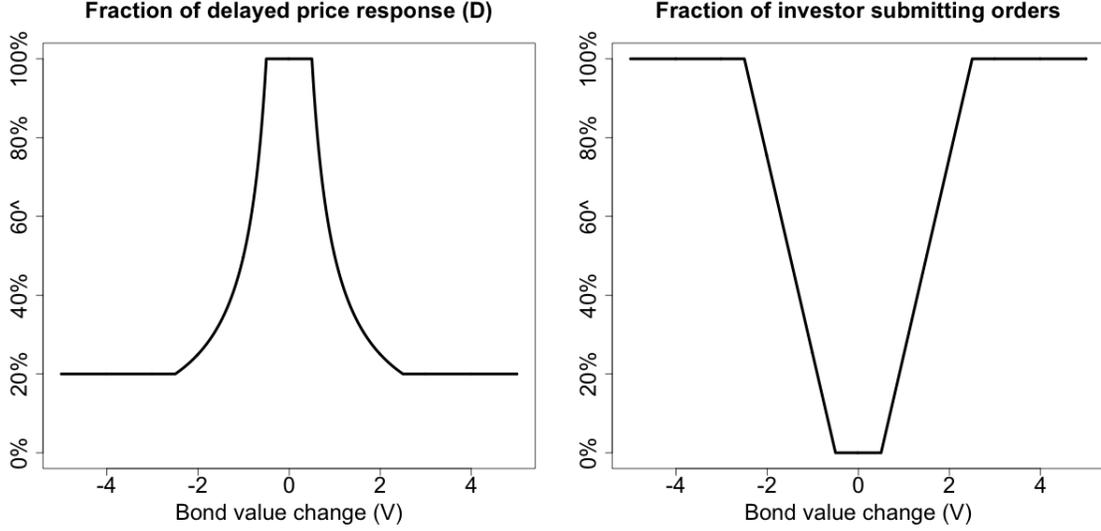


Figure 11. Illustration of the model with transaction cost. Left panel plots the fraction of delayed price response and right panel plots the fraction of investors submitting orders. When value change $|V|$ is too small to overcome fixed cost of transactions, investors do not trade and there is full price underreaction $D = 100\%$. When $|V|$ is very large, all investors trade and price underreaction reaches the lowest level. For intermediate values of $|V|$, investors choose mixed strategies and only a fraction of investors submit orders. Parameters: bond value volatility $\sigma_V = 1$, fixed transaction cost $c = 0.5$, variable transaction cost $\psi = 0.25$, and price impact coefficient $\lambda = 1$.

submitting orders. The equilibrium is given by:

1. When $|V| \leq \sqrt{2\psi c}$, $k(V) = 0$, and $D(V) = 1$.
2. When $|V| \in \left(\sqrt{2\psi c}, \frac{\psi+\lambda}{\psi} \cdot \sqrt{2\psi c} \right]$, $k(V) = \frac{\psi}{\lambda} \cdot \frac{V - \sqrt{2\psi c}}{\sqrt{2\psi c}}$, and $D(V) = \frac{\sqrt{2\psi c}}{V}$.
3. When $|V| > \frac{\psi+\lambda}{\psi} \cdot \sqrt{2\psi c}$, $k(V) = 1$, and $D(V) = \frac{\psi}{\psi+\lambda}$.

Proof. I discuss each of the three regions. Note that, if an investor submits an order, he will trade $q = \frac{V-p}{\psi}$ units and obtain expected utility $\underbrace{\frac{D^2(V) \cdot V^2}{2\psi}}_{\text{trading gains}} - \underbrace{c}_{\text{transaction cost}}$.

1. When $|V| \leq \sqrt{2\psi c}$, trading gains does not justify transaction cost c , so investors choose not to trade.
2. When $|V| \in \left(\sqrt{2\psi c}, \frac{\psi+\lambda}{\psi} \cdot \sqrt{2\psi c} \right]$, $R = D(V) \cdot V = \sqrt{2\psi c}$, so trading gains exactly offset transaction cost, so investors are indifferent so they can adopt a mixed strategy.
3. When $|V| > \frac{\psi+\lambda}{\psi} \cdot \sqrt{2\psi c}$, trading gains are higher than transaction cost, so all investors submit orders $q = \frac{D(V) \cdot V}{\psi}$. Solving for $D(V)$ using $p = \lambda \cdot q$ gives $D(V) = \frac{\psi}{\psi+\lambda}$.

□

Difference from the rational inattention model When asset value decomposes into two components ($V = X_1 + X_2$), this transaction cost mechanism qualitatively generates the “main effect” in the rational inattention model. This is because $D(V)$ decreases in $|V|$. When $|X_1|$ is larger, $|V|$ is on average larger, so price response to X_1 is faster.

However, for the exact same reason, this transaction-cost model generates the *opposite* of the “distraction effect”: larger $|V|$ speeds up the incorporation of X_2 as well. The opposite is true in the rational inattention model where larger $|X_1|$ consumes investor attention and slows down incorporation of X_2 .

B Additional details on data and measurement

B.1 Determinants of the payoff-relevance measures

To check the validity of the payoff-relevance measures computed in Section 3.3, I check whether they are systematically related to other corporate bond sensitivity measures. I sort bonds by these two payoff-relevance measures using common breakpoints into 5×5 portfolios.³⁰ Panel B in Table 6 shows the average duration and credit spreads for those portfolios. Bonds in the bottom σ_i^{trsy} quintile have an average duration in the 2.28 to 4.46 year range, and that rises monotonically to 7.42 - 13.42 years in the top quintile. Similarly, bonds with higher σ_i^{stock} have higher credit spreads.

B.2 Details on GMM-based underreaction estimation

I give more details about the GMM-based estimation here. To restrict the underreaction parameters to be between 0 and 1, I apply logistic transforms:

$$\text{Underreaction}_{i,p}^{\text{stock}} = \frac{1}{1 + e^{-\eta_{i,p}^{\text{stock}}}},$$

$$\text{Underreaction}_{i,p}^{\text{trsy}} = \frac{1}{1 + e^{-\eta_{i,p}^{\text{trsy}}}}$$

where $\eta_{i,p}^{\text{stock}}, \eta_{i,p}^{\text{trsy}}$ are unconstrained real parameters. For each bond i in each period p , I estimate $(\eta_{i,p}^{\text{stock}}, \eta_{i,p}^{\text{trsy}}, b_{i,p}^{\text{stock}}, b_{i,p}^{\text{trsy}})$ jointly. Because the estimation uses overlapping windows, I compute autocorrelation-consistent standard errors with truncated kernel of eight lags (Zeileis (2004)). I then use the delta method to derive standard errors for $\text{Underreaction}_{i,p}^{\text{stock}}$ and $\text{Underreaction}_{i,p}^{\text{trsy}}$.

³⁰That is, the break-points are the 20%, 40%, 60%, and 80% percentiles of all the payoff-relevance measures $\{\sigma_1^{\text{stock}}, \dots, \sigma_n^{\text{stock}}, \sigma_1^{\text{trsy}}, \dots, \sigma_n^{\text{trsy}}\}$ where n is the number of bonds.

Panel A: Payoff-relevance proxies											
Relevance of interest rate risk (σ^{trsy})						Relevance of default risk (σ^{stock})					
σ^{stock}	σ^{trsy} quintile					σ^{stock}	σ^{trsy} quintile				
quintile	Low	2	3	4	High	quintile	Low	2	3	4	High
Low	1.24%	2.17%	3.23%	4.27%	7.23%	Low	0.83%	1.03%	1.25%	1.28%	1.39%
2	1.40%	2.23%	3.26%	4.42%	6.88%	2	2.19%	2.19%	2.20%	2.22%	2.27%
3	1.41%	2.26%	3.27%	4.58%	7.02%	3	3.24%	3.21%	3.20%	3.26%	3.24%
4	1.44%	2.28%	3.25%	4.64%	6.94%	4	4.70%	4.56%	4.64%	4.64%	4.71%
High	1.36%	2.34%	3.29%	4.65%	8.40%	High	7.52%	7.63%	8.30%	8.29%	10.17%

Panel B: Bond characteristics											
Duration (years)						Yield spread to treasuries					
σ^{stock}	σ^{trsy} quintile					σ^{stock}	σ^{trsy} quintile				
quintile	Low	2	3	4	High	quintile	Low	2	3	4	High
Low	2.28	2.69	4.26	6.18	13.42	Low	0.88%	0.90%	1.18%	1.14%	1.39%
2	3.67	3.52	4.45	6.13	13.00	2	2.58%	2.11%	1.70%	1.49%	1.46%
3	4.06	3.77	4.50	6.97	12.67	3	3.70%	2.76%	2.19%	2.14%	1.67%
4	4.46	4.10	4.86	6.59	11.34	4	3.97%	3.44%	3.01%	2.53%	2.08%
High	3.53	4.22	5.00	5.94	7.42	High	4.95%	4.46%	4.65%	4.23%	4.34%

Table 6. Characteristics of bonds by payoff-relevance of risks. Bonds are sorted into quintiles by the payoff-relevance of default risk (σ^{stock}) and interest rate risk (σ^{trsy}). These payoff-relevance measures, shown in Panel A, are computed as the (annualized) volatility of bond returns explained by stock and Treasury returns, respectively. Panel B shows that bonds with higher σ^{trsy} have longer duration and those with higher σ^{stock} have larger credit spreads.

Discarding inaccurate estimates

I discard cases where $\hat{b}_{i,p}^{\text{risk}} < 0$ because negative bond return sensitivity to interest rate and default risk shocks is hard to justify in theory. I also require the estimated standard error of $\text{Underreaction}_{i,p}^{\text{risk}}$ to be below 1. The cases of very large standard errors tend to happen when $b_{i,p}^{\text{risk}}$ is low, in which case the underreaction parameter is not well identified, as shown in Figure 12. Out of a total of 1,2421 bond-periods with at least 26 observations, 7,480 (60%) stock estimates and 9,014 (73%) Treasury estimates survive this filter. Figure 12 provides details about the distribution of estimates kept versus discarded. Figure 13 plots the histogram of underreaction estimates used.

GMM-based and regression-based estimates are similar

To verify the accuracy of GMM-based bond-level underreaction estimates, I check to see if they are similar to the regression-based, portfolio-level estimates. Because the latter rely on weaker parametric assumptions, this helps us check if our GMM parametric restrictions

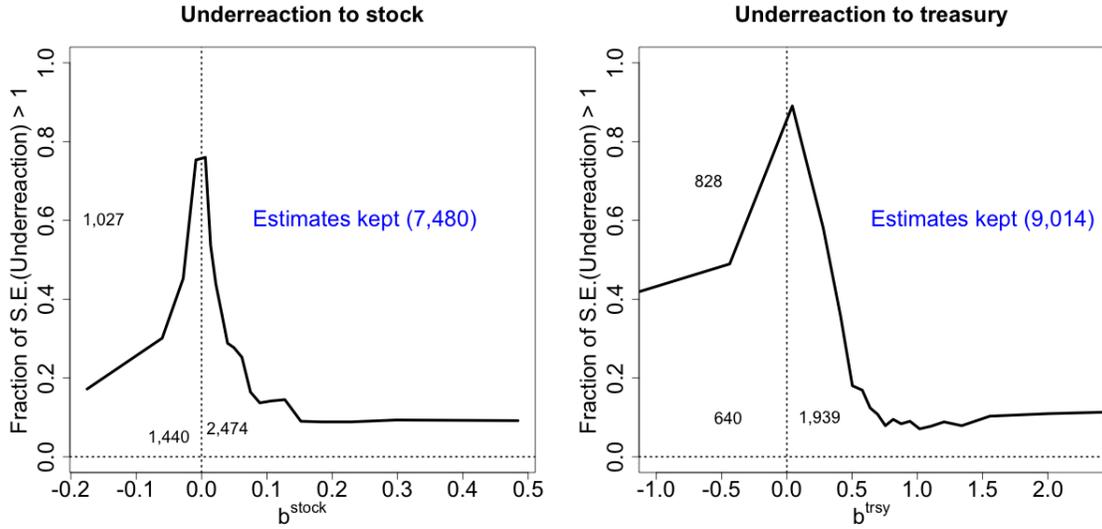


Figure 12. Illustration of the GMM estimates kept versus discarded. The figures plot the fractions of $\hat{\sigma}(\text{Underreaction}) > 1$ as a function of \hat{b} for stock risk (left) and Treasury risk (right). In the paper, I retain a GMM estimate only if $b > 0$ and $\hat{\sigma}(\text{Underreaction}_i) < 1$. Counts of retained estimates are marked in blue and those discarded are marked in black. For instance, in the left figure, 1,027 bonds have $b^{\text{stock}} < 0$ and $\hat{\sigma}(\text{Underreaction}^{\text{stock}}) < 1$; 1,440 bonds have $b^{\text{stock}} < 0$ and $\hat{\sigma}(\text{Underreaction}^{\text{stock}}) > 1$, etc.

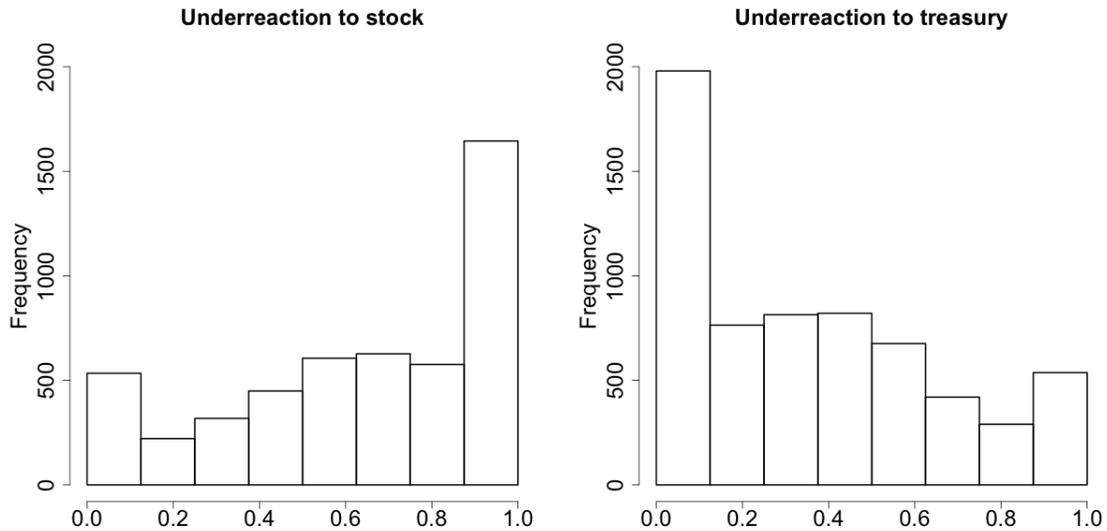


Figure 13. Histogram of GMM-estimated price underreaction.

are valid.

In Figure 14, I plot the average of GMM estimates for the 25 portfolios against the regression-based estimates. The green solid line is the 45% degree line and the dotted green lines are two standard deviation error bands. In almost all cases, the average of GMM

estimates are within two standard deviations of the portfolio-level estimates, indicating that estimates from the two methods are similar.

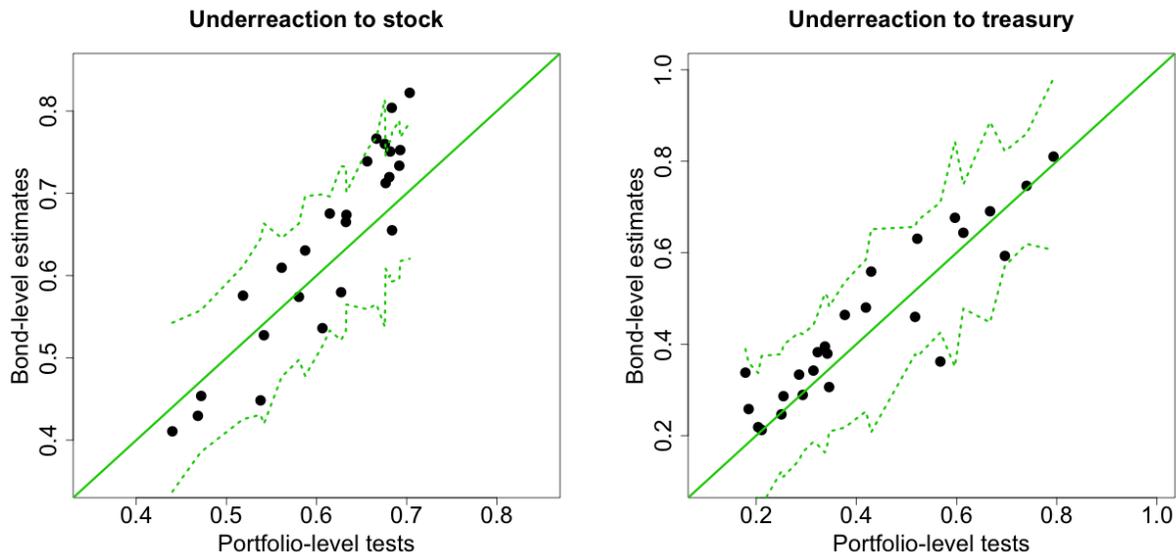


Figure 14. Comparing GMM-based and regression-based underreaction estimates. Bonds are double sorted into 5×5 portfolios using the payoff-relevance proxies (σ^{stock} and σ^{trsy}). For each of the 25 portfolios, I plot the average of bond-level GMM-based underreaction estimates in section 4.2 against the panel regression-based estimates in section 4.1. The green solid line is the 45% degree line and the green dotted lines are two standard deviation bands from the portfolio-level regressions.

C Additional results on alternative explanations

C.1 Testing for “distraction effect” using ex-post realized shock sizes

As explained in Appendix A.5, the key difference between rational inattention and transaction costs is that they generate opposite predictions about how importance of one risk impacts the price response to another. Under rational inattention, larger shocks from one risk consumes attention and increases underreaction to the other risk (“distraction effect”). Transaction cost generates the opposite.

To test this prediction, for each bond i in each week t , I separately measure the shocks

from two sources:

$$\text{shock}_{i,t}^{\text{stock}} = \hat{b}_i^{\text{stock}} \cdot \text{StockRet}_{i,t} \quad (39)$$

$$\text{shock}_{i,t}^{\text{trsy}} = \hat{b}_i^{\text{trsy}} \cdot \text{TrsyRet}_{i,t} \quad (40)$$

where $\hat{b}_i^{\text{stock}}, \hat{b}_i^{\text{trsy}}$ are bond-specific price response coefficients estimated using GMM. I then sort the sample by the sizes of the two shocks, $|\text{shock}_{i,t}^{\text{stock}}|$ and $|\text{shock}_{i,t}^{\text{trsy}}|$, into 3×3 bins. I use the same breakpoints to form the bins: $\{< 25\text{bp}, 25 - 75\text{bp}, > 75\text{bp}\}$.

After dividing observations into the 3×3 bins, each individual bond has too few observations. Therefore, I group together every 4 bonds with similar price coefficients into a “superbond.”³¹ I use GMM to estimate price underreactions for each superbond in each of the 3×3 bins, and then regress the underreaction estimates on shock size bin indicator:³²

$$\begin{aligned} \text{Underreaction}_{I,\Delta_s,\Delta_t}^{\text{stock}} &= \sum_{\Delta_s=1}^3 \sum_{\Delta_t=1}^3 \gamma_{\Delta_s,\Delta_t}^{\text{stock}} \cdot \left(\mathbf{1}_{|\text{shock}^{\text{stock}}| \text{ in bin } \Delta_s} \times \mathbf{1}_{|\text{shock}^{\text{trsy}}| \text{ in bin } \Delta_t} \right) \\ &+ \sum_I \eta_I^{\text{stock}} \cdot \mathbf{1}_{\text{superbond } I} + \epsilon_I^{\text{stock}} \end{aligned} \quad (41)$$

$$\begin{aligned} \text{Underreaction}_{I,\Delta_s,\Delta_t}^{\text{trsy}} &= \sum_{\Delta_s=1}^3 \sum_{\Delta_t=1}^3 \gamma_{\Delta_s,\Delta_t}^{\text{trsy}} \cdot \left(\mathbf{1}_{|\text{shock}^{\text{stock}}| \text{ in bin } \Delta_s} \times \mathbf{1}_{|\text{shock}^{\text{trsy}}| \text{ in bin } \Delta_t} \right) \\ &+ \sum_I \eta_I^{\text{trsy}} \cdot \mathbf{1}_{\text{superbond } I} + \epsilon_I^{\text{trsy}} \end{aligned} \quad (42)$$

where I denotes superbonds and $\Delta_s, \Delta_t \in \{1, 2, 3\}$ denote the stock- and Treasury-shock size bins. I cluster standard errors at the superbond level. The main coefficients of interest, $\gamma_{\Delta_s,\Delta_t}^{\text{stock}}$ and $\gamma_{\Delta_s,\Delta_t}^{\text{trsy}}$, are shown in the two panels of Table 7. The evidence supports the inattention explanation and rejects the transaction cost explanation. When $|\text{shock}^{\text{trsy}}|$ is larger, the price underreaction to Treasury is lower rather than higher. The same is true with how underreaction to stock returns varies with $|\text{shock}^{\text{trsy}}|$.

C.2 Can time-varying risk premium explain findings?

Time-varying risk premium can also potentially explain return predictability patterns. However, the price underreaction I document imply too high a Sharpe ratio for such expla-

³¹Specifically, let n denote the number of bonds. I double sort bonds by \hat{b}_i^{stock} and \hat{b}_i^{trsy} into $m \times m$ “superbond” portfolios where $m = \lfloor \sqrt{n/4} \rfloor$ so that there are roughly 4 bonds per “superbond.”

³²I apply the same data quality filters to the GMM estimates: the price reaction coefficient needs to be positive and the standard error of underreaction needs to be below 1.

Panel A. Underreaction to stock					
$ \text{shock}^{\text{stock}} $	$ \text{shock}^{\text{trsy}} $ bin				
bin	1	2	3	3 minus 1	t-stat
1	64%	65%	62%	-2%	(-0.95)
2	58%	62%	63%	5%***	(2.67)
3	50%	52%	57%	7%***	(4.33)
3 minus 1	-14%***	-12%***	-5%***		
t-stat	(-11.46)	(-7.55)	(-2.61)		

Panel B. Underreaction to Treasury					
$ \text{shock}^{\text{stock}} $	$ \text{shock}^{\text{trsy}} $ bin				
bin	1	2	3	3 minus 1	t-stat
1	65%	54%	55%	-10%***	(-9.13)
2	69%	56%	56%	-13%***	(-7.19)
3	80%	66%	63%	-17%***	(-6.16)
3 minus 1	15%***	13%***	8%***		
t-stat	(6.53)	(6.15)	(4.18)		

Table 7. The effects of stock and Treasury shock sizes on price underreaction. For each bond in each week, I estimate $\text{shock}^{\text{stock}}$ and $\text{shock}^{\text{trsy}}$ which are the bond value changes due to stock and Treasury returns, respectively. I then double sort weekly observations into 3×3 bins by $|\text{shock}^{\text{stock}}|$ and $|\text{shock}^{\text{trsy}}|$. The table reports average price underreaction for these 3×3 bins. Details are explained in regressions (41) and (42); Panels A and B show the estimated coefficients $\gamma_{\Delta_s, \Delta_t}^{\text{stock}}$ and $\gamma_{\Delta_s, \Delta_t}^{\text{trsy}}$, respectively.

nation to be plausible.

To see this, I consider a very simple trading strategy that uses the return predictions constructed in section 5.1. I only use lagged stock returns to forecast returns because lagged treasuries returns are shown to be not useful. In each week, the strategy goes long on bonds with top 10% return forecasts and short on bonds in bottom 10%. I equal weight all positions and adjust exposure such that the long and short sides each have a gross exposure of 100%.

Table 8 shows that this simple, equal-weighted strategy generates an annualized return of 7.8% with a Sharpe ratio of 2.61. Moreover, the returns are not spanned by well-known asset pricing factors. After adjusting for the Fama-French five factors (Fama and French (2015)) plus the momentum factor, the equal-weighted strategy alpha is still as large as the raw return. The results are only slightly weaker if I construct a value-weighted strategy, in which case the annualized return is 6.5% with a Sharpe ratio of 1.94.

Section 5.2 shows that the return predictability is within bid-ask spreads, so this hypothetical trading strategy cannot be implemented. However, time-varying risk premium explanations are about the variation of equilibrium mid-prices, so the fact that the implied

Sharpe ratio is too high is still strong evidence against such explanations.

Risk adjustment	Annualized alpha			
	Return	Volatility	T-statistic	Sharpe ratio
No Adjustment	7.82%	2.99%	8.68	2.61
CAPM	8.13%	2.92%	9.26	2.79
Fama-French 3 Factor	8.15%	2.90%	9.32	2.80
Fama-French-Cahart	8.20%	2.90%	9.40	2.83
Fama-French 5 Factor + Momentum	8.05%	2.89%	9.25	2.78

Table 8. Pre-transaction cost profits when trading on lagged stock information.

I use lagged stock returns to forecast bond returns and form a weekly rebalancing strategy that goes long on bonds with top 10% predicted returns and short on the bottom 10%. The first row reports the raw annualized strategy return, and the remaining rows report alphas when adjusting with various factor risk models.

D Why isn't it costless to learn interest rates?

In the paper, I argue that corporate bond investors do not adjust bond valuation one-to-one with duration-matched treasuries, and the exact adjustment sensitivity depends on contemporaneous shocks to credit risk. As empirical evidence for this view, I first show that corporate bond sensitivity to treasuries varies with the market level credit risk shock. To proxy for the amount of contemporaneous credit-related information, I use the size of the weekly change of five-year Markit CDX investment grade (IG) index spread. Because our Markit data starts in 2006, I only use data during 2006 - 2014. I then sort the sample period into quintiles by CDX spread change sizes, and regress bond returns on stock and duration-matched Treasury returns for each quintile:

$$\text{BondRet}_{i,t} = a + \beta^{\text{stock}} \cdot \text{StockRet}_{i,t} + \beta^{\text{trsy}} \cdot \text{TrsyRet}_{i,t} + \sum_i \gamma_i \cdot \mathbf{1}_{\text{bond } i} + u_{i,t}. \quad (43)$$

I then plot the variation of β^{trsy} across quintiles in the right panel of Figure 15. Standard errors are clustered by week and bond.

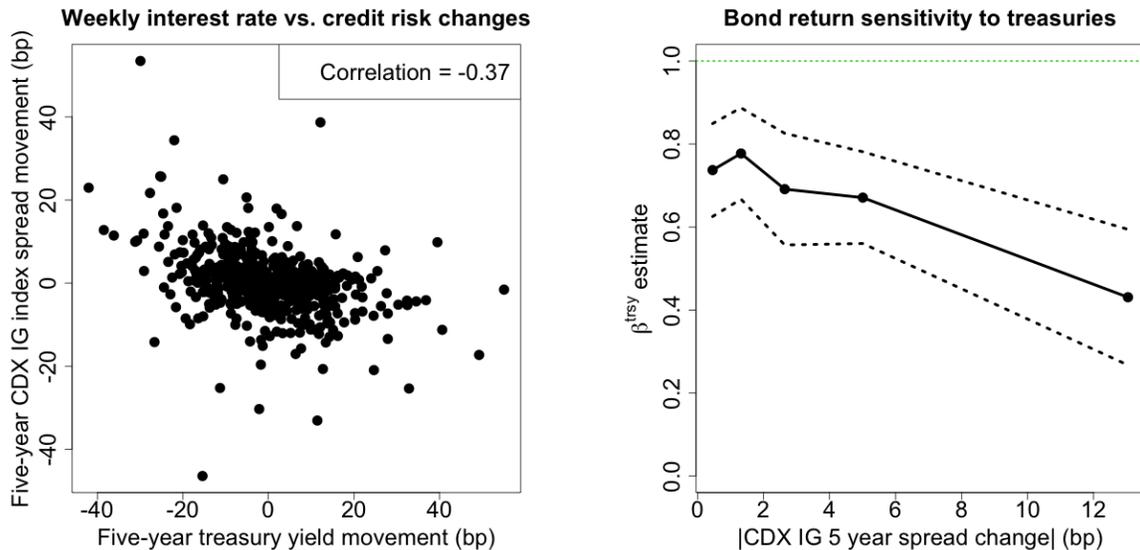


Figure 15. Explaining why bond returns do not move one-to-one with treasuries. **Left:** weekly changes in five-year CDX IG index spreads plotted against weekly changes in five-year Treasury yields. **Right:** bond return sensitivity to duration-matched treasuries as a function of CDX movements. I sort the sample into quintiles using the absolute value of the changes in five-year CDX IG index spreads. For each quintile, I regress weekly corporate bond returns on duration-matched Treasury returns and stock returns, and report the regression coefficients on treasuries. Dotted lines are two standard deviation bands and standard errors in the regression are clustered by week and bond.

As is clearly shown in the right panel of Figure 15, bond price sensitivity to treasuries is lower in periods when credit risk changes are larger. The difference between the first and last quintile is statistically significant with a t-statistic of 3.09. The fact that β^{trsy} is lower than 1 is consistent with the negative correlation between Treasury yields and credit spreads, shown in left panel of Figure 15. When Treasury yields rise (fall), credit spread narrows (widens), so bond yield moves less than one-for-one with Treasury yields.

While this is slightly outside the scope of our paper, I also use an approximation to show that this mechanism generates roughly the right magnitude of β^{trsy} deviation from unity. Consider a bond whose yield decomposes into r_t , the risk-free interest rate, plus c_t , the credit spread. Then, a regression of bond return on duration-matched Treasury return gives a regression coefficient of:

$$\frac{\text{Cov}(\Delta r_t + \Delta c_t, \Delta r_t)}{\text{Var}(\Delta r_t)} = 1 - \underbrace{\frac{\text{Cov}(\Delta c_t, \Delta r_t)}{\text{Var}(\Delta r_t)}}_{\text{Deviation from unity}} \quad (44)$$

Using five-year Treasury yield to proxy for r_t and five-year CDX IG index spread to proxy for c_t , I estimate the deviation term in (44) to be 0.23. The total deviation is 0.39 over the full sample. The remaining $0.39 - 0.23 = 0.16$ is explained away by slow price reaction to Treasury movements.³³

The negative correlation between Treasury yields and credit spreads also helps explain the variation of deviation across the five quintiles sorted by CDX spread changes. To see this, in Figure 16, I plot the actual deviation of β^{trsy} from unity against the estimated deviation using equation (44) across deciles.

³³In section 3.2, I estimate that 27% of Treasury-induced bond movements are delayed. This means that underreaction can explain approximate $0.61 \cdot \frac{27\%}{1-27\%} = 0.226$ of deviation.

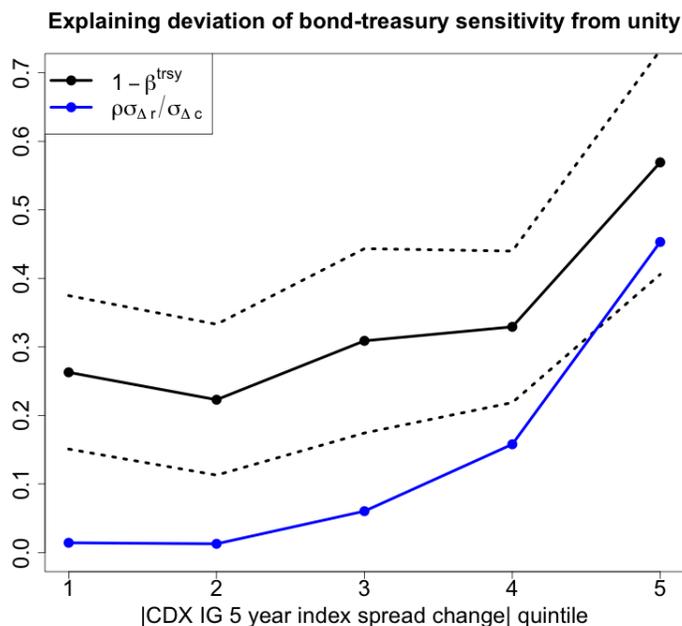


Figure 16. Explaining deviation of bond-Treasury sensitivity from unity. Data is sorted into quintiles using the size of five-year CDX IG spread changes. Black line plots one minus the estimated bond price sensitivity to duration-matched treasuries for each quintile, and the dotted lines are two standard error bands. The blue line plots the amount of deviation explained by the negative correlation between Treasury yields and credit spreads.