Learning, Price Discovery, and Macroeconomic Announcements

Haozhe Han, Grace Xing Hu, Calvin Dun Jia*

December 14, 2023

Abstract. We examine price discovery in China's equity market by leveraging a unique feature: unscheduled macroeconomic announcements that are often released outside of regular trading hours. This feature allows us to separate the impacts of investor learning and trading on price discovery. Our research shows that investor learning, proxied by the duration of non-trading periods and social media activity before trading, significantly speeds up post-announcement price discovery. However, announcements made during trading hours slow this process down, as learning and trading activities occur simultaneously. Our findings support theories that better-informed investors enhance the informativeness of prices, thereby underscoring the importance of investor learning in the process of price discovery.

Keywords: Learning, Macroeconomic Announcements, Price Informativeness, Market Efficiency

JEL Classification: G12, G14

---

*Han (haozhehan@tju.edu.cn) is from College of Management and Economics at Tianjin University; Hu (hux@pbcsf.tsinghua.edu.cn) is from PBC School of Finance at Tsinghua University; Jia (dun.jia@phbs.pku.edu.cn) is from HSBC Business School at Peking University. We benefit from discussions with Hengjie Ai, Zhi Da, Vivian Fang, Wei Jiang, Matthias Lassak, Kai Li, Dong Lou, Pavel Savor, Liyan Yang, Xiaoyan Zhang, and Hong Zhang. Errors are ours.

Conflict-of-interest disclosure statement

Haozhe Han

I have nothing to disclose.

Grace Xing Hu

I have nothing to disclose.

Calvin Dun Jia

I have nothing to disclose.

1 Introduction

Efficient price discovery matters the most for ensuring the informativeness of asset prices and for a well-functioning capital market (Fama, 1970). The existing literature predominantly emphasizes the central role of trading for facilitating the price discovery process in the stock market.[1] The key benefit of market trading is that equilibrium prices are formed through trading which aggregates public and private information across investors of varied types. Nonetheless, investors are also assumed to observe and learn from the market prices in addition to other informative signals in order to pin down their trading positions that feed into the market equilibrium (Grossman and Stiglitz (1980); Kyle (1985); Biais et al. (1999); and more recently, Goldstein and Yang (2015)). Therefore, studies on price discovery do not explicitly deal with the potential circularity that learning affects market trading that shifts asset prices which factor into the process of learning, i.e., a "reflection problem" as dubbed by Morris and Shin (2018). In other words, across a wide class of models and empirical explorations, both channels of learning and trading are jointly imposed and studied, which naturally results in their identified impacts on price discovery largely intertwined.[2]

Our paper, however, isolates and explicitly studies the channel of learning for shifting the process of price discovery in the stock market. Specifically, we examine the types of learning other than "learning from the price" and break down the reflection problem. We focus on the contribution of learning to price discovery before market trading has yet to generate the new equilibrium price that aggregates information across investors. To serve this purpose, we frame our study by focusing on China's stock market and exploit its important and unique institutional features. Macroeconomic announcements in China are often unscheduled and

[1] Boehmer and Wu (2012) and Beber and Pagano (2013) show that trading frictions such as constrained short sales prevent efficient price discovery. Barclay and Hendershott (2003, 2008) highlight that price discovery consistently benefits from trading during and outside the trading sessions and during the pre-opening period. Brogaard et al. (2014, 2019) identify the contributions of high-frequency traders for installing efficient price discovery.

[2] For example, to rationalize the role of trading for aggregating information, the framework of rational expectations equilibrium (REE) models with costly information acquisition is derived from a critical assumption that investors *immediately* act upon prices and informative signals through trading. In addition, the empirical evidence on price discovery is established based on the specification that asset prices could potentially react to important macro news within minutes upon data releases (e.g., Hu et al. (2017)).

arrive to the market with significant timing variations.[3] Many of these macroeconomic announcements fall within non-trading hours, during which investors do not have immediate access to trading.[4] In the absence of the real-time market prices as signals, investors could make an effort to learn during these information-sensitive hours – right after seeing an important macroeconomic announcement but before trading is possible. Therefore, China's market setting gives us the leverage to explore the impacts of "learning without trading" on price discovery through cross-event studies for causal inference. In advanced markets, however, the effects of learning and trading on price discovery are mixed, as investors trade immediately, often in high volume, following the release of macroeconomic news, utilizing either the equity or the almost around-the-clock derivatives markets. Therefore, our paper is of important general interest at least for, if not beyond, the emerging market countries when their financial market developments are still in progress.[5]

We provide several pieces of important empirical evidence in this paper. We first document significant variations in the timing of Chinese macroeconomic announcements. In our sample of 20 indicators, which includes the 15 most tracked Chinese macroeconomic indicators by Bloomberg and 5 monetary policy announcements by the People's Banks of China (PBOC) from 2009 to 2020, a total of 854 announcements were made outside trading hours, making up 45% of all our sampled events. The variations in the releasing time arise

[3]First, unlike the U.S. market, macro announcements in China that regularly release economic and financial statistics can fall outside regular trading hours. Second, most macro announcements in China do not follow a fixed and pre-scheduled timetable, and the day and time of data release vary substantially across announcements. For example, the National Bureau of Statistics of China publishes PMI data as early as 9:00 a.m. and as late as 8:00 p.m. China's central bank, the People's Bank of China (PBOC) may release China's monetary and financial statistics such as monetary aggregates and total social financing data after market close, before market open, or during the weekends (Guo et al., 2023).

[4]Brokerage firms and the stock exchanges in China do not accept, execute, or clear orders when markets are closed. In addition, China's derivatives markets are still considered underdeveloped, meaning that these markets are closed while the stock markets are closed. For example, the trading hours of stock exchanges in China and for trading stock index futures in China are the same.

[5]An incomplete list of emerging market countries that do not have a 24-hour domestic derivative market includes Argentina, Brazil, Chile, Colombia, Egypt, India, Indonesia, Mexico, Nigeria, Philippines, Russia, South Africa, Thailand, Turkey among many others. In addition, we see the statistical agencies in Brazil, India, South Africa, Turkey and others routinely release their GDP and CPI data outside the regular trading hours of their domestic stock exchanges.

from differences between macroeconomic indicators as well as changes within indicators, as a result of the lack of a fixed announcement release schedule. The release time of M2 announcements by the PBOC, for example, ranges from as early as 8:00 a.m. to as late as 8:00 p.m. Most importantly, the variations in the timing of these macroeconomic releases are not related to their information content (Guo et al., 2023). There is no correlation between announcement timing and the surprise component of macroeconomic announcements, nor does timing affect post-announcement market returns.

Exploiting the timing variations, we show that quicker and more efficient price discovery can be realized upon market opening if macro announcements are released overnight. Following Boguth et al. (2023), we rely on the R-squared (R_t^2) of an unbiased regression model that regresses the total announcement returns on the cumulative announcement returns to capture the price informativeness up to the time stamp t. For announcements made during non-trading hours, R_t^2 jumps immediately by 28% at the market opening time (9:30 a.m.), whereas the increase is substantially smaller, around 5%, in the first minute of trading following the release of macroeconomic news during trading hours. Clearly, announcements made during non-trading hours generate a significantly accelerated price discovery process as compared to the case when announcements are made during trading hours.

We then demonstrate that the faster price discovery associated with news releases made during non-trading hours is not driven by differences in the information content between groups nor by the presence of information aggregation in the pre-opening sessions. Rather, the speed and efficiency of price discovery are directly related to distance-to-trading, the calendar time elapsed between the release time and the market opening time, which equals zero for trading-hours news. In a regression setup that controls the overall market impact, we find that a one-day increase in distance-to-trading corresponds to a 7.0% increase in price informativeness at the first minute of trading, with a significant t-value of 6.57. The results remain robust if we restrict our sample to exclusively non-trading-hours news: a one-day increase in distance-to-trading leads to a 3.0% increase in price informativeness at 9:30 a.m., with a significant t-value of 3.16.

We further show that investors' learning, particularly among retail investors who can learn from the market prices when there is market trading, is an important mechanism explaining why longer distance-to-trading could speed up price discovery. Theoretically, a wide class of models highlight the importance of investors' learning as triggered by increased attention, which improves price informativeness in a market equilibrium (Peng and Xiong (2006); Kacperczyk et al. (2016); and more recently, Kacperczyk et al. (2023)). Following Ai et al. (2023) that highlights the active role of learning among less informed investors, we therefore hypothesize that the intensity of retail investors' learning increases with the distance-to-trading. In the spirit of Da et al. (2011), Ben-rephael et al. (2021) and Fisher et al. (2022), we construct measures of retail investors' learning efforts counting the number of posts, fans, and the interaction intensity related to the released macroeconomic news on China's largest social media platform, Weibo Inc. We confirm that a longer duration of distance-to-trading is associated with the degree of pre-trading learning intensity. We then show that allowing investors to digest and process newly released news information overnight before trading leads to faster post-announcement price discovery once the market resumes trading. Specifically, at the first minute when investors can trade on the news, a one standard deviation increase in pre-trading posts (78 posts), fans following these posts (138 million), and retweets, comments, and likes following these posts (128 retweets, comments, and likes) are associated with 5.2%, 5.8%, and 4.1% faster price discovery.

Leveraging the results of Dugast and Foucault (2018) and Ai et al. (2022) that learning takes time and information quality may improve over time because of learning, we further test the hypotheses related to the outcomes of a learning channel. We examine the theoretical predictions that retail investors' learning can effectively enhance average information quality in the market upon opening and shrink the information gaps between informed and less informed investors.[6] Indeed, we find that both realized return volatility and turnover jump

[6]In the Appendix, we extend the baseline model with three periods per Vayanos and Wang (2012) by allowing less informed investors to carry more informative signals over time. Two different types of liquidity supplying traders differ in their level of sophistication in processing the macro announcement, and they jointly accommodate the trading needs of liquidity demanding traders. A longer distance-to-trading over non-trading hours gives less informed traders more time to learn and extract useful information from a macro announcement before trading. As a result, upon market opening, a larger number of liquidity supplying traders, who have already incorporated the news announcements, are better informed for trading.

upon market opening, and the size of the jumps is correlated with the intensity of pre-trading retail investors' learning. After the initial jumps, both return volatility and turnover become lower for announcements with more pre-trading investor learning. This evidence supports the volatility trade-off driven by improvement in market information quality, as noted in Ai et al. (2023). In addition, we document that pre-trading learning increasingly mitigates the degree of information asymmetry across investors, as measured by the average bid-ask spreads across stocks Garfinkel (2009). All of this evidence emphasizes the importance of pre-trading learning among retail investors, through which greater price informativeness can be achieved once market trading resumes.

Additionally, we provide an extensive set of additional evidence and show that our empirical findings are robust. Similar results are obtained for 1) a matched sample of macroeconomic announcements that have comparable post-announcement market impacts but differ in the timing of releases; and 2) a sub-sample of announcements with fixed types of macroeconomic indicators. These results further mitigate the concern that the release time of macroeconomic indicators is endogenously determined by their information content. We also show that China's stock market surpasses its U.S. counterpart by exhibiting quicker price discovery upon market opening if the macro announcements are made during off-hours. Hence, in light of all the frictions that negatively affect the Chinese stock market performance (Allen et al., 2023), our findings suggest that retail investors can still significantly benefit from the timing arrangements of macro announcements in China. Overnight learning can therefore effectively reduce the sizeable noise component in the trading of a large number of retail investors. As retail investors are more likely to learn about market fundamentals, this reinstates at least partially the social welfare of the Chinese economy (Brunnermeier et al., 2021).

Our paper underscores the unique impact of a pure learning channel on price discovery during market closure. By exploiting the institutional features of China's market setting, our

In contrast, if an announcement is made within trading sessions, only those sophisticated traders who are immediately well informed of the macro news are able to respond through trading right after the arrival of the announcement.

paper makes an important contribution in terms of our identification scheme. As all learning investors in our setting won't observe the equilibrium market prices until the market reopens, our identifications are immune from the confounding effects on price discovery driven by "learning from the price" or the "information paradox" that increased price informativeness can deter continuous learning among less informed investors (Grossman and Stiglitz, 1980). Hence, our paper demonstrates that pre-trading learning is a separate and under-explored mechanism that prepares for efficient price discovery once trading is resumed.

Related Literature. Our paper is related to several strands of literature. First, the existing literature stresses the critical importance of market trading for installing the information aggregation process in order to obtain efficient price discovery. Biais et al. (1999) identify the benefits of having pre-opening sessions in the stock market for efficient price discovery. Barclay and Hendershott (2003) find that the largest fraction of price discovery is achieved through day trading, though price discovery can be quicker and more efficient during the pre-opening sessions. Barclay and Hendershott (2008) document that it is important to have high trading volume in the pre-trading sessions so that the opening price is more efficient and will result in a greater degree of price discovery before market opening. Brogaard et al. (2014, 2019) find that price discovery benefits from high-frequency traders who can submit limit orders with and information advantage. While the channels of learning and trading are jointly studied in the literature, little is known if information improvement before pre-opening sessions is good enough for generating sizable price discovery upon market opening. Our paper exploits the unique market setting in China and is the first to identify and highlight the importance of the learning channel on price discovery conditional on closed market trading. In particular, our empirical setting well isolates the effects of learning from all information sources other than market prices. Our paper also provides a learning-based interpretation on why pre-opening hours are critical for enhancing price discovery.

Second, an important stream of literature is devoted to studying the risk and return profiles of stocks in response to macroeconomic announcements. Savor and Wilson (2013, 2014) first document that the U.S. equity market exhibits larger excess returns on days of data releases for inflation, unemployment, and various interest rates. Lucca and Moench (2015) detect a pre-announcement drift of the equity premium before the FOMC statement release. Ai and Bansal (2018) and Ai et al. (2022) theoretically show the asset pricing implications of recursive preferences in windows of macroeconomic announcements. Cieslak et al. (2019) find that the equity premium realized before and on FOMC days is part of a larger FOMC premium cycle. Hu et al. (2022) emphasize the heightening and subsequent reduction of market uncertainty before the FOMC announcements. Brusa et al. (2020) show that the stock markets of 35 countries all exhibit strong reactions to the FOMC announcements. Boguth et al. (2023) find that equity prices following FOMC announcements are less informative about future indicative prices. Guo et al. (2023) first exploit the fact that macro announcements in China can randomly arrive to the markets with significant timing variations. Our paper is the first to show that post-macro announcement stock market dynamics can be significantly affected by the length of duration of an information-sensitive period, which is after the announcement arrival but before market opening (i.e., the distance-to-trading). Our results show that, in our setting, investors learning during non-trading hours – rather than the trading itself – is what generates sizable price discovery upon the release of macroeconomic data once market trading resumes.

Lastly, our paper contributes to the series of work that evaluates the efficiency of financial markets in China. Based on earlier data, Allen et al. (2005) and Allen et al. (2012) provide comprehensive overviews of China's financial system and conclude that China's stock markets are less efficient given its stock prices are not reflective of fundamental values of listed firms. Geng and Pan (2019) find improved price efficiency in China's bond markets but at a cost of the increasing the divergence of the cost of borrowing between state-owned and non-state-owned firms. Carpenter et al. (2021) document that China's stock markets have become increasingly efficient in the way that stock prices are informative about firms' future profits and the market is effectively aggregating the relevant information. Liu et al. (2019) construct the relevant size and value factors for valuing stocks in China. He and Wei (2023) and Hu and Wang (2022) provide very detailed reviews of major markets in China using more recent data. They show that China's capital markets have experienced significant growth and development in recent years. In addition to studying price discovery for general implications, our paper shows that the institutional design for releasing important economic and financial data in China can be exploited to evaluate the general price informativeness in China's stock markets. Interestingly, our results suggest that a longer distance-to-trading helps increase the information advantage among those otherwise less informed investors, which results in enhanced price efficiency upon market opening. In addition, we highlight that our paper is also the first to systematically summarize the impacts of a comprehensive list of macro announcements in China on its stock markets.

The rest of the paper is structured as follows. We summarize our data on macroeconomic announcements in China and introduce the institutional background in Section 2. In Section 3, we document our main findings regarding the impacts of distance-to-trading on price discovery. Section 4 discusses why our findings can be rationalized by the channel of retail investors' learning. We provide additional robustness tests in Section 5. Finally, Section 6 concludes. In the Appendix and Internet Appendix, we provide additional theoretical results and empirical evidence.

2 Data

2.1 Macroeconomic Announcements in China

We consider the release of major macroeconomic indicators in China. To pinpoint important market-moving macroeconomic indicators in China, we first include 15 Chinese macroeconomic indicators covered by the Bloomberg relevance scores. These scores track the number of subscriptions on the Bloomberg terminal and thereby represent the top China macroeconomic indexes that investors pay attention to. These 15 economic indicators are: the consumer price index and producer price index (CPI/PPI), gross domestic production (GDP), purchasing managers' index (PMI), Caixin China purchasing managers' index (Caixin), industrial production (IP), broad money supply (M2), trade balance (Trade), foreign exchange reserves (FER), required reserve ratio (RRR), profit of industrial enterprises (PI), foreign direct investment (FDI), balance of payments (BOP), Swift global payments CNY (Swift), sales prices of residential buildings (SPRB), and foreign exchange settlement and sales by banks (FESS). We supplement the set of indicators with 5 additional monetary policy announcements from the PBOC: open market operations of medium-term lending facility (OMO), monthly summary of standing lending facility/medium-term lending facility/pledged supplementary lending operations (SLF/MLF/PSL), central treasury cash management (CTCM), central bank bills swap (CBS), and loan prime rate (LPR).[7]

We collect the release date and time of the macroeconomic announcements between January 2009 and December 2020 from the Bloomberg terminal and the website of the PBOC. The announcement times are with minute-level time stamps. Unlike many developed countries, most macroeconomic announcements in China do not follow a fixed timetable, and the actual release calendar date and time may vary substantially between announcements. In Table 1, we report the timing distribution of the 20 major macroeconomic indicators covered in our sample. It shows that the earliest macroeconomic announcement in our sample arrives on a day at 6:40 a.m., the median is at 10:00 a.m., and the latest is at 9:07 p.m.

Based on the release time, we divide the macroeconomic announcements into two groups: those within regular trading hours and those outside of regular trading hours. In our sample, 1,044 announcements are made during regular trading hours, while 854 announcements are made during non-trading hours. There are three scenarios for announcements made outside of trading hours: 1) 217 announcements are made before the stock market opens (9:30 a.m.) on a trading day; 2) 421 announcements are made after the stock market closes (3:00 p.m.) on a trading day; and 3) 216 announcements are made on a non-trading day, which includes both weekends and holidays.[8]

[7]The People's Bank of China employs many monetary policy tools, and we only include the ones that have a regular releasing schedule and contain information that can potentially move the market.

[8]Fifteen macroeconomic announcements are released during the noon break, between 11:30 a.m. and 1:00 p.m., in our sample period. We exclude these announcements in our analysis.

Table 1: Release Time of Major Macroeconomic Indicators in China

Announcement	MinT	MedT	MaxT	#Trd	#NonTrd	#Open	#Close	#Weekend/Holiday	Score	Source
CPI/PPI	9:30	9:30	13:30	122	22	0	0	22	98	National Bureau of Statistics
GDP	10:00	10:00	15:00	47	1	0	1	0	96	National Bureau of Statistics
PMI	9:00	9:00	20:00	0	179	113	1	65	94	Federation of Logistics & Purchasing
Caixin	9:45	9:45	10:30	308	11	0	0	11	92	Markit
IP	10:00	10:00	15:40	109	14	0	3	11	88	National Bureau of Statistics
M2	8:00	16:00	20:00	33	111	6	94	11	86	The People's Bank of China
Trade	9:32	10:58	17:30	96	33	0	3	30	82	General Administration of Customs
FER	8:00	16:00	18:27	11	77	6	51	20	69	The People's Bank of China
PI	9:30	9:30	11:00	85	28	0	0	28	51	National Bureau of Statistics
RRR	12:12	18:06	20:01	0	26	0	18	8	41	The People's Bank of China
FDI	6:40	10:16	20:30	90	46	2	41	3	36	Ministry of Commerce
BOP	14:30	16:46	19:05	1	37	0	37	0	35	State Administration of Foreign Exchange
Swift	9:00	9:00	21:07	1	60	58	2	0	33	SWIFT
SPRB	9:30	9:30	9:30	30	3	0	0	3	29	National Bureau of Statistics
FESS	10:00	15:53	19:52	10	34	0	34	0	27	State Administration of Foreign Exchange
OMO	9:10	9:46	9:46	46	8	8	0	0	-	The People's Bank of China
SLF/MLF/PSL	9:23	15:51	19:12	11	62	1	57	4	-	The People's Bank of China
CTCM	7:24	16:31	19:45	27	83	4	79	0	-	The People's Bank of China
CBS	9:00	9:00	9:00	0	19	19	0	0	-	The People's Bank of China
LPR	9:30	9:30	9:30	17	0	0	0	0	-	The People's Bank of China
All	6:40	10:00	21:07	1044	854	217	421	216		

Notes: This table reports the summary statistics on the release time of 20 major macroeconomic indicators in China: 15 indicators with Bloomberg subscription scores and 5 monetary policy tools of the People's Bank of China. "MinT", "MedT", and "MaxT" refer to the minimum, median, and maximum of the release time. "#Trd" refers to the number of announcements released during trading hours. "#NonTrd" refers to the number of announcements released during non-trading hours. "#Open" refers to the number of announcements released before the stock market opens (9:30 a.m.) on a trading day. "#Close" refers to the number of announcements released after the stock market closes (3:00 p.m.) on a trading day. "#Weekend/Holiday" refers to the number of announcements released on weekends and holidays. "Score" refers to the subscription scores by Bloomberg. "Source" is the official releasing entity of the indicator. The sample period is from January 2009 to December 2020.

2.2 The Financial Markets and Trading Hours

Along with its fast economic development, China's financial markets have grown tremendously in recent years. Established in 1990, the two stock exchanges, the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE), are now second globally in terms of market capitalization, second only to the United States. Despite their large size and growing importance, China's financial markets, the derivatives market in particular, remain largely underdeveloped. The Chinese financial futures and options markets, launched in 2010 and 2015, respectively, have a very short history and are significantly smaller compared to its own stock market and the derivatives market in other developed countries.[9]

Stock trading can only take place on the two stock exchanges during regular trading hours in China. Regular trading hours include two sessions: the morning session from 9:30 a.m. to 11:30 a.m., and the afternoon session from 1:00 p.m. to 3:00 p.m. The total trading session is therefore only four hours (240 minutes) per day, shorter than most of the developed markets. Before the market opens, there is a pre-opening auction session that runs from 9:15 to 9:25 a.m., during which orders are placed in advance and an opening price of the stock is decided based on a call auction process. There is no after-hours trading session in China.[10] Moreover, financial futures and options can only be traded during regular trading hours in China, in contrast to many developed countries (including the U.S.), where the derivatives markets are open almost 24 hours around the clock.[11]

[9] Interested readers can refer to Hu and Wang (2022) for a review on the development and characteristics of the financial derivatives market in China.

[10] In addition to the main board on Shanghai and Shenzhen stock exchanges, ChinNext and The Science and Technology Innovation Board (STAR) are two separate boards of the stock markets in China established in 2009 and 2019, respectively. As part of the Shenzhen Stock Exchange, ChinNext is focused on small and medium-sized enterprises in innovative and high-growth industries. The companies listed on ChinNext are often in emerging industries. STAR is part of the Shanghai Stock Exchange and is designed to support the development of innovative companies in the country. The STAR market focuses on high-tech and strategic emerging industries. For stocks listed on ChinNext and STAR, there is an after-hours fixed-price trading session from 3:05 p.m. to 3:30 p.m.

[11] CSI 300 index futures were launched on April 16, 2010, and were traded from 9:15 a.m. to 11:30 a.m. and 1:00 p.m. to 3:15 p.m. from 2010 to 2015. After 2016, the trading hours for the CSI 300 index futures were changed to 9:30 a.m. to 11:30 a.m. and 1:00 p.m. to 3:00 p.m. CSI 300 index options were launched much later, on December 23, 2019, and are traded from 9:30 a.m. to 11:30 a.m. and 1:00 p.m. to 3:00 p.m.

The under-developed financial markets, coupled with short trading hours, make China a unique laboratory for studying the information transmission mechanism when trading is not available at the time information arrives to the market. For macroeconomic announcements released during non-trading hours, Chinese investors have to wait until the stock market opens for trading because both the stock and derivatives markets are close. This setup is very different from the way macroeconomic announcements are released in other developed countries. In the U.S., for example, important macroeconomic announcements are released either shortly before the stock market opens (non-farm payroll, GDP, CPI, etc.), from 8:30 a.m. to 9:15 a.m. eastern time, or within regular trading hours (FOMC, ISM, CSI, etc.). U.S. investors could therefore immediately trade on the news using, for instance, the market index futures contracts that are open to trade almost around-the-clock or the market index ETFs that are actively traded during both trading and pre-trading hours.

In summary, we can characterize the environment of interest in which a macroeconomic announcement arrives in between regular trading hours using a timeline. Specifically, according to Figure 1, an announcement is released outside the trading session. That is, it is released to the public after a regular trading day $t-1$ and before the next trading day t. In addition, we highlight a period of time to denote the duration of time between the arrival of an announcement and the beginning of the next trading session (i.e., distance-to-trading). The distance-to-trading is of great interest of our paper as it captures the duration of hours during which investors cannot observe and learn from asset prices for investment decisions. In particular, while the market is yet to incorporate the information of the newly arrived macroeconomic announcements released overnight, it is important to examine the impact of investors' learning other from the price overnight on price discovery once market resumes trading on the next trading day.

2.3 Market Responses on Macroeconomic Announcement Days

Our main empirical results are based on the return of the CSI 300 index, the capitalization-weighted index tracking the performance of the 300 largest stocks listed on the Shanghai

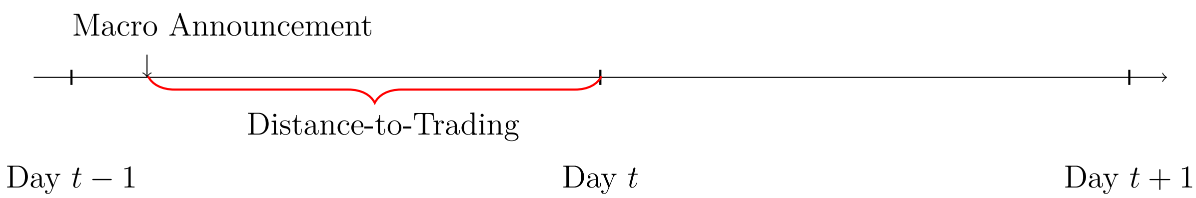


Figure 1: A Timeline

Stock Exchange and the Shenzhen Stock Exchange. We obtain the high-frequency intraday tick data of the index, including both price and volume, from the RESSET Financial Database. The CSI 300 index data are with second-level time stamps and available for every five-second time interval from January 2009 to December 2020. We complement the index data with high-frequency tick data for its constituent stocks, provided by RESSET and with minute-by-minute time stamps for each one-minute interval from January 2009 through December 2020. Based on the high-frequency data, we calculate the returns for the CSI 300 index and its constituent stocks at a daily frequency as well as for different time windows around the release of macroeconomic announcements. We winsorize returns at the 1% and 99% levels to mitigate the potential large impact driven by a few extreme values.

We report the summary statistics of the daily market returns on the macroeconomic announcement days in Table 2. For the announcements released during regular trading hours, the daily market returns are calculated as the percentage returns of the CSI 300 index from the close (3 p.m.) of the previous trading day to the close (3 p.m.) of the announcement day. For the announcements made outside trading hours, the daily market returns are calculated as the percentage returns of the CSI 300 index from the close of the previous trading day to the close of the next trading day after the release time. When multiple announcements occur within the same trading window, we consolidate them and report a single daily return for these announcements. In total, we have 1,318 unique daily market returns for the 1,898 announcements from January 2009 to December 2020.

The average market return is 5.15 basis points (bps) on the announcement days, which is positive but not statistically significant with a t-value of 1.35. Out of the 20 announcements tracked in our sample, the PMI announcement days elicit the most positive market reactions, yielding an average of 34.20 bps, which is significant at the 5% level. While the market returns are large on several announcement days, specifically positive on SLF/MLF/PSL, LPR, and Trade release days, and negative on OMO, RRR, and Swift release days, none of these results are statistically significant. The overall daily market returns are close to zero on the non-announcement days, with an average of -0.77 bps with a small t-value of -0.22.

Table 2: Daily Market Returns and Distance-to-Trading on Macroeconomic Announcement Days

		Market Return (daily)			Distance-to-Trading (Dur)	
	NDays	Mean	Std	TStat	Mean	Std
Ann Days	1318	5.15	138	1.35	0.49	0.95
CPI/PPI	144	-1.50	142	-0.13	0.29	0.76
GDP	48	-12.95	150	-0.60	0.02	0.11
PMI	179	34.20	150	3.04	0.96	1.78
Caixin	319	14.43	144	1.79	0.20	1.16
IP	123	-1.41	131	-0.12	0.19	0.61
M2	144	12.84	134	1.15	1.15	1.30
Trade	129	16.89	138	1.39	0.42	0.83
FER	88	4.56	135	0.32	1.01	1.09
PI	113	-1.43	145	-0.10	0.39	0.83
RRR	26	-24.99	161	-0.79	1.73	1.85
FDI	136	-11.85	140	-0.99	0.40	0.81
BOP	38	7.08	163	0.27	2.34	2.14
Swift	61	-20.12	110	-1.43	0.04	0.10
SPRB	33	5.72	155	0.21	0.18	0.58
FESS	44	4.59	129	0.24	1.23	1.11
OMO	54	-28.44	121	-1.72	0.00	0.00
SLF/MLF/PSL	73	26.64	135	1.68	1.06	1.35
CTCM	110	-19.04	138	-1.44	0.64	0.71
CBS	19	3.97	133	0.13	0.02	0.00
LPR	17	23.01	124	0.77	0.00	0.00
Non-ann Days	1600	-0.77	139	-0.22		

Notes: This table reports the summary statistics of market returns and distance-to-trading on the macroeconomic announcement days and other days in China. The market returns are the average log returns of CSI 300 and are in basis points. Dur is the time (in unit of calendar days) between announcement time and the first trading time after the announcement. "Ann Days" refers to the announcement days, and "Non-ann Days" refers to the trading day without announcements. The sample period is from January 2009 to December 2020.

Table 2 also shows large variations in the distance-to-trading among macroeconomic announcements in China. We measure the distance-to-trading (Dur) for each announcement as the number of calendar days between the actual release time and the first instance that investors can trade, which equals zero for announcements released during trading hours. For the 1,318 macroeconomic announcement days in our sample, the average distance-to-trading is 0.49, with a large standard deviation of 0.95. The large variations in distance-to-trading are a result of differences across announcement types as well as variations within fixed macroeconomic announcement types.

To further pin down the magnitudes of market reactions on macroeconomic announcement days, we group announcement days based on the surprise component of the releases and investigate the corresponding market returns. The index surprise δ is calculated as the difference between the actual release and the median of Bloomberg economists' forecasts, normalized by its full-sample standard deviation. Because of Bloomberg coverage limits, we can only calculate index surprises for CPI/PPI, GDP, PMI, Caixin, IP, M2, Trade, FER, and FDI announcements.[12] Based on the index surprise, we divide the announcement days into three different groups: the bad news group with δ less than -0.5, the neutral news group with δ between -0.5 and 0.5, and the good news group with δ larger than 0.5.

We observe significant market movement on the macroeconomic announcement days that deliver unexpected information, as presented in Panel A of Table 3. The average returns on these announcement days are -20.27 bps for the bad news group and 26.14 bps for the good news group. The return differences between groups of good and neutral news and between groups of neutral and bad news are significant. The majority of these market reactions occur after the release of the macro news. The average post-announcement return, $R^{[0,239]}$, which captures the 240-minute (4 trading hours) return after the regular trading hours announcement and the 240-minute return (4 trading hours) after 9:30 a.m. of the following trading day for after-hours announcements, is -23.69 bps and 23.66 bps for the bad and good news groups, respectively. Both numbers and their between group differences are statistically significant.

[12]For the CPI/PPI announcement days, we calculate the index surprises based on the surprise component of the CPI release and forecast, considering that the CPI receives high Bloomberg relevance scores. It is also worth noting that the market may not always view higher than expected CPI as good news. In unreported robustness tests, we exclude the CPI/PPI announcement days in our sample and obtain results similar to those reported in Table 3.

Conversely, no significant pre-announcement returns are observed before macroeconomic announcements. The average pre-announcement return, $R^{[-240,-1]}$, measuring the 240-minute return before the release time for announcements made during regular trading hours and the 240-minute return before 3 p.m. of the previous trading day for after-hours announcements, stands at -3.49 bps and 11.44 bps for the bad and good news groups, respectively. Neither of these numbers are statistically significant. As expected, for neutral announcements with no surprising information, no significant market movements are observed either prior to or following the announcements.

Putting all the evidence together, it is clear that macroeconomic announcements in China carry important informational content and could result in substantial price movement in the equity market. This observation aligns with the findings from our regression analysis where the returns of different announcement windows are regressed on the surprise component of macroeconomic announcements. As shown in Panel B of Table 3, a one-unit increment in the index surprise δ is associated with an increase of approximately 15 bps in the post-announcement returns, yet it exhibits no significant effect on the pre-announcement returns.

It's also worth emphasizing that we observe no significant difference in the release timing among macroeconomic announcements with varying information content.[13] For both the bad and good news groups, the average distance-to-trading is around 0.59 days (or 850 minutes), and it does not statistically differ from that of the neutral group. In a regression setup, we also find that the distance-to-trading does not affect the announcement returns of macroeconomic releases. In other words, the information content and the market response to macroeconomic announcements are not directly associated with the timing of their release.

[13]When multiple announcements are made on the same day, we only consider the distance-to-trading for the announcements with the highest Bloomberg relevance score.

Table 3: Macroeconomic Index Surprises and Announcement Returns

Panel A: Market returns for announcements sorted based on index surprises

	Bad $\delta < -0.5$	Neutral $-0.5 \leq \delta \leq 0.5$	Good $\delta > 0.5$	B−N	G−N
δ	-1.15***	0.01	1.07***		
	[-28.23]	[0.54]	[15.16]		
Dur	0.59***	0.64***	0.59***	-0.05	-0.05
	[7.25]	[11.25]	[6.42]	[-0.46]	[-0.47]
R^{daily}	-20.27**	8.86	26.14***	-29.13***	17.28*
	[-2.17]	[1.57]	[2.98]	[-2.69]	[1.67]
$R^{[-240,-1]}$	-3.49	8.07	11.44	-11.56	3.37
	[-0.40]	[1.54]	[1.44]	[-1.15]	[0.35]
$R^{[0,239]}$	-23.69***	0.93	23.66***	-24.61**	22.73**
	[-2.64]	[0.16]	[2.60]	[-2.21]	[2.10]
N	182	485	207		

Panel B: The impact of index surprise on market returns

	R^{daily}		$R^{[-240,-1]}$		$R^{[0,239]}$	
	(1)	(2)	(3)	(4)	(5)	(6)
δ	11.63***	11.55***	3.83	3.97	14.56***	14.43***
	[2.60]	[2.60]	[0.88]	[0.91]	[3.17]	[3.20]
Dur	5.71	4.83	0.03	0.52	7.96*	5.13
	[1.33]	[1.03]	[0.01]	[0.17]	[1.83]	[1.09]
Constant	3.19	3.74	6.38	6.07	-3.93	-2.20
	[0.70]	[0.81]	[1.44]	[1.34]	[-0.83]	[-0.46]
Index FE		Yes		Yes		Yes
R^2	0.01	0.02	0.00	0.00	0.02	0.03
N	874	874	874	874	874	874

Notes: Panel A reports the summary statistics of market returns for sorted groups on macroeconomic announcement days. The announcement days are sorted into bad news ($\delta < -0.5$), neutral news ($-0.5 \leq \delta \leq 0.5$), and good news ($\delta > 0.5$) grouped by index surprise δ. The index surprise δ is calculated as the difference between the actual release and the median of Bloomberg economists' forecasts, normalized by its standard deviation. The market returns are the average log returns of CSI 300 and are in basis points. "B-N" and "G-N" indicate the difference between the bad and neutral group and the good and neutral group, respectively. R^{daily} refers to the daily return on announcement days. $R^{[-240,0)}$ and $R^{[0,239)}$ refer to returns from the beginning of minute "-240" to the end of minute "-1", and from the beginning of minute "0" to the end of minute "239". The minute "0" is the opening time of the stock market (9:30 a.m.) for announcements released during non-trading hours and is the actual announcement time for announcements released during trading hours. Dur is the time (in unit of calendar days) between announcement time and the first trading time after the announcement. Panel B reports the regression results of $R_i = \alpha + \beta_1 \delta_i + \beta_2 Dur_i + \epsilon_i$. The sample period is from January 2009 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

3 Baseline Results

3.1 The speed of price discovery: non-trading-hours and trading-hours announcements

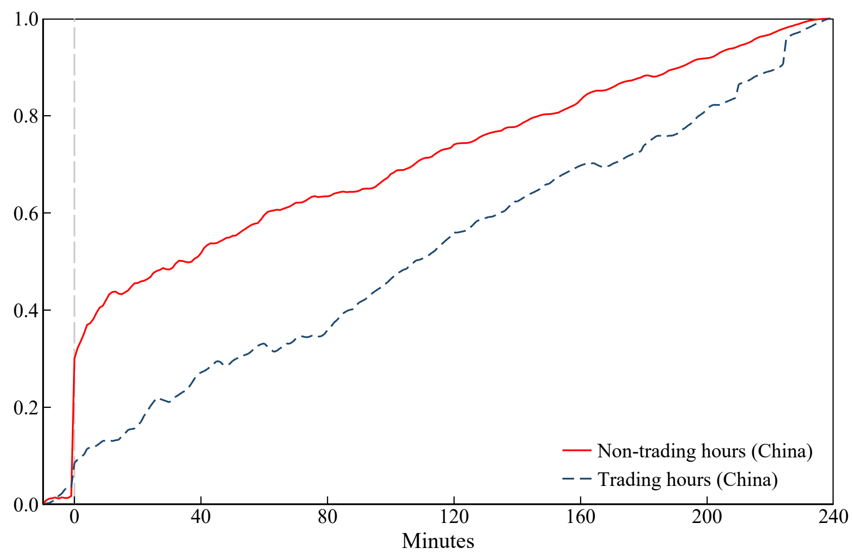
To investigate the impact of distance-to-trading on the speed of price discovery, we rely on an unbiased regression model similar to Biais et al. (1999) and Boguth et al. (2023). Formally, for given $-10 \leq t \leq 239$, we regress the total returns surrounding the announcements' release time on the cumulative announcement returns ending at time t:

$$R_i^{[-10,239]} = \alpha_t + \beta_t R_i^{[-10,t]} + \epsilon_{i,t}, \tag{1}$$

where $R_i^{[-10,t]}$ denotes the cumulative return of the CSI 300 index from 10 minutes before the release to the time t after the release for the announcement i; $R_i^{[-10,239]}$ denotes the cumulative return of the CSI 300 index from 10 minutes before the release time to the 240th minute after the release time for the announcement i. Time 0 is the market opening time at 9:30 a.m. on the following trading day after the release for announcements released during non-trading hours and is the actual release time for announcements released during trading hours.

Following Boguth et al. (2023), we focus on the R-squared of the regression (1), denoted as R_t^2, which measures the price informativeness at time t. By construction, R_t^2 always starts from zero and coverages toward one as t moves from the beginning to the end of the time window. The path of R_t^2, however, provides useful information on the speed of price discovery. Figure 2 compares the R_t^2 of the unbiased regressions (1) for macroeconomic announcements released during non-trading and trading hours. For announcements made during non-trading hours, R_t^2 jumps immediately by 28% at the market opening time at 9:30 a.m. and stays above the R_t^2 of trading hours announcements for the entire post-announcement time window. R_t^2 also increases at the time of release for announcements made during trading hours, but the increase is much smaller, only about 5% in size. Clearly, the announcements made during non-trading hours experience a much quicker price discovery process than those made during trading hours.

Figure 2: Cumulative R-squared around Chinese Macroeconomic Announcements



Notes: This figure shows the R-squared R_t^2 of the unbiased regressions. The dependent variables are the macroeconomic announcements window t returns of the CSI 300 index from 10 minutes prior to the announcement to the end of the 240th minute after the announcement, $R_i^{[-10,239]}$, and the independent variables are the returns of the partial announcement window from 10 minutes prior to the announcement to minute t around the announcement i, $R_i^{[-10,t]}$: $R_i^{[-10,239]} = \alpha_t + \beta_t R_i^{[-10,t]} + \epsilon_{i,t}$. "Non-trading hours (China)" refers to the Chinese macroeconomic announcements released during non-trading hours; "Trading hours (China)" refers to the Chinese macroeconomic announcements released during regular trading hours. The time "0" is the opening time of the stock market (9:30 a.m.) for announcements released during non-trading hours and is the actual announcement time for announcements released during trading hours. The sample period is from January 2009 to December 2020.

To further pin down the effect of announcement time on the speed of price discovery, we regress the returns at different post-announcement windows on the information content of individual news announcements, proxied by the total return in the window $[-10, 239]$ around the release time. The regression is specified as follows:

$$R_i^t = \alpha^t + \beta_1^t R_i^{[-10,239]} + \beta_2^t R_i^{[-10,239]} \times Non_i + \beta_3^t Non_i + \epsilon_i^t, \quad (2)$$

where R_i^t denotes the return of the CSI 300 index on a given time interval t around the release of the announcement i, Non_i is a dummy variable that equals one if the announcement i is released during non-trading hours, $R_i^{[-10,239]}$ denotes the total return of the CSI 300 index from 10 minutes prior to the release time to the end of the 240th minute after the announcement. The regression coefficient β_1^t measures the average proportion of price discovery occurring in time window t for announcements released during trading hours, and $\beta_1^t + \beta_2^t$ measures the average proportion of price discovery occurring in time window t for the announcement released during non-trading hours. Our focus is therefore on the coefficient β_2^t, which captures the difference in the speed of price discovery between the announcements within non-trading hours and trading hours. In particular, we split the four-hour trading window after the announcement into the following five periods.

Time window "0": The initial one minute of trading right after the announcement. If the announcement is released during trading hours, for example at 10:00 a.m. on Monday, the initial one-minute return, which can be denoted by R_i^0, is calculated using the last transaction prices at the end of 9:59 a.m. and the end of 10:00 a.m., respectively. If the announcement is released during non-trading hours, for example, at 5:00 p.m. on Friday, the initial one-minute return is calculated using the last transaction prices at the market closing time of 3:00 p.m. on Friday and the end of 9:30 a.m. on the following Monday.

Time window "$[1, 59]$", "$[60, 119]$", "$[120, 179]$", "$[180, 239]$": Following the initial minute, we label the 59-minute trading window from the beginning of minute "1" to the end of minute "59" by "[1,59]" and denote the market return by $R_i^{[1,59]}$; the 60-minute trading window from the beginning of minute "60" to the end of minute "119" by "[60,119]" and denote the market return by $R_i^{[60,119]}$; the 60-minute trading window from the beginning of minute "120" to the end of minute "179" by "[120,179]" and denote the market return by $R_i^{[120,179]}$; and the last 60-minute trading window from the beginning of minute "60" to the end of minute "239" by "[180,239]" and denote the market return by $R_i^{[180,239]}$.

Panel A of Table 4 shows that 5.2% of the price discovery occurs at the first minute of trading ("0") for announcements released during trading hours, whereas 27.9% (0.052+0.227) of the price discovery occurs at the first minute of trading for announcements released during non-trading hours. The difference in the speed of price discovery in the first minute is 22.7% with a significant t-stat of 8.62. In terms of economic magnitudes, the coefficient implies that the speed of price discovery is about five times faster for announcements released during non-trading hours. Consistent with faster price discovery in the first minute of trading, the coefficients of β_2^t are negative and significant for trading windows "[180,239]".

Overall, our results show that the speed of price discovery is faster for announcements released during non-trading hours, compared to the ones released during trading hours. The latter is an extreme case where the distance-to-trading equals zero, as investors can trade immediately after the release of this news.

Table 4: The Impact of Distance-to-trading on the Speed of Price Discovery

	Post-announcement time windows				
	(1)	(2)	(3)	(4)	(5)
	0	$[1,59]$	$[60,119]$	$[120,179]$	$[180,239]$
Panel A: Non-trading-hours vs. trading-hours announcements					
$R^{[-10,239]}$	0.052***	0.170***	0.195***	0.186***	0.396***
	[4.33]	[11.11]	[14.41]	[9.45]	[19.30]
$R^{[-10,239]} \times Non$	0.227***	0.032	-0.017	-0.033	-0.209***
	[8.62]	[1.32]	[-0.86]	[-1.37]	[-7.91]
Non	-1.779	3.092	-0.045	1.981	-3.250
	[-0.68]	[1.21]	[-0.02]	[0.86]	[-1.32]
Constant	-1.262	2.703*	-1.615	-2.739*	2.913
	[-1.22]	[1.79]	[-1.03]	[-1.68]	[1.64]
R^2	0.240	0.229	0.249	0.222	0.468
N	1673	1673	1673	1673	1673
Panel B: The impact of distance-to-trading					
$R^{[-10,239]}$	0.106***	0.182***	0.188***	0.184***	0.341***
	[8.16]	[13.85]	[17.29]	[12.43]	[19.96]
$R^{[-10,239]} \times Dur$	0.070***	0.004	-0.002	-0.017***	-0.054***
	[6.57]	[0.49]	[-0.41]	[-2.74]	[-4.56]
Dur	-0.087	0.253	2.652**	-1.355	-1.463
	[-0.04]	[0.20]	[2.57]	[-1.37]	[-1.12]
Constant	-1.704	4.001***	-2.882**	-1.266	1.851
	[-1.39]	[3.05]	[-2.25]	[-0.97]	[1.28]
R^2	0.208	0.227	0.251	0.225	0.442
N	1673	1673	1673	1673	1673

Notes: Panel A reports the regression results of $R_i^t = \alpha^t + \beta_1^t R_i^{[-10,239]} + \beta_2^t R_i^{[-10,239]} \times Non_i + \beta_3^t Non_i + \epsilon_i^t$. Panel B reports the regression results of $R_i^t = \alpha^t + \beta_1^t R_i^{[-10,239]} + \beta_2^t R_i^{[-10,239]} \times Dur_i + \beta_3^t Dur_i + \epsilon_i^t$. The dependent variables are the log returns of CSI 300 for the respective time intervals and are in basis points. Dummy variable Non_i equals 1 if the announcement is released during non-trading hours. Dur_i is the time between announcement time and the first trading time after the announcement and is in days (i.e., distance-to-trading). The sample period is from January 2009 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

3.2 The speed of price discovery: high and low distance-to-trading announcements

In this section, we further investigate the quantitative relation between the speed of price discovery and the distance-to-trading, utilizing the rich variations in the release time of macroeconomic announcements in China. We divide the announcements released during non-trading hours into high and low groups according to the distance-to-trading, then estimate the unbiased regressions, respectively. Figure 3 shows the R-squared R_t^2 from the unbiased regression (1). "High Distance-to-trading" refers to the announcements with distance-to-trading above the median; "Low Distance-to-trading" refers to the announcements with distance-to-trading below the median. The path of R_t^2 shows that the price discovery speed is faster for announcements with high distance-to-trading.

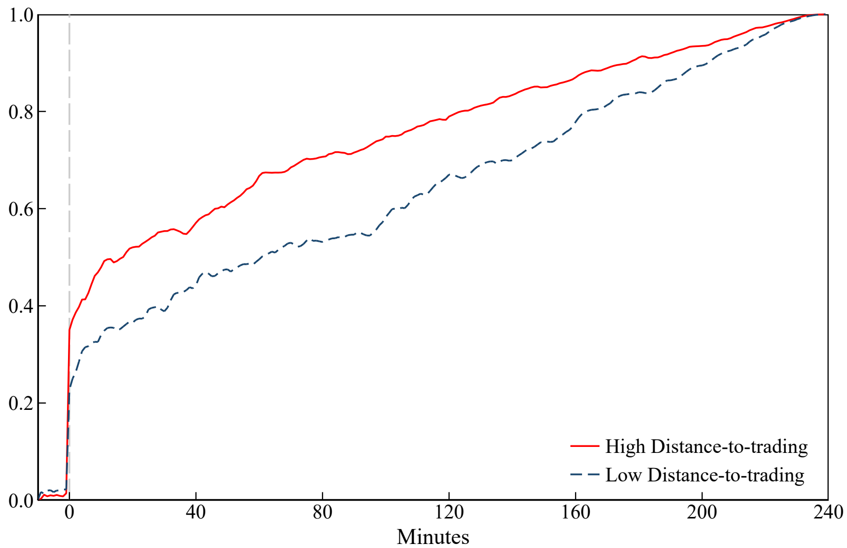
To further pin down the effect of distance-to-trading on the speed of price discovery, we replace the dummy variable Non_i with the distance-to-trading Dur_i in Equation (2). In particular, Dur_i is zero for announcements released during trading hours and is the time between the release time and the market opening time (9:30 a.m.) on the next trading day for announcements released during non-trading hours. We estimate the impact of the distance-to-trading on the speed of price discovery through the following regression:

$$R_i^t = \alpha^t + \beta_1^t R_i^{[-10,239]} + \beta_2^t R_i^{[-10,239]} \times Dur_i + \beta_3^t Dur_i + \epsilon_i^t, \tag{3}$$

where our object of interest is the coefficient β_2^t, which captures the impact of distance-to-trading on the the speed of price discovery.

We report the estimation results in Panel B of Table 4. The results paint a clear picture that the speed of price discovery is faster when the distance-to-trading increases. The proportion of price discovery occurring in time window "0" increases by 7.0% with a significant t-value of 6.57 when the distance-to-trading increases by one day. Consistent with the faster price discovery at time "0", there is less price discovery in the subsequent time windows"[120,179]" and "[180,239]" as the distance-to-trading increases. The coefficients β_2^t

Figure 3: The Impact of Distance-to-trading on Cumulative R-squared around Macroeconomic Announcements



Notes: This figure shows R-squared R_t^2 from unbiased regressions based on the sample of announcements released during non-trading hours. The dependent variables are the macroeconomic announcements window t returns of CSI 300 index from 10 minutes prior to the announcement to 240 minutes after the announcement, $R_i^{[-10,239]}$, and the independent variables are the returns of the partial announcement window from 10 minutes prior to the announcement to minute t around the announcement i, $R_i^{[-10,t]}$: $R_i^{[-10,239]} = \alpha^t + \beta^t R_i^{[-10,t]} + \epsilon_i^t$. "High Distance-to-trading" refers to the Chinese macroeconomic announcements released during non-trading hours with high distance-to-trading; "Low Distance-to-trading" refers to the Chinese macroeconomic announcements released during non-trading hours with low distance-to-trading. The time "0" is the opening time of stock market (9:30 a.m.) for announcements released during non-trading hours. The sample period is from January 2009 to December 2020.

are significantly negative for these time windows.

We further test the robustness of the above results using only the announcements released during non-trading hours and report the results in Table 5. The relation between the speed of discovery and distance-to-trading remains robust. Time window "close-to-open" denotes the close-to-open return. For example, if the announcement is released at 5:00 p.m. on Friday, the close-to-open return is calculated using the last transaction prices at the market closing time of 3:00 p.m. on Friday and the market opening price on the following Monday.

As the distance-to-trading increases by one day, the proportion of price discovery occurring in time window "close-to-open" increases by 3.0% and is statistically significant with a t-value of 3.16. The results confirm that the strong relation between price discovery speed and distance-to-trading is not driven entirely by the differences between non-trading-hours and trading-hours announcements. In fact, within the sample of announcements all released during non-trading hours, the calendar time duration between the announcement time and the market opening time still has a considerable impact on the speed of price discovery. Importantly, these results also demonstrate that the faster price discovery associated with announcements released during non-trading hours as shown in Subsection 3.1 is not simply determined by the presence of the pre-opening sessions.

Table 5: The Impact of Distance-to-trading on the Speed of Price Discovery (Non-Trading Hours Only)

	Post-announcement time windows					
	(1) close-to-open	(2) [open, 0]	(3) $[1, 59]$	(4) $[60, 119]$	(5) $[120, 179]$	(6) $[180, 239]$
$R^{[-10,239]}$	0.175***	0.060***	0.209***	0.173***	0.175***	0.209***
	[6.85]	[4.98]	[8.33]	[9.80]	[10.06]	[9.70]
$R^{[-10,239]} \times Dur$	0.030***	0.002	-0.004	0.002	-0.014**	-0.015*
	[3.16]	[0.44]	[-0.52]	[0.32]	[-2.14]	[-1.83]
Dur	0.574	0.192	-0.750	3.674***	-2.641**	-1.049
	[0.30]	[0.32]	[-0.51]	[3.09]	[-2.31]	[-0.93]
Constant	-4.792*	0.753	6.615**	-5.555**	2.117	0.861
	[-1.87]	[0.68]	[2.55]	[-2.50]	[1.00]	[0.41]
R^2	0.265	0.136	0.231	0.239	0.220	0.280
N	717	717	717	717	717	717

Notes: This table reports the regression results of $R_i^t = \alpha^t + \beta_1^t R_i^{[-10,239]} + \beta_2^t R_i^{[-10,239]} \times Dur_i + \beta_3^t Dur_i + \epsilon_i^t$ based on the sample of announcements released during non-trading hours. The dependent variables are the log returns of CSI 300 for the respective time intervals and are in basis points. Dur_i is the time between announcement time and the first trading time after the announcement and is in days (i.e., distance-to-trading). The sample period is from January 2009 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

4 Discussion and Additional Evidence: The Learning Channel

Our previous empirical results show that longer distance-to-trading is associated with faster price discovery once the market resumes trading. As motivated by these observations, we discuss why our findings can be rationalized by the channel of retail investors' learning. We then derive a few testable hypotheses and provide direct evidence in support of the learning channel.

4.1 Theoretical Discussions and Hypotheses

We first provide theoretical discussions on the important implications of allowing for distance-to-trading to affect the price informativeness upon market opening. The classic rational expectations equilibrium (REE) model framework predicts that the equilibrium price informativeness depends on the mass of informed relative to uninformed investors (Grossman and Stiglitz (1980); Kyle (1985)). Informed investors learn from the market price and their own informative signals for investment in risky asset whereas less informed investors extract information by directly learning from the market price. Learning from the price affects trading decisions that determine the equilibrium prices through market clearing. When trading and the market price are absent, public and private signals enter the information set of investors only. Kacperczyk et al. (2016) show that investors being informed or less informed are endogenously determined by varying information processing capabilities across investors. With a greater capacity for processing information, sophisticated investors draw and carry more informative signals for investment decisions after turning themselves into informed investors. Less sophisticated investors, however, invest with noisier information as a result of tighter capacity constraints.

While a macroeconomic announcement as a public signal perfectly reveals the measurable macroeconomic fundamentals, it only partially reveals the relevant information that determines the asset payoff. Processing information as derived from the macroeconomic announcements is therefore critical for generating informative signals for investment decisions among investors. As sophisticated investors like institutional investors are better informed, the capacity constraint for information processing is more relevant for less sophisticated investors like retail investors. Recent evidence suggests that dynamically, retail investors do increase their attention and start learning about the macroeconomic and financial news over time (Da et al. (2011); Ben-rephael et al. (2021); Fisher et al. (2022)). Dugast and Foucault (2018) and Ai et al. (2022) further show that information processing takes some time, by which investors signals may become more precise later.

Therefore, it is very likely that longer distance-to-trading potentially gives less sophisticated investors more time to process the relevant information, pick up more precise signals, and eventually reduce their information disadvantage. Hence, we hypothesize that the intensity of learning, especially among retail investors, may increase with distance-to-trading. Importantly, when market trading is absent overnight, both informed and less informed investors do not observe the equilibrium asset prices that could have aggregated the information across investors once trading is resumed. Retail investors' learning thus does not come from learning from the prices nor from the quotes for bids and asks. In addition, the "information paradox" that increasing price efficiency deters extra learning, as in Grossman and Stiglitz (1980), is not applied in our setting. Hence, we test whether our measure of the distance-to-trading proxies for the degree of retail investors' pre-trading learning in absence of the market price signal per the following hypothesis:

Hypothesis 1. *If announcements fall outside trading hours, the degree of retail investors' learning before trading increases with the duration of distance-to-trading.*

Next, if the speed of price discovery once the market resumes is affected by the increased retail learning before trading, we should see faster price discovery associated with greater learning overnight. We then test the following hypothesis:

Hypothesis 2. *If announcements fall outside trading hours, greater pre-trading learning among retail investors leads to faster price discovery upon market opening. Comparatively, price discovery is slower for announcements falling within trading hours as the result of a lack of pre-trading learning.*

Our paper further examines the additional market outcomes driven by the retail investors' learning overnight. For example, if the increased retail learning before trading could make an increasing number of less sophisticated investors become more informed over time before market opening, this should ultimately lead to more precise average information in the market across investors upon market opening. To test for the presence of improvements of information quality once the market resumes trading, we resort to a variance decomposition identity of the following equation in the spirit of Ai et al. (2023) and provide additional intuition:

$$Var(r_t) = Var(\mathbb{E}(r_t|S_{t^o})) + \mathbb{E}(Var(r_t|S_{t^o})), \tag{4}$$

where r_t is the market return to be realized by the end of the first trading day after the macroeconomic announcement and S_{t^o} refers to the average signal across all investors upon market opening given $t^o < t$ and this average signal is partially informative about r_t. The identity of Equation (4) says that the total variance of returns, $Var(r_t)$ can be decomposed into two components: the return variance realized *upon market opening* as of the time stamp t^o, $Var(\mathbb{E}(r_t|S_{t^o}))$, and the variance of returns that will realize by the end of the day $\mathbb{E}(Var(r_t|S_{t^o}))$. Suppose the pre-trading learning indeed improves the precision of less sophisticated investors' signals. The market will be filled with more precise information upon opening. As a result, the variance of returns conditional on announcements (i.e., the first term of Equation (4)) should be larger. While the first term captures the realized return volatility driven by the quality of information contained in the stock prices, we should expect such volatility to jump upon market opening, and the degree of information quality improvement, if any, should increase with distance-to-trading in which learning is increasingly stronger.

Also, as derived from Kyle (1985) and Kim and Verrecchia (1991), trading volume is driven by liquidity trading and portfolio rebalancing needs, which is a result of investors acting upon their informative signals. Such trading motive generates a positive co-movement between trading volume and return volatility. We therefore expect trading volume to jump upon market opening if return volatility spikes up. Plus, the degree of jumps should be correlated with the length of distance-to-trading if the realized return volatility jumps are triggered by the improved information quality in the market.

Additionally, while the learning effects could have alleviated the information disadvantages of retail investors, the information gaps, or disagreement between different investor types, should be narrowed. We further test whether retail investors' learning over the distance-to-trading helps reduce the information asymmetry across investors upon market opening. In particular, as noted in Bollerslev et al. (2018), rising disagreement breaks down the volatility-volume co-movement. Therefore, we test for shrinkage in the information asymmetry across investors at market opening if we observe jumps in both return volatility and trading volume. We therefore have the following hypothesis:

Hypothesis 3. *If announcements fall outside trading hours, greater pre-trading learning among retail investors leads to increases in realized return volatility and trading volume and lowered information asymmetry upon market opening. Comparatively, post-announcement jumps in volatility and trading volume and the reduction in information asymmetry are smaller for announcements falling within trading hours as a result of a lack of pre-trading learning.*

Finally, according to the identity of Equation (4), we should also see that a larger $Var(\mathbb{E}(r_t|S_{t^o}))$ leads to lower future volatility $\mathbb{E}(Var(r_t|S_{t^o}))$ in the post-announcement returns by the end of the day, given the total variance of returns is pre-fixed. Hence, we cast the additional hypothesis in the following, again by bringing up the positive co-movement between volatilities and trading volume.

Hypothesis 4. *Greater learning among retail investors before trading leads to lower return volatility and trading volume later on, after the initial jumps upon market opening. Comparatively, such reduction in volatility and volume is smaller for announcements falling within trading hours as a result of a lack of pre-trading learning.*

We will demonstrate that all these hypotheses align well with our additional empirical evidence to be shown later. It is therefore safe to argue that the learning channel, especially learning among retail investors, helps rationalize our main findings. Specifically, as retail investors' learning increases with distance-to-trading before the market opens once an announcement falls outside trading hours, the effects of learning mitigate the information asymmetry across investors and improve the average information quality in the market. Once the market resumes trading, faster price discovery follows and greater price informativeness is achieved.

On the other hand, we also note that when a macroeconomic announcement just falls in the middle of trading hours, efficient learning among retail investors may not yet take place. The information gaps can be untimely significant and the average information quality is less than precise. Therefore, the price discovery process is comparatively slower and the market prices are less informative.

In Appendix A, we also present a simple and static model and demonstrate that less sophisticated investors benefit from a longer distance-to-trading, which causes more of them to become better informed about more precise signals upon market opening. As a result, price informativeness increases with the distance-to-trading. Importantly, we also show in the model that trading volume and return volatility in equilibrium are shifted by the effect of both increased price informativeness and reduced posterior market uncertainty. The results based on our numerical exercises suggest that the effects of increased price informativeness dominate upon market opening, which later on drives down market noise and leads to lowered trading volume and return volatility as time evolves. These results also accord well with the dynamic volatility trade-off, as demonstrated by Equation (4).

4.2 The Impact of Investor Learning on Speed of Price Discovery

To test our hypotheses, we rely on social media posts related to macroeconomic announcements to proxy for retail investors' learning activities around the news announcements' release time. Our web-scrapped data are from Weibo, a Chinese micro-blogging website and one of China's largest social media platforms, for the period from January 2018 to December 2020. We focus on three measures that proxy for investors' learning activities before they are able to trade on the news: 1) the total number of posts (*posts*) with headline keywords matched with the macroeconomic indexes; [14] 2) the number of fans (*fans*) following the bloggers who post the posts with headline keywords matched with the macroeconomic indexes; and 3) the total number of retweets, comments, and likes (*interactions*) of the posts with headline keywords matched with the macroeconomic indexes. The numbers of posts, fans, and interactions are calculated over the period of 72 hours prior to the time when investors can trade on the news – the actual release time for announcements made during trading hours or 9:30 am of the next trading day for announcements made outside trading hours.[15]

In Hypothesis 1, we expect non-trading-hours news announcements to have higher *posts*, *fans*, and *interactions* resulting from the news coverage of the just-released announcements, which provides the information and time for investors, especially retail investors, to digest the news before they trade on the following day. Indeed, as reported in Table 6, the number of *posts*, *fans*, and *interactions* are all positively related to the distance-to-trading (Dur). For the full sample of non-trading-hours and trading-hours announcements, an increase of one day in the distance-to-trading is associated with 21 more related Weibo posts, 40 million more fans who tracked these posts, and 23 more retweets, comments, and likes following these posts. The results remain similar if we focus only on non-trading-hours news announcements: an increase of one day in the distance-to-trading is associated with 21 more related Weibo posts, 39 million more fans who tracked these posts, and 19 more retweets, comments, and likes following these posts.

Next, we turn to Hypothesis 2 and investigate the impact of investor learning on the speed of price discovery. Similar to our previous discussions, we regress the log returns of the CSI 300 for different post-announcement time intervals on our three measures – *posts*, *fans*, and *interactions* – which all serve as proxies for investor learning before they trade

[14] In Internet Appendix Table IA.X, we report the keywords that we use to match the 20 major macroeconomic indicators.

[15] We choose the fixed 72-hour pre-trading window because the longest distance-to-trading is 66.5 hours for announcements made after 3:00 pm (market close time) on Friday, with the exception of a few public holidays. We have experimented with other pre-trading window horizons such as 48-hour and 24-hour windows. The results remain similar.

Table 6: Distance-to-trading and Investor Learning Before Trading

	Non-trading-hours and Trading-hours News			Non-trading-hours News Only		
	(1) *posts*	(2) *fans*	(3) *interactions*	(4) *posts*	(5) *fans*	(6) *interactions*
Dur	0.021***	0.040***	0.023***	0.021***	0.039***	0.019**
	[4.12]	[4.38]	[2.90]	[3.50]	[3.77]	[2.03]
Constant	0.032***	0.051***	0.045***	0.033***	0.053***	0.054***
	[11.21]	[9.90]	[9.49]	[5.81]	[5.66]	[5.41]
R^2	0.052	0.059	0.023	0.044	0.053	0.013
N	584	584	584	305	305	305

Notes: This table reports the regression results of $Learning_i = \alpha + \beta_1 Dur_i + \epsilon_i$. The dependent variables $Learning_i$ are proxied by the number of related Weibo posts or fans of these posts or the retweets, comments, and likes of these posts during the 72 hours before the announcement time (for trading-hours news) or the market opening time of the next trading day (for non-trading-hours news). The units for the number of posts, retweets, comments, and likes are in thousands and the unit for the number of fans is in billions. The sample period is from January 2018 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

on the news. As shown in Table 7, more investor learning before trading is associated with faster post-announcement price discovery. For the sample pooling non-trading-hours and trading-hours news together, an increase of one standard deviation in the pre-trading posts (78 posts), the fans following these posts (138 million), and retweets, comments, and likes following these posts (128 interactions) is associated with 5.2%, 5.8%, and 4.1% faster price discovery at the first minute (Time "0") when investors can trade on the news. The impact of investors' learning on the price discovery speed is similar if we focus on the sample of non-trading-hours news only. An increase of one standard deviation in the pre-trading posts (94 posts), the fans following these posts (161 million), and retweets, comments, and likes following these posts (156 interactions) is associated with 4.4%, 4.9%, and 3.7% faster price discovery at the market opening time (9:30 am) of the next trading day.

In Internet Appendix Table IA.I, we show our estimation results related to Hypotheses 1 and 2 are robust if the degree of retail investors' learning is measured differently. In specific, the learning intensity is instead measured by the number of posts not only with the headline keywords matched with the macroeconomic data indicators but also with the keyword "stocks" mentioned either in the main text or the headline of a post. The results again suggest that learning from these Weibo posts increases with the distance-to-trading and price discovery upon market opening is faster if more relevant Weibo posts are associated with posters' discussions of the stocks.

Lastly, it's important to clarify that we do not asume that retail investors' pre-trading learning comes exclusively from the Weibo posts. Rather, we view these posts as a reflection of the collective efforts made by retail investors to extract and understand information relevant to stock investment, especially when macroeconomic news arrives during periods when trading is not active.

4.3 The Impact of Investor Learning on Information Outcomes

In this subsection, we explore the impact of investors' pre-trading learning activities on the averaged information quality and the information asymmetry across investors as reflected by the underlying stocks after the stock market opens for trade. Apart from looking into the return volatility and stock turnovers, we use the bid-ask spreads of the CSI 300 index component stocks to proxy for the information asymmetry across investors.[16] For each of the CSI 300 component stocks, we first calculate their return volatility, average turnover, and average bid-ask spreads, based on the high-frequency minute-by-minute quotes from January 2009 to December 2020, during different time windows after the news releases. We then aggregate these three measures to the index level, weighted by the respective index weights of the component stocks.[17] Since the return volatility, turnover rate, and bid-ask spreads have strong intra-day patterns and a time-series trend, we further normalize them by their respective mean and standard deviations of the same calendar time in the previous month, making them comparable across different trading windows.

We estimate the following regressions to quantify the impact of distance-to-trading and

[16]The return volatility of each time window is calculated as the squared root of the sum of return squared, $\sqrt{\sum_{i=1}^{N}(lnP_{i,s} - lnP_{i-1,s})^2}$, where $P_{i,s}, (0 \leq i \leq N)$ denote the $N+1$ number of minute-end prices within the given time window for stock s. The spreads of each time window are calculated as $\frac{ask_s - bid_s}{(ask_s + bid_s)/2}$, where ask_s and bid_s are the order-volume weighted averages of ask and bid prices for stock s.

[17]For the volatility measure, we have also tested the results using the return volatility of the CSI 300 index itself. The results remain similar to our baseline results based on the weighted average volatility of the CSI 300 component stocks.

Table 7: The Impact of Pre-Trading Investor Learning Activities on the Speed of Price Discovery

	Non-trading-hours and Trading-hours News					Non-trading-hours News Only				
	(1) 0	(2) $[1, 59]$	(3) $[60, 119]$	(4) $[120, 179]$	(5) $[180, 239]$	(6) 0	(7) $[1, 59]$	(8) $[60, 119]$	(9) $[120, 179]$	(10) $[180, 239]$
Panel A: Investor Learning Proxied by the Number of Related Posts 72 hours Before Trading										
$R^{[-10,239]}$	0.183***	0.184***	0.163***	0.143***	0.248***	0.316***	0.201***	0.169***	0.138***	0.136***
	[7.40]	[10.91]	[10.54]	[10.07]	[11.46]	[10.62]	[7.69]	[7.52]	[6.99]	[5.54]
$R^{[-10,239]} \times Learning$	0.673***	-0.032	-0.040	-0.094	-0.377**	0.470***	-0.021	-0.101	-0.198**	-0.063
	[5.78]	[-0.25]	[-0.38]	[-0.79]	[-2.15]	[4.33]	[-0.17]	[-0.94]	[-2.29]	[-0.49]
Learning	19.750	-26.726	10.063	8.049	1.647	30.410	-19.108	1.759	-3.469	-1.390
	[1.11]	[-1.14]	[0.58]	[0.48]	[0.08]	[1.61]	[-0.70]	[0.09]	[-0.20]	[-0.08]
Constant	-1.767	1.776	-1.435	-3.962**	2.151	-4.646	0.473	-0.385	-0.278	1.355
	[-0.76]	[0.81]	[-0.76]	[-2.04]	[0.98]	[-1.36]	[0.15]	[-0.14]	[-0.10]	[0.54]
R^2	0.304	0.258	0.254	0.200	0.358	0.507	0.275	0.248	0.174	0.216
N	584	584	584	584	584	305	305	305	305	305
Panel B: Investor Learning Proxied by the Number of Fans Following Related Posts 72 Hours Before Trading										
$R^{[-10,239]}$	0.179***	0.185***	0.163***	0.144***	0.249***	0.309***	0.206***	0.166***	0.142***	0.136***
	[7.53]	[11.59]	[10.83]	[10.40]	[11.60]	[10.59]	[8.47]	[7.55]	[7.26]	[5.54]
$R^{[-10,239]} \times Learning$	0.423***	-0.030	-0.020	-0.072	-0.220**	0.306***	-0.055	-0.026	-0.132***	-0.042
	[5.74]	[-0.37]	[-0.33]	[-1.16]	[-2.02]	[4.44]	[-0.64]	[-0.45]	[-3.64]	[-0.60]
Learning	18.513*	-14.508	7.726	-5.455	1.943	19.755*	-14.861	7.281	1.610	-8.604
	[1.73]	[-1.03]	[0.67]	[-0.50]	[0.12]	[1.73]	[-0.89]	[0.59]	[0.16]	[-0.85]
Constant	-2.137	1.659	-1.561	-3.254*	2.035	-4.753	0.755	-0.922	-0.650	2.058
	[-0.95]	[0.78]	[-0.84]	[-1.71]	[0.90]	[-1.40]	[0.24]	[-0.34]	[-0.24]	[0.82]
R^2	0.312	0.258	0.254	0.200	0.359	0.513	0.276	0.248	0.177	0.218
N	584	584	584	584	584	305	305	305	305	305
Panel C: Investor Learning Proxied by the Number of Retweets, Comments, and Likes Following Related Posts 72 Hours Before Trading										
$R^{[-10,239]}$	0.198***	0.185***	0.170***	0.134***	0.240***	0.329***	0.196***	0.174***	0.130***	0.135***
	[7.86]	[10.67]	[11.36]	[9.96]	[11.37]	[11.31]	[7.12]	[8.06]	[7.13]	[5.54]
$R^{[-10,239]} \times Learning$	0.322**	-0.043	-0.172*	0.069	-0.183	0.239**	0.064	-0.181**	-0.062	-0.047
	[2.29]	[-0.34]	[-1.84]	[0.66]	[-1.48]	[2.06]	[0.44]	[-2.07]	[-0.72]	[-0.48]
Learning	27.246**	-11.278	-7.376	2.315	-1.011	25.723**	-12.141	-15.630*	3.099	2.803
	[2.37]	[-0.74]	[-0.77]	[0.23]	[-0.09]	[2.24]	[-0.71]	[-1.69]	[0.32]	[0.27]
Constant	-2.623	1.269	-0.652	-3.672**	2.357	-5.160	0.362	0.864	-0.559	1.111
	[-1.15]	[0.61]	[-0.36]	[-1.99]	[1.12]	[-1.56]	[0.12]	[0.33]	[-0.22]	[0.46]
R^2	0.294	0.257	0.258	0.199	0.353	0.503	0.276	0.256	0.169	0.216
N	584	584	584	584	584	305	305	305	305	305

Notes: This table reports the regression results of $R_i^t = \beta_0 + \beta_1 R_i^{[-10,239]} + \beta_2 R_i^{[-10,239]} \times Learning_i + \beta_3 Learning_i + \epsilon_i$. The dependent variables are the log returns of CSI 300 for the respective time intervals and are in basis points. $Learning_i$ is the number of related Weibo posts (or the fans of these posts, or the retweets, comments, and likes of these posts) during the 72-hour period prior to announcement time (for trading hours news) and the market opening time of the next trading day (for non-trading hours). The units for the number of posts, retweets, comments, and likes are in thousands and the unit for the number of fans is in billions. The sample period is from January 2018 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

investors' learning activities on return volatility, turnover, and bid-ask spreads:

$$Y_i^t = \alpha^t + \beta_1^t Dur_i + \epsilon_i^t, \tag{5}$$

$$Y_i^t = \alpha^t + \beta_1^t Learning_i + \epsilon_i^t, \tag{6}$$

where Y_i^t denotes the normalized average return volatility, turnover rate, and bid-ask spreads for time window t after announcement i, Dur_i is the calendar time between the release time and the market opening time (9:30 am) on the next trading day for non-trading-hours announcements and is zero for trading-hours announcements, and $Learning_i$ is the total number of posts with headline keywords matched with the macroeconomic indexes. In Internet Appendix Table IA.II, we present the regression results based on the number of fans and the total number of retweets, comments, and likes.

Table 8 reports the estimation results of Equation (5) and Equation (6).[18] Consistent with the predictions of Hypothesis 3, we find that longer distance-to-trading and more pre-trading investors' learning activities lead to higher volatility and turnover at the time when the market opens for trade. At the first minute (time "0") when investors can trade on the news, a one-day increase in distance-to-trading leads to a 0.176 unit increase in normalized volatility and a 0.157 unit increase in normalized turnover, while an increase of one standard deviation in posts (78 posts) leads to a 0.103 unit increase in normalized volatility and a 0.141 unit increase in normalized turnover. This impact on volatility and turnover decreases in magnitude, but remains statistically higher, for the next trading hour $[1, 59]$.

Consistent with Hypothesis 4, we also find that, after the initial spike at time "0", the impact on stock volatility and turnover starts to decrease. Starting from the second hour $[60, 119]$, stock volatility and turnover are no longer significantly higher for news announcements with longer distance-to-trading and more pre-trading learning. In fact, during the time window $[180, 239]$, or the fourth trading hour post-announcement, both stock volatility

[18] In this section, we only present the regression results based on the full sample of non-trading-hours and trading-hours announcements. The regression results based on the sample of non-trading-hours announcements remain similar.

and turnover become significantly lower for announcements with longer distance-to-trading and more investor learning: a one-day increase in distance-to-trading leads to a 0.032 and 0.028 unit drop in normalized volatility and turnover, and an increase of one standard deviation in posts (78 posts) leads to a 0.052 and 0.069 unit drop in normalized volatility and turnover. In Internet Appendix Table IA.III, we also show the estimation results regarding the learning effects on the direct measure of stock market trading volume. The results on trading volume are well consistent with those if trading is measured by stock turnovers, which again lend support to Hypotheses 3 and 4.

In addition, we find that the empirical patterns of bid-ask spreads are consistent with the predictions of Hypotheses 3 and 4. Longer distance-to-trading and more investor learning are generally associated with narrower bid-ask spreads for trading windows one hour after the announcement time, that is, $[60, 199]$, $[120, 179]$, and $[180, 239]$. For the first hour (i.e., time "0" and $[1, 59]$), the impact of distance-to-trading and learning on bid-ask spreads is negative but not statistically significant.

5 Robustness Tests and Additional Results

5.1 Analysis based on a matched sample

One may be concerned that the release time of macroeconomic news is related to their information content. In particular, regulators may intentionally release important news during non-regular trading hours and hope to minimize the news impact on the market. If the release time is endogenously related to their market impact, our previous analysis might be biased. In this section, we address this concern by using a matched sample of macroeconomic announcements that share a similar overall market impact but differ in release time. We use the 250-minute returns of the CSI 300 index as the measure of total market impact and require the differences between the matched pairs of announcements to be within the 1 bps threshold. For announcements released during non-trading hours, the 250-minute return is from 2:50 p.m. of the previous trading day to 2:59 p.m. of the following trading

Table 8: The Impact of Distance-to-trading and Pre-Trading Investor Learning on Volatility, Turnover, and Bid-ask Spreads

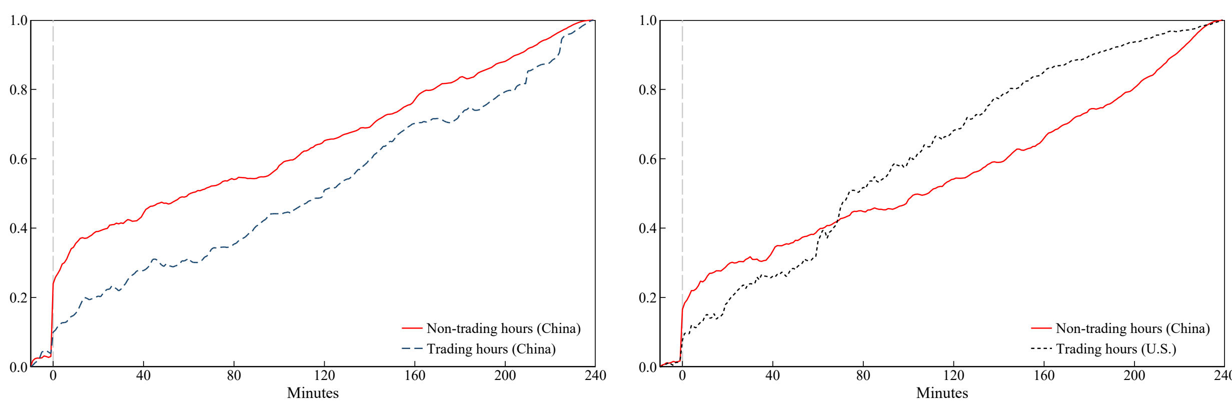
	0		[1, 59]		[60, 119]		[120, 179]		[180, 239]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: The Impact of Distance-to-Trading and Learning Activities on Volatility										
Dur	0.176***		0.032**		-0.010		-0.018		-0.032**	
	[5.73]		[2.34]		[-0.89]		[-1.46]		[-2.47]	
Learning		1.323**		0.091		-0.111		-0.295		-0.665***
		[2.07]		[0.33]		[-0.49]		[-1.24]		[-2.82]
Constant	-0.011	0.012	-0.003	0.042*	-0.021	0.013	-0.015	0.015	-0.011	0.029
	[-0.43]	[0.24]	[-0.26]	[1.73]	[-1.60]	[0.53]	[-1.15]	[0.61]	[-0.80]	[1.12]
R^2	0.031	0.011	0.004	0.000	0.000	0.000	0.001	0.002	0.004	0.009
N	1673	584	1673	584	1673	584	1673	584	1673	584
Panel B: The Impact of Distance-to-Trading and Learning Activities on Turnover										
Dur	0.157***		0.059***		0.020		0.006		-0.028**	
	[5.05]		[2.82]		[1.22]		[0.35]		[-1.99]	
Learning		1.812***		1.114**		0.489		0.081		-0.884***
		[2.84]		[2.15]		[1.31]		[0.21]		[-3.03]
Constant	-0.004	-0.036	0.008	0.007	-0.001	0.022	0.007	0.021	0.006	0.036
	[-0.13]	[-0.70]	[0.42]	[0.19]	[-0.07]	[0.68]	[0.41]	[0.57]	[0.36]	[1.07]
R^2	0.019	0.017	0.006	0.012	0.001	0.003	0.000	0.000	0.002	0.009
N	1673	584	1673	584	1673	584	1673	584	1673	584
Panel C: The Impact of Distance-to-Trading and Learning Activities on Bid-Ask Spreads										
Dur	-0.018		-0.030		-0.049**		-0.049**		-0.050**	
	[-0.68]		[-1.27]		[-2.15]		[-1.99]		[-2.05]	
Learning		-0.101		-0.821**		-0.892*		-0.689		-0.586
		[-0.20]		[-1.99]		[-1.74]		[-1.33]		[-1.15]
Constant	-0.087***	-0.038	-0.117***	-0.037	-0.135***	-0.082	-0.156***	-0.098*	-0.131***	-0.098**
	[-2.74]	[-0.65]	[-4.19]	[-0.76]	[-4.92]	[-1.62]	[-5.62]	[-1.93]	[-4.91]	[-2.00]
R^2	0.000	0.000	0.001	0.004	0.002	0.004	0.002	0.002	0.002	0.002
N	1673	584	1673	584	1673	584	1673	584	1673	584

Notes: Columns (1), (3), (5), (7), and (9) report the regression results of $Y_i^t = \alpha^t + \beta_1^t Dur_i + \epsilon_i^t$, where the sample period is from January 2009 to December 2020. Columns (2), (4), (6), (8), and (10) report the regression results of $Y_i^t = \alpha^t + \beta_1^t Learning_i + \epsilon_i^t$, where the sample period is from January 2018 to December 2020. The dependent variables are the normalized weighted averages of return volatility, turnover, and bid-ask spreads of CSI 300 components stocks for the respective time intervals. Dur_i is the time between announcement time and the first trading time after the announcement and is in days. $Learning_i$ is the number of related Weibo posts during the 72-hour period prior to announcement time (for trading-hours news) and the market opening time of the next trading day (for non-trading-hours news). The unit for the number of posts is in thousands. The *t*-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

day. For announcements released during trading hours, the 250-minute post-announcement return is from 10 minutes prior to the actual announcement time to the end of the 240th minute afterward. For the 717 non-trading-hours announcements, we perform a one-to-one matching with replacement based on the 956 trading-hours announcements and obtain 639 pairs of matched announcements. As shown in Internet Appendix Table IA.IV, the matched announcements are comparable in terms of market impact: the average 250-minute return is 4.06 bps for announcements released during both non-trading and trading hours.

Panel (a) of Figure 4 shows the R-squared R_t^2 from unbiased regressions based on the matched samples. Table 9 reports the results of the regression estimation of Equations (2) and (3) based on the matched samples of macroeconomic announcements. Overall, the results are robust and similar to the baseline results discussed in Section 3.

Figure 4: Cumulative R-squared around Matched Macroeconomic Announcements



(a) Trading hours (China)

(b) Trading hours (U.S.)

Notes: This figure shows R-squared R_t^2 from unbiased regressions using matched samples with Chinese and U.S. announcements released during trading hours in panels (a) and (b). The dependent variables are the macroeconomic announcements window t returns of CSI 300 index and E-mini S&P 500 index futures from 10 minutes prior to the announcement to 240 minutes after the announcement, $R_i^{[-10,239]}$, and the independent variables are the returns of the partial announcement window from 10 minutes prior to the announcement to minute t around the announcement i, $R_i^{[-10,t]}$: $R_i^{[-10,239]} = \alpha^t + \beta^t R_i^{[-10,t]} + \epsilon_i^t$. "Non-trading hours (China)" refers to the matched Chinese macroeconomic announcement released during non-trading hours; "Trading hours (China)" refers to the matched Chinese macroeconomic announcements released during regular trading hours; "Trading hours (U.S.)" refers to the matched U.S. macroeconomic announcements released during regular trading hours. The time "0" is the opening time of stock market (9:30 a.m.) for announcements released during non-trading hours and is the actual announcement time for announcements released during trading hours. The sample period is from January 2009 to December 2020.

Table 9: The Impact of Distance-to-trading on the Speed of Price Discovery: Matched Sample

	Post-announcement time windows									
	(1) 0	(2) $[1,59]$	(3) $[60,119]$	(4) $[120,179]$	(5) $[180,239]$	(6) 0	(7) $[1,59]$	(8) $[60,119]$	(9) $[120,179]$	(10) $[180,239]$
	Panel A: Non-trading-hours vs. trading-hours (China)					**Panel C: Non-trading-hours vs. trading-hours (U.S.)**				
$R^{[-10,239]}$	0.086***	0.197***	0.195***	0.193***	0.329***	0.043***	0.198***	0.360***	0.236***	0.163***
	[4.23]	[8.57]	[8.41]	[9.87]	[14.14]	[4.42]	[10.90]	[15.50]	[11.25]	[10.73]
$R^{[-10,239]} \times Non$	0.207***	0.001	-0.015	-0.043	-0.150***	0.250***	0.015	-0.161***	-0.134***	0.030
	[5.44]	[0.04]	[-0.48]	[-1.62]	[-4.66]	[6.48]	[0.42]	[-4.56]	[-4.12]	[0.95]
Non	-2.156	2.375	2.852	2.916	-5.987**	-0.257	2.707	-4.238*	-0.285	2.073
	[-0.78]	[0.85]	[1.05]	[1.16]	[-2.40]	[-0.11]	[1.19]	[-1.83]	[-0.13]	[1.08]
Constant	0.108	2.520	-4.324**	-3.865**	5.561***	0.708	0.101	1.662	-0.526	-1.946**
	[0.08]	[1.35]	[-2.16]	[-2.05]	[2.96]	[1.23]	[0.10]	[1.18]	[-0.43]	[-1.98]
R^2	0.187	0.160	0.152	0.150	0.298	0.144	0.154	0.265	0.140	0.161
N	1278	1278	1278	1278	1278	1130	1130	1130	1130	1130
	Panel B: The impact of distance-to-trading (China)					**Panel D: The impact of distance-to-trading (U.S.)**				
$R^{[-10,239]}$	0.155***	0.193***	0.180***	0.182***	0.291***	0.118***	0.211***	0.293***	0.190***	0.188***
	[7.49]	[9.41]	[10.29]	[12.19]	[15.74]	[6.01]	[10.88]	[14.78]	[10.77]	[11.93]
$R^{[-10,239]} \times Dur$	0.058***	0.008	0.010	-0.016	-0.060***	0.082***	-0.008	-0.023*	-0.033**	-0.018
	[2.93]	[0.62]	[0.91]	[-1.56]	[-5.56]	[3.22]	[-0.51]	[-1.79]	[-2.45]	[-1.63]
Dur	0.271	-0.055	2.877**	-1.316	-1.777	2.861	-0.288	-0.063	-2.356**	-0.154
	[0.12]	[-0.04]	[2.50]	[-1.17]	[-1.51]	[1.33]	[-0.21]	[-0.06]	[-2.12]	[-0.15]
Constant	-1.403	3.693**	-4.408***	-1.657	3.775***	-1.016	1.615	-0.378	0.574	-0.795
	[-0.94]	[2.40]	[-2.91]	[-1.16]	[2.62]	[-0.84]	[1.38]	[-0.30]	[0.50]	[-0.80]
R^2	0.163	0.160	0.156	0.150	0.295	0.132	0.153	0.244	0.132	0.162
N	1278	1278	1278	1278	1278	1130	1130	1130	1130	1130

Notes: Panels A and C report the regression results of $R_i^t = \beta_0 + \beta_1 R_i^{[-10,239]} + \beta_2 R_i^{[-10,239]} \times Non_i + \beta_3 Non_i + \epsilon_i$. Panels B and D report the regression results of $R_i^t = \beta_0 + \beta_1 R_i^{[-10,239]} + \beta_2 R_i^{[-10,239])} \times Dur_i + \beta_3 Dur_i + \epsilon_i$. Panels A and B use the matched sample of Chinese announcements released during non-trading hours and trading hours. Panels C and D use the matched sample of Chinese announcements released during non-trading hours and U.S. announcements released during trading hours. The dependent variables are the log returns of CSI 300 (China) and S&P 500 futures (U.S.) for the respective time intervals and are in basis points. Dummy variable Non_i equals 1 if the announcement is released in non-trading hours. Dur_i is the time between announcement time and the first trading time after the announcement and is in days (i.e., distance-to-trading). The sample period is from January 2009 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

5.2 Fixed macroeconomic announcement type

Our baseline regressions are based on a pooled sample of the 20 most important macroeconomic announcements in China. One concern is that different sorts of macroeconomic announcements may be fundamentally different from one another, and this cross-type difference may have an impact on our baseline findings. To address this concern, we examine the robustness of our results by restricting ourselves to macroeconomic announcements in the same index but with various release times.

Panel A of Table 10 reports the results for PMI announcements. Given that all 179 PMI announcements are released during non-trading hours, we focus on Equation (3) to estimate the relation between the price discovery speed and the duration of distance-to-trade. In Panels B to D of Table 10, we report the estimation results per Equation (3) using only the Trade Balance (Trade), Industrial Production (IP), or Central Treasury Cash Management (CTCM) announcements. In our sample period, there are 109 IP, 96 Trade, and 27 CTCM announcements released during regular trading hours, and 14 IP, 33 Trade, and 83 CTCM announcements released during non-trading hours. The large variations in the announcement time of IP, Trade, and CTCM provide the identification we need for the estimation. Overall, the results are robust for both PMI, IP, Trade, and CTCM. For announcements with longer distance-to-trade, the speed of price discovery is significantly faster. In unreported results, we have also performed the tests on foreign exchange settlement and sales by banks (FESS) and profit of industrial enterprises (PI), and the relation between the price discovery speed and distance-to-trade remains robust.

5.3 Comparison with U.S. macroeconomic announcements

In this section, we test the prediction of Hypothesis 2 by investigating the relation between the speed of price discovery and distance-to-trading using the macroeconomic announcements in China and their counterparts in the U.S. The very deep and extremely liquid financial markets in the U.S., including various market index ETFs, futures, and options, allow investors to trade immediately following the release of important macroeconomic indi-

Table 10: The Impact of Distance-to-trading on the Speed of Price Discovery: For Fixed Macroeconomic Announcement Type

	Post-announcement time windows									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	0	$[1,59]$	$[60,119]$	$[120,179]$	$[180,239]$	0	$[1,59]$	$[60,119]$	$[120,179]$	$[180,239]$
	Panel A: PMI (Non-Trading Hours Only)					**Panel C: Trade**				
$R^{[0,239])}$	0.249***	0.201***	0.148***	0.144***	0.232***	0.007	0.157***	0.201***	0.332***	0.262***
	[5.49]	[4.68]	[5.27]	[5.67]	[6.02]	[0.67]	[3.78]	[4.98]	[7.54]	[6.35]
$R^{[0,239]} \times Dur$	0.034***	-0.005	0.008	-0.015**	-0.018*	0.150***	0.017	-0.017	-0.056**	-0.092***
	[3.48]	[-0.48]	[1.27]	[-2.28]	[-1.88]	[6.94]	[0.55]	[-0.70]	[-2.34]	[-4.42]
Dur	2.212	-3.556	3.510**	-1.875	0.361	-1.281	-6.022	6.857	3.979	-6.450
	[0.67]	[-1.43]	[2.32]	[-1.05]	[0.22]	[-0.23]	[-0.98]	[1.19]	[0.81]	[-1.32]
Constant	-9.932*	18.680***	-7.461*	-0.856	-6.407	1.839	4.120	-1.700	-13.423**	9.039*
	[-1.85]	[3.61]	[-1.72]	[-0.21]	[-1.31]	[0.96]	[0.79]	[-0.40]	[-2.51]	[1.80]
R^2	0.374	0.262	0.257	0.182	0.353	0.480	0.199	0.314	0.450	0.344
N	179	179	179	179	179	129	129	129	129	129
	Panel B: IP					**Panel D: CTCM**				
$R^{[0,239])}$	-0.010	0.140***	0.289***	0.202***	0.338***	0.013	0.177***	0.198***	0.299***	0.264***
	[-0.76]	[3.81]	[4.89]	[6.04]	[5.72]	[0.54]	[4.79]	[4.27]	[8.36]	[9.24]
$R^{[0,239]} \times Dur$	0.103**	0.066***	-0.068*	-0.001	-0.098***	0.224***	-0.014	0.002	-0.130**	-0.051
	[2.57]	[3.17]	[-1.96]	[-0.07]	[-2.85]	[4.74]	[-0.36]	[0.05]	[-2.06]	[-0.81]
Dur	-9.622	-8.468*	-1.623	6.072	9.354	-12.949*	6.259	3.462	-3.517	3.691
	[-1.24]	[-1.80]	[-0.21]	[1.20]	[1.36]	[-1.87]	[1.15]	[0.70]	[-0.50]	[0.51]
Constant	2.623*	-1.960	2.751	-0.235	-2.681	-0.249	-3.743	-4.839	11.646*	-2.458
	[1.76]	[-0.47]	[0.45]	[-0.05]	[-0.48]	[-0.05]	[-0.64]	[-0.84]	[1.84]	[-0.47]
R^2	0.289	0.282	0.312	0.320	0.377	0.244	0.224	0.276	0.388	0.437
N	123	123	123	123	123	110	110	110	110	110

Notes: The table reports the regression results of $R_i^t = \beta_0 + \beta_1 R_i^{[-10,239]} + \beta_2 R_i^{[-10,239]} \times Dur_i + \beta_3 Dur_i + \epsilon_i$. The dependent variables are the log returns of CSI 300 for the respective time intervals and are in basis points. Dummy variable Non_i equals 1 if the announcement is released in non-trading hours. Dur_i is the time between announcement time and the first trading time after the announcement and is in days (i.e., distance-to-trading). Panels A, B, C, and D are based on the sample of PMI announcements, industrial production announcements, trade balance announcements, and central treasury cash management announcements, respectively. The sample period is from January 2009 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

cators. As a result, although many important indexes, such as non-farm payroll and GDP, are announced at 8:30 eastern time, one hour before the stock market opens, the distance-to-trading is close to zero for investors in the U.S. market. We argue that, despite the apparent institutional differences between the two markets – such as investor composition and overall market efficiency – a comparison of the two markets can still offer some insights into how investors' distance-to-trading affects the speed of price discovery.

For each Chinese macroeconomic announcement released during non-trading hours, we match it with a U.S. announcement with a similar total market impact, proxied by the 250-minute returns of the CSI 300 index and the S&P 500 index. We obtain 565 pairs of matched announcements. Panels C of Internet Appendix Table IA.IV shows that the matched Chinese and U.S. announcements have a very similar total market impact.

For this matched sample of announcements, we perform similar tests specified in the previous sections and report the results in Panel (a) of Figure 4 and Panels C and D of Table 9. The results are consistent and all suggest that the Chinese announcements made during non-trading hours have a faster speed of price discovery than their counterparts in the U.S. market, once the market opens to trade. While we acknowledge that many institutional differences between the two markets might affect the estimation, we think the effect of distance-to-trading is more likely to be underestimated because the Chinese stock market is often considered as being less efficient than the U.S. market. The fact that the price discovery process is still faster for the Chinese non-regular-hours announcements highlights the important role of pre-trading learning on the speed of price discovery.

In the Internet Appendix, we further explore the robustness of our main results. First, we examine if our results are driven by other potential heterogeneities across macroeconomic announcements in China. According to Table IA.V, we note that some macroeconomic data, e.g., CPI/PPI, GDP, and Caixin among others are released regularly under a pre-scheduled timetable of which investors are pre-informed. Other announcements such as M2, Trade, and FDI among others arrive in the market as random as if they are unscheduled. In tables IA.VI, IA.VII, and IA.VIII, we present the estimation results by differentiating pre-scheduled and unscheduled macro news. We show that our baseline results are not affected by the exact scheduling of macroeconomic announcements. Second, we are concerned if the channel of "learning from the price" is still present overnight in China as China domestic A-share investors may check and learn from the real-time prices of "China Concepts Stocks" listed in the U.S. market while the U.S. market is however open for trades. To rule out this possibility, we rerun our estimations by focusing on the sample of Chinese macroeconomic announcements which fall between Saturday 4:00 a.m. and Monday 9:30 a.m. in Beijing time when the U.S. market is closed. Our results in Table IA.IX again suggest that the price discovery process in China is faster if there is greater pre-trading learning. In particular, we show that the pre-trading learning among domestic investors in China is not sourced from the market prices in the U.S. market.

6 Conclusion

Exploiting a unique institutional feature of China's capital market, which exhibits significant timing heterogeneity with respect to its macro announcements, we isolate and identify the impacts of learning on shifting the process of price discovery in absence of the confounding factor of market trading. We show that investors' learning before trading, as proxied by the length of the non-trading period (distance-to-trading) or the related posts, fans, and interactions on China's largest social media platform, could lead to faster and more efficient price discovery once the market opens for trading. Our documented empirical facts are closely consistent with the models that predict increased learning among less informed investors enhances the price informativeness of equity prices in equilibrium. As the timing arrangements of macroeconomic announcements in China can benefit retail investors by triggering learning and reducing noise trading, this contributes to the overall social welfare of the Chinese economy. Our paper is also the first to systematically study the impacts of a comprehensive list of macroeconomic announcements in China on its stock markets. Importantly, our findings may also be of interest to other emerging market countries seeking to enhance the price efficiency of their financial markets.

Appendix

A An Illustrative Model

In a simple framework, we examine the asset pricing implications of an environment when an important macro news announcement is made outside trading hours. Our model extends Vayanos and Wang (2012) by allowing for two types of liquidity-supplying traders who differ in their level of sophistication in processing a macro announcement, though they jointly accommodate the trading needs of liquidity-demanding traders. Specifically, if an announcement is made within trading sessions, only those sophisticated traders who are informed of the macro news announcement would immediately respond through trading in a short window. In contrast, those non-trading hours after the announcement but before trading give less informed traders more time to learn from the public announcement without trading. Therefore, a significantly larger fraction of liquidity-supplying traders can be well informed of the macro news announcement once the market reopens for trading.

A.1 Environment

We consider a financial market with one risky asset (e.g., a stock market portfolio) and a risk-free asset. The risky asset pays off a random dividend D to be realized in period 2 with $D \sim \mathbf{N}(0, \sigma^2)$. The supply of the risky asset is normalized to be a unit share. The return on the risk-free asset is normalized for $r = 0$ such that the risk-free asset has a constant price of 1 and serves as the numeraire. The financial market opens for trading in three periods for $t = 0, 1, 2$. The price of the risky asset is denoted by p_t and is endogenously determined in equilibrium through market clearing. In case of no uncertainty regarding the risky asset payoff once period 2 unfolds, $p_2 = D$. We therefore focus on the equilibrium price of the risky asset in period 1 (i.e., p) for which we suppress the period index for simplicity.

We assume there is a unit measure of investors who have CARA preferences and derive utility from wealth in period 2. Without loss of generality, we simply set the risk aversion to one for ease of notation. All traders are homogeneous in period 0 and become heterogeneous in period 1, which then justifies the need for trading in period 1. The heterogeneity across traders is driven by the information heterogeneity and the realizations of endowment shocks. Specifically, a total fraction $1-\pi \in [0,1]$ of traders receive an endowment of $z \cdot D$ in period 2 with shocks to endowment $z \sim \mathbf{N}(0, \sigma_z^2)$, which is independent of the asset payoff D and realized in period 1. On the other hand, the fraction π of traders receive no endowment. Traders receiving endowment shocks would initiate trading in period 1 and demand market liquidity, and thus they are considered liquidity demanders. Therefore, z shocks can be interpreted as liquidity shocks. In equilibrium, traders receiving no such liquidity shocks would accommodate the risk-sharing trading needs and provide liquidity. They are considered the liquidity suppliers.

In addition, our model entertains the fact that an announcement may fall outside the trading session. Without loss of generality, we allow for an announcement to arrive *before* the trading period at $t = 1$. In the following, we model a macro announcement as a signal s that partially reveals the dividend payoff D subject to a noise ϵ:

$$s = D + \epsilon, \quad \text{s.t.} \quad \epsilon \sim \mathbf{N}(0, \sigma_\epsilon^2), \tag{A.1.E}$$

where σ_ϵ^2 denotes the variance of the signal noise.

A.2 Market Equilibrium with Different Types of Learning Investors

Importantly, in line with Kacperczyk et al. (2016), we further assume that a fraction $\delta \in (0,1)$ of liquidity-supplying traders of mass π are less sophisticated. It takes more time for them to be attentive to public signals and then extract information from the macro announcement (i.e., these traders are indexed as n). On the other hand, a fraction $1-\delta$ of liquidity suppliers would be sophisticated enough to immediately process macro announcements and draw useful information (i.e., sophisticated traders are indexed as a). For simplicity, we denote a constant hazard rate of "learning from the signal" per an infinitesimal time among the less sophisticated investors (i.e., $\alpha > 0$). Hence, the fraction of less sophisticated investors who stay uninformed from the macro announcement upon market trading follows that

$$\delta(\lambda) = \delta e^{-\alpha\lambda} \in [0, \delta], \tag{A.2.E}$$

A faster learning ratio α with a more extended duration of distance-to-trading λ triggers a stronger learning effect, which increases the total size of informed investors among those unsophisticated ones upon market trading. Therefore, our derived hazard rate of learning implies that information processing takes time before trading but eventually reduces the number of less informed traders over time in the market, which is in line with Dugast and Foucault (2018). Therefore, across announcement events, we should see a greater fraction of less informed liquidity-supplying traders in case the investors are left with distance-to-trading that is too short. As λ is larger, they are given more time to turn themselves into more informed liquidity suppliers. We then label the liquidity-demanding traders as type d. Following Vayanos and Wang (2012), we assume that all liquidity demanders are informed of the macro announcement. This holds because these traders can infer s perfectly from the price because they also observe their own liquidity shock realizations. Therefore, investors' sophistication matters for those less informed liquidity-supplying traders only.

For trader type $i = n, a, d$, upon the opening of trading in period 1, a trader optimizes the demand q of the risky asset conditional on her information set adjusted for learning, $\mathcal{I}$. It follows that investors solve the following utility maximization problem in a general form:

$$\max_{q_i} U(W_{2,i}) = -\mathbb{E}_{\mathcal{I}} e^{-W_{2,i}}, \tag{A.3.E}$$

$$\text{s.t.} \quad W_{2,i} = W_1 + q_i(D - p) + \mathbb{I}_{i=d} \cdot zD. \tag{A.4.E}$$

Accordingly, W_1 is the initial wealth that is common to all traders. Here, $\mathbb{I}_{i=d} = 1$ if the trader is a liquidity demander conditional on receiving a shock realization z in period 1, and $\mathbb{I}_{i=d} = 0$ for liquidity-supplying traders. Once trading is open for period 1, sophisticated liquidity-supplying traders carry the signal s in their information set whereas less informed traders are taking the market price as given. Therefore, the information heterogeneity and the liquidity shocks are driving traders apart, which delivers the following optimal asset demands of trader types:

$$q_n(p) = \frac{\mathbb{E}(D|p) - p}{\sigma^2_{D|p}}, \tag{A.5.E}$$

$$q_a(p) = \frac{\mathbb{E}(D|s) - p}{\sigma^2_{D|s}}, \tag{A.6.E}$$

$$q_d(p) = \frac{\mathbb{E}(D|s) - p}{\sigma^2_{D|s}} - z. \tag{A.7.E}$$

The financial market equilibrium is then determined given the asset demands of different types of traders, and the market equilibrium price in period 1 satisfies the market-clearing condition such that

$$\pi\delta(\lambda)q_n(p) + \pi(1 - \delta(\lambda))q_a(p) + (1 - \pi)q_d(p) = 1. \tag{A.8.E}$$

A.3 Equilibrium Solutions

We then solve for the model equilibrium via the guess and verify method. We first conjecture the equilibrium price is in linear form such that

$$p = a + b(s - c \cdot z), \tag{A.9.E}$$

where a, b, and c are constants to be determined.

Applying Bayes' rule, we have the conditional expectations and posterior variances regarding dividends among sophisticated investors who act upon the signal:

$$\mathbb{E}(D|s) = \gamma_S \cdot s, \tag{A.10.E}$$

$$\sigma^2_{D|s} = \frac{1}{1/\sigma^2 + 1/(\sigma^2_\epsilon)} = (1 - \gamma_S)\sigma^2, \tag{A.11.E}$$

where $\gamma_S = \frac{\sigma^2}{\sigma^2 + \sigma^2_\epsilon}$. Given the equilibrium price of the risky asset as in Equation (A.9.E), the expectation and variances regarding D among those uninformed investors are

$$\mathbb{E}(D|p) = \gamma_N(s - cz), \tag{A.12.E}$$

$$\sigma^2_{D|p} = \frac{1}{1/\sigma^2 + 1/(\sigma^2_\epsilon + c^2\sigma^2_z)} = (1-\gamma_N)\sigma^2, \tag{A.13.E}$$

where $\gamma_N = \frac{\sigma^2}{\sigma^2+\sigma^2_\epsilon+c^2\sigma^2_z}$.

Plugging these equations into Equation (A.8.E), we can solve for the coefficients such that

$$a = -\frac{\sigma^2_{D|s}\sigma^2_{D|p}}{\pi\delta(\lambda)\sigma^2_{D|s} + (1-\pi\delta(\lambda))\sigma^2_{D|p}} = -(1-b)\sigma^2, \tag{A.14.E}$$

$$b = \frac{\gamma_N\pi\delta(\lambda)\sigma^2_{D|s} + \gamma_S(1-\pi\delta(\lambda))\sigma^2_{D|p}}{\pi\delta(\lambda)\sigma^2_{D|s} + (1-\pi\delta(\lambda))\sigma^2_{D|p}}, \tag{A.15.E}$$

$$c = \frac{1-\pi}{1-\pi\delta(\lambda)}\sigma^2_\epsilon. \tag{A.16.E}$$

A.4 Model Predictions

Next, we examine the asset pricing implications of varying the distance-to-trading λ given some learning ratio of α among unsophisticated investors. Intuitively, we see that $\lambda \to \infty$ captures the extreme scenario when all less informed traders eventually are aware of the informative signals delivered through macro announcements upon trading. To the other extreme, when an announcement exactly falls within the trading session in period 1, $\lambda \to 0$ gives that a sizable fraction of liquidity-supplying traders are not able to incorporate the macro announcements immediately for trading (i.e., $\delta(\lambda) \to \delta$). We then derive the theoretical results by focusing on a few market statistics: first, the measures of the price informativeness, which capture the degree of efficiency for the equilibrium price to reflect the fundamentals; second, the quantity of trading volume; and third, the volatility of returns.

We evaluate the stock price informativeness (PI) using the following two metrics. Both measures are to reflect the degree of information quality of stock prices revealing the dividend payoff:

$$PI_1 = \frac{1}{\sigma^2_{D|p}} = \frac{1}{\sigma^2} + \frac{1}{\sigma^2_\epsilon + c^2\sigma^2_z}, \tag{A.17.E}$$

$$PI_2 = \frac{cov(P,D)}{\sigma_P} = \frac{\sigma}{\sqrt{\sigma^2 + \sigma^2_\epsilon + c^2\sigma^2_z}}. \tag{A.18.E}$$

Given that $\delta'(\lambda) < 0$, it's easy to show that $\frac{\mathrm{d}PI_1}{\mathrm{d}\lambda} > 0$ and $\frac{\mathrm{d}PI_2}{\mathrm{d}\lambda} > 0$ as c decreases in λ. That is, when a macro announcement arrives much earlier before a trading session begins, traders are given more time to start processing the news and extract new information. As a result, upon market opening, more traders would be able to supply liquidity by trading on the macro announcement. Hence, the noise component is reduced in the equilibrium price. Stock prices become more informative about the asset payoffs.

Second, since we have normalized the total number of shares of the risky asset to one, the trading volume denotes the turnover rate. We take as given that the trading takes place between the liquidity suppliers of two types on one side and the liquidity demanders on the other side. For tractability, we examine the variance of the directional quantity trading of the liquidity demanders and see how the distance-to-trading affects this measure:

$$Volume = Var\{(1-\pi)[\frac{(\gamma_s - b)s - a + bcz}{\sigma^2_{D|s}} - z]\} \tag{A.19.E}$$

$$= [\frac{1-\pi}{\sigma^2_{D|s}}]^2[(\gamma_s - b)^2(\sigma^2 + \sigma^2_\epsilon) + [bc - \sigma^2_{D|s}]^2\sigma^2_z]. \tag{A.20.E}$$

Equation (A.20.E) suggests that the depth of trading volume is driven by two forces. In case of greater distance-to-trading for a larger λ, the first term in the brackets reflects the degree of noise reduction in which uncertainty about the asset payoff is lowered. This mitigates the trading volume driven by price uncertainty. The second term captures the market liquidity increase driven by liquidity-supplying trades due to increased price responsiveness to the dividend payoff.

Third, we examine the variance of gross returns of the risky asset $D - p$, which reflects the return volatility ($RetVol$) upon market opening. Specifically, it gives that

$$RetVol = Var(D - p) = (1 - b)^2\sigma^2 + b^2[\sigma_\epsilon^2 + c^2\sigma_z^2]. \tag{A.21.E}$$

Equation (A.21.E) suggests that the return volatility is driven by two forces. In case of greater distance-to-trading for larger λ, the first term reflects the degree of noise reduction in which uncertainty about asset payoff σ^2 is scaled down, as we can show that $\frac{\mathrm{d}b}{\mathrm{d}\lambda} > 0$. The second term captures the return volatility driven by price responsiveness to the dividend payoff for increased price informativeness. This precisely defines the trade-off between a reduction in dividend uncertainty (i.e., noise reduction), and increased price informativeness for greater volatility (i.e., information-driven volatility).

In sum, we have shown within a simple model structure that the duration of information processing before trading as measured by distance-to-trading significantly affects the price informativeness of the payoff fundamentals. In addition, trading volume and return volatility in equilibrium are shifted by the effect of both the increased price informativeness and the reduced market uncertainty given that learning can happen before trading. It may well be that the effect of increased price informativeness dominates through market clearing upon market opening, which later on drives down market noise and leads to lowered trading volume and return volatility as time evolves.

References

Ai, Hengjie, and Ravi Bansal, 2018, Risk preferences and the macroeconomic announcement premium, *Econometrica* 86, 1383–1430.

Ai, Hengjie, Ravi Bansal, and Leyla Jianyu Han, 2022, Information acquisition and the pre-announcement drift, *Available at SSRN 3964349* .

Ai, Hengjie, Leyla Jianyu Han, and Lai Xu, 2023, Information-driven volatility, *Available at SSRN 3961096* .

Allen, Franklin, Jun, Chenying Zhang, and Mengxin Zhao, 2012, *China's Financial System: Opportunities and Challenges*, 63–143 (University of Chicago Press).

Allen, Franklin, Jun Qian, and Meijun Qian, 2005, Law, finance, and economic growth in china, *Journal of Financial Economics* 77, 57–116.

Allen, Franklin, Jun Qian, Chenyu Shan, and Julie Zhu, 2023, Dissecting the long-term performance of the chinese stock market, *The Journal of Finance, Forthcoming* .

Barclay, Michael J., and Terrence Hendershott, 2003, Price Discovery and Trading After Hours, *The Review of Financial Studies* 16, 1041–1073.

Barclay, Michael J., and Terrence Hendershott, 2008, A comparison of trading and non-trading mechanisms for price discovery, *Journal of Empirical Finance* 15, 839–849.

Beber, Alessandro, and Marco Pagano, 2013, Short-selling bans around the world: Evidence from the 2007–09 crisis, *The Journal of Finance* 68, 343–381.

Ben-rephael, Azi, Bruce I. Carlin, Zhi Da, and Ryan Israelsen, 2021, Information consumption and asset pricing, *Journal of Finance* 76, 357–394.

Biais, Bruno, Pierre Hillion, and Chester Spatt, 1999, Price Discovery and Learning during the Preopening Period in the Paris Bourse, *Journal of Political Economy* 107, 1218–1248.

Boehmer, Ekkehart, and Juan (Julie) Wu, 2012, Short Selling and the Price Discovery Process, *The Review of Financial Studies* 26, 287–322.

Boguth, Oliver, Adlai J. Fisher, Vincent Gregoire, and Charles Martineau, 2023, Noisy fomc returns: Information, price pressure, and post-announcement reversals, *Available at SSRN 4131740* .

Bollerslev, Tim, Jia Li, and Yuan Xue, 2018, Volume, Volatility, and Public News Announcements, *The Review of Economic Studies* 85, 2005–2041.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2014, High-Frequency Trading and Price Discovery, *Review of Financial Studies* 27, 2267–2306.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2019, Price discovery without trading: Evidence from limit orders, *The Journal of Finance* 74, 1621–1658.

Brunnermeier, Markus K, Michael Sockin, and Wei Xiong, 2021, China's Model of Managing the Financial System, *The Review of Economic Studies* 89, 3115–3153.

Brusa, Francesca, Pavel Savor, and Mungo Wilson, 2020, One central bank to rule them all, *Review of Finance* 24, 263–304.

Carpenter, Jennifer N., Fangzhou Lu, and Robert F. Whitelaw, 2021, The real value of china's stock market, *Journal of Financial Economics* 139, 679–696.

Cieslak, Anna, Adair Morse, and Annette Vissing-Jorgensen, 2019, Stock returns over the fomc cycle, *Journal of Finance* 74, 2201–2248.

Da, Zhi, Joseph Engelberg, and PengJie Gao, 2011, In search of attention, *The Journal of Finance* 66, 1461–1499.

Dugast, Jérôme, and Thierry Foucault, 2018, Data abundance and asset price informativeness, *Journal of Financial Economics* 130, 367–391.

Fama, Eugene F., 1970, Efficient capital markets: A review of theory and empirical work, *The Journal of Finance* 25, 383–417.

Fisher, Adlai, Charles Martineau, and Jinfei Sheng, 2022, Macroeconomic Attention and Announcement Risk Premia, *The Review of Financial Studies* 35, 5057–5093.

Garfinkel, Jon A., 2009, Measuring Investors' Opinion Divergence, *Journal of Accounting Research* 47, 1317–1348.

Geng, Zhe, and Jun Pan, 2019, The soe premium and government support in china's credit market, Working Paper 26575, National Bureau of Economic Research.

Goldstein, Itay, and Liyan Yang, 2015, Information diversity and complementarities in trading and information acquisition, *The Journal of Finance* 70, 1723–1765.

Grossman, S. J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *The American Economic Review* 70, 393–408.

Guo, Rui, Dun Jia, and Xi Sun, 2023, Information acquisition, uncertainty reduction, and pre-announcement premium in china, *Review of Finance* 27, 1077–1118.

He, Zhiguo, and Wei Wei, 2023, China's financial system and economy: A review, *Annual Review of Economics* 15, 451–483.

Hu, Grace Xing, Jun Pan, and Jiang Wang, 2017, Early peek advantage? efficient price discovery wiht tiered information disclosure, *The Journal of Financial Economics* 126, 399–421.

Hu, Grace Xing, Jun Pan, Jiang Wang, and Haoxiang Zhu, 2022, Premium for heightened uncertainty: Explaining pre-announcement market returns, *Journal of Financial Economics* 145, 909–936.

Hu, Grace Xing, and Jiang Wang, 2022, A review of china's financial markets, *Annual Review of Financial Economics* 14, 465–507.

Kacperczyk, Marcin, Jaromir Nosal, and Savitar Sundaresan, 2023, Market power and price informativeness, *Available at SSRN 3137803* .

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2016, A rational theory of mutual funds' attention allocation, *Econometrica* 84, 571–626.

Kim, Oliver, and Robert E. Verrecchia, 1991, Trading volume and price reactions to public announcements, *Journal of Accounting Research* 29, 302–321.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *The Econometrica* 53, 1315–1335.

Liu, Jianan, Robert F. Stambaugh, and Yu Yuan, 2019, Size and value in china, *Journal of Financial Economics* 134, 48–69.

Lucca, David, and Emanuel Moench, 2015, The pre-fomc announcement drift, *Journal of Finance* 70, 329–371.

Morris, Stephen, and Hyun Song Shin, 2018, Central bank forward guidance and the signal value of market prices, *AEA Papers and Proceedings* 108, 572–77.

Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics* 80, 563–602.

Savor, Pavel, and Mungo Wilson, 2013, How much do investors care about macroeconomic risk? evidence from scheduled economic announcements, *Journal of Financial and Quantitative Analysis* 48, 343–375.

Savor, Pavel, and Mungo Wilson, 2014, Asset pricing: A tale of two days, *Journal of Financial Economics* 113, 171–201.

Vayanos, Dimitri, and Jiang Wang, 2012, Liquidity and asset returns under asymmetric information and imperfect competition, *The Review of Financial Studies* 25, 1339–1365.

Internet Appendix for "Learning, Price Discovery, and Macroeconomic Announcements"

In this document, we report additional results of robustness. First, we examine our measured learning intensity based on the related Weibo posts are relevant to the stock prices and stock market in general. In specific, we focus on those macro-related Weibo posts that also mention the keyword "stock" in its title or in the main text of the posts. As shown in Table IA.I, we run the estimation and find the results are robust and similar to the baseline results.

Second, we show that using data on fans and interaction across posts to proxy for the learning intensity does not alter our baseline results as well. Table IA.II reports the impact of investor learning, as measured by the number of fans and the total number of retweets, comments, and likes, on investors' average information quality.

Third, we establish the robustness of our results if trading volume is not normalized by outstanding shares. Table IA.III reports the impact of distance-to-trading and pre-trading investor learning on trading volume.

Fourth, we show the details of our matched and refined announcement samples for demonstrating the robustness of our main results. Table IA.IV reports the summary statistics of the post-announcement returns for the full sample of macroeconomic announcements in China and the U.S., as well as the matched sample of macroeconomic announcement days.

In addition, one may be concerned that the speed of price discovery is related to the release schedule of macroeconomic news. We explore if our results are driven by the fact that some of the macroeconomic announcements in China have their data release schedule pre-fixed and pre-informed to domestic investors. For example, the release time of major macroeconomic indicators from the National Bureau of Statistics is pre-scheduled. If the release time is endogenously related to their release schedule and market impact, our empirical results might be biased.

Table IA.V first reports the median release time and the release schedule of the major macroeconomic indicators in China. The pre-scheduled announcements include those for the

Table IA.I: The Impact of Pre-Trading Investor Learning Activities on the Speed of Price Discovery: Weibo Posts Mentioning "Stock"

	posts (Mentioning "Stock")	Post-announcement time windows				
	(1)	(2)	(3)	(4)	(5)	(6)
	learning	0	$[1, 59]$	$[60, 119]$	$[120, 179]$	$[180, 239]$
Panel A: Non-trading-hours and Trading-hours News						
Dur	0.003***					
	[2.87]					
$R^{[-10,239]}$		0.199***	0.182***	0.162***	0.142***	0.242***
		[8.31]	[11.15]	[11.02]	[10.69]	[11.55]
$R^{[-10,239]} \times Learning$		2.395***	-0.004	-0.141	-0.492	-1.780**
		[2.70]	[-0.01]	[-0.26]	[-0.98]	[-2.18]
Learning		129.257	-110.118	0.762	82.720	-19.880
		[1.36]	[-0.77]	[0.01]	[0.92]	[-0.21]
Constant	0.005***	-1.736	1.334	-1.009	-4.155**	2.294
	[8.74]	[-0.78]	[0.64]	[-0.55]	[-2.25]	[1.08]
R^2	0.020	0.293	0.257	0.254	0.201	0.357
N	584	584	584	584	584	584
Panel B: Non-trading-hours News Only						
Dur	0.002**					
	[2.40]					
$R^{[-10,239]}$		0.327***	0.201***	0.169***	0.133***	0.132***
		[11.71]	[7.94]	[8.06]	[7.24]	[5.65]
$R^{[-10,239]} \times Learning$		2.302***	-0.149	-0.937**	-0.776**	-0.033
		[4.72]	[-0.22]	[-2.42]	[-2.01]	[-0.06]
Learning		237.222**	-111.223	-130.920	38.792	21.511
		[2.31]	[-0.64]	[-1.58]	[0.39]	[0.18]
Constant	0.005***	-4.749	0.288	0.609	-0.758	1.156
	[4.51]	[-1.44]	[0.10]	[0.23]	[-0.29]	[0.48]
R^2	0.016	0.506	0.275	0.252	0.172	0.216
N	305	305	305	305	305	305

Notes: Column (1) reports the regression results of $Learning_i = \alpha + \beta_1 Dur_i + \epsilon_i$. The dependent variables $Learning_i$ are proxied by the number of related Weibo posts mentioning "Stock" during the 72 hours before the announcement time (for trading-hours news) or the market opening time of the next trading day (for non-trading-hours news). The unit for the number of posts is in thousands. Columns (2) to (6) report the regression results of $R_i^t = \beta_0 + \beta_1 R_i^{[-10,239]} + \beta_2 R_i^{[-10,239]} \times Learning_i + \beta_3 Learning_i + \epsilon_i$. The dependent variables are the log returns of CSI 300 for the respective time intervals and are in basis points. The sample period is from January 2018 to December 2020. The *t*-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table IA.II: The Impact of Pre-Trading Investor Learning on Volatility, Turnover, and Bid-ask Spreads

	Number of Fans					Number of Retweets, Comments, and Likes				
	(1) 0	(2) [1, 59]	(3) [60, 119]	(4) [120, 179]	(5) [180, 239]	(6) 0	(7) [1, 59]	(8) [60, 119]	(9) [120, 179]	(10) [180, 239]
Panel A: Volatility										
Learning	1.046***	0.228	0.148	-0.025	-0.269*	0.649	0.070	0.002	-0.024	-0.350**
	[3.18]	[1.32]	[0.96]	[-0.16]	[-1.90]	[1.57]	[0.48]	[0.01]	[-0.17]	[-2.29]
Constant	-0.007	0.029	-0.002	0.004	0.020	0.031	0.042*	0.008	0.004	0.021
	[-0.15]	[1.22]	[-0.10]	[0.16]	[0.78]	[0.67]	[1.79]	[0.35]	[0.15]	[0.83]
R^2	0.021	0.004	0.002	0.000	0.005	0.007	0.000	0.000	0.000	0.007
N	584	584	584	584	584	584	584	584	584	584
Panel B: Turnover										
Learning	1.360***	0.716**	0.392	0.149	-0.386**	1.040**	0.655**	0.309	0.255	-0.427***
	[4.03]	[2.42]	[1.61]	[0.66]	[-2.09]	[2.42]	[2.00]	[1.32]	[1.11]	[-2.64]
Constant	-0.056	0.003	0.015	0.014	0.026	-0.018	0.017	0.025	0.010	0.023
	[-1.12]	[0.09]	[0.46]	[0.38]	[0.78]	[-0.37]	[0.47]	[0.79]	[0.27]	[0.70]
R^2	0.029	0.015	0.006	0.001	0.006	0.015	0.011	0.003	0.002	0.006
N	584	584	584	584	584	584	584	584	584	584
Panel C: Bid-ask Spread										
Learning	-0.072	-0.386*	-0.563**	-0.442	-0.274	-0.291	-0.639**	-0.774**	-0.709**	-0.735**
	[-0.24]	[-1.70]	[-2.02]	[-1.52]	[-0.96]	[-0.88]	[-2.21]	[-2.35]	[-2.27]	[-2.54]
Constant	-0.037	-0.045	-0.079	-0.096*	-0.103**	-0.025	-0.036	-0.075	-0.087*	-0.081*
	[-0.64]	[-0.92]	[-1.61]	[-1.91]	[-2.15]	[-0.45]	[-0.75]	[-1.56]	[-1.77]	[-1.70]
R^2	0.000	0.002	0.005	0.003	0.001	0.001	0.006	0.008	0.007	0.008
N	584	584	584	584	584	584	584	584	584	584

Notes: This table reports the regression results of $Y_i^t = \alpha^t + \beta_1^t Learning_i + \epsilon_i^t$, where the sample period is from January 2018 to December 2020. The dependent variables are the normalized weighted averages of return volatility, turnover, and bid-ask spreads of CSI 300 components stocks for the respective time intervals. $Learning_i$ is the number of fans and the total number of retweets, comments, and likes following related Weibo posts during the 72-hour period prior to announcement time (for trading-hours news) and the market opening time of the next trading day (for non-trading-hours news). The units for the number of retweets, comments, and likes are in thousands and the unit for the number of fans is in billions. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table IA.III: The Impact of Distance-to-trading and Pre-Trading Investor Learning on Trading Volume

	Post-announcement time windows									
	(1) 0	(2) [1, 59]	(3) [60, 119]	(4) [120, 179]	(5) [180, 239]	(6) 0	(7) [1, 59]	(8) [60, 119]	(9) [120, 179]	(10) [180, 239]
Dur	0.160***	0.062***	0.013	0.011	-0.017					
	[4.56]	[2.84]	[0.83]	[0.67]	[-1.27]					
Posts						1.563***	0.966**	0.423	0.039	-0.726***
						[2.73]	[2.23]	[1.36]	[0.13]	[-2.64]
Constant	-0.005	0.018	0.020	0.019	0.027	-0.046	-0.001	0.027	0.018	0.044
	[-0.15]	[1.00]	[1.18]	[1.13]	[1.64]	[-0.86]	[-0.02]	[0.88]	[0.56]	[1.28]
R^2	0.020	0.008	0.000	0.000	0.001	0.011	0.011	0.003	0.000	0.006
N	1673	1673	1673	1673	1673	584	584	584	584	584
Fans	1.245***	0.652***	0.329*	0.122	-0.278*					
	[3.96]	[2.59]	[1.68]	[0.71]	[-1.68]					
Interactions						0.825**	0.610**	0.268	0.184	-0.399**
						[2.48]	[2.35]	[1.41]	[1.03]	[-2.58]
Constant	-0.069	-0.006	0.021	0.011	0.033	-0.027	0.006	0.030	0.009	0.035
	[-1.33]	[-0.18]	[0.70]	[0.35]	[0.98]	[-0.52]	[0.18]	[1.01]	[0.30]	[1.09]
R^2	0.023	0.015	0.005	0.001	0.003	0.008	0.012	0.003	0.001	0.005
N	584	584	584	584	584	584	584	584	584	584

Notes: This table reports the regression results of $Volume_i^t = \alpha^t + \beta_1^t Dur_i + \epsilon_i^t$, where the sample period is from January 2009 to December 2020, and the regression results of $Volume_i^t = \alpha^t + \beta_1^t Learning_i + \epsilon_i^t$, where the sample period is from January 2018 to December 2020. The dependent variables are the normalized weighted averages of trading volume of CSI 300 components stocks for the respective time intervals. Dur_i is the time between announcement time and the first trading time after the announcement and is in days (i.e., distance-to-trading). $Learning_i$ is the number of related Weibo posts (or the fans of these posts, or the retweets, comments, and likes of these posts) during the 72-hour period prior to announcement time (for trading-hours news) and the market opening time of the next trading day (for non-trading-hours news). The units for the number of posts, retweets, comments, and likes are in thousands and the unit for the number of fans is in billions. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table IA.IV: Post-announcement Market Returns on Macroeconomic Announcement Days in China

Post-Ann Return$^{[-10,239]}$	Obs	Mean	Std.	Min	P25	P50	P75	Max
Panel A: Full Sample								
Non-Trd (China)	717	12.40	144.50	-555.76	-58.04	8.90	82.21	642.64
Trd(China)	956	-2.11	147.13	-745.77	-78.93	-5.00	73.36	633.73
Trd(U.S.)	1899	-0.70	61.12	-305.81	-31.02	1.50	32.98	554.79
Panel B: Matched Sample with Chinese Announcements within trading hours								
Non-Trd (China)	639	4.06	108.67	-457.19	-53.63	4.20	66.09	362.57
Trd(China)	639	4.06	108.66	-456.43	-53.80	4.40	65.79	363.57
Panel C: Matched Sample with U.S. Announcements within trading hours								
Non-Trd (China)	565	4.65	78.82	-228.22	-42.62	3.94	53.51	251.85
Trd(U.S.)	565	4.65	78.82	-228.24	-42.62	3.99	53.60	252.30

Notes: We match the sample of announcements released during non-trading hours with announcements within regular trading hours (matching with replacement) based on the $[-10, 239]$ returns of the CSI 300 index and E-mini S&P 500 index futures. The $[-10, 239]$ trading window is extended from 2:50 p.m. of the previous trading day to 2:59 p.m. of the following trading day for non-trading-hours announcements and is from 10 minutes prior to the actual announcement time to the end of the 240th minute afterward for trading-hours announcements. The distribution of the $[-10, 239]$ returns is reported for the full sample of announcements in Panel A. The distribution of the $[-10, 239]$ returns is reported for the matched sample of announcements with Chinese and U.S. announcements released during trading hours in Panels B and Panel C, respectively. The sample period is from January 2009 to December 2020.

CPI/PPI, Caixin, GDP, IP, LPR, PI, PMI, SLF/MLF/PSL, and SPRB, for some of which the announcement dates are fixed but the actual announcement time varies throughout the day. Table IA.VI and Table IA.VII report the impact of distance-to-trading on the speed of price discovery with additional controls for pre-scheduled announcements. Overall, the results are robust and similar to the baseline results discussed in Section 3. And as shown in Table IA.VIII, the improvement in the speed of price discovery is not significantly different between groups of pre-scheduled and non-scheduled announcements.

Finally, we notice that the U.S. listed Chinese concept stocks can still be traded in an open market while the Chinese domestic stock market is closed. One may be concerned that investors overnight learning may well be happening through the channel of learning from prices of other markets. It may be the case that the price formation as reflected in the pricing of U.S. listed Chinese concept stocks in advance after the release of Chinese macroeconomic announcements can pass along to Chinese domestic market investors. Such learning effects will then affect the price discovery process once the market reopens. We

Table IA.V: Release Time and Schedule of Major Macroeconomic Indicators in China

Announcement	MedT	Pre-scheduled
CPI/PPI	9:30	Pre-scheduled since 2012
GDP	10:00	Pre-scheduled since 2012
PMI	9:00	Pre-scheduled since 2012
Caixin	9:45	✓
IP	10:00	Pre-scheduled since 2012
M2	16:00	×
Trade	10:58	×
FER	16:00	×
PI	9:30	Pre-scheduled since 2012
RRR	18:06	×
FDI	10:16	×
BOP	16:46	×
Swift	9:00	Pre-scheduled with a few exceptions
SPRB	9:30	✓
FESS	15:53	×
OMO	9:46	×
SLF/MLF/PSL	15:51	×
CTCM	16:31	×
CBS	9:00	×
LPR	9:30	✓

Notes: This table reports whether the release time of 20 major macroeconomic indicators is pre-scheduled. "MedT" refers to the median of the release time. "Pre-scheduled" indicates whether the announcement time is pre-scheduled. "Since 2012" indicates, for announcements released by the National Bureau of Statistics, that only the announcement date was pre-scheduled before 2012, and both the announcement date and time are pre-scheduled from 2012 onward. "Few exceptions" indicates, for "Swift", that indicators are generally released at 9:30 a.m. on the fourth Thursday of each month, with nine exceptions. The sample period is from January 2009 to December 2020.

show that this is not the case in the following. Table IA.IX reports the results for the sample of China's macroeconomic announcements made between Saturday 4:00 a.m. to Monday 9:30 a.m. (i.e., the period when the U.S. stock market is closed). This is to completely shut down the potential confounding channel that learning effects can be sourced from the U.S. listed Chinese concept stocks. Clearly the price discovery process in China is little affected by whether or not the U.S. market is running or not.

Table IA.X reports the keywords matched with the major macroeconomic indicators in China from Weibo.

Table IA.VI: The Impact of Distance-to-trading on the Speed of Price Discovery: Controlling for Pre-Scheduled News

	Post-announcement time windows				
	(1)	(2)	(3)	(4)	(5)
	0	$[1, 59]$	$[60, 119]$	$[120, 179]$	$[180, 239]$
Panel A: Non-trading-hours vs. trading-hours announcements					
$R^{[-10,239]}$	0.008	0.181***	0.196***	0.228***	0.387***
	[0.64]	[9.22]	[10.49]	[10.40]	[16.40]
$R^{[-10,239]} \times Non$	0.245***	0.028	-0.017	-0.050**	-0.205***
	[10.09]	[1.10]	[-0.84]	[-2.19]	[-7.75]
Non	-2.051	3.268	0.569	2.955	-4.742*
	[-0.79]	[1.22]	[0.23]	[1.21]	[-1.84]
$R^{[-10,239]} \times Pre$	0.071***	-0.018	-0.002	-0.069***	0.017
	[3.30]	[-0.75]	[-0.08]	[-2.80]	[0.62]
Pre	-1.896	0.852	2.131	4.274*	-5.360**
	[-0.82]	[0.33]	[0.87]	[1.74]	[-2.04]
Constant	-0.269	2.222	-2.934	-5.214**	6.195**
	[-0.18]	[0.99]	[-1.33]	[-2.24]	[2.55]
R^2	0.249	0.230	0.249	0.232	0.469
N	1673	1673	1673	1673	1673
Panel B: The impact of distance-to-trading					
$R^{[-10,239]}$	0.090***	0.196***	0.186***	0.217***	0.312***
	[4.78]	[11.75]	[12.62]	[12.66]	[16.44]
$R^{[-10,239]} \times Dur$	0.071***	0.002	-0.002	-0.020***	-0.052***
	[6.56]	[0.33]	[-0.38]	[-3.48]	[-4.59]
Dur	-0.186	0.190	2.967***	-1.198	-1.773
	[-0.09]	[0.15]	[2.83]	[-1.21]	[-1.38]
$R^{[-10,239]} \times Pre$	0.029	-0.025	0.004	-0.063**	0.055*
	[1.16]	[-1.06]	[0.18]	[-2.57]	[1.91]
Pre	-1.653	-0.097	3.162	2.973	-4.385*
	[-0.66]	[-0.04]	[1.33]	[1.25]	[-1.67]
Constant	-0.848	4.084**	-4.593**	-2.792	4.149**
	[-0.44]	[2.11]	[-2.57]	[-1.48]	[2.08]
R^2	0.210	0.228	0.251	0.232	0.446
N	1673	1673	1673	1673	1673

Notes: Panel A reports the regression results of $R_i^t = \alpha^t + \beta_1^t R_i^{[-10,239]} + \beta_2^t R_i^{[-10,239]} \times Non_i + \beta_3^t Non_i + \beta_4^t R_i^{[-10,239]} \times Pre_i + \beta_5^t Pre_i + \epsilon_i$. Panel B reports the regression results of $R_i^t = \alpha^t + \beta_1^t R_i^{[-10,239]} + \beta_2^t R_i^{[-10,239]} \times Dur_i + \beta_3^t Dur_i + \beta_4^t R_i^{[-10,239]} \times Pre_i + \beta_5^t Pre_i + \epsilon_i$. The dependent variables are the log returns of CSI 300 for the respective time intervals and are in basis points. Dummy variable Non_i equals 1 if the announcement is released during non-trading hours. Dur_i is the time between announcement time and the first trading time after the announcement and is in days (i.e., distance-to-trading). Dummy variable Pre_i equals 1 if the announcement is pre-scheduled. The sample period is from January 2009 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table IA.VII: The Impact of Distance-to-trading on the Speed of Price Discovery: Controlling for Pre-Scheduled News (Non-trading-hours News Only)

	Post-announcement time windows					
	(1) close-to-open	(2) open-to-0	(3) $[1, 59]$	(4) $[60, 119]$	(5) $[120, 179]$	(6) $[180, 239]$
$R^{[-10,239]}$	0.162***	0.056***	0.223***	0.167***	0.178***	0.215***
	[5.11]	[3.59]	[8.29]	[8.50]	[8.60]	[9.43]
$R^{[-10,239]} \times Dur$	0.029***	0.001	-0.004	0.002	-0.014**	-0.014*
	[3.24]	[0.39]	[-0.43]	[0.26]	[-2.16]	[-1.79]
Dur	0.620	0.189	-0.755	3.680***	-2.594**	-1.140
	[0.33]	[0.31]	[-0.51]	[3.06]	[-2.27]	[-1.01]
$R^{[-10,239]} \times Pre$	0.040	0.012	-0.043	0.019	-0.009	-0.019
	[1.06]	[0.74]	[-1.03]	[0.63]	[-0.35]	[-0.51]
Pre	-0.832	-0.787	2.424	-0.924	2.152	-2.033
	[-0.20]	[-0.46]	[0.55]	[-0.24]	[0.64]	[-0.56]
Constant	-4.508	1.032	5.759*	-5.229**	1.346	1.601
	[-1.41]	[0.70]	[1.90]	[-2.07]	[0.54]	[0.64]
R^2	0.267	0.137	0.233	0.240	0.221	0.281
N	717	717	717	717	717	717

Notes: This table reports the regression results of $R_i^t = \alpha^t + \beta_1^t R_i^{[-10,239]} + \beta_2^t R_i^{[-10,239]} \times Dur_i + \beta_3^t Dur_i + \beta_4^t R_i^{[-10,239]} \times Pre_i + \beta_5^t Pre_i + \epsilon_i$ based on the sample of announcements made within non-trading hours. The dependent variables are the log returns of CSI 300 for the respective time intervals and are in basis points. Dur_i is the time between announcement time and the first trading time after the announcement and is in days (i.e., distance-to-trading). Pre_i equals 1 if the announcement is pre-scheduled. The sample period is from January 2009 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table IA.VIII: The Impact of Distance-to-trading on the Speed of Price Discovery: Unscheduled and Pre-Scheduled News

	0		0		close-to-open	
	(1) Unscheduled	(2) Pre-scheduled	(3) Unscheduled	(4) Pre-scheduled	(5) Unscheduled	(6) Pre-scheduled
$R^{[-10,239]} \times Non$	0.257***	0.232***				
	[8.27]	[6.05]				
Non	-1.797	-2.286				
	[-0.57]	[-0.53]				
$R^{[-10,239]} \times Dur$			0.069***	0.074***	0.022*	0.036***
			[4.73]	[4.38]	[1.73]	[3.29]
Dur			0.140	-0.588	1.671	-0.558
			[0.05]	[-0.18]	[0.70]	[-0.19]
$R^{[-10,239]}$	0.002	0.083***	0.092***	0.118***	0.170***	0.191***
	[0.62]	[4.48]	[4.86]	[6.70]	[4.85]	[5.91]
Constant	-0.478	-2.080	-1.028	-2.401	-5.626	-4.263
	[-0.49]	[-1.34]	[-0.54]	[-1.51]	[-1.64]	[-1.19]
Empirical p-value	0.329		0.420		0.284	
R^2	0.234	0.266	0.181	0.244	0.214	0.380
N	844	829	844	829	481	236

Notes: This table reports the regression results of the impact of distance-to-trade on the speed of price discovery. Columns (5) and (6) report the regression results based on the sample of announcements released during non-trading hours. Columns (1), (3), and (5) report the regression results based on the sample of unscheduled announcements. Columns (2), (4), and (6) report the regression results based on the sample of pre-scheduled announcements. The dependent variables are the log returns of CSI 300 components stocks for the respective time intervals. Non_i equals 1 if the announcement is released during non-trading hours. Dur_i is the time between announcement time and the first trading time after the announcement and is in days (i.e., distance-to-trading). Empirical p-values of Fisher's permutation test for the differences in coefficients of $R^{[-10,239]} \times Non$ or $R^{[-10,239]} \times Dur$ between unscheduled and pre-scheduled announcements are calculated by a 1000 bootstrapping procedure. The sample period is from January 2009 to December 2020. Standard errors are clustered at the announcement and stock level. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table IA.IX: The Impact of Distance-to-trading on the Speed of Price Discovery: Excluding the Impact of China Concepts Stock

	Post-announcement time windows					
	(1) close-to-open	(2) open-to-0	(3) $[1, 59]$	(4) $[60, 119]$	(5) $[120, 179]$	(6) $[180, 239]$
$R^{[-10,239]}$	0.176***	0.056***	0.228***	0.219***	0.172***	0.149***
	[3.78]	[4.01]	[6.47]	[7.38]	[7.31]	[4.41]
$R^{[-10,239]} \times Dur$	0.033**	0.003	-0.001	-0.010	-0.017***	-0.008
	[2.56]	[0.49]	[-0.13]	[-1.30]	[-3.06]	[-0.72]
Dur	-1.161	0.135	-3.991	4.175**	1.548	-0.704
	[-0.31]	[0.11]	[-1.64]	[2.12]	[0.97]	[-0.35]
Constant	-0.632	4.484	10.759*	-6.981	-8.219*	0.590
	[-0.09]	[1.47]	[1.77]	[-1.25]	[-1.85]	[0.12]
R^2	0.353	0.153	0.352	0.375	0.317	0.236
N	168	168	168	168	168	168

Notes: This table reports the regression results of $R_i^t = \alpha^t + \beta_1^t R_i^{[-10,239]} + \beta_2^t R_i^{[-10,239]} \times Dur_i + \beta_3^t Dur_i + \epsilon_i$ based on the sample of announcements made within Saturday 4:00 a.m. to Monday 9:30 a.m (after the U.S. stock market close on Friday). The dependent variables are the log returns of CSI 300 for the respective time intervals and are in basis points. Dur_i is the time between announcement time and the first trading time after the announcement and is in days (i.e., distance-to-trading). The sample period is from January 2009 to December 2020. The t-statistics are reported in square brackets, and *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Table IA.X: Keywords for Matching Macroeconomic Indexes

Announcement	Keywords
CPI/PPI	CPI, PPI, Consumer Price Index, Producer Price Index
GDP	GDP, Gross Domestic Product
PMI	Manufacturing Purchasing Managers' Index, Bureau of Statistics Manufacturing Purchasing Managers' Index, Purchasing Managers' Index
Caixin	Caixin PMI, Caixin China PMI, Caixin China Manufacturing PMI, Caixin China Manufacturing Purchasing Managers' Index
IP	Industrial Value Added, Industrial Production, Industrial Value Added of Enterprises Above Designated Size
M2	M2, Broad Money, Aggregate Financing to the Real Economy, Money Supply, Monetary Aggregate
Trade	Export, Import, Balance of Trade, Trade Surplus, Total Value of Imports and Exports
FER	Foreign Exchange Reserves, Foreign Exchange Reserve Balance, Foreign Exchange Reserve Size
PI	Profit of Industrial Enterprises, Profit of Industrial Enterprises Above Designated Scale
RRR	Required Reserve Ratio, Required Reserve Ratio for RMB, Required Reserve Ratio for Deposit Financial Institution, Required Reserve Ratio for Financial Institution
FDI	FDI, Foreign Direct Investment
BOP	Balance of Payments, Balance of Payments Trade in Services, Balance of Payments Trade in Goods and Services
Swift	RMB Payment, RMB Settlement, Active Currency
SPRB	Sales Prices of Residential Buildings, Sales Price of Second-hand Residential Buildings, Sales Price of New Constructed Commodity Residential Buildings
FESS	Foreign Exchange Sales, Foreign Exchange Settlement, Foreign Exchange Sales by Banks, Foreign Exchange Settlement by Banks, Foreign Exchange Settlement and Sales by Banks
OMO	Open Market Operations
SLF/MLF/PSL	SLF, MLF, PSL, Standing Lending Facility, Medium-term Lending Facility, Pledged Supplementary Lending
CTCM	Central Treasury Cash Management, Central Treasury Cash Management Operation Announcement
CBS	Central Bank Bills Swap
LPR	LPR, Loan Prime Rate

Notes: This table reports the keywords that we use to match the 20 major macroeconomic indexes. The sample period is from January 2018 to December 2020.