

Federal Reserve Bank of New York  
Staff Reports

**Price and Size Discovery in  
Financial Markets:  
Evidence from the U.S. Treasury  
Securities Market**

Michael Fleming  
Giang Nguyen

Staff Report No. 624  
August 2013  
Revised August 2018



This paper presents preliminary findings and is being distributed to economists and other interested readers solely to stimulate discussion and elicit comments. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

**Price and Size Discovery in Financial Markets:  
Evidence from the U.S. Treasury Securities Market**

Michael Fleming and Giang Nguyen

*Federal Reserve Bank of New York Staff Reports*, no. 624

August 2013; revised August 2018

JEL classification: G01, G12, G14, G18

**Abstract**

We study the workup protocol, an important size discovery mechanism in the U.S. Treasury securities market. We find that shocks in workup order flow explain 6-8 percent of the variation of returns on benchmark notes and, across maturities, contribute 10 percent to the variation of the yield curve level factor. Information related to proprietary client order flow is more likely to show up in workup trades, whereas information derived from public announcements is more likely to come through pre-workup (or “lit”) trades. Our findings highlight how the nature of information affects the trade-off between speed and execution price as informed traders choose between the lit and workup channels.

Key words: dark pool, liquidity, price impact, information share, fixed income market

---

Fleming: Federal Reserve Bank of New York (email: michael.fleming@ny.frb.org). Nguyen: Pennsylvania State University (email: giang.nguyen@psu.edu). This paper was previously circulated under the title “Order Flow Segmentation and the Role of Dark Pool Trading in the Price Discovery of U.S. Treasury Securities.” The authors thank Thierry Foucault (the editor) and two anonymous referees for helpful comments. They also thank Leslie Boni, Andrew Ellul, Larry Harris, Joel Hasbrouck, Charles Jones, Pete Kyle, Ingrid Werner, Haoxiang Zhu, participants at the First Annual SEC Conference on Regulation of Financial Markets, the 10th Annual Central Bank Workshop on the Microstructure of Financial Markets, the 2014 FMA Annual Meeting, the 2015 AFA Annual Meeting, the 2015 Annual Fixed Income Conference, the 2017 WFA Annual Meeting, and seminar participants at Tulane University, Penn State University, the University of Oregon, the Federal Reserve Board, and the Office of the Comptroller of the Currency. In addition, the authors thank Arthur D’Arcy, Dan Cleaves, and Stuart Wexler from ICAP and James Greco from Jefferies for helpful information on the trading of U.S. Treasury securities, as well as Weiling Liu and Collin Jones for excellent research assistance. Giang Nguyen acknowledges the financial support from Smeal Small Research Grants program. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve.

The ability to hide trading intention is valuable to market participants. [Harris \(1997\)](#) writes: “the art of trading lies in knowing when and how to expose trading interest.” Revealing the full extent of one’s trading interest can result in adverse price impact, which is particularly severe for traders who need to trade large quantities. As a result, many trading venues provide features allowing traders to manage order exposure and mitigate price impact, including iceberg orders and size discovery mechanisms. Iceberg orders allow traders to hide part of an order’s size, while size discovery mechanisms allow traders to exchange large quantities of an asset at a fixed transaction price.

In this paper, we analyze the workup protocol, a size discovery mechanism used in the trading of U.S. Treasury securities – a market that is both important and sizeable.<sup>1</sup> We show that the protocol uncovers a large portion of market liquidity, and that its importance has been growing over time. As of 2016, trades in workups account for about 60% of trading volume in on-the-run Treasury notes and about 50% in the 30-year bond. Our central question is whether this size discovery mechanism also plays a role in price discovery. Although the protocol is designed to mitigate instantaneous price impact, it is not necessarily an information-free mechanism. Informed traders naturally prefer to trade as large a quantity as possible to exploit their information advantage. The workup mechanism allows them to do so in a way that minimizes information free-riding and achieves the best price for their trades. Even in the absence of private information related to macroeconomic fundamentals, workups could reveal important information about market demand and supply, which can affect traders’ private valuations.<sup>2</sup>

We quantify the price discovery role of the workup protocol using [Hasbrouck \(1991\)](#)’s information share measure, computed from a vector auto-regression (VAR) of return and order flow

---

<sup>1</sup>U.S. Treasury securities serve vital functions in the financial system, including providing pricing benchmarks for other assets, serving as collateral and hedging instruments in numerous financial transactions, and being a key instrument of monetary policy and store of value, especially during times of market turmoil. The market is sizeable, with \$12.8 trillion in marketable debt outstanding as of December 2015, compared to the \$19.8 trillion size of the U.S. equity market. The trading volume of the six benchmark securities on the BrokerTec electronic trading platform alone averages about \$100-140 billion per day, compared to roughly \$200 billion per day across all U.S. equity markets.

<sup>2</sup>[Burdett and O’Hara \(1987\)](#) argue that information about the existence of a large trading interest can be valuable. [Pancs \(2013\)](#) provides a theoretical model of workups in which information asymmetry arises as one party tries to learn the true block size of its counterparty to decide whether to front run (outside of workups). In [Duffie and Zhu \(2017\)](#), workup participants are subject to a winner’s curse as they learn about the true extent of the other side’s trading intention.

segmented by the workup mechanism into pre-workup trades and workup trades. The model accounts for feedback effects between trades and prices, thereby allowing us to isolate the impact of exogenous order flow shocks on returns. We estimate the model using BrokerTec data from 2006 to 2016 – high quality, tick-by-tick data that allow us to cleanly estimate measures of price discovery and identify activities that occur in the normal trading process versus those that occur during the size discovery process. We show that workup order flow is informative, but less informative than the pre-workup (or “lit”) order flow: the former explains between 6 to 8% of the total return variance of the Treasury notes, compared to the 11-15% contribution of the latter. Collectively, all order flow variables explain roughly 19 to 24% of variation in Treasury returns.

We next examine whether the price discovery role of the workup protocol changes on days when we expect private information to be most intense, in order to gain insight into the nature of private information and how traders might exploit it. We focus on two sets of days: 1) days of important macroeconomic, monetary policy, and auction information releases, and 2) days without public announcements but with extreme order imbalances. On the first set of days, private information is more likely derived from public information. We find that workup trades are relatively less informative on such days, suggesting that informed traders are more concentrated in the lit trading channel. In contrast, the second set of days captures private information that is more likely associated with proprietary client order flow information.<sup>3</sup> Interestingly, workup trades are more informative on these days, implying that traders in possession of proprietary client order flow information are more likely to trade in workups. A possible explanation for the differences between the two sets of days is that private information derived from public information is more likely to be short-lived, as everyone has access to public information, whereas confidential client order flow information is only known to the dealer who possesses it. For such a dealer, the workup mechanism enables the

---

<sup>3</sup>BrokerTec is an interdealer market, where non-fundamental trading needs most likely arise from dealers’ market making activities with their clients. Even though principal trading firms constitute a major group of participants on BrokerTec, as shown in the [Joint Staff Report \(2015\)](#), they operate with a business model that does not typically take on inventory positions overnight. Therefore, if the market finishes a trading day with an extreme order imbalance, such an imbalance can be reasonably attributed to dealers.

search for liquidity at the best price and, in the case of insufficient counter trading interest, allows the dealer to move to the next price point with the information likely remaining intact.

Finally, we expand our empirical framework to model the joint dynamics of returns and order flow across the benchmark yield curve, to account for correlated trading across the yield curve and the tight link across prices of various maturities that arises through the no-arbitrage principle. This results in more accurate estimates of price discovery, and allows us to consider the role of the workup protocol in the price discovery of factors that underpin the yield curve. We show that trading activity in the six benchmark securities explains roughly 20% of the variation in the level factor of the yield curve and that workup order flow is almost as important as lit order flow. We conclude that the workup protocol, while primarily helping discover latent liquidity, also conveys information that is relevant to participants' private valuations. Our findings support the idea in [O'Hara \(2015\)](#) that information in today's high-frequency trading world need not be fundamental.

Our paper contributes to the literature that links microstructure issues with asset prices, in particular with respect to the Treasury market. In addition to studies that examine how the bond markets respond to public information around announcements (see e.g., [Fleming and Remolona, 1997](#)), a large literature focuses on price discovery through trading and finds that order flow reveals private information. [Brandt and Kavajecz \(2004\)](#) find that net order flow explains up to 26% of the variation in daily yield changes on non-announcement days. [Green \(2004\)](#) provides evidence of increased private information revealed through trades right after macroeconomic announcements. [Pasquariello and Vega \(2007\)](#) find that unanticipated order flow has a significant impact on daily yield changes on both announcement and non-announcement days. [Jiang and Lo \(2014\)](#) show that the intensity of private information is higher for longer-term bonds and on announcement days, and conclude that private information derives from heterogeneity in processing and interpreting public information. Building on this literature, we show that the workup protocol, a key market design feature facilitating the exchange of large trading interests, plays a non trivial role in price discovery. Treating this market as one where such a workup option is not available and trade flow is homogeneous may give rise to a less-than-accurate characterization of how trading

affects price dynamics. Indeed, we provide evidence that traders with short-lived information (e.g., private valuations based on public information) prefer to trade through the lit channel for immediate execution, while those with longer-lived information (such as confidential client order flow information) prefer to trade in workups.

Our paper is also among the first to contribute empirical evidence to the nascent literature on size discovery, even though it is a ubiquitous mechanism in today's financial markets (see [Duffie and Zhu, 2017](#)). For example, block trading facilities and crossing networks in equity markets match buy and sell orders at a fixed reference price, thereby allowing traders to tap into a valuable block of liquidity without suffering large price impact. [Zhu \(2014\)](#) estimates that equity dark pools account for about 15% of U.S. equity trading volume. [Collin-Dufresne, Junge, and Trolle \(2016\)](#) show that both workup and mid-market matching trading protocols used in the interdealer market for index credit default swaps (CDS) collectively account for over 70% of the trading volume in the two most liquid CDS contracts. Our study of the Treasury size discovery mechanism and its role in price discovery is a timely addition to this literature and provides important evidence for furthering theoretical work in this area, especially given that the data needed to study size discovery mechanisms in other markets are not easily accessible. Moreover, to the extent that workups share some common features of a dark pool, we contribute evidence from the bond market to the equity-heavy empirical literature on dark pools and hidden liquidity (see e.g., [Bessembinder, Panayides, and Venkataraman, 2009](#); [Ready, 2012](#); [Comerton-Forde and Putnins, 2015](#); [Menkveld, Zhu, and Yueshen, 2017](#)).

Our paper also complements earlier work on the workup protocol in the Treasury market. [Dungey, Henry, and McKenzie \(2013\)](#) model trade duration in the Treasury market and provide evidence that workups affect trade duration, which is also a measure of information arrival in the market. [Boni and Leach \(2004\)](#) study the workup protocol used in voice-assisted interdealer trading, based on GovPX data for October 1997, focusing on explaining why the protocol is used. Neither of these papers specifically examines the role of workups in price discovery. Our paper is the first to study the informational importance of the components of order flow segmented by the protocol,

both individually for each security and jointly for the yield curve. Our results shed further light on the nature of private information and how informed traders optimize their exposure strategy to exploit their advantage, given the protocol’s availability.

The paper proceeds as follows. We describe the workup protocol in Section 1. We then discuss the data and present key descriptive statistics of workups in Section 2. In Section 3, we present a microstructure model for the dynamics of segmented order flow and price, and analyze the information contribution of workups. In Section 4, we present our model of multiple-security price discovery and analyze the informational role of order flow in aggregating private information about the Treasury yield curve factors. We summarize key findings and conclude in Section 5.

# 1 The workup process

## 1.1 Institutional background

This paper focuses on the interdealer segment of the secondary market for U.S. government securities.<sup>4</sup> Trading in this segment, especially in the on-the-run securities, occurs mostly on two electronic trading platforms, BrokerTec and eSpeed. A comparison of daily trading volume on BrokerTec and eSpeed for the period from June 2015 to May 2016 reveals that BrokerTec has roughly 75% market share in the notes and just over 85% in the bond.<sup>5</sup> Coupled with its dominant market share, BrokerTec also accounts for more price discovery.<sup>6</sup>

Both BrokerTec and eSpeed operate as electronic limit order markets in which trading is completely anonymous. Our empirical analysis is based on BrokerTec data and our discussion of how the market works is for the BrokerTec platform. Without access to eSpeed tick data, we

---

<sup>4</sup>In this interdealer market, participation was originally limited to Treasury securities dealers, many of which are primary dealers with obligations to participate in Treasury securities auctions. The paper uses the term “dealers” interchangeably with “traders” and “market participants,” even though the market has expanded in recent years to include other participants such as hedge funds and principal trading firms.

<sup>5</sup>To gain a sense for how market shares evolve, we also compare BrokerTec trading statistics with those of eSpeed reported in [Dungey, Henry, and McKenzie \(2013\)](#) (for 2006) and [Luo \(2010\)](#) (for January 2006 to May 2008). The comparisons show that BrokerTec has a roughly 60% market share in the 2-, 5- and 10-year notes and 30-year bond for these earlier periods, implying that BrokerTec’s market share has increased significantly in recent years.

<sup>6</sup>[Dunne, Li, and Sun \(2015\)](#) compare price discovery on the two platforms using non-contiguous data for 2002, 2004 and 2005 and find that more price discovery takes place in the more active but less transparent BrokerTec platform (the authors argue that BrokerTec is less transparent because it allows hidden orders, whereas eSpeed does not).

are not able to analyze cross-platform activities nor study determinants of the platform choice. In the following discussion, we abstract from the platform choice and focus on traders' decisions conditioning on their choice to participate on BrokerTec.

On BrokerTec, as in any other electronic limit order market, the limit order book – the collection of limit orders at various price levels submitted by market participants – provides liquidity to the market. Traders demanding liquidity can use marketable orders.<sup>7</sup> When such an order arrives, it is matched with one or more limit orders standing on the opposite side of the book at the requested or better price(s), following the price and time priority rule.<sup>8</sup> The execution priority is slightly more nuanced when the execution involves a partially hidden limit order (i.e., “iceberg order”). In that case, the displayed portion of the iceberg order is executed, followed by other displayed orders behind it in the same price queue, before the hidden part of the iceberg order becomes visible and available for execution. Once a trade is complete, and assuming the iceberg order has not been executed in full, the replenished amount from the hidden part of the iceberg order keeps its original position in the queue, as opposed to moving to the end of the queue.

BrokerTec charges a fee for order execution, and this fee is not order-type specific (but is firm-specific).<sup>9</sup> That is, a participating firm is charged the same fee whether it supplies or consumes liquidity, and whether its order is executed in the pre-workup or workup stage. Therefore, for each firm, the fee is not a consideration when it comes to the choice of order type and exposure. However, firms might be subject to different fee structures, perhaps reflecting different levels of activity.

---

<sup>7</sup>A marketable order is an order to buy a given quantity at or above the current best ask price, or an order to sell at or below the current best bid price. In this paper, we use the terms “marketable orders” and “aggressive orders” interchangeably.

<sup>8</sup>For example, consider an aggressive order to buy \$100 million at a price of 25580/256 when there is \$30 million available at the best ask price of 25578/256 and another \$100 million at 25580/256. The first \$30 million of the order will be matched with limit sell orders at 25578/256. Assuming there is no hidden depth at that price, the remaining \$70 million will be executed at 25580/256. If a limit sell order is an iceberg order, then upon the execution of the displayed portion, the next portion becomes visible and available for execution. Continuing with the above example, assume that there is \$15 million hidden depth at the best ask price of 25578/256. The order will be executed as follows: \$45 million at 25578/256 (\$30 million displayed + \$15 million hidden), and \$55 million at 25580/256. For simplicity, our example follows the convention in BrokerTec's database of reporting prices in 256ths of a point, where a point equals one percent of par.

<sup>9</sup>We do not have information on the exact fees, which are proprietary.



## 1.2 The workup process

The workup is a distinctive trading feature in the Treasury market. In a typical limit order market, transactions occur and conclude almost instantaneously when a marketable order arrives and matches up with limit orders on the other side of the market.<sup>10</sup> In the Treasury market, however, there is a time dimension to each transaction. The matching of the initial marketable order is just the beginning of a transaction. Once all possible matches have been made (against displayed and hidden depth in the book, if any), a short time window opens, allowing market participants to “work up” volume (i.e., transact additional volume) at the last price as long as counter trading interest exists. Conceptually, the whole process is characterized as “a single deal extended in time.”<sup>11</sup>

Historically, the workup process consisted of two distinct phases: 1) the private phase, which gave an exclusive right of first refusal to the original parties to the transaction, and 2) the public phase, which is open to all other market participants. However, in 2006, the private phase was replaced by a public phase, making the workup a double-public process. As a result, a transaction progresses straight to the first public workup phase after all possible matches with the limit order book have been completed. During this phase, additional trading interest can come from either side of the market, and these extra trades are conducted on a first-come, first-served basis.<sup>12</sup>

The first public workup phase is open for a pre-specified duration (it was 4 seconds from early 2006 until July 2011 when the duration was shortened to 3 seconds). If there is no trading interest, the workup process automatically expires at the end of this duration. However, if and when a new execution occurs during this time window, the second public phase commences and a new duration opens up. It is then re-settable each time a new execution occurs. This protocol allows the workup to

---

<sup>10</sup>To be clear, we refer to each individual match as a “trade”, and the aggregate of all trades resulting from the execution as a “transaction”.

<sup>11</sup>The workup protocol in the Treasury market predates electronic trading. Prior to 2000, interdealer trading was by voice. Interdealer brokers facilitate trades, mediate quantity negotiations beyond quoted depth, and search for additional counter parties if needed. The quantity negotiation on BrokerTec is governed by a set of precise rules stipulating the window of opportunity for workup trades, replacing the role of human brokers in going back and forth between counter parties working up the size of a trade.

<sup>12</sup>For more information, see “System and Method for Providing Workup Trading,” patent number US 7,831,504 B1, dated November 9, 2010, and “System and Method for Providing Workup Trading without Exclusive Trading Privileges,” patent number US 8,005,745 B1, dated August 23, 2011.

last as long as there is trading interest at the same price point, or to terminate after a predetermined time period if no such interest exists so that the market can move forward.

All trades during a workup – triggered by the initial execution of a marketable order – occur at the same price as that of the original order. With the price being fixed, the workup protocol does not have an instantaneous effect on price, and is merely a liquidity discovery facility. In many ways, the workup process can be characterized as a dark pool trading mechanism, and more precisely, a crossing network.<sup>13</sup> As in a crossing network, workup trades are matched on a first-come, first-served basis at a reference price derived from the initial order execution. The key departure from a typical crossing network in equity markets is that this Treasury crossing network operates as an integral part of the trading process on the BrokerTec platform.

The workup protocol provides traders with an important tool to manage the exposure of their trading intentions. As discussed in [Harris \(1997\)](#), the exposure of a large trading interest conveys information useful to other traders who might use it against the exposing trader (e.g., by front-running, or refusing to trade on the other side). With the workup protocol, traders can submit orders of smaller size than their desired quantity, and then expand the quantity during the workup if appropriate. They can even avoid exposing their intentions altogether by waiting until someone else exposes a trading opportunity that interests them, at which point they can participate in the ensuing workup. Thus, the workup protocol offers a higher degree of control over whether and when to trade the desired additional quantity, whereas iceberg orders are subject to the risk of being adversely executed before the order submitter has a chance to modify or cancel. However, a major cost for the traders hoping to expand volume in a workup is that the incremental quantity they want to transact may not materialize if counter trading interest is lacking. Thus, for those who need immediate execution, non-execution risk can be an important deterrent to using workups.

---

<sup>13</sup>[Buti, Rindi, and Werner \(2011\)](#) characterize dark pools as having “limited or no pre-trade transparency, anonymity and derivative pricing.” The workup process has precisely these characteristics. First, the workup process enables execution of additional trading interest not observed by market participants before each transaction. Second, all trades through interdealer brokers in the Treasury market are anonymous. During a workup, market participants can send in orders, which are then matched by the system. Any unmatched volume is held in the system waiting for subsequent counter trading interest in the workup. Third, the price for these workup executions derives from the execution of the initial marketable order that triggers the workup.

## 2 Data and descriptive statistics

### 2.1 BrokerTec data

Our tick data is from BrokerTec and contains records of all market activity on the platform, from limit order submission, cancellation, and modification, to matching with incoming marketable orders, time-stamped to the millisecond. The database clearly marks when a marketable order arrives, as well as when the workup phase starts and finishes. As a result, we can identify the complete sequence of activities pertaining to each transaction and attribute the activities to the pre-workup and workup phases without ambiguity.

The trade direction of the original marketable order, i.e., whether the aggressive side is a buy or a sell, is also recorded, thus providing unambiguous signing of all pre-workup trades. The signing of trades executed during the workup process warrants some further discussion. BrokerTec considers these trades as an extension (time-wise and volume-wise) of the original execution and therefore signs these trades with the same direction of the original trade. However, to correctly attribute price discovery to the different parts of order flow, we sign trades during workups by the side from which they are initiated. Consider the following example. A buy marketable order of \$100 million at the current best ask price of 25596/256 arrives and immediately exhausts the \$90 million limit orders standing at the best ask. The market then enters the workup stage with the \$10 million left-over buy order resting in the system. Subsequently, a \$15 million sell order arrives, resulting in a workup trade of \$10 million and a resting sell order of \$5 million. Finally, a \$5 million buy order arrives and completely matches with the resting sell order, resulting in a \$5 million trade. The pre-workup trade of \$90 million is classified as buyer-initiated, the \$10 million workup trade is classified as seller-initiated, and the \$5 million workup trade is classified as buyer-initiated.

We treat the whole process from the initial execution to the expiration of the ensuing workup as a single transaction. Each transaction can involve multiple trades or order matches. Those trades that are part of the initial execution are referred to as “pre-workup” or “lit” trades, and represent the lit part of the order flow. Those that occur during the workup phase are referred to as “workup”

or “dark” trades, and constitute the dark part of the order flow. With the workup trade signing procedure above, we are able to split the workup order flow further into two parts. The first, which we refer to as “aggressive workup trades”, consist of trades initiated from the same side as the initial aggressive side. The second, referred to as “passive workup trades”, are trades initiated from the other side (which essentially expand the limit order book). Our order flow variables are constructed at the transaction frequency and reflect the total par value of trades corresponding to the pre-workup, aggressive workup, and passive workup categories described above. Our price discovery estimation and analysis is at the transaction frequency, and not at the trade frequency, because price updates in response to trades occur at the former. For ease of reference, we refer to transactions with at least one trade during the workup phase as “transactions with workup”, and those that do not

Our sample covers 11 years from 2006 to 2016 and all six on-the-run nominal coupon Treasury securities (2-, 3-, 5-, 7-, and 10-year notes and 30-year bond).<sup>14</sup> The minimum tick size for the 2-, 3-, and 5-year notes is  $2/256$  (0.78125 cents per \$100 par value), and that for the 7-, 10-, and 30-year securities is  $4/256$ . We do not have access to comparable tick data for any other Treasury securities, nor for activities in related markets (e.g., the eSpeed platform or the futures market).

## **2.2 Trading and workup activities**

We report descriptive statistics of trading and workup activities in Table 1. Panel A shows average daily trading volume and number of transactions. The 5- and 10-year notes are the most actively traded, averaging well over \$30 billion daily, followed by the 2-year note with about \$26 billion traded per day. This far exceeds the average of roughly \$7 billion in daily trading volume for the 30-year bond. The average number of transactions per day also varies across securities, from roughly 1,000 for the 2- and 3-year notes to over 2,700 for the 5- and 10-year notes. In total, the trading volume of these six securities on the BrokerTec platform averages over \$130 billion per day. Importantly, over half of trading volume – and over 60% in the case of the 5- and 10-year notes – arises in workups. Likewise, the average probability of a transaction with at least one trade during

---

<sup>14</sup>The on-the-run 3- and 7-year notes do not have full data coverage due to discontinuity in issuance. Issuance of the 3-year note was suspended between May 2007 and November 2008, and issuance of the 7-year note was suspended between April 1993 and February 2009.

the workup phase (henceforth “transaction with workup”) ranges from 51-60% for the notes, and is 45% for the bond. These statistics demonstrate that the workup mechanism discovers a substantial portion of market liquidity.<sup>15</sup>

To complement the sample average statistics on workups, Figure 1 shows the time series of workup probability (upper chart) and workup volume share (lower chart). They both reveal upward trends, indicating that the workup mechanism continues to grow in importance in this market. By the end of 2016, the workup probability ranges between 55 and 65% in each of the notes and accounts for 60-70% of trading volume. Workup intensity and volume share are markedly lower in the 30-year bond, but also trend upward over the sample period, nearing or reaching 50% for workup probability and share of volume by the end of 2016.

Conditional statistics at the transaction level for transactions with workup are presented in Panel B of Table 1. Notably, the 2-year note has the largest transaction size, at about \$45 million. This is more than double the size of an average transaction in the 5- and 10-year notes and about nine times the size of a typical transaction in the 30-year bond. The average number of trades per transaction is around 10 for the notes and about 4 for the bond. Conditional on there being a workup, roughly two thirds of trades and volume occur in the workup phase. Moreover, within the workup phase, a greater share is initiated from the same side as the initial aggressor. To see this more clearly, we plot in Figure 2 a breakdown of transactions in which workup trades expand: 1) the aggressive side only, 2) the passive side only, and 3) both sides. Aggressive-only workups account for the largest share of workups (44% and above), whereas passive-only workups are much less prevalent (occurring in about 15-18% of workups, except for the 29% observed for the bond). The remaining 22-38% of workups expand both sides.

Panel C of Table 1 reports statistics for transactions with no trades during the workup phase (henceforth “transactions without workup”). These transactions tend to be much smaller in size, in

---

<sup>15</sup>We compare workups on BrokerTec with those on eSpeed as reported in [Dungey, Henry, and McKenzie \(2013\)](#) for the period from January 2006 to October 2006. The likelihood of workups is a few percentage points higher on BrokerTec than on eSpeed. However, BrokerTec workups typically discover a slightly smaller proportion of transaction volume. Accordingly, the overall share of workup volume does not seem to differ greatly between the two platforms, at least for the period of the comparison.

terms of both dollar volume and trade count. Moreover, transactions without workup tend to be somewhat smaller than even just the pre-workup portions of transactions with workup. For example, the 2-year note's average size of transactions without workup is about \$12 million, compared with a \$16 million pre-workup size and \$45 million total size for transactions with workup. Small trades can be absorbed more easily by outstanding limit orders, and hence are less likely to be executed with a workup, consistent with the intuition in [Harris \(1997\)](#) that small traders are usually not concerned with exposure issues.

Lastly, workups are more prevalent during New York trading hours. We plot in Figure 3 the probability of workup over the course of a typical trading day, from 18:30 of the previous calendar day to 17:30 of the current day (Eastern Time – ET).<sup>16</sup> There is a clear jump in workups as trading passes from Tokyo to London, and another bigger jump as trading starts in New York. The probability of workup in the 10-year note, for example, is about 25% during Tokyo trading hours, increases to about 40% after the London market open, and stands at about 60% between the core New York trading hours of 8:00 and 15:00.

### **2.3 Determinants of workups**

Given that the workup is a key trading feature in this market, we now ask how its importance varies with market conditions. This question is motivated by the need for more empirical evidence given the nascent literature on size discovery mechanisms. Existing research focuses on equity dark pools, but such work has limited generalizability to the Treasury workup protocol.<sup>17</sup>

---

<sup>16</sup>As described in [Fleming \(1997\)](#), the global trading day in U.S. Treasury securities starts at 8:30 local time in Tokyo, which is 18:30 EST (or 19:30 EDT) the previous calendar day (Japan has not adopted daylight saving time). Trading then passes on to London at 8:00 local time, i.e., 3:00 ET. New York trading then starts at 7:30 and continues until 17:30. Statistics for the hour from 18:30-19:30 of the previous day are based on the periods over which the U.S. is on standard time.

<sup>17</sup>Previous theoretical studies of dark pools and crossing networks include [Hendershott and Mendelson \(2000\)](#), [Degryse, Van Achter, and Wuyts \(2009\)](#), [Ye \(2012\)](#), [Zhu \(2014\)](#), and [Buti, Rindi, and Werner \(2017\)](#). The first four models consider a dark pool or crossing network that exists alongside a dealer market, a setting that differs significantly from the Treasury workup protocol considered here. The workup mechanism is not concurrent with the “lit” exchange but instead an integral and sequential part of the trading process. Moreover, the lit exchange is a limit order book and not a dealer market, so the ability to submit limit orders (which is not captured in these models) can result in different equilibrium outcomes with respect to the decision to trade in workups. The last paper models the interaction between a dark pool and a limit order book, and is therefore closest to our setting. Even though the model is for a concurrent dark pool, it provides important intuition to understanding how the decision to trade in the workup is determined by the state of the limit order book.

The most relevant theoretical work with which to interpret the workup determinants studied here is that by [Duffie and Zhu \(2017\)](#), who model the workup as a mechanism for traders to reduce inventory imbalances before commencing a regular auction market. How likely and how much they trade in the workup balances two countervailing considerations: the cost to liquidate leftover inventory in the post-workup market and the winner's curse from transacting too much in the workup. Because the model does not feature a pre-workup limit order book, we can only infer how the prevailing state of the limit order market enters traders' expectation with respect to the liquidation cost and winner's curse. Variables that raise the expected post-workup liquidation cost, such as volatility and trading rate, are expected to increase workup usage and volume.<sup>18</sup> Variables such as (higher) depth, revelation of hidden depth, and (narrower) bid-ask spread, increase workup activities by reducing the winner's curse.

Earlier empirical studies of the Treasury market, e.g., [Fleming, Mizrach, and Nguyen \(2018\)](#) and [Engle, Fleming, Ghysels, and Nguyen \(2018\)](#), have highlighted two important features of this market: 1) the spread is almost always one tick, but widens around public announcements, and 2) depth disappears prior to, but quickly returns after, public announcements, which represent a key source of fundamental information in the market. As such, periods with high depth and narrow spreads are generally considered as having a low asset value information arrival rate, which in [Duffie and Zhu \(2017\)](#)'s framework is predicted to increase workup activities. We also draw some parallel between the workup protocol and the auction conducted by NYSE floor brokers at the specialist posts empirically studied in [Sofianos and Werner \(2000\)](#). The floor brokers are part of the exchange's trading process, and while the auction is in progress, the market stops. [Sofianos and Werner \(2000\)](#) find that floor brokers participate more when volatility is high and in liquid stocks, seemingly consistent with the prediction of a positive relationship of workup activities with liquidity and volatility.

---

<sup>18</sup>Volatility increases liquidation cost by raising the price impact cost of liquidating leftover undesirable inventory. A higher transaction rate implies potentially greater availability of trading opportunities in the future, which reduces traders' incentives to trade aggressively in the post-workup auction market, thereby raising the liquidation cost and making it more beneficial to unload more of the undesirable inventory during the workup. For more detailed discussion, see [Duffie and Zhu \(2017\)](#).

We incorporate the above predictions in a logistic regression, similar to [Boni and Leach \(2004\)](#), linking the workup probability with 1) order book depth, 2) spread, 3) pre-workup volume, 4) revelation of hidden depth, 5) volatility, 6) inter-transaction duration, 7) indicators for time periods around public announcements, and 8) other variables to account for trading hour effects and potential positive auto-correlation of workup activities (as intuited in [Buti, Rindi, and Werner, 2017](#)).<sup>19</sup> To supplement the workup probability model, we also estimate a generalized linear mixed model with a logit link function for the fraction of volume discovered in workups – this represents a composite variable that encapsulates both the probability and the extent of workup volume. The results for the workup choice model appear in Table 2, and for the workup volume share model in Table 3.

We first highlight coefficients that convey a consistent story across all securities and that are in line with the predictions previously discussed. Workups are more likely and uncover a greater volume share the higher the liquidity in the order book as captured by higher same-side book depth, narrower spread, and the revelation of hidden order in pre-workup trading. These results are also consistent with [Buti, Rindi, and Werner \(2017\)](#)'s predictions; even though the concurrent dark pool in their model is not the same as the workup, the intuition that the increased competition in liquidity provision might encourage traders to join workups instead of posting limit orders to already long queues in a deep order book is fully applicable here.

Workups are also more likely and achieve a greater volume share the lower the inter-transaction duration (equivalently, the higher the transaction rate) and the higher the volatility. The positive relationship between market volatility and workups is consistent with earlier findings in setups similar to the current Treasury workup protocol (e.g., [Boni and Leach, 2004](#); [Sofianos and Werner, 2000](#)), and in line with [Duffie and Zhu \(2017\)](#)'s theory on size discovery as previously discussed. This finding, however, contrasts with the negative relationship between volatility and dark pool activity predicted by dark pool theory models and widely documented in empirical studies of equity dark pools. As discussed in footnote 17, many important considerations in the decision to trade in workups are not captured by these dark pool theories. Perhaps the non-concurrent nature of the

---

<sup>19</sup>See Appendix A for a description of how these variables are constructed.



workup and the larger choice set afforded by the limit order book market affect traders' optimization outcome, resulting in this deviation from the equity dark pool literature.

Other variables have a less clear impact on workup activity, and the small number of securities considered prevents us from drawing concrete conclusions. However, some discussion on the announcement effects is warranted, given that the arrival of value-relevant information is an important determinant of workups in [Duffie and Zhu \(2017\)](#)'s model. That is, a higher information arrival rate reduces the delay in achieving optimal inventory position post-workup, and thus predicts less workup activity as traders opt to liquidate more inventory imbalances afterward. The public announcements represent the arrival of fundamental information for Treasury securities and are expected to lower workup activity. The evidence is weakly supportive of this prediction, but mainly in the five-minute period after announcements.<sup>20</sup> We only find strong and consistent results in the 30-year bond that workup probability and volume share are lower before, during, and after announcements.<sup>21</sup>

### 3 Informational value of workup trades

#### 3.1 A microstructure model of price and trade

We proceed to specify a microstructure model for the dynamics of price and order flow, built upon the general framework described in [Hasbrouck \(2007\)](#). The notable feature of our model is

---

<sup>20</sup>[Menkveld, Zhu, and Yueshen \(2017\)](#) also find that the market share of equity dark trading declines after macroeconomic and earnings announcements.

<sup>21</sup>Our results seem to suggest a clientele effect in the 30-year bond whereby the bond is less predisposed to trading in workups, especially during periods of increased information such as around announcements. There are likely many reasons for why the bond is different from the notes, two of which are discussed here. One is that the supply of the benchmark 30-year bond is much lower than that of other securities, as indicated by the average sample period auction size of \$13 billion as compared to around \$30 billion typically auctioned for the short- and medium-term notes (see Table IA1 in the Internet Appendix). In addition, the ownership base of the 30-year bond is tilted more heavily toward investment funds. Institutional investors tend to acquire the bond at auction much more so than they do other securities, which reduces their need to trade in the secondary market. Both of these factors contribute to a lower level of liquidity supply and demand in the market for the bond, which in turn implies a lower execution probability and participation in workups. Perhaps the lower usage and volume of workups around announcements does not arise as a matter of choice but is the consequence of lacking countertrading interest. The unconditional statistics of inside market depth available for trading and transaction volume in Table 1 support this observation.

that it accounts explicitly for the segmentation of order flow due to the workup feature. From this model, we derive a structural VAR representation in (irregular) transaction time for estimation.

Let  $t$  index the  $t^{th}$  marketable order. We distinguish events occurring during the pre-workup and workup phases of the  $t^{th}$  transaction by the subscripts  $t^-$  and  $t^+$  respectively.  $P_{t^-}$  denotes the best bid-ask midpoint (logged) observed as of the  $t^{th}$  transaction, and  $m_{t^-}$  the unobservable efficient price. Let  $LT_{t^-}$  be the signed volume of pre-workup (or “lit”) trading, and  $DT_{t^+}$  the signed volume of workup (or “dark”) trading (positive if buyer-initiated and negative if seller-initiated). The basic building blocks of the model are:

$$m_{t^-} = m_{t^- - 1} + w_{t^-} \quad (1)$$

$$P_{t^-} = m_{t^-} + cLT_{t^-} \quad (2)$$

$$LT_{t^-} = v_{1,t^-} + \beta_1 v_{1,t^- - 1} \quad (3)$$

$$w_{t^-} = u_t^- + \lambda_1 v_{1,t^-} + \lambda_2 v_{2,t^+ - 1} \quad (4)$$

$$DT_{t^+} = v_{2,t^+} + \beta_2 v_{2,t^+ - 1} + \alpha_1 v_{1,t^-} + \alpha_2 u_{t^-}, \quad (5)$$

where the efficient price  $m_{t^-}$  is specified to follow a random walk as in equation (1).  $w_{t^-}$  is the efficient price increment and the subscript  $t^-$  indicates that the price updating takes place with the execution of the  $t^{th}$  marketable order, but before the workup begins. The observed price  $P_{t^-}$ , as expressed in equation (2), consists of the permanent component  $m_{t^-}$  as well as a component reflecting trading frictions ( $cLT_{t^-}$ ). Since workup trades are conducted at the price determined in the pre-workup trading round,  $P_{t^-}$  depends contemporaneously on the lit trade flow  $LT_{t^-}$  but not on the dark trade flow  $DT_{t^+}$ .

To allow for the positive auto-correlation of transaction sign as predicted by theory (e.g., [Parlour, 1998](#)) and observed in the data, a moving average (MA) model, MA(1), is specified for the lit trade flow as in equation (3), where  $v_{1,t^-}$  is a white noise process and captures the pre-workup trade innovation. Likewise, equation (5) for the dark trade flow  $DT_{t^+}$  has an MA(1) error structure with the error term  $v_{2,t^+}$ . However, it also includes the contemporaneous effect of the innovation in lit

trade flow  $v_{1,t-}$  that precedes and initiates the workup process, as well as public information that arrives at the time of the trade  $u_{t-}$ .

Equation (4) models the efficient price increment  $w_{t-}$  as consisting of both a public information component  $u_{t-}$  that is unrelated to trade and a trade-related private information component.<sup>22</sup> The latter component consists of non-public information inferred from the lit trade flow, as well as the lagged dark trade flow. While workup trades have no immediate price effect as they are executed at an established price, traders can observe the workup trade flow after each transaction and update their belief about security value in subsequent transactions. Therefore, innovations in the dark component of order flow enter the efficient price increment equation with a lag. Finally, the model's innovation terms, namely  $u_{t-}$ ,  $v_{1,t-}$  and  $v_{2,t+}$ , are assumed to be uncorrelated.

From this point, we simplify the notation by suppressing the plus and minus superscripts of  $t$ . With this setup, we can derive a vector moving average of order 2 (VMA(2)) for  $Y_t \equiv \begin{bmatrix} LT_t & \Delta P_t & DT_t \end{bmatrix}'$  with an error vector  $\epsilon_t$ , where  $\epsilon_t$  relates to the model's exogenous variables through the following expression:  $\epsilon_t = B \begin{bmatrix} v_{1,t} & u_t & v_{2,t} \end{bmatrix}'$ , with:

$$B = \begin{bmatrix} 1 & 0 & 0 \\ \lambda_1 + c & 1 & 0 \\ \alpha_1 & \alpha_2 & 1 \end{bmatrix}. \quad (6)$$

Assuming the invertibility condition, a VAR representation (of infinite order) exists for  $Y_t$  with the error vector  $\epsilon_t$  and a covariance matrix  $\Omega \equiv \text{Var}(\epsilon_t)$ . The matrix  $B$  accordingly captures the contemporaneous dynamic structure of the model. It is a lower triangular matrix, reflecting our key assumption with respect to the causal ordering in the model. Specifically, the ordering goes from pre-workup trades to price updating to workup trades (which also corresponds to the way we deliberately stack up the vector  $Y_t$ ). Price revision following the pre-workup trade variable reflects the commonly adopted assumption in the literature that traders monitor order flow to update

---

<sup>22</sup>It should be kept in mind that trade-related private information here refers to information as revealed from BrokerTec order flow, and public information refers to any non-BrokerTec-trade-related information.

their beliefs about the asset's value. That the pre-workup trade variable and price revision precede the workup trade variable in the ordering is natural given the way the workup process works: a marketable order (i.e., pre-workup, or originating, trade) must arrive and execute against standing limit orders before the workup process opens at the established price point. Formulated this way, the model implies that price updates incorporate two sources of information: 1) public information unrelated to trades ( $u_t$ ), and 2) private information inferred from contemporaneous trade flow innovation ( $v_{1,t}$ ) and lagged workup trade flow innovation ( $v_{2,t-1}$ ).

One important consideration is that workup trades on BrokerTec can be initiated from either side, even though they are all recorded in the BrokerTec database as having the same aggressive side as the initial workup-triggering marketable order. To allow for the possibility that they can have different information impact, we split the workup order flow variable  $DT$  into two parts: 1) trades initiated by the same side  $DTS$ , and 2) trades initiated by the opposite side  $DTO$ . Equations (4)-(5) can be easily expanded to accommodate this split and give a VAR representation. Thus, the vector of endogenous variables in the VAR is now  $Y_t \equiv \left[ LT_t \quad \Delta P_t \quad DTS_t \quad DTO_t \right]'$ . The ordering of the variables remains as previously discussed, except for the ordering between  $DTS_t$  and  $DTO_t$  which is ambiguous. Since price discovery constructs (to be discussed next) are sensitive to variable ordering, we compute them for both orderings (i.e.,  $Y_t \equiv \left[ LT_t \quad \Delta P_t \quad DTS_t \quad DTO_t \right]'$  and  $Y_t \equiv \left[ LT_t \quad \Delta P_t \quad DTO_t \quad DTS_t \right]'$ ) and then take the average.

### 3.2 Empirical estimation

For empirical implementation, we estimate a VAR(5) model separately for each trading day:

$$\begin{aligned}
LT_t &= \alpha_1 + \sum_{j=1}^5 \beta_{1,1,j} LT_{t-j} + \sum_{j=1}^5 \beta_{1,2,j} \Delta P_{t-j} + \sum_{j=1}^5 \beta_{1,3,j} DTS_{t-j} + \sum_{j=1}^5 \beta_{1,4,j} DTO_{t-j} + \epsilon_{1,t} \\
\Delta P_t &= \alpha_2 + \sum_{j=1}^5 \beta_{2,1,j} LT_{t-j} + \sum_{j=1}^5 \beta_{2,2,j} \Delta P_{t-j} + \sum_{j=1}^5 \beta_{2,3,j} DTS_{t-j} + \sum_{j=1}^5 \beta_{2,4,j} DTO_{t-j} + \epsilon_{2,t} \\
DTS_t &= \alpha_3 + \sum_{j=1}^5 \beta_{3,1,j} LT_{t-j} + \sum_{j=1}^5 \beta_{3,2,j} \Delta P_{t-j} + \sum_{j=1}^5 \beta_{3,3,j} DTS_{t-j} + \sum_{j=1}^5 \beta_{3,4,j} DTO_{t-j} + \epsilon_{3,t} \\
DTO_t &= \alpha_4 + \sum_{j=1}^5 \beta_{4,1,j} LT_{t-j} + \sum_{j=1}^5 \beta_{4,2,j} \Delta P_{t-j} + \sum_{j=1}^5 \beta_{4,3,j} DTS_{t-j} + \sum_{j=1}^5 \beta_{4,4,j} DTO_{t-j} + \epsilon_{4,t}.
\end{aligned}$$

Given the assumed ordering discussed earlier, we can fully identify the structural dynamics (i.e., the matrix of contemporaneous effects and variances of orthogonalized shocks) from the covariance matrix of model residuals  $[\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4]'$ . From the model, we compute the permanent price impact and information share for each order flow variable.

The permanent price impact, which is the long-run cumulative response of price to order flow shocks, reflects the permanent change in price after transitory effects have washed out (see [Hasbrouck, 1991](#)). It corresponds to the increment in the efficient price  $w_t$ :

$$\mathbb{E}[\Delta P_t + \Delta P_{t+1} + \dots | \epsilon_t] = \Psi_{\infty, P} \epsilon_t. \quad (7)$$

We approximate  $\Psi_{\infty, P}$  by truncating the cumulative impulse response function (IRF) of price at 50 lags.<sup>23</sup> The price impact is measured with units in hundredths of a percent of par value (basis points), which is equivalent to cents per \$100 par value.

The information share of a variable measures the extent to which its variation contributes to the variance of the efficient price update  $w_t$ . From equation (7), this variance can be approximated by:

$$\tilde{\sigma}_w^2 = \Psi_{h, P} \Omega \Psi'_{h, P}, \quad (8)$$

---

<sup>23</sup>As is standard in the literature, we compute the IRF for price from the estimated VAR model by forecasting the system recursively forward to the chosen horizon, assuming that the system is initially at rest, i.e., all variables are set to 0 except for the shocked variable. Inspection of the path of the estimated IRFs indicates that the cumulative IRF converges well before the 50th transaction following a shock. Thus  $\Psi_{50, P}$  provides a reasonable approximation of the permanent price impact  $\Psi_{\infty, P}$ .

where  $h = 50$  as discussed above. Given the structure of the system, it is easy to show that the right-hand side of equation (8) is a linear combination of orthogonalized shock variances  $\sigma_u^2$ ,  $\sigma_{v_1}^2$ ,  $\sigma_{v_{2,S}}^2$ , and  $\sigma_{v_{2,O}}^2$ .<sup>24</sup> Each of these terms can then be expressed as a percentage of  $\sigma_w^2$ . The percentage attributable to  $\sigma_u^2$  indicates the extent to which public information drives the variation in the efficient price update, whereas those attributable to  $\sigma_{v_1}^2$ ,  $\sigma_{v_{2,S}}^2$ , and  $\sigma_{v_{2,O}}^2$  show the contribution of non-public information revealed through pre-workup trades, same-side workup trades, and other-side workup trades respectively. Our price discovery analysis is mainly based on information shares because this construct accounts for the different degree of variation of the variables in the system and allows for comparison over time and across maturities.

### 3.3 Information content of workup trades

We report estimates of price impact per unit shock in each order flow type in Table 4, with supplementary results in Tables IA2 and IA3 in the Internet Appendix.<sup>25</sup> In general, price impact of same-side workup trades is only slightly lower, or even higher in the case of the 2- and 3-year notes, than that of lit trades. This is especially the case when we account for the differences in the average size of shocks in each order flow type (see Panel B of Table IA2). Another important result is that other-side workup trades have price impact in the opposite direction and of significantly smaller magnitude as compared to same-side workup trades – a result that is confirmed by statistical tests in Table IA3. The smaller magnitude of price impact of other-side workup trades makes sense because such trades are not crossing the bid-ask spread, but are instead executing on the favorable side of the spread. The literature heretofore has assumed that order flow is signed as BrokerTec reports it and that all of these trades have the same positive impact. Our estimation, by taking into account the nuances of the workup protocol, offers a clear improvement on past literature. We also find that the price impact of both types of workup trades in the 30-year bond are statistically insignificant

<sup>24</sup>The subscripts  $S$  and  $O$  in  $\sigma_{v_2}^2$  denote same-side and other-side workup trades.

<sup>25</sup>Table IA2 shows the duration-adjusted price impact estimates per unit shock in each of the order flow variables in Panel A, and duration-adjusted price impact estimates per average-sized shock in each of the order flow variables in Panel B. Table IA3 reports the results of tests comparing price impact across order flow types.

and much smaller in magnitude than that of lit trades, albeit with point estimates larger than those for the notes.<sup>26</sup>

Next, we present information share estimates in Table 5. Lit trade flow contributes between 11% and 18% to return variation, on average. At the upper bound of the 90% range, this part of order flow explains about 17-26% of the total variation in the efficient price innovations. Workup trade flow has lower information share, and a larger part of its informational value comes from workup trades initiated from the same side as the initial aggressor. Specifically, same-side workup trades contribute around 6-8% and other-side workup trades contribute less than 3% to price discovery of the notes, with both types contributing about 7-10% in total. We also observe that the 30-year bond exhibits a markedly lower role of workup trades in price discovery; both forms of workup trades contribute collectively less than 2%, implying that price discovery through trades in this maturity occurs mainly through the lit trade channel.<sup>27</sup> In total, innovations in trades explain between 19 to 24% of efficient return variance, highlighting the importance of trading in aggregating heterogeneous private valuations. The remaining 76-81% reflects price adjustment due to public information.<sup>28</sup>

Figure 4 provides the time series of the information structure of the six securities. It shows that the composition of information revealed through trades (pre-workup, same-side workup, and other-side workup trades) remains quite stable over time. Among the three trade types, lit trades clearly contribute the largest share, especially for the 30-year bond, followed by same-side workup trades. Workup trades initiated by the other side contribute little to price discovery (formally confirmed in tests reported in Table IA3 in the Internet Appendix).

---

<sup>26</sup>The stark difference in the magnitude of the permanent price impact of trades in the 30-year bond from that of the notes can be partially explained by the differing duration and average size of pre-workup, same-side workup, and other-side workup trades. After adjusting for these factors, the magnitude of price impact of lit trades in the bond is not too different from that of the notes, as shown in Table IA2. Moreover, the duration-adjusted price impact of an average-sized workup is much lower for the bond, reflecting the fact that the bond has a markedly lower workup volume on average. Without data on market participants, we are not able to formally test for any clientele effects that might help explain why traders in the bond are less inclined to use workups. Footnote 21 provides a discussion of possible explanations.

<sup>27</sup>As discussed in footnote 21, there seems to be a distinct clientele effect in the 30-year bond that is less inclined to workup trading due to lower liquidity supply and demand. The bond is also most sensitive to interest rate risk, giving rise to a more urgent trading demand that can naturally be more easily satisfied in lit trading.

<sup>28</sup>As noted earlier, what we call public information includes information from announcements as well as everything not related to BrokerTec order flow, including order flow information from eSpeed and the futures market.

### 3.4 Information structure on special days

We now analyze price discovery on special days, when the information environment might be different, to understand whether the role of the workup protocol as an information transmission channel changes. We specifically look at days with important announcements, days when the market is highly volatile, and days when the market experiences extreme order flow imbalances (an indicator of possible supply/demand shocks). From the daily information share estimates for various components of order flow and public information, we construct two aggregate information share measures. The first is a measure of total price discovery, reflecting the private-vs-public (or trade-vs-nontrade) information split. It is the sum of information shares of all trade types (lit, same-side workup, and other-side workup). The second measures the relative informational importance of the workup protocol, and is computed as the sum of information shares of same-side workup trades and other-side workup trades, divided by the total trade-related information share. These are the two dependent variables (with logit transformation, following [Mizrach and Neely, 2008](#)) in our regression analysis below.

#### 3.4.1 Announcement days

We consider three types of announcements: 1) FOMC rate decisions, 2) important macroeconomic releases, and 3) Treasury auction results.<sup>29</sup> These announcements are shown to be important to Treasury price formation (see e.g., [Fleming and Remolona, 1997](#); [Balduzzi, Elton, and Green, 2001](#); [Green, 2004](#); [Pasquariello and Vega, 2007](#)). In addition, [Jiang and Lo \(2014\)](#) show that the intensity of private information is highest on announcement days. Accordingly, these announcements provide an opportunity to study how private information derived from public announcements gets incorporated into prices. To see this, we run the following regression and report the results in Table 6:

$$\ln \frac{IS_t}{1-IS_t} = \alpha + \beta_1 \times FOMC_t + \beta_2 \times Macro_t + \beta_3 \times Auction_t + \beta_4 \times Trend + \epsilon_t, \quad (9)$$

---

<sup>29</sup>See Appendix B for the list of announcements considered.



where *Trend* is a linear time trend and the dependent variable *IS* is: 1) the fraction of trade-related information shares (Panel A), or 2) the fraction of workup information shares standardized by the total trade-related information share (Panel B).

The mostly negative coefficients on the announcement day variables in Panel A indicate that trade flow (including both the transparent pre-workup part as well as the less transparent workup parts) collectively contributes less to price discovery on announcement days. The evidence is strongest for FOMC days (as indicated by the consistently negative and significant coefficients) and weakest for auction days (for which only three of the six coefficients are significant). The findings imply that the public information share is greater on announcement days, consistent with the release of public information on such days, which can move prices without trades, as shown in [Fleming and Remolona \(1999\)](#). [Brandt and Kavajecz \(2004\)](#) and [Pasquariello and Vega \(2007\)](#) also find that order flow is less important on announcement days.

Interestingly, Panel B shows barely any evidence indicating widespread change in the private information structure on announcement days as compared to non-announcement days. The coefficients on FOMC days are all insignificant, as are the majority of those on macro announcement days and auction days. If anything, it seems that workup trades are relatively more informative on macro days for the 2- and 3- notes, and on auction days for the two least liquid securities (7- and 30-year ones). Nevertheless, we can reasonably conclude that the mix of information content of lit and dark trade flow remains quite similar.

Looking at all announcements can mask the fact that announcements can differ in how they create opportunities for producing an information advantage. If market participants have similar beliefs about a given announcement, there is little scope for profiting from forecasting and interpreting the news well. On the other hand, a high degree of disagreement in the market toward an upcoming announcement provides an incentive for creating an information advantage. Therefore, by focusing on those announcements with the highest information content, we can have a cleaner analysis of how market participants take advantage of private information derived from public information. If the information content of trade flow increases on days with diverse opinions, it provides evidence that

traders exploit their information advantage by placing marketable orders for immediate execution. In contrast, an increase in non-trade information share suggests that better informed traders are more likely choose to adjust their limit order positions, resulting in price updates without requiring trades. Finally, any change in the relative information content of workups would indicate which of the two trade channels is more likely chosen by informed participants.

In the spirit of [Pasquariello and Vega \(2007\)](#), we construct an index of belief dispersion using Bloomberg economist survey data for all regularly scheduled macroeconomic announcements during the sample period. For each survey (pertaining to a given announcement at a given time), we use the survey standard deviation as the measure of disagreement among market participants toward the corresponding announcement. We demean and standardize this measure using its sample mean and standard deviation to achieve comparability across different announcements. Next, to account for days with multiple announcements, we compute the average dispersion across all announcements on a given day, weighted by the number of economists covering each announcement. As a result, we obtain a dispersion index for each announcement day. We classify announcement days into those with high dispersion of beliefs if the dispersion index is above the 95th percentile, and low dispersion of beliefs if the index is below the 5th percentile.

We add these two indicators, interacted with the macro announcement dummy, to the regression model in Equation 9, and report the corresponding results in Table 7. Similar to the results in Panel A of Table 6, we continue to see that prices are more likely to update via non-trade channels (resulting in lower information content of trades) on announcement days, especially those with high dispersion of beliefs. On low belief dispersion days, however, the sum of the coefficients on Macro and Macro x Low Belief Dispersion is positive across maturities, but mostly insignificant in unreported tests on the sums. Furthermore, for announcements with high dispersion of beliefs, we find that the less transparent workup order flow carries relatively less information. The coefficients on Macro x High Belief Dispersion in Panel B of Table 7 are negative and significant for all the notes. Meanwhile, there is no significant change in the information value of each component of order flow on announcement days with low dispersion of beliefs. The evidence presented here thus

indicates that, around periods when private information generated from public news is expected to be high, better informed traders are more likely to trade on their information through the transparent trade channel (pre-workup trades) than through the less transparent one (workup trades) for which the risk of failed execution could easily outweigh the benefits of trading in workups.

### 3.4.2 Days with extreme net order flow

The second type of private information that might arise in the Treasury interdealer market is client order flow information possessed by dealers. While unrelated to the fundamentals of Treasury securities themselves, demand/supply shocks have been shown to affect Treasury security values (see e.g., [Lou, Yan, and Zhang, 2013](#)). If dealers have valuable information about impending large orders of their clients, such information would eventually manifest in the order flow in the interdealer market. As noted earlier, the heavy presence of principal trading firms on BrokerTec does not necessarily affect the quality of daily order imbalance as a proxy for client order flow information.

We now test whether the information share composition in the interdealer market changes on days with increased information about demand and supply shocks. We proxy for the level of this information by the extent of one-sidedness of order flow, denoted by  $MagOIV$  and computed as the absolute order flow imbalance on a given day divided by total trading volume on that day. To avoid capturing instances of extreme order imbalances induced by public information, we run the following regression only on non-announcement days:<sup>30</sup>

$$\ln \frac{IS_t}{1 - IS_t} = \alpha + \beta_1 \times MagOIV_t + \beta_2 \times Trend + \epsilon_t, \quad (10)$$

where the dependent variable  $IS$  is as previously discussed. The results of this regression are reported in Table 8. Panel A shows that the information share of trades of the 2-, 3-, 5-, and 10-year notes increases on days with extreme order flow imbalance (positive coefficient for  $MagOIV$ ),

<sup>30</sup>Previous studies, e.g., [Fleming and Remolona \(1997\)](#) and [Green \(2004\)](#), show that public information shocks dissipate quickly in the minutes after announcements. Thus, it is not likely that the effect of an announcement carries over to the next non-announcement day.

although only the coefficient on the 5-year is significant. The reverse is true for the less liquid 7- and 30-year securities. One possible explanation for the differences between the two groups is that the former typically has a much deeper limit order book that could absorb these order flow shocks more easily. The lower liquidity supply of the latter group implies that large order flow shocks might take longer to be absorbed. In the process, the limit order book gets revised more often, which would lead to a greater extent of efficient return variance driven by non-trade updates.

An interesting result emerges from Panel B for the 2-, 5-, and 30-year securities (results for other benchmark securities are insignificant). Workup trades reveal relatively more information on extreme order flow days, when private information related to confidential client order flow is expected to be high. This is a departure from the lower informational role of workups on announcement days with high dispersion of beliefs. We argue that private information derived from public announcements is likely to be short-lived, as everyone has access to public information and market participants in this interdealer market are highly sophisticated. On the other hand, confidential client order flow information is known only to the dealer that possesses it. It is natural to expect that an informed trader in the latter scenario faces a lower risk of information expiry if he fails to find counter trading interest in workups. Accordingly, he is more likely to rely on workups to source liquidity for his trades and achieve better execution prices as a result.

### 3.4.3 Volatile days

Finally, we examine whether volatility changes the information structure in the market. We show in the descriptive section that volatility increases the probability and usage of the workup protocol. Does volatility encourage informed traders to trade more outside of workups due to increased risk of non-execution, as predicted by the model in [Zhu \(2014\)](#) for dark liquidity pools? For this test, we estimate the following regression:

$$\ln \frac{IS_t}{1 - IS_t} = \alpha + \beta_1 \times \text{LogVol}_t + \beta_2 \times \text{Trend} + \beta_3 \times X_t + \epsilon_t, \quad (11)$$

where  $LogVola_t$  is the log realized volatility on day  $t$ , computed as the sum of squared one-second log returns for the day over the trading hours of 7:00 to 17:30 ET. Because market volatility subsumes the presence of information and effects of illiquidity, it is important to control for the liquidity level in order to isolate the effect of the information-driven component of volatility. In the regression above,  $X_t$  includes the inside market depth (logged) and total number of transactions (logged). We report in Table 9 the results for two specifications: one without liquidity controls and another with liquidity controls.

As Panel A shows, the informativeness of trades decreases with volatility, with or without controls for liquidity. An increase in market volatility, controlling for the liquidity level, likely signals the arrival of information. It seems that such information gets impounded into price more quickly through non-trade channels (e.g., revisions of limit orders) than through trades. [Cohen and Shin \(2003\)](#) show that in times of volatility, trades are likely motivated by price movements. Since our empirical framework accounts for such a feedback effect, it makes sense to see a lower information share of truly unanticipated shocks in trades.

More importantly, we show in Panel B that within the trade channel, there is less information transmission through the workup channel relative to the lit trade channel when the market is highly volatile. The coefficient on log volatility remains negative and significant. As we control for liquidity, the effect of volatility on workups' informativeness turns positive for the 7-year note and 30-year bond, suggesting that liquidity is important in the price discovery of these less liquid securities. Furthermore, liquidity has a positive impact on the informativeness of workups. If volatility deters trading in dark trading channels due to increased risk of non-execution (as informed traders tend to pile on one side), the availability of ample liquidity eases this risk and should have an opposite effect on informed traders' choice of trading venue. That is what we find. Our results thus provide strong empirical support for the intuition in [Zhu \(2014\)](#).

## 4 Price discovery across the Treasury yield curve

### 4.1 A model of trade and price dynamics of multiple correlated securities

Thus far, our analysis is done for each security individually in order to focus on the important dynamics of trades and prices. However, it is important to realize that these securities are highly correlated, because they all reflect risk-free interest rates. Therefore, trading activity and price movements in the 2-year note, for example, likely affect the dynamics of other securities as well. The public information shares documented in the previous section for each security might include trade-related information in the other maturities along the yield curve. A more accurate characterization of the information shares of trade and public information thus calls for a model that considers jointly the dynamics of trades and prices of all securities.

It is well understood that Treasury yields are tightly linked through the no-arbitrage principle and that they can be summarized by three factors representing the level, slope, and curvature of the yield curve. Thus, instead of including all six Treasury return series in our empirical model, we extract from them more economically important constructs, namely the level and the slopes at three maturity ranges of the benchmark yield curve.<sup>31</sup> Apart from econometric advantages (reducing the dimensionality of the model and the redundancy from modeling six return series that are highly correlated), we are able to study the role of maturity-specific order flow on the discovery of the level and slope factors of the benchmark Treasury yield curve.

We extract the factors as follows. The return of the level factor, denoted by  $LEV$ , is the average of the six securities' standardized returns (using the respective mean and standard deviation of returns of each security each day).<sup>32</sup> The short-term slope factor  $SLS$  is the difference in standardized returns between the 3- and 2-year notes, the medium-term slope factor  $SLM$  is

---

<sup>31</sup>By extracting the slope factor at three different maturity segments, we implicitly account for the curvature of the yield curve while making the results easier to interpret.

<sup>32</sup>This standardization is important to remove duration effects, because returns on longer-dated securities have significantly higher variation than returns on shorter-dated securities. Without this standardization, the variation of returns on the 30-year bond would dominate the variation of the level factor.

between the 7- and 5-year notes, and the long-term slope factor  $SLL$  is between the 30-year bond and 10-year note. Thus, the VAR model we estimate is for the vector:

$$Y_t \equiv \left[ LT_t^{(2)} \dots LT_t^{(30)} \quad LEV_t \quad SLS_t \quad SLM_t \quad SLL_t \quad DTS_t^{(2)} \dots DTS_t^{(30)} \quad DTO_t^{(2)} \dots DTO_t^{(30)} \right]',$$

where  $t$  indexes one-second intervals throughout a trading day. As before, after estimating the model (separately for each day), we can compute the information share of lit trade and dark trade flows of the six securities with respect to the common price innovations (i.e., variance of innovation in the level factor), as well as innovations to the term premium in the short-, medium-, and long-term maturities. This analysis will enable us to see which security is most important in the price discovery of the yield curve, and compare the information role of lit and dark trade flows. Our analysis in this section is for the sub-sample period of February 27, 2009 to the end of 2016 over which all six securities are present in the data.

Our VAR model accounts for: 1) feedback effects between yield curve factors and trades, 2) possible correlated trading across the yield curve, 3) feedback effects between lit and workup trade types across maturities, and 4) cross effects among yield curve factors. The regressions of yields in [Brandt and Kavajecz \(2004\)](#) take net order flow variables as exogenous, and thus do not account for the important dynamics 1) and 2) above, and 3) is not a feature in their analysis. In addition, our model is estimated using one-second data and can capture with more precision the complex interactions in the trading process. As a result, we can provide a detailed attribution of the information contribution to each order flow type.

There are several important considerations in the estimation of our model. First is the issue of non-synchronicity of trades across securities and the need to line up trade and price observations on the same frequency for joint modeling. A natural approach is to adopt the frequency of transaction in any security. For example, suppose that there is a transaction in the 2-year note at 10:00:00 and 10:01:00, and a transaction in the 5-year note at 10:00:35. The data is aggregated as a transaction at 10:00:00 (for which trade order flows are set to 0 for all other securities except the 2-year note), a

transaction at 10:00:35 (for which trade order flows are set to 0 for all except the 5-year note), and a transaction at 10:01:00 (for which trade order flows are set to 0 for all except the 2-year note). Complication arises because of the time dimension: a transaction does not finish instantaneously but is extended in time.<sup>33</sup> We overcome this by synchronizing all securities to a fine enough granularity – at the one-second frequency as in [Hasbrouck \(1995\)](#). If a trade is part of the initial execution of a marketable order, it is classified as a lit trade. If a trade is part of the workup and is on the same side as the initial market order, it is classified as a same-side workup trade. If a trade is part of the workup but is on the other side, it is classified as an other-side workup trade. We then aggregate these trade variables over each second. We measure prices of the securities as the mid-point prevailing at the end of each one-second interval.

The second issue is that the causality ordering of variables is no longer clear-cut with synchronized data. Information share calculations are sensitive to variable ordering (a variable will have the greatest information share if it is placed first, and smallest information share if placed last). We overcome this issue by computing information shares for many possible variable orderings, and then taking the average. It is impossible to do the calculation for all permutations of orderings given the dimension of the model (22 variables), so we resort to performing the calculation on selective orderings as follows. We permute the four blocks of variables (lit trade, return, same-side workup, other-side workup). Within each block, we consider six orderings of the maturities (where each maturity takes turn to be the first in the ordering) and assume the same maturity ordering for all four blocks. This results in 144 orderings over which to compute the information shares.<sup>34</sup>

## 4.2 Price discovery of yield curve factors

To assess the extent to which each order type drives the variation of innovations in the yield curve level and term premia, we aggregate the information shares across the six maturities for pre-workup

---

<sup>33</sup>For example, suppose the 10:00:00 transaction in the 2-year note does not finish until 10:00:50, and part of the order flow occurs after 10:00:35. Attaching workup order flow variables of the 2-year note to the 10:00:00 time stamp and the order flows of the 5-year note to the 10:00:35 time stamp would result in out-of-sequence trade flow variables, which might prevent us from correctly attributing the information content of trades in the 2- and 5-year notes.

<sup>34</sup>[Hasbrouck \(1995\)](#) finds that at the one-second sampling frequency, the lower and upper bounds of information shares estimated from different variable orderings are very tight. Thus, our one-second sampling frequency helps ease the concern over information shares' sensitivity to variable ordering.



trades, same-side workup trades, and other-side workup trades, and report the results in Table 10. We uncover two key findings. First, the workup protocol is nearly as informative to the yield curve level as the transparent trade channel, and that informativeness comes mostly from same-side (or aggressive) workups. This is qualitatively consistent with the results from the single-security model. Second, trades are most informative about the level factor. The information share of trades (aggregated across trade types and maturities) is 21%, implying that 79% of the variation in the level factor innovations is due to public information (more precisely, non-BrokerTec-trade-related information). On the other hand, trade innovations explain only 7%, 3%, and 4% of the variation in the innovations of the short-, medium-, and long-term slope factors.

These results imply that information related to the Treasury term premium is mostly impounded into prices via revisions in the limit order book and not via trades. [Brandt and Kavajecz \(2004\)](#), [Green \(2004\)](#), and [Pasquariello and Vega \(2007\)](#) all find that Treasury order flow is most important for the level factor and much less so for the slope and curvature factors. It is well-documented in the term structure literature that the level factor has much lower noise than the slope and curvature factors (see e.g., [Joslin, Singleton, and Zhu, 2011](#); [Joslin, Le, and Singleton, 2013](#)). Therefore, new information about the level factor can be much easier to discern by the market. It also means that traders with valuable information about the level factor are better off exploiting it as quickly as possible. In contrast, given the high level of noise surrounding the slope and curvature factors, it is harder for the market to decipher if an innovation is truly a shock or just noise.

Zooming in on the level factor, we analyze how the price discovery of the factor changes on macroeconomic and FOMC announcement days. Table IA4 in the Internet Appendix presents the trade-related information shares of each security with respect to the level factor, and Table IA5 reports the regression of these information shares on announcement dummies. We consider the information shares aggregated across the three trade flow types, as well as the fraction of the information share of workup trades (both same and opposite sides) to that of all trades. The findings are broadly similar with the single-security analysis presented earlier. First, total trade-related information shares are lower on announcement days, because public information takes on a more

important role. Second, workup trades become relatively less informative on days of macroeconomic announcements around which there is a high level of belief dispersion, indicating that informed traders who derive information advantage from public information prefer the lit trading channel more than usual.

### **4.3 Which maturities are the most informative?**

Having shown that trading and workup activity is most informative about the level factor, we next assess which security contributes the greatest to the price discovery of the level factor.<sup>35</sup> We plot the time series of the contributions of pre-workup, same-side workup, and other-side workup order flow by securities in a stacked area plot in Figure 5. The figure shows that the 5- and 10-year notes are most important. The time-series averages of these contributions, reported in Table IA4 in the Internet Appendix, indicate that the 5- and 10-year notes each contributes about 5% to the variance of the level factor's innovations, followed by the 30-year bond with a 3.4% contribution (the contributions by the other securities are small in comparison). Among order flow types, pre-workup and same-side workup order flow have similar contributions to the discovery of the level factor – 9.8% and 9.6% respectively – whereas other-side workup order flow contributes only 1.2%.

The finding that information shows up first in the 5- and 10-year notes, and to a slightly lesser extent the 30-year bond, is novel and consistent with the fact that these instruments are in high demand for hedging medium- and long-term interest rate risk. Indeed, the 30-year bond is most sensitive to interest rate risk, so despite its lower liquidity supply and demand, its order flow contributes significantly to the price discovery of the yield curve level factor.

## **5 Conclusion**

We study the workup protocol, an important price discovery mechanism in the U.S. Treasury securities market, and examine its role in the price formation process. We show that, while having no instantaneous price impact by design, workup trades contribute significantly to price discovery,

---

<sup>35</sup>Detailed decomposition of the contribution of the six securities to the price discovery of the level factor and three slope factors is provided respectively in Tables IA4 and IA6-IA8 in the Internet Appendix.

albeit less than lit trades, explaining between 6-8% of variation in efficient returns of the five benchmark Treasury notes. Most of workups' price discovery value comes from workups that expand volume on the aggressive side. The 30-year bond is an exception for which almost all price discovery is through the lit trade channel, a finding that is likely connected to its lower liquidity.

We also find that the balance of contribution to price discovery between lit and dark order flow tilts toward the lit channel around periods when private information production from publicly available data is expected to be most intense. In contrast, workup trades become more informative relative to lit trades when private information is most likely linked to confidential client order flow. We argue that this difference is probably due to the different nature of private information in the two cases. Private information derived from public information is likely short-lived given the level of sophistication of participants and the wide access to public information, whereas private information related to proprietary client order flow is more likely to last longer and affords the informed trader the patience to search for counter trading interest in workups to secure better execution prices.

When considering the joint dynamics of prices and trades across the yield curve, we find that workup order flow has more informational value than the single-security analysis indicates. Shocks in workup order flow explain over 9% of the variation of level factor innovations, just slightly below that of lit order flow. Moreover, order flow (both lit and dark) is most important for the discovery of the level factor, and much less so for the slope and curvature factors. Along the yield curve, trades in the 5- and 10-year notes convey the most private information about the level of Treasury yields.

Our paper contributes to research showing the important role of trading in aggregating private information about asset prices. Furthermore, we add to the discussion of market design in today's electronic limit order markets in response to the rise in high frequency trading. The continuous limit order market design has shown certain limitations (e.g., encouraging an arms race in trading technology and front-running by fast traders at the expense of slow traders). Naturally, these limitations invite further research into alternative market design features and possible implications. Our paper presents an interesting case study of a continuous limit order market combined with periodic call auctions, and offers empirical implications on trading patterns, exposure choice, and

price discovery in such a market setup. Potentially, these call auctions help reduce the speed of fast traders, allow all market participants the chance to take part in a transaction, and help slow traders better manage their order exposure and the risk of being picked off by faster participants. We leave the study of such benefits in the context of high frequency trading to future research.

## Appendix A List of workup explanatory variables

The workup probability and volume share models described in Section 2 in the paper contain the following explanatory variables to capture the trade-off among non-execution risk, increased control over one's trading activities, and the ability to avoid adverse price impact:

- prevailing inside depth on the same side of the transaction (logged).
- prevailing inside depth on the opposite side of the transaction (logged).
- prevailing relative spread in basis points  $\left(10,000 \frac{P_A - P_B}{(P_A + P_B)/2}\right)$ .
- pre-workup volume (the volume transacted before the workup starts) (logged).
- whether pre-workup trading has revealed any iceberg orders.
- average duration between transactions over the last five transactions (in seconds, logged).
- prevailing market volatility, measured by the high-low range of the logged mid-quote over the preceding five-minute interval.
- pre-trade probability of workup, calculated over the last five transactions.
- pre-trade volume share of workups, averaged over the last five transactions with workup.
- Tokyo and London trading hour dummies (equals 1 for transactions between 18:30 EST, or 19:30 EDT, the previous day and 3:00, and between 3:00 and 7:30 respectively).
- Pre-announcement, announcement, and post-announcement dummies (equals 1 for transactions occurring in the [-5min:0), [0:1min), and [1min:5min) interval around an announcement. The announcements are listed in Appendix B.)

## **Appendix B Economic announcements**

### **B.1 Macroeconomic announcements**

We consider macroeconomic announcements that are classified as "Market Moving" by Bloomberg: 1) Employment Report, 2) Consumer Price Index, 3) Durable Goods Orders, 4) GDP, 5) Housing Starts, 6) Initial Jobless Claims, 7) Personal Income and Outlays, 8) Producer Price Index, 9) Retail Sales, 10) Trade Balance, 11) Industrial Production, 12) Existing Home Sales, 13) ISM Manufacturing, 14) New Home Sales, and 15) Philadelphia Fed Business Outlook Survey.

### **B.2 Monetary policy announcements**

The monetary policy announcements considered are FOMC rate decision announcements, which typically occur after regularly scheduled FOMC meetings (eight times per year). In addition, there were two rate changes announced after unscheduled meetings during our sample period (January 22, 2008 and October 8, 2008).

### **B.3 Treasury auction result announcements**

We collect from <https://www.treasurydirect.gov> the dates and times of auctions of the six nominal coupon securities over the 2006-2016 period. Auction results are announced shortly after the closing time of competitive bidding, typically at 13:00 ET. The 2- and 5-year notes are newly issued every month. The 3- and 7-year notes are also on the regular monthly issuance schedule from November 2008 and February 2009 respectively. The 3-year was briefly discontinued between May 2007 and November 2008, and – prior to that period – was on a quarterly issuance schedule. The 10-year note is newly issued every quarter, with re-openings in the following month and – since November 2008 – two months. Starting in May 2009, the 30-year bond is also on a quarterly issuance cycle with two re-openings. Earlier, the 30-year bond was newly issued once a year with re-openings after six months (2006), or twice a year with re-openings after three months (2007 and 2008).

## References

- Balduzzi, P., E. Elton, and C. Green, 2001, “Economic News and Bond Prices: Evidence from the U.S. Treasury Market,” *Journal of Financial and Quantitative Analysis*, 36, 523–543.
- Bessembinder, H., M. Panayides, and K. Venkataraman, 2009, “Hidden Liquidity: An Analysis of Order Exposure Strategies in Automated Markets,” *Journal of Financial Economics*, 94, 361–383.
- Boni, L., and C. Leach, 2004, “Expandable Limit Order Markets,” *Journal of Financial Markets*, 7, 145–185.
- Brandt, M., and K. Kavajecz, 2004, “Price Discovery in the U.S. Treasury Market: the Impact of Order Flow and Liquidity on the Yield Curve,” *Journal of Finance*, 59, 2623–2654.
- Burdett, K., and M. O’Hara, 1987, “Building blocks: an introduction to block trading,” *Journal of Banking and Finance*, 11, 193–212.
- Buti, S., B. Rindi, and I. Werner, 2011, “Diving Into Dark Pools,” University of Toronto, Bocconi University and Ohio State University Working Paper.
- Buti, S., B. Rindi, and I. Werner, 2017, “Dark Pool Trading Strategies, Market Quality, and Welfare,” *Journal of Financial Economics*, 124, 244–265.
- Cohen, B., and H. S. Shin, 2003, “Positive Feedback Trading under Stress: Evidence from the U.S. Treasury securities market,” BIS Working Papers No. 122.
- Collin-Dufresne, P., B. Junge, and A. Trolle, 2016, “Market Structure and Transaction Costs of Index CDSs,” EPFL and Swiss Finance Institute Working Paper.
- Comerton-Forde, C., and T. Putnins, 2015, “Dark Trading and Price Discovery,” *Journal of Financial Economics*, 118, 70–92.

- Degryse, H., M. Van Achter, and G. Wuyts, 2009, “Dynamic order submission strategies with competition between a dealer market and a crossing network,” *Journal of Financial Economics*, 91, 319–338.
- Duffie, D., and H. Zhu, 2017, “Size Discovery,” *Review of Financial Studies*, 30, 1095–1150.
- Dungey, M., O. Henry, and M. McKenzie, 2013, “Modeling Trade Duration in U.S. Treasury Markets,” *Quantitative Finance*, 13, 1431–1442.
- Dunne, P., Y. Li, and Z. Sun, 2015, “Price Discovery in the Dual-Platform U.S. Treasury Market,” *Global Finance Journal*, 28, 95–110.
- Engle, R., M. Fleming, E. Ghysels, and G. Nguyen, 2018, “Liquidity and Volatility in the U.S. Treasury market,” Working Paper.
- Fleming, M., 1997, “The Round-the-Clock Market for U.S. Treasury Securities,” *FRBNY Economic Policy Review*, July 1997, 9–32.
- Fleming, M., B. Mizrach, and G. Nguyen, 2018, “The Microstructure of a U.S. Treasury ECN: The BrokerTec Platform,” *Journal of Financial Markets*, (forthcoming).
- Fleming, M., and E. Remolona, 1997, “What Moves the Bond Market?,” *Federal Reserve Bank of New York Economic Policy Review*, December, 31–50.
- Fleming, M., and E. Remolona, 1999, “Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information,” *Journal of Finance*, 54, 1901–1915.
- Green, C., 2004, “Economic News and the Impact of Trading on Bond Prices,” *Journal of Finance*, 59, 1201–1233.
- Harris, L. E., 1997, “Order Exposure and Parasitic Traders,” USC Working Paper.
- Hasbrouck, J., 1991, “The Summary Informativeness of Stock Trades: An Econometric Analysis,” *Review of Financial Studies*, 4, 571–595.



- Hasbrouck, J., 1995, “One Security, Many Markets: Determining the Contributions to Price Discovery,” *Journal of Finance*, 50, 1175–1199.
- Hasbrouck, J., 2007, *Empirical Market Microstructure*, Oxford University Press.
- Hendershott, T., and H. Mendelson, 2000, “Crossing networks and dealer markets: competition and performance,” *Journal of Finance*, 55, 2071–2115.
- Jiang, G., and I. Lo, 2014, “Private Information Flow and Price Discovery in the U.S. Treasury Market,” *Journal of Banking and Finance*, 47, 118–133.
- Joint Staff Report, 2015, “The U.S. Treasury Market on October 15, 2014,” U.S. Department of the Treasury, Board of Governors of the Federal Reserve System, Federal Reserve Bank of New York, U.S. Securities and Exchange Commission, and U.S. Commodity Futures Trading Commission.
- Joslin, S., A. Le, and K. Singleton, 2013, “Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs,” *Journal of Financial Economics*, 109, 604–622.
- Joslin, S., K. Singleton, and H. Zhu, 2011, “A New Perspective on Gaussian Dynamic Term Structure Models,” *Review of Financial Studies*, 24, 926–970.
- Lou, D., H. Yan, and J. Zhang, 2013, “Anticipated and Repeated Shocks in Liquid Markets,” *Review of Financial Studies*, 26, 1891–1912.
- Luo, H., 2010, “Profitable Opportunities around Macroeconomic Announcements in the U.S. Treasury Market,” Brock University Master Thesis.
- Menkveld, A., H. Zhu, and B. Yueshen, 2017, “Shades of Darkness: A Pecking Order of Trading Venues,” *Journal of Financial Economics*, 124, 503–534.
- Mizrach, B., and C. Neely, 2008, “Information shares in the U.S. Treasury market,” *Journal of Banking and Finance*, 32, 1221–1233.

- O'Hara, M., 2015, "High Frequency Market Microstructure," *Journal of Financial Economics*, 116, 257–270.
- Pancs, R., 2013, "Workup," *Review of Economic Design*, 18, 37–71.
- Parlour, C., 1998, "Price Dynamics in Limit Order Markets," *Review of Financial Studies*, 11, 789–816.
- Pasquariello, P., and C. Vega, 2007, "Informed and Strategic Order Flow in the Bond Markets," *Review of Financial Studies*, 20, 1975–2019.
- Ready, M., 2012, "Determinants of Volume in Dark Pools," University of Wisconsin-Madison Working Paper.
- Sofianos, G., and I. Werner, 2000, "The trades of NYSE floor brokers," *Journal of Financial Markets*, 3, 139–176.
- Ye, M., 2012, "A Glimpse into the Dark: Price Formation, Transaction Costs, and Market Share in the Crossing Network," UIUC Working Paper.
- Zhu, H., 2014, "Do Dark Pools Harm Price Discovery?," *Review of Financial Studies*, 27, 747–789.

**Table 1: Summary Statistics of Trading and Workup Activities**

	2-Year	3-Year	5-Year	7-Year	10-Year	30-Year
<b>Panel A: Day-level statistics</b>						
Volume (\$B)	25.9	15.7	38.3	11.9	33.5	6.9
<i>Pre-workup %</i>	48.6	45.4	38.0	40.2	39.9	52.0
<i>Workup %</i>	51.4	54.6	62.0	59.8	60.1	48.0
Transaction Count	917	1,030	2,701	1,517	2,779	2,241
<i>% with Workup</i>	51.4	54.2	60.4	56.7	59.2	44.6
<i>% with Iceberg Order Match</i>	11.1	10.2	8.7	5.3	8.8	6.9
Pre-trade Depth (Same Side, \$M)	300.45	121.72	52.75	55.26	44.31	4.49
Pre-trade Depth (Opposite Side, \$M)	102.30	35.11	19.21	11.03	16.36	2.44
<b>Panel B: Transactions with workup</b>						
Transaction Size (\$M)	45.4	24.3	21.5	12.7	18.4	5.2
<i>Pre-workup</i>	16.4	8.2	6.3	3.9	5.6	1.8
<i>Workup (same side)</i>	21.2	11.8	9.7	5.8	8.2	2.0
<i>Workup (other side)</i>	7.8	4.4	5.5	3.0	4.5	1.4
Trade Count	11.0	9.9	11.3	9.1	11.1	4.3
<i>Pre-workup</i>	3.3	3.1	3.3	2.8	3.4	1.5
<i>Workup (same side)</i>	5.6	4.9	5.1	4.1	4.9	1.7
<i>Workup (other side)</i>	2.1	1.9	2.9	2.2	2.8	1.1
<b>Panel C: Transactions without workup</b>						
Transaction Size (\$M)	11.7	6.2	4.5	2.7	4.0	1.5
Trade Count	2.4	2.3	2.3	2.0	2.4	1.3
<b>Panel D: Sample size</b>						
Number of Observations	2,523,594	2,595,614	7,431,732	2,975,486	7,627,895	6,166,652
Number of Trading Days	2,752	2,520	2,752	1,962	2,745	2,752

This table provides summary statistics of trading and workup activities in the on-the-run nominal coupon Treasury securities on the BrokerTec platform. The sample period is 2006-2016 (with incomplete coverage for the 3- and 7-year notes due to issuance discontinuity, and the 10-year note due to data unavailability). A transaction refers to a complete sequence of order executions that starts with the arrival of a market order and ends when all workup activities associated with the original market order complete. A trade refers to a single paired order matching. Numbers reported in Panel A are daily averages. Numbers in Panel B are averages across transactions with at least one trade during the workup phase and numbers in Panel C are averages across transactions without any trades during the workup phase.

**Table 2: Determinants of Workup Probability**

	2-Year	3-Year	5-Year	7-Year	10-Year	30-Year
Intercept	-0.951***	-1.336***	-1.105***	-1.672***	-1.067***	-1.133***
Same-side depth	0.151***	0.242***	0.252***	0.288***	0.266***	0.173***
Opposite-side depth	-0.137***	-0.014***	0.013***	0.254***	-0.034***	0.896***
Spread	-0.029***	-0.021***	-0.010***	-0.015***	-0.009***	-0.003***
Pre-workup log volume	0.133***	0.060***	0.177***	0.108***	0.191***	-0.269***
Hidden depth revealed	0.155***	0.254***	0.241***	0.345***	0.199***	0.844***
Market volatility	7.469***	4.366***	3.059***	1.515***	2.316***	0.124***
Inter-trade duration	-0.073***	-0.125***	-0.155***	-0.120***	-0.145***	-0.170***
Recent workup probability	0.724***	0.782***	0.602***	0.573***	0.524***	0.383***
Recent workup volume share	0.425***	0.468***	0.361***	0.329***	0.377***	0.206***
Tokyo trading hour dummy	-0.419***	-0.038***	-0.075***	0.021***	-0.241***	-0.250***
London trading hour dummy	-0.249***	0.014***	-0.062***	0.019***	-0.097***	-0.069***
Pre-announcement	0.487***	0.282***	-0.020*	0.221***	-0.057***	-0.051***
Announcement	0.076***	0.018	-0.013	0.059***	-0.031***	-0.036***
Post-announcement	-0.018	-0.050***	0.015*	-0.081***	0.016*	-0.107***
Max-rescaled R-squared	0.085	0.071	0.077	0.080	0.084	0.096

This table reports the coefficients of the logistic regression for whether traders participate in workups. The dependent variable is equal to 1 for transactions with workup and 0 for transactions without workup. Depth (logged) and spread variables are measured immediately before each transaction. Pre-workup log volume is the log of volume transacted before a workup starts. Hidden depth revealed is an indicator for whether trading in the pre-workup stage uncovers hidden liquidity. Market volatility is measured over the five-minute interval immediately before each transaction. It is measured by the high-low range of logged price over the interval, scaled by  $\sqrt{4\log 2}$  and annualized by a factor of  $\sqrt{12 \times 24 \times 250}$ . Recent inter-trade duration, workup probability, and workup volume share conditioning on workup are computed from the preceding five transactions. Tokyo trading hour dummy is equal to 1 for the period from 18:30 EST (or 19:30 EDT) the previous day to 3:00 ET. London trading hour dummy is equal to 1 for the period from 3:00 ET to 7:30 ET. Announcement is equal to 1 if the transaction occurs during the same minute as an announcement. Pre-announcement (Post-announcement) is equal to 1 if the transaction occurs in the five-minute period before (after) an announcement. The model is estimated using BrokerTec data for on-the-run 2-, 3-, 5-, 7-, 10-, and 30-year Treasury securities over the period 2006-2016. Statistical significance at the 1%, 5%, and 10% levels is marked with \*\*\*, \*\*, and \* respectively.

**Table 3: Determinants of Workup Volume Share**

	2-Year	3-Year	5-Year	7-Year	10-Year	30-Year
Intercept	-1.419***	-1.764***	-1.453***	-1.867***	-1.433***	-1.268***
Same-side depth	0.182***	0.266***	0.275***	0.283***	0.294***	0.174***
Opposite-side depth	-0.032***	0.108***	0.152***	0.363***	0.129***	0.813***
Spread	-0.024***	-0.018***	-0.010***	-0.014***	-0.009***	-0.002***
Pre-workup log volume	-0.366***	-0.537***	-0.514***	-0.645***	-0.554***	-1.013***
Hidden depth revealed	0.218***	0.308***	0.295***	0.399***	0.268***	0.727***
Market volatility	7.605***	4.505***	2.873***	1.240***	2.096***	0.163***
Inter-trade duration	-0.059***	-0.095***	-0.123***	-0.109***	-0.117***	-0.138***
Recent workup probability	0.617***	0.653***	0.496***	0.514***	0.427***	0.333***
Recent workup volume share	0.430***	0.436***	0.341***	0.340***	0.356***	0.200***
Tokyo trading hour dummy	-0.412***	-0.105***	-0.133***	-0.026***	-0.264***	-0.296***
London trading hour dummy	-0.264***	-0.053***	-0.109***	-0.028***	-0.136***	-0.108***
Pre-announcement	0.508***	0.297***	0.045***	0.198***	0.040***	-0.055***
Announcement	0.095***	0.018**	-0.012**	0.053***	-0.019***	-0.037***
Post-announcement	-0.014	-0.024**	0.026***	-0.079***	0.026***	-0.098***

This table reports the coefficients of the generalized linear mixed model for workup volume share. The dependent variable is the fraction of a transaction's volume executed in the workup phase, and is modeled with a logit link function. Depth (logged) and spread variables are measured immediately before each transaction. Pre-workup log volume is the log of volume transacted before a workup starts. Hidden depth revealed is an indicator for whether trading in the pre-workup stage uncovers hidden liquidity. Market volatility is measured over the five-minute interval immediately before each transaction. It is measured by the high-low range of logged price over the interval, scaled by  $\sqrt{4\log 2}$  and annualized by a factor of  $\sqrt{12 \times 24 \times 250}$ . Recent inter-trade duration, workup probability, and workup volume share conditioning on workup are computed from the preceding five transactions. Tokyo trading hour dummy is equal to 1 for the period from 18:30 EST (or 19:30 EDT) the previous day to 3:00 ET. London trading hour dummy is equal to 1 for the period from 3:00 ET to 7:30 ET. Announcement is equal to 1 if the transaction occurs during the same minute as an announcement. Pre-announcement (Post-announcement) is equal to 1 if the transaction occurs in the five-minute period before (after) an announcement. The model is estimated using BrokerTec data for on-the-run 2-, 3-, 5-, 7-, 10-, and 30-year Treasury securities over the period 2006-2016. Statistical significance at the 1%, 5%, and 10% levels is marked with \*\*\*, \*\*, and \* respectively.

**Table 4: Permanent Price Impact of Segmented Order Flow**

	Workup Trades								
	Lit Trades			Same Side			Other Side		
	Mean	90% Range		Mean	90% Range		Mean	90% Range	
	$\times 10^{-3}$								
2-Year	2.79*	[0.56	7.84]	3.11*	[0.70	7.96]	-2.34	[-7.00	0.99]
3-Year	7.87*	[1.58	22.25]	8.73*	[2.00	21.63]	-4.67	[-14.02	2.53]
5-Year	16.95*	[5.10	42.88]	10.02*	[4.46	19.20]	-5.25*	[-12.11	-0.09]
7-Year	42.72*	[14.11	98.88]	38.19*	[14.56	80.89]	-16.56	[-49.47	10.71]
10-Year	39.33*	[13.36	93.23]	23.83*	[11.51	43.64]	-13.44*	[-28.81	-0.80]
30-Year	395.50*	[153.09	771.34]	79.13	[-27.70	214.51]	-46.64	[-197.03	66.00]

This table reports the permanent price impact per unitary shock in each order flow variable, derived from a VAR(5) model of pre-workup trade flow, return, workup trade flow initiated from the same side, and workup trade flow initiated from the other side. Estimation is based on BrokerTec data for on-the-run nominal coupon Treasury securities over the period 2006-2016. Observations outside the [7:00-17:30] time window are excluded. The model is estimated separately for each day. The mean and 90% range (5th and 95th percentiles) are computed from the time series of daily price impact estimates. Statistical significance at the 5% level (one-sided) is marked with an asterisk (\*).

**Table 5:** Information Shares of Pre-workup and Workup Order Flow

	Pre-workup			Workup (Same Side)			Workup (Other Side)			Total		
	Mean	90% Range		Mean	90% Range		Mean	90% Range		Mean	90% Range	
2-Year	11.82	[5.92	18.86]	7.64	[2.88	14.81]	2.79	[0.56	7.55]	22.24	[13.82	33.80]
3-Year	11.47	[6.00	18.13]	7.18	[2.17	12.70]	1.71	[0.43	3.79]	20.36	[12.27	29.37]
5-Year	13.30	[8.45	17.53]	6.56	[2.12	11.35]	1.68	[0.45	3.39]	21.54	[13.50	28.84]
7-Year	11.21	[5.74	17.43]	5.91	[2.19	9.88]	1.40	[0.32	2.92]	18.52	[10.82	26.21]
10-Year	15.16	[10.30	19.56]	7.09	[2.63	11.39]	1.74	[0.60	3.46]	24.00	[17.20	30.37]
30-Year	18.26	[10.15	26.10]	1.32	[0.31	3.04]	0.51	[0.10	1.28]	20.09	[11.83	28.48]

This table reports the information share (%) of three components of order flow: pre-workup trades, workup trades initiated from the same side, and workup trades initiated from the other side. The estimates are derived from a VAR(5) model of pre-workup trade flow, return, workup trade flow initiated from the same side, and workup trade flow initiated from the other side. Estimation is based on BrokerTec data for on-the-run nominal coupon Treasury securities over the period 2006-2016. Observations outside the [7:00-17:30] time window are excluded. The model is estimated separately for each day. The mean and 90% range (5th and 95th percentiles) are computed from the time series of daily information share estimates.

**Table 6: Price Discovery on Announcement Days**

Variable	2-Year	3-Year	5-Year	7-Year	10-Year	30-Year
Panel A: $IS = IS_{prewk} + IS_{wkSame} + IS_{wkOther}$						
Auction	-0.072**	-0.123***	0.033	0.022	-0.010	-0.093**
FOMC	-0.144***	-0.301***	-0.357***	-0.338***	-0.199***	-0.318***
Macro	0.005	-0.051***	-0.066***	-0.066***	-0.025***	-0.095***
Trend	0.058***	0.055***	0.064***	0.140***	0.041***	0.091***
Intercept	-1.501***	-1.575***	-1.515***	-2.155***	-1.301***	-1.690***
Adj. $R^2$	0.115	0.138	0.218	0.350	0.167	0.349
Panel B: $IS = \frac{IS_{wkSame} + IS_{wkOther}}{IS_{prewk} + IS_{wkSame} + IS_{wkOther}}$						
Auction	0.048	0.016	0.017	0.095*	0.060	0.222**
FOMC	0.055	0.019	0.038	-0.025	0.046	-0.005
Macro	0.058**	0.091***	0.010	0.011	0.018	-0.041*
Trend	0.054***	0.065***	0.081***	0.006	0.067***	-0.113***
Intercept	-0.415***	-0.605***	-0.871***	-0.496***	-0.862***	-1.969***
Adj. $R^2$	0.039	0.065	0.134	0.000	0.100	0.158
N	2730	2515	2751	1960	2744	2750

This table reports estimates of the regression:  $\ln \frac{IS_t}{1-IS_t} = \alpha + \beta_1 \times FOMC_t + \beta_2 \times Macro_t + \beta_3 \times Auction_t + \beta_4 \times Trend + \epsilon_t$ , where *FOMC*, *Macro*, and *Auction* are dummies that equal 1 on days of FOMC rate decision announcements, macroeconomic announcements, and auctions of new securities of a given maturity, and *Trend* is a linear time trend. In Panel A, the dependent variable *IS* is the total information share of pre-workup (*prewk*), same-side workup (*wkSame*), and other-side workup (*wkOther*) trades. In Panel B, the dependent variable *IS* is the fraction of the trade-related information share attributable to workup trades (i.e., the sum of information shares of same-side workup and other-side workup trades, divided by the total trade-related information share). The daily information shares are computed from a VAR(5) model of pre-workup trade flow, return, same-side workup trade flow, and other-side workup trade flow. The model is estimated separately for each day and is based on BrokerTec data for on-the-run nominal coupon Treasury securities over the period 2006-2016. Observations outside the [7:00-17:30] time window are excluded. Statistical significance at the 1%, 5%, and 10% levels is marked with \*\*\*, \*\*, and \* respectively.



**Table 7: Price Discovery on Announcement Days with High Dispersion of Beliefs**

Variable	2-Year	3-Year	5-Year	7-Year	10-Year	30-Year
Panel A: $IS = IS_{prewk} + IS_{wkSame} + IS_{wkOther}$						
Auction	-0.073**	-0.120***	0.028	0.026	-0.007	-0.099**
FOMC	-0.137***	-0.293***	-0.347***	-0.334***	-0.191***	-0.314***
Macro	0.007	-0.044***	-0.062***	-0.065***	-0.023***	-0.098***
Macro x High Belief Dispersion	-0.098**	-0.205***	-0.141***	-0.169***	-0.105***	-0.023
Macro x Low Belief Dispersion	0.086*	0.116**	0.117***	0.108**	0.092***	0.130***
Trend	0.057***	0.053***	0.062***	0.137***	0.040***	0.090***
Intercept	-1.497***	-1.566***	-1.509***	-2.140***	-1.297***	-1.687***
Adj. $R^2$	0.117	0.148	0.225	0.355	0.174	0.351
Panel B: $IS = \frac{IS_{wkSame} + IS_{wkOther}}{IS_{prewk} + IS_{wkSame} + IS_{wkOther}}$						
Auction	0.050	0.019	0.004	0.096*	0.064	0.227**
FOMC	0.061	0.025	0.059	-0.019	0.061	-0.001
Macro	0.066***	0.106***	0.025	0.024	0.028	-0.033
Macro x High Belief Dispersion	-0.142**	-0.281***	-0.355***	-0.396***	-0.263***	-0.107
Macro x Low Belief Dispersion	-0.042	-0.036	0.078	0.022	0.087	-0.084
Trend	0.053***	0.062***	0.078***	0.000	0.065***	-0.114***
Intercept	-0.411***	-0.596***	-0.859***	-0.469***	-0.854***	-1.967***
Adj. $R^2$	0.040	0.070	0.148	0.012	0.108	0.158
N	2730	2515	2751	1960	2744	2750

This table reports estimates of the regression:  $\ln \frac{IS_t}{1-IS_t} = \alpha + \beta_1 \times FOMC_t + \beta_2 \times Macro_t + \beta_3 \times Auction_t + \beta_4 \times Trend + \beta_5 \times (Macro_t \times HiDif_t) + \beta_6 \times (Macro_t \times LoDif_t) + \epsilon_t$ , where  $FOMC$ ,  $Macro$ , and  $Auction$  are dummies that equal 1 on days of FOMC rate decision announcements, macroeconomic announcements, and auctions of new securities of a given maturity,  $Trend$  is a linear time trend, and  $HiDif$  and  $LoDif$  are dummies for macroeconomic announcement days with high (top 5%) and low (bottom 5%) dispersion of beliefs among market participants. Dispersion of beliefs is measured by the standard deviation of Bloomberg economist forecasts for each announcement. On days with multiple announcements, the dispersion measure is averaged across announcements, weighted by the number of participating economists. In Panel A, the dependent variable  $IS$  is the total information share of pre-workup ( $prewk$ ), same-side workup ( $wkSame$ ), and other-side workup ( $wkOther$ ) trades. In Panel B, the dependent variable  $IS$  is the fraction of the trade-related information share attributable to workup trades (i.e., the sum of information shares of same-side workup and other-side workup trades, divided by the total trade-related information share). The daily information shares are computed from a VAR(5) model of pre-workup trade flow, return, same-side workup trade flow, and other-side workup trade flow. The model is estimated separately for each day and is based on BrokerTec data for on-the-run nominal coupon Treasury securities over the period 2006-2016. Observations outside the [7:00-17:30] time window are excluded. Statistical significance at the 1%, 5%, and 10% levels is marked with \*\*\*, \*\*, and \* respectively.

**Table 8:** Price Discovery on Non-Announcement Days with Extreme Order Imbalances

Variable	2-Year	3-Year	5-Year	7-Year	10-Year	30-Year
Panel A: $IS = IS_{prewk} + IS_{wkSame} + IS_{wkOther}$						
MagOIV	0.199	0.072	1.238***	-1.697***	0.329	-0.517*
Trend	0.064***	0.043***	0.067***	0.131***	0.046***	0.080***
Intercept	-1.536***	-1.532***	-1.559***	-2.054***	-1.331***	-1.636***
Adj. $R^2$	0.124	0.075	0.231	0.445	0.200	0.328
Panel B: $IS = \frac{IS_{wkSame} + IS_{wkOther}}{IS_{prewk} + IS_{wkSame} + IS_{wkOther}}$						
MagOIV	1.362***	-0.633	2.251***	-0.928	0.710	1.618**
Trend	0.051***	0.063***	0.086***	0.004	0.069***	-0.091***
Intercept	-0.511***	-0.569***	-0.953***	-0.456***	-0.888***	-2.099***
Adj. $R^2$	0.064	0.060	0.134	0.001	0.098	0.124
N	1015	910	1065	769	1057	1092

This table reports estimates of the regression:  $\ln \frac{IS_t}{1-IS_t} = \alpha + \beta_1 \times MagOIV_t + \beta_2 \times Trend + \epsilon_t$ , where  $MagOIV_t$  is the magnitude of order flow imbalance standardized by total trading volume on day  $t$ , and  $Trend$  is a linear time trend. The regression is estimated on non-announcement dates only. In Panel A, the dependent variable  $IS$  is the total information share of pre-workup ( $prewk$ ), same-side workup ( $wkSame$ ), and other-side workup ( $wkOther$ ) trades. In Panel B, the dependent variable  $IS$  is the fraction of the trade-related information share attributable to workup trades (i.e., the sum of information shares of same-side workup and other-side workup trades, divided by the total trade-related information share). The daily information shares are computed from a VAR(5) model of pre-workup trade flow, return, same-side workup trade flow, and other-side workup trade flow. The model is estimated separately for each day and is based on BrokerTec data for on-the-run nominal coupon Treasury securities over the period 2006-2016. Observations outside the [7:00-17:30] time window are excluded. Statistical significance at the 1%, 5%, and 10% levels is marked with \*\*\*, \*\*, and \* respectively.

**Table 9: Price Discovery and Volatility**

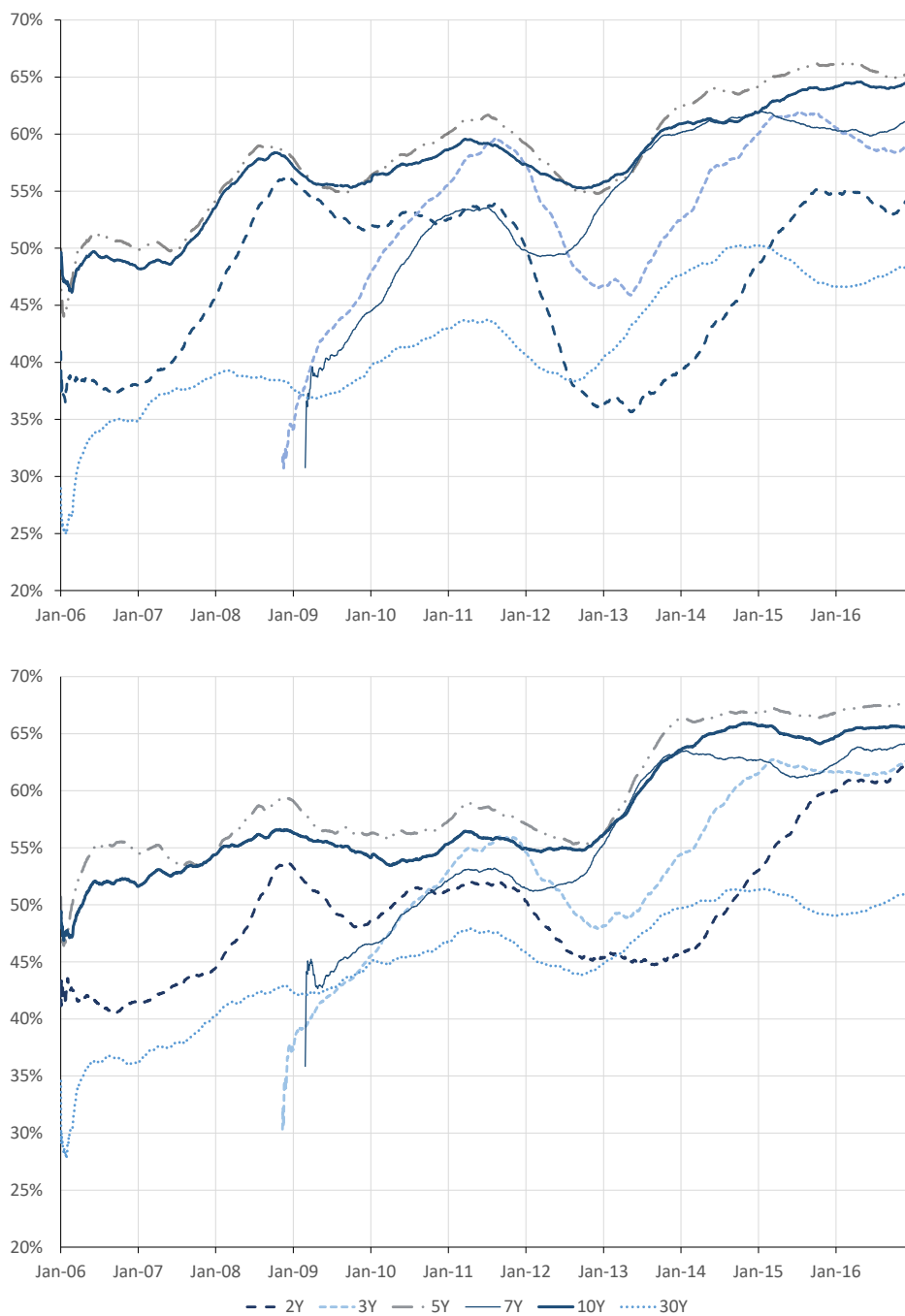
Variable	2-Year	3-Year	5-Year	7-Year	10-Year	30-Year
Panel A: $IS = IS_{prewk} + IS_{wkSame} + IS_{wkOther}$						
Log Volatility	-0.212***	-0.274***	-0.310***	-0.271***	-0.182***	-0.289***
Trend	0.032***	0.042***	0.048***	0.124***	0.037***	0.091***
Intercept	-2.255***	-2.568***	-2.449***	-2.837***	-1.747***	-2.306***
Adj. $R^2$	0.187	0.219	0.333	0.391	0.213	0.432
Log Volatility	-0.171***	-0.134***	-0.252***	-0.263***	-0.103***	-0.240***
LogDepth	-0.012	0.096***	0.145***	0.096***	0.131***	0.098***
LogNtrans	-0.069**	0.038	0.244***	0.169***	0.187***	0.061***
Trend	0.030***	0.045***	0.043***	0.113***	0.037***	0.102***
Intercept	-1.646***	-2.084***	-3.543***	-3.618***	-2.439***	-2.094***
Adj. $R^2$	0.188	0.244	0.362	0.404	0.267	0.436
Panel B: $IS = \frac{IS_{wkSame} + IS_{wkOther}}{IS_{prewk} + IS_{wkSame} + IS_{wkOther}}$						
Log Volatility	-0.321***	-0.155***	-0.469***	-0.211***	-0.497***	-0.375***
Trend	0.014**	0.058***	0.057***	-0.006	0.055***	-0.112***
Intercept	-1.512***	-1.092***	-2.205***	-0.981***	-2.011***	-2.703***
Adj. $R^2$	0.109	0.073	0.307	0.026	0.238	0.225
Log Volatility	-0.464***	-0.159***	-0.033	0.096***	-0.023	0.113***
LogDepth	0.140***	0.334***	0.577***	0.657***	0.649***	0.888***
LogNtrans	0.456***	0.814***	0.618***	0.608***	0.804***	0.371***
Trend	0.033***	0.021***	0.060***	-0.044***	0.065***	0.020**
Intercept	-4.996***	-5.661***	-3.261***	-1.901***	-4.305***	0.540**
Adj. $R^2$	0.128	0.204	0.546	0.294	0.583	0.369
N	2730	2515	2751	1960	2744	2750

This table reports estimates of the regression:  $\ln \frac{IS_t}{1-IS_t} = \alpha + \beta_1 \times \text{LogVola}_t + \beta_2 \times \text{Trend} + \epsilon_t$ , where  $\text{LogVola}$  is the daily realized volatility (logged) computed as the daily sum of squared one-second returns, and  $\text{Trend}$  is a linear time trend. We also report the estimates of the regression that includes liquidity variables:  $\text{LogDepth}_t$ , which is the log of inside market depth averaged over day  $t$ , and  $\text{LogNtrans}_t$ , which is the log number of transactions on day  $t$ . In Panel A, the dependent variable  $IS$  is the total information share of pre-workup ( $prewk$ ), same-side workup ( $wkSame$ ), and other-side workup ( $wkOther$ ) trades. In Panel B, the dependent variable  $IS$  is the fraction of the trade-related information share attributable to workup trades (i.e., the sum of information shares of same-side workup and other-side workup trades, divided by the total trade-related information share). The daily information shares are computed from a VAR(5) model of pre-workup trade flow, return, same-side workup trade flow, and other-side workup trade flow. The model is estimated separately for each day and is based on BrokerTec data for on-the-run nominal coupon Treasury securities over the period 2006-2016. Observations outside the [7:00-17:30] time window are excluded. Statistical significance at the 1%, 5%, and 10% levels is marked with \*\*\*, \*\*, and \* respectively.

**Table 10:** Information Shares of Trades for Yield Curve Factors

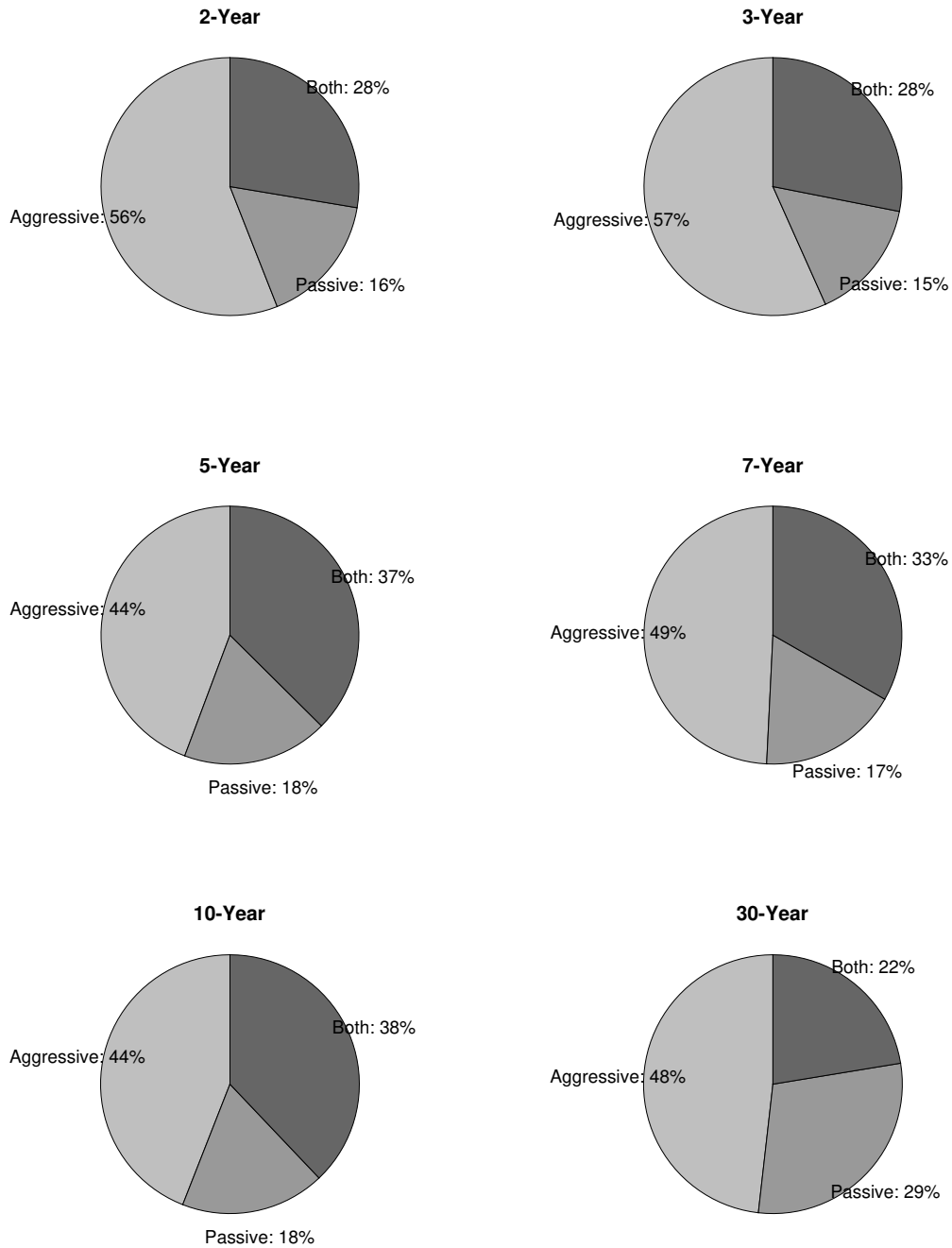
	Statistic	Pre-workup	Workup (same side)	Workup (other side)	Total
Level Factor	Mean	9.834	9.625	1.211	20.669
	StdErr	(0.044)	(0.052)	(0.017)	(0.097)
Short-term Slope Factor	Mean	2.677	3.695	0.742	7.113
	StdErr	(0.025)	(0.037)	(0.020)	(0.063)
Medium-term Slope Factor	Mean	1.647	1.423	0.323	3.393
	StdErr	(0.012)	(0.015)	(0.005)	(0.023)
Long-term Slope Factor	Mean	2.964	0.890	0.424	4.278
	StdErr	(0.014)	(0.007)	(0.004)	(0.018)

This table reports the contributions of order flow variables (in percentage points) to the variance of the four factors of the Treasury benchmark yield curve, derived from a VAR(5) model of pre-workup trade flow, same-side workup trade flow, other-side workup trade flow, and four yield curve factors extracted from one-second returns of six nominal on-the-run coupon Treasury securities. Estimation is based on BrokerTec data over the period from February 27, 2009 to the end of 2016 (1953 days in total). The model is estimated and information shares are computed separately for each trading day. Observations outside the [7:00-17:30] time window are excluded.



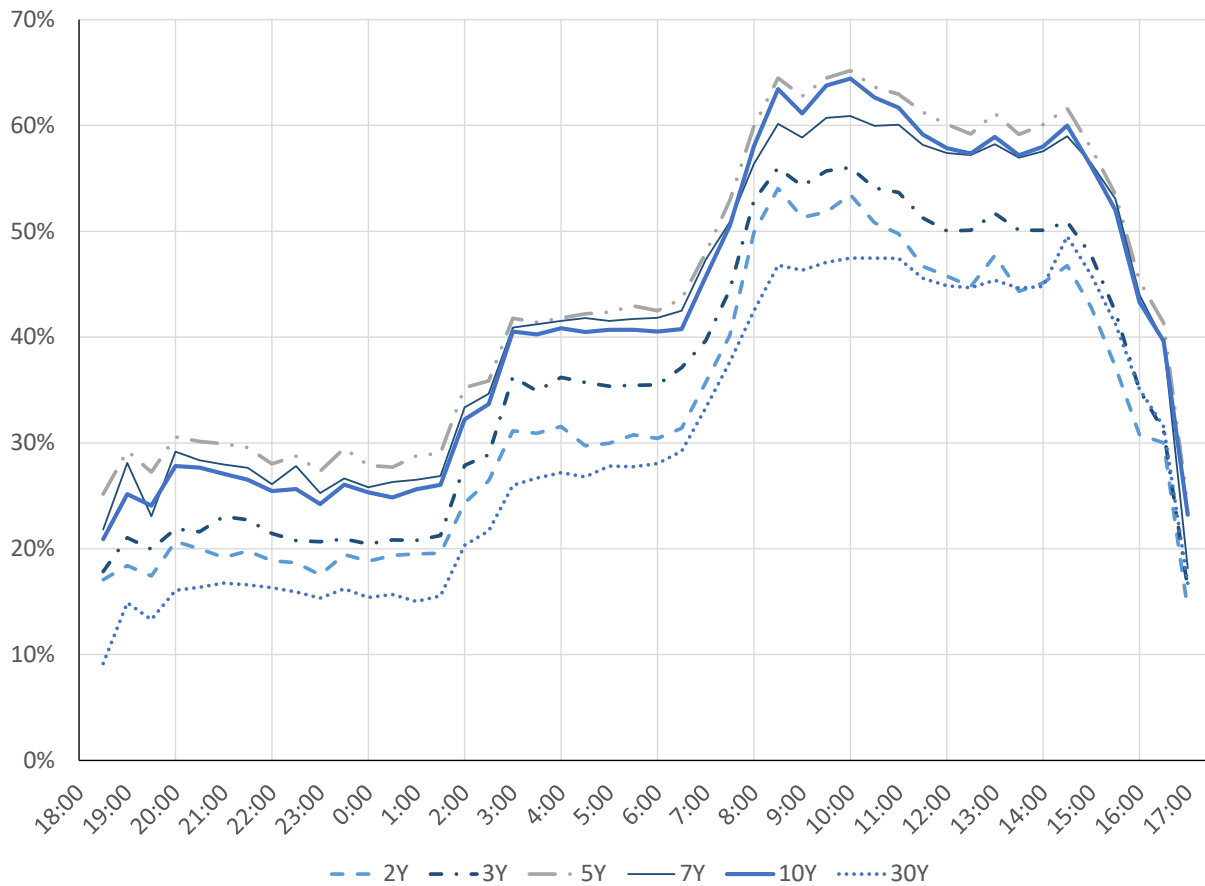
**Figure 1: Workup Activity over Time**

This figure shows the 250-day moving average of daily workup probability (upper plot) calculated as the percent of transactions for which there is at least one trade during the workup phase, and daily workup volume share (lower plot) calculated as the percent of trade volume transacted during the workup phase. The calculations are based on BrokerTec data for on-the-run nominal coupon securities over the 2006-2016 period.



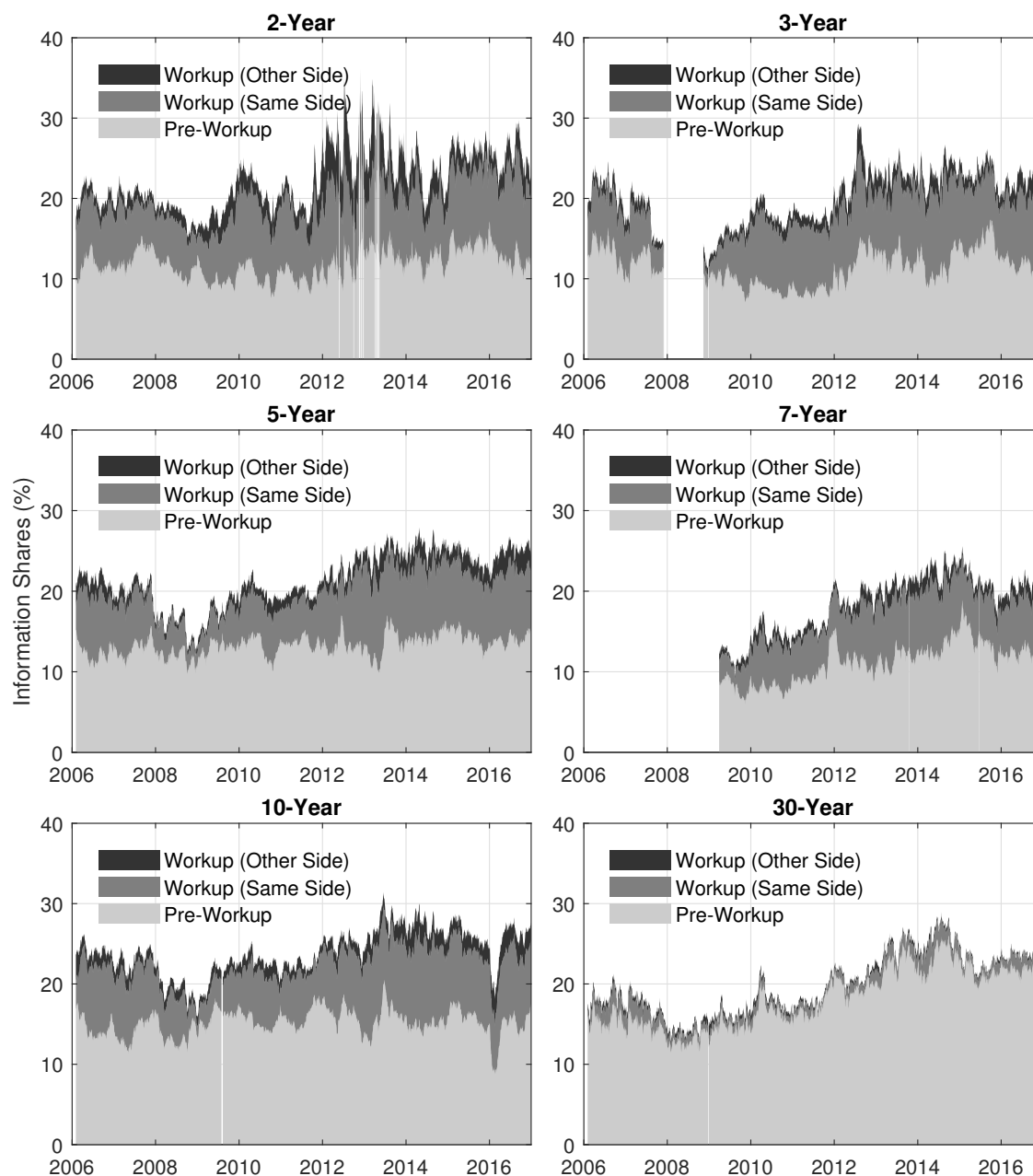
**Figure 2: Which Side Do Workups Expand?**

This figure shows the percentages of transactions in which there is at least one trade during the workup phase and in which workup trades are initiated from: 1) the aggressive side only, 2) the passive side only, and 3) both sides. The percentages are based on counts of transactions of the three types. The calculations are based on BrokerTec data for on-the-run nominal coupon securities over the 2006-2016 period.



**Figure 3: Intraday Pattern of Workup Probability**

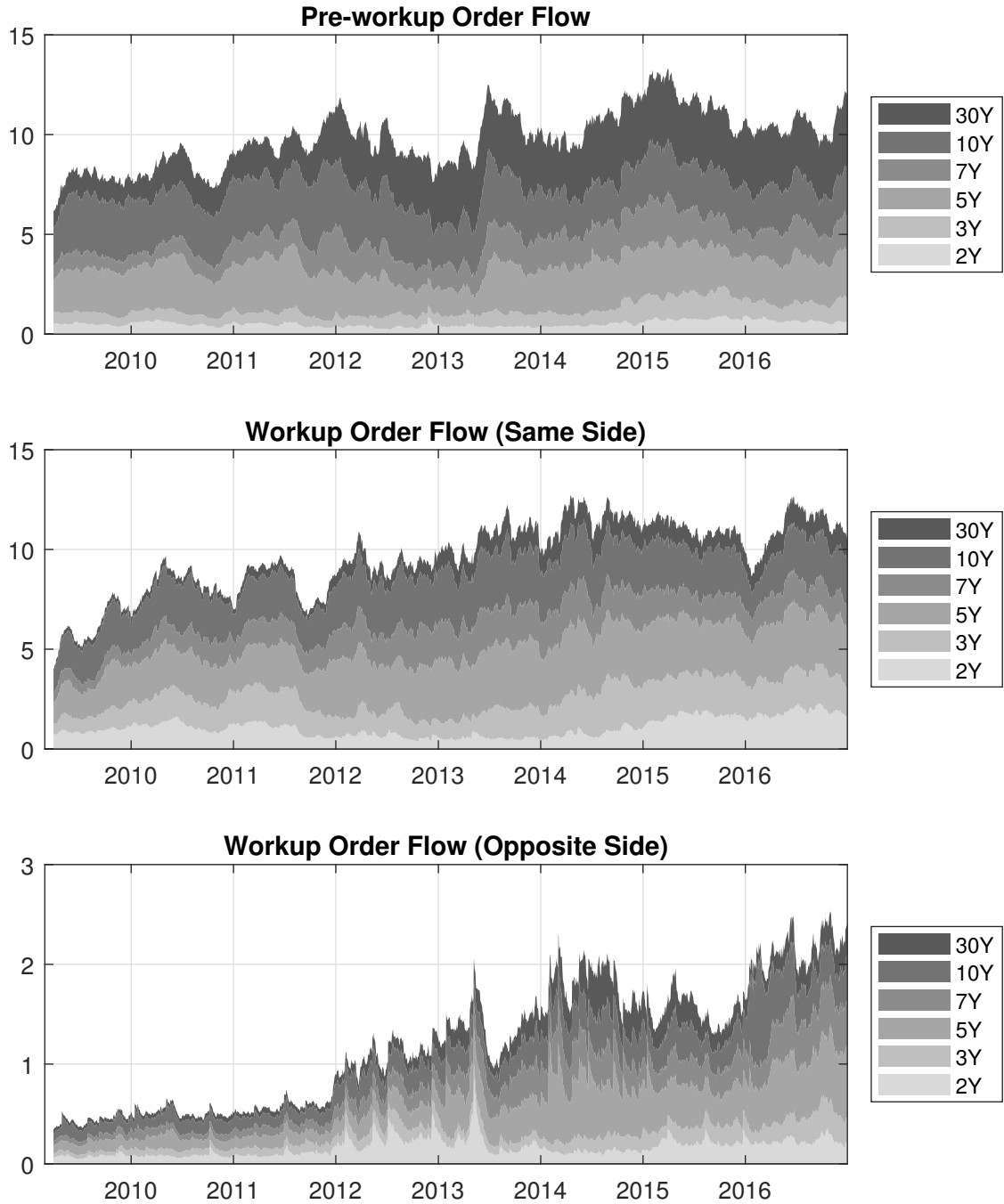
This figure shows the pattern of workup probability over the global trading day (Eastern Time). The plot starts at 18:30 of the previous day and ends at 17:30 of the current day. The numbers are first calculated for each 30-minute interval and day as the percent of transactions in which there is at least one trade during the workup phase, and then averaged across days. The calculations are based on BrokerTec data for on-the-run nominal coupon securities over the 2006-2016 period.



**Figure 4: Information Share of Pre-Workup and Workup Order Flow**

This figure shows the 22-day moving average of the information share of three order flow variables, using [Hasbrouck \(1991\)](#)'s information share approach. The information share measures are computed separately for each day from a VAR(5) model of pre-workup trade flow, return, workup trade flow initiated from the same side, and workup trade flow initiated from the other side, and then averaged over rolling 22-day intervals. Discontinuities on the plots indicate periods of auction discontinuance in the 3- and 7-year notes or days for which the VAR model cannot otherwise be estimated reliably. Estimation is based on BrokerTec data for on-the-run nominal coupon Treasury securities over the 2006-2016 period. Observations outside the [7:00-17:30] time window are excluded.





**Figure 5: Information Shares of Order Flow Variables on Level Factor**

This figure shows the 22-day moving average of the information shares of three order flow variables with respect to the level factor of the Treasury benchmark yield curve, using [Hasbrouck \(1991\)](#)'s information share approach. The information share measures are computed separately for each day from a VAR(5) model of pre-workup trade flow, same-side workup trade flow, other-side workup trade flow, and four yield curve factors extracted from one-second returns of six nominal on-the-run coupon Treasury securities. Estimation is based on BrokerTec data over the period from February 27, 2009 to the end of 2016 (when all six securities are present). Observations outside the [7:00-17:30] time window are excluded.