Journal of Corporate Finance 52 (2018) 143–167

Contents lists available at ScienceDirect

Journal of Corporate Finance

journal homepage: www.elsevier.com/locate/jcorpfin

Institutional trading and Abel Noser data

Gang Hu*, Koren M. Jo, Yi Alex Wang, Jing Xie

Hong Kong Polytechnic University, School of Accounting and Finance, Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Keywords:
Institutional Investor
Trading
Abel Noser
ANcerno
Accounting
Corporate Finance
Investments
Market Microstructure

JEL classification:
G12
G14
G23
M40

ABSTRACT

We survey the growing academic literature using Abel Noser data, including 55 publications thus far. We analyze publication patterns to explore how the availability of a specialized microstructure dataset propagates across different areas within finance and into other disciplines such as accounting. Of note, we identify corporate finance and accounting as the most under-researched areas that offer promising opportunities for future academic research using the data. To provide guidance for researchers interested in using Abel Noser data, we analyze institutional trading using transaction-level data spanning more than 12 years and covering 233 million transactions with $37 trillion traded. We provide background information on the origin and history of the data, offer suggestions for cleaning and using the data, and discuss (dis)advantages of Abel Noser compared to other data sources for institutional trading. We also document two simple facts: 1) institutional trade sizes decline dramatically over time, rendering trade size-based inferences of institutional trades problematic; 2) we estimate that Abel Noser data cover 12% of CRSP volume over our sample period and 15% for 1999–2005, significantly higher than the estimate in Puckett and Yan (2011) for 1999–2005: 8%, a widely quoted number in the literature. This background should prove useful for researchers seeking to address a number of as yet unexplored issues, especially in corporate finance and accounting research.

1. Introduction

Institutional investors play an increasingly important role in all aspects of financial markets. In this paper, we focus on a proprietary dataset of institutional trading transactions: Abel Noser data (also known as Abel/Noser or ANcerno data). Academic research using Abel Noser data dates back to Blume (1993).[1] We survey the growing academic literature, including 47 published papers and eight forthcoming papers so far, using Abel Noser data to address various research questions in market microstructure, corporate finance, investments, and accounting.

We analyze publication patterns and show how the availability of a specialized microstructure dataset propagates across different areas within finance and into other disciplines such as accounting. Based on our analysis of publication patterns, we discuss potentially fruitful directions for future research, and identify corporate finance and accounting as two under-researched areas that offer the most promising opportunities for future research using Abel Noser data. We believe that the trend of finance studies using Abel Noser data will remain strong in the foreseeable future, as evidenced by the ten 2018 or forthcoming publications. Within finance, corporate finance appears to be the most under-researched area given that it is a large field with rich literature and various interesting

* Corresponding author.
E-mail addresses: gang.hu@polyu.edu.hk (G. Hu), koren.jo@polyu.edu.hk (K.M. Jo), alex.yi.wang@connect.polyu.hk (Y.A. Wang), jingxie@polyu.edu.hk (J. Xie).

[1] Table 5 contains the list of publications using Abel Noser data. An up-to-date list of publications can be found at the Abel Noser (ANcerno) Data Page: http://ganghu.org/an

https://doi.org/10.1016/j.jcorpfin.2018.08.005
Received 18 December 2017; Received in revised form 8 August 2018; Accepted 10 August 2018
Available online 16 August 2018
0929-1199/

corporate events and contexts. However, only 11 out of the 55 publications so far are in corporate finance, even less than 15 in microstructure. This may mean that corporate finance, and similarly accounting, could offer the best potential for future research using Abel Noser data.

Moreover, Abel Noser data is well-suited for studying corporate finance topics. One of the main advantages of Abel Noser data is its high frequency. This feature makes Abel Noser an ideal source for in-depth examinations of institutional investors' trading behavior around various corporate events. An example of this line of academic inquiries is Bethel et al. (2009), one of the earliest publications using Abel Noser data, and it is published in the *Journal of Corporate Finance*. The authors examine institutional trading around mergers and shareholder voting outcomes, and find evidence for an active market for voting rights around merger events. In a recent working paper, Li and Schwartz-Ziv (2018) examine how shareholder votes and trades are related in a broader context also using Abel Noser data. Both studies highlight the advantage of using detailed institutional trading records to gain further insights into important corporate finance issues raised in prior studies, such as Gillan and Starks (2000), and Bethel and Gillan (2002).

Several other studies also highlight the fact that Abel Noser data are uniquely well-suited to tackle various corporate finance research topics, compared to alternative data sources such as SEC 13F filings, which are only quarterly. Chemmanur et al. (2009) explicitly identify SEO (seasoned equity offering) allocations using Abel Noser institutional trading records for the first time in the literature. Ahern and Sosyura (2015) study sensationalism in media coverage amid merger rumors. They find that institutional investors are net sellers in the target firm with merge rumors, suggesting that institutions provide liquidity to individuals who buy targets upon merger rumors. Henry and Koski (2017) examine whether skilled institutions indeed exploit positive abnormal ex-dividend returns, and find that institutions concentrate trading around certain ex-dates and are capable of identifying ex-day events with higher trading profit.

To provide guidance for researchers interested in using Abel Noser data, we first provide background information on the origin and history of Abel Noser data. We describe the data and answer important questions such as why institutional investors provide their trading data to Abel Noser, and why Abel Noser pr`ovides institutional trading data to academia. We discuss various data issues and offer suggestions for cleaning and using the data. We then discuss advantages and disadvantages of Abel Noser data compared to other potential data sources for institutional trading: Plexus, SEI, SEC 13F filings, CRSP/Thomson mutual fund holdings, TAQ, CAUD, TORQ, NASDAQ clearing records data, and NASDAQ PostData. The main disadvantage of Abel Noser database is that it covers trades for a subset of institutional investors. On the other hand, Abel Noser data offer clear advantages over the above alternative data sources for institutional trading. The main advantages include high frequency, long and continuous time series, detailed and accurate information on specific transactions such as side (buy versus sell), and type and identities of sample institutions allowing for cross-sectional and institution-specific analyses.

We further analyze institutional investors' trading behavior using Abel Noser data. Our sample contains 232.6 million institutional trading transactions, 1.26 trillion shares and $37.5 trillion traded from January 1999 to September 2011. Abel Noser provides institutional trading data from 1997 to 2015. However, data for the first two years are very small and a few variables were removed after September 2011, including a key institution identifier. Section 3.1 provides further discussions about sample start and end. Using Abel Noser data, we document two simple facts. First, we document that institutional trade sizes decline dramatically over time, consistent with the findings in Chordia et al. (2011). This significant decline renders trade size-based inferences of institutional trades problematic, supporting the findings in Cready et al. (2014).

Second, we put an upper bound and four progressively tighter lower bounds on Abel Noser data's coverage of CRSP volume. The lowest lower bound follows the same methodology in Puckett and Yan (2011), and we are able to arrive at a similar estimate as that in Puckett and Yan (2011). Our other three lower bounds modify the methodology and offer higher (tighter) estimates. Using our new method, we estimate that Abel Noser data cover 12.3% ~ 12.6% of CRSP volume over the entire sample period from January 1999 to September 2011, and 14.9% ~ 15.3% for 1999–2005. These estimates are significantly higher than the 8% estimate in Puckett and Yan (2011) for 1999–2005, a widely quoted number in the academic literature (see, e.g., Brown et al. (2014), Cready et al. (2014), Ljungqvist and Qian (2016), and Henry and Koski (2017)).

Overall, our paper should prove useful for researchers seeking to address a number of unexplored issues using Abel Noser data, especially in corporate finance and accounting research. The rest of the paper is organized as follows. Section 2 describes Abel Noser data and compares it with various other data sources for institutional trading. Section 3 discusses data issues and presents empirical results. Section 3.1.11 presents an analysis of publication patterns using Abel Noser data. Section 4 contains a brief survey of the growing academic literature using Abel Noser data, including 55 publications so far and selected working papers. Section 5 concludes with a discussion of directions for future research.

2. Data sources for institutional trading

In this section, we first describe Abel Noser institutional trading data. We then discuss the advantages and disadvantages of using Abel Noser data to study institutional trading compared to other potential data sources such as Plexus, SEI, SEC 13F filings, CRSP/Thomson mutual fund holdings, TAQ, CAUD, TORQ, NASDAQ clearing records data, and NASDAQ PostData.

2.1. Abel Noser data

The descriptions and discussions in this paper are based on knowledge accumulated through the first author spending more than two years working inside Fidelity's Global Equity Trading group on similar trading data and directly interacting with Abel Noser as an outside consulting service. It is also based on numerous company visits and conversations with Abel Noser over a period of more than ten years as an academic subscriber of the data later.

2.1.1. Who is Abel Noser? (Abel/Noser, ANcerno, and Abel Noser)

Located in New York City, Abel Noser is a brokerage firm that also provides transaction cost analysis to institutional clients. The firm was co-founded by Stanley Abel and Gene Noser in 1975 with a goal of providing low cost trading to institutional clients. Confusion often exists regarding the terms of Abel/Noser, ANcerno, and Abel Noser. The firm was originally named Abel/Noser Corp. and it provides both brokerage services as a sell-side firm and trading cost measurement services to buy-side institutional investors. Some of their transaction cost analysis clients raised concerns about confidentiality of their trading data and potential misuse of their data by the brokerage side of Abel Noser. In response to this concern and the growth of transaction cost analysis services, the firm set up Abel Noser Holdings LLC as the parent company, and created a wholly-owned subsidiary named ANcerno Ltd., separate from its brokerage business to specialize in transaction cost analysis (AN for Abel Noser, and *cerno*, Latin for examine and discern). However, this name change caused confusion in the marketplace in terms of the relationship between ANcerno and Abel Noser. So ANcerno changed its name again to Abel Noser Solutions, rendering ANcerno defunct. Abel Noser Solutions, just like its predecessor ANcerno, is still wholly-owned by Abel Noser Holdings and separate from the brokerage business. The above is why we refer to the data as Abel Noser data.

2.1.2. Why do institutional investors provide their trading data to Abel Noser?

The answer to this question is twofold, different for the two main types of institutions who provide their trading data to Abel Noser: investment managers and plan sponsors. Investment managers are mutual fund families such as Fidelity Investments, Putnam Investments, and Lazard Asset Management. Examples of plan sponsors include the California Public Employees' Retirement System (CalPERS), the Commonwealth of Virginia, and United Airlines.

Plan sponsors, who fall under the jurisdiction of the Department of Labor (DoL) rather than the Securities and Exchange Commission (SEC), are subject to the Employment Retirement Income Security Act (ERISA), which mandates that plan sponsors need to demonstrate “best execution” for their trading transactions conducted on behalf of plan participants. In order to fulfill this requirement, many plan sponsors subscribe to an external transaction cost analysis service, such as Abel Noser or Plexus. Therefore, one might characterize plan sponsors' presence in Abel Noser data as “semi-voluntary.”

On the other hand, investment managers subscribe to Abel Noser's services on a voluntary basis. For each individual trade, transaction cost savings may appear to be small. On the other hand, total dollar trading costs, especially for large investment managers, can be significant. In addition, since it is well-documented that investment managers' abnormal return performance or alpha on average is minimal or non-existent, potential trading cost savings may become more important relative to investment alpha. So investment managers also subscribe to external transaction cost analysis service providers such as Abel Noser. Some large investment managers, such as Fidelity and State Street, even created their own internal groups dedicated to analyzing and lowering trading costs. For example, Fidelity created a Trading Techniques & Measurement (TT&M) group around 2000, part of its Global Equity Trading, which handles all Fidelity mutual funds' equity trading transactions. At its peak, TT&M had more than ten full-time employees, including several Ph.D.’s and an ex-MIT statistics professor.

2.1.3. Why does Abel Noser provide institutional trading data to academia?

Why is Abel Noser friendly to academia and positively inclined to provide their proprietary data to academic researchers? Perhaps it has something to do with its academically-connected roots. Gene Noser, one of the co-founders and Chairman Emeritus of the firm, is one of the authors of a *Journal of Finance* publication on trading costs: Berkowitz et al. (1988).

Being a consulting service provider, it is possible that Abel Noser values the credibility and publicity brought about by the association with academia. For example, Plexus, one of Abel Noser's main competitors of in the transaction cost analysis service provider space, also used to provide data for academia research. In earlier years, however, the provision of data for academic research was quite ad hoc and the sample periods provided were short and sometimes scattered: a few weeks or months' data sometimes with breaks within the sample periods, for example, the three-month Abel Noser data used in Blume (1993).

When and why did Abel Noser start supplying comprehensive institutional trading data to academic researchers systematically, and for continuous and long sample periods? This is largely due to the efforts by the first author of this paper. During 2000 to 2003, the first author was an employee at Fidelity's TT&M group, who was the largest client of Abel Noser. In fact, Fidelity's TT&M group helped Abel Noser setting up its institutional trading data structure and trading costs measurement framework (“specs”), which were significantly different from their competitor: Plexus'. With the strong support of TT&M group's head, who holds a Ph.D. from MIT himself, the first author successfully convinced Abel Noser to provide “all” their data for his doctoral dissertation, after replacing identities of sample institutions with numeric codes. This happened in 2002, after Abel Noser and the first author signed a confidentiality agreement that year. All three chapters of the first author's dissertation use Abel Noser data. The dissertation, Hu (2005), became public after a successful defense at Boston College in 2005 and was the basis for three eventual publications: Hu et al. (2008), Hu (2009), and Chemmanur et al. (2010).

It is worth noting that there were non-trivial initial costs involved for Abel Noser in providing its trading data for academic research, e.g., extracting, anonymizing, uploading, and updating the data, and providing additional information such as a “data dictionary” and cross-reference tables so that researchers can understand and make good use of the data. Having incurred these significant initial costs mainly at the urging by the first author, Abel Noser opened up and started providing the data to other academic researchers.

Unlike most other data vendors, it is clear that Abel Noser's motivation in this endeavor is not financial – in early years, the data

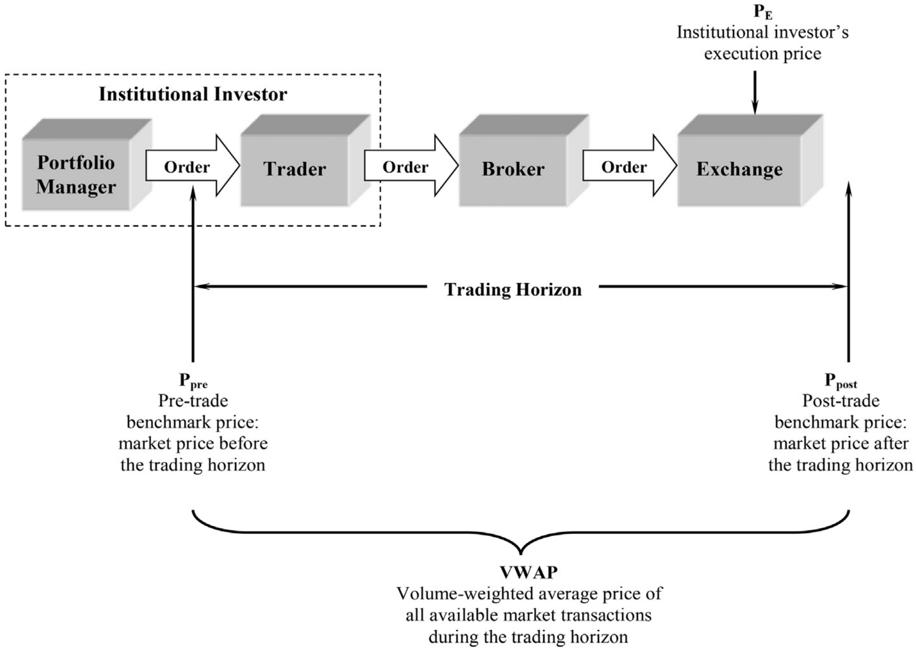


Fig. 1. Typical institutional trading process. This figure depicts a typical institutional trading process. It is a modified reproduction of Fig. 3 in Hu (2009). The portfolio manager and the trader are inside the same buy-side institutional investor. P_{pre}, VWAP, and P_{post} are market prices, whereas P_E is the institutional investor's own execution price.

were provided free of charge when asked by any academic researcher, only conditional on signing a confidentiality agreement. Abel Noser soon realized there is a higher demand than they expected. In order to cover overhead and other costs, they started charging a nominal $500 annual subscription fee which includes all historical data. Later the subscription fee gradually increased over the years, but still only to cover costs involved.

2.1.4. Description of Abel Noser data

For each client trading transaction, the Abel Noser database contains 107 different variables. Numbers of variables available vary slightly over time since Abel Noser added and removed variables in certain years. In addition, there are about ten cross-reference files that contain additional information about the identities of client, manager, broker, and the stock. For brevity, we do not list all 107 variables and these reference files here. We focus our discussions on commonly used variables.

Before we discuss variables provided by Abel Noser, it is useful to elaborate on the trading process to better understand how Abel Noser collects these variables. Fig. 1 depicts a typical institutional trading process. This process starts with the portfolio manager (PM) making investment decisions, which stock to buy or sell, and the quantity to be traded. Then the PM sends the order to a trader inside the same institution. With inputs from the trader, the PM also determines a trading horizon, which is often a trading day. The trading horizon can also be shorter than a trading day, or it can span multiple trading days. Fig. 1 and our description of the investment and trading process is a modified version of Fig. 3 and the accompanying description in Hu (2009), as cited in Henry and Koski (2017). Trading data for investment manager clients are received directly from these clients' Order Delivery System. Abel Noser uses a *lognumber* to identify each batch of trades received from clients. To facilitate trading cost analysis of trades by clients, Abel Noser provides information on various benchmark prices and volumes corresponding to each trade. This is why Abel Noser data have so many columns or variables.

Each client trading execution corresponds to one observation in Abel Noser data, which include three main categories of variables: first, identities of market participants including: *clientcode*, *clienttypecode*, *clientmgrcode*, *clienttdrcode*, and *clientbkrcode*. The *clientcode* is a unique numeric ID for each of Abel Noser's institutional clients. The *clienttypecode* identifies the type of institutional clients: investment managers (*clienttypecode* = 2), plan sponsors (*clienttypecode* = 1), and brokers (*clienttypecode* = 3). Abel Noser data do not contain many broker clients and the data coverage tends to be sporadic. Thus, empirical analysis of institutional trading should normally exclude observations with *clienttypecode* = 3 and missing *clienttypecode*. Other codes are discussed later in Section 3.1. The second category of variables contains specific information for each transaction, including the *symbol*, *CUSIP*, *side*, *price*, *volume*, and *commissionusd*. *Symbol* and *CUSIP* identify the stock traded. *Side*, *price*, *volume*, and *commissionusd* specify whether the trade is a buy or sell, the execution price, the number of shares traded, and dollar commissions paid on the transaction. The third category of variables contains reference and market benchmark information (market price and market volume) for various time horizons and aggregation levels for each transaction.

2.2. Other data sources

In this sub-section, we describe various other potential data sources for institutional trading: Plexus, SEI, SEC 13F filings, CRSP/Thomson mutual fund holdings, TAQ, CAUD, TORQ, NASDAQ clearing records data, and NASDAQ PostData. We also discuss the advantages and disadvantages of using Abel Noser data to study institutional trading compared to these alternative data sources.

2.2.1. Plexus and SEI

Several early microstructure studies use institutional trading data provided by Plexus, for example, Keim and Madhavan (1995, 1997) and Conrad et al. (2001, 2003). Irvine et al. (2007) use Plexus data innovatively to study tipping. Plexus data are perhaps the closest in nature to Abel Noser data. In fact, Abel Noser and Plexus were competitors in the marketplace of providing transaction cost analysis services. As mentioned before, the main advantage of Abel Noser data is the long and continuous time series which allows for analyses of many interesting corporate finance and investment research topics. On the other hand, Plexus data were typically provided on a much more ad hoc basis and for much shorter and scattered sample periods. Importantly, Plexus has long stopped providing its data to academia since its acquisition by Investment Technology Group (ITG) in 2005. Chan and Lakonishok (1993, 1995) use data provided by SEI Corporation, which are similar to Plexus and Abel Noser data as well, but it also stopped providing data to academia a long time ago.

2.2.2. SEC 13F filings

Under the 1934 Securities Act, all institutional investors with more than $100 million in 13(f) securities (mostly publicly traded equity, but also convertible bonds and options) must file SEC Form 13F. The disclosure of portfolio holdings is at the management company level on a quarterly basis, e.g., mutual fund management companies and hedge funds. A fund management company (e.g., Fidelity and Vanguard) managing multiple individual funds will report aggregated holdings by all of these individual funds.

The date when the Form 13F is filed with the SEC is referred to as the “filing date,” and the quarter-end date on which the portfolio is disclosed as the “quarter-end portfolio date.” Normally, the SEC requires that the maximum lag between the two dates should be 45 calendar days. However, subject to approval by SEC, institutions can delay or prevent disclosure of certain holdings, usually up to one year, from the date required for the original 13F filings. Such “hiding” of holdings is disclosed in an amendment to the original 13F filing after a request is denied, or after the confidentiality period expires. For further detailed discussions of hiding holdings by 13F investors, refer to Agarwal et al. (2013) and Aragon et al. (2013).

The advantages of Abel Noser data compared to 13F filings is highlighted in Chemmanur et al. (2009) for institutional trading around corporate events such as seasoned equity offerings (SEOs): with daily trading records from Abel Noser, the authors are able to identity allocations and trading of SEOs by each institution. The advantage of Abel Noser is also highlighted by Puckett and Yan (2011) for investments topics: they find that institutional investors earn significant abnormal returns on their trades within the trading quarter (namely interim trading), which are undetected by 13F data. The main disadvantage of Abel Noser database is that it does not cover all institutional investors in the market, while the 13F database has more complete coverage of institutional investors since these investors are required to report their holdings on a quarterly basis.

2.2.3. CRSP/Thomson mutual fund holdings

The Securities Exchange Act of 1934 and the Investment Company Act of 1940 mandated mutual funds to report their holdings to the SEC of dates (“report dates”). These disclosures must include all of the fund's portfolio positions, including, but not limited to common equities, preferred equities, options, bonds, and short positions. Prior to May 2004, the SEC only required mutual funds to file their portfolio holdings twice a year using the semi-annual N-30D form. Since May 2004, the Investment Company Act of 1940 mandates that individual mutual funds disclose their portfolio holdings quarterly in Forms N-CSR and N-Q with a delay of no longer than 60 days. The date this portfolio information is filed with EDGAR is known as the “file date,” and the difference between the file date and report date is often known as the “reporting delay.” Agarwal et al. (2015) and Schwarz and Potter (2016) provide detailed descriptions of reporting requirements for mutual funds.

The most widely used mutual fund holding data is the Thomson Mutual Fund Holdings (Thomson) database, which contains portfolio holdings since January 1979. This data has been used by majority of existing mutual fund research papers (e.g., Zheng (1999) and Kacperczyk et al. (2008)). Prior to its ownership by Thomson, this database was known as CDA Investment Technologies data (Daniel et al. (1997)), as well as the CDA Spectrum holdings database (Wermers (1999)). It contains many mutual fund portfolios reported as of dates not mandated by the 1940 Investment Company Act. Starting in the second half of 2003, CRSP Mutual Fund Database's (CRSP) portfolios became available. The CRSP Mutual Fund Database's coverage is from January 2001 to present.

The main advantage of Abel Noser data compared to CRSP or Thomson Mutual Fund Holdings is its daily frequency of reporting. With semi-annual or quarterly portfolio disclosures from CRSP or Thomson Mutual Fund Holdings, researchers cannot observe the interim trading between two consecutive report dates. Kacperczyk et al. (2008) conduct a thorough analysis of unobserved actions of mutual funds leading to return gap—the difference between the reported mutual fund return and the return on a portfolio that invests in the previously disclosed fund holdings. The main disadvantage of Abel Noser database is that it does not cover all funds in the market, while CRSP and Thomson Mutual Fund Holdings have better coverage in terms of mutual funds since these funds are required to report their holdings. In addition, Abel Noser's fund-level identifiers and information can be problematic at times, though they have been used in prior academic studies.

2.2.4. TAQ

The Trade and Quote (TAQ) database contains intraday transactions data (trades and quotes) for all securities listed on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), as well as NASDAQ National Market System (NMS) and SmallCap issues.

The main disadvantage of TAQ compared with Abel Noser is that TAQ database does not identify the trader and does not classify direction of trades (buys versus sells). Pioneered by Lee and Ready (1991), research studies analyzing these data apply an algorithm to assign the side of the trade depending on whether the trade price is closer to bid or ask. However, the Lee and Ready (1991) algorithm has become unreliable in more recent years due to sub-penny trading and shallow depths at the quoted bid and ask. On the other hand, Abel Noser data contain accurate information on whether each transaction is a buy or a sell by sample institutions. Thus, no inference algorithm is needed.

When using TAQ, to proxy for investor type, researchers commonly use trades cutoffs by dollar size (Lee and Radhakrishna (2000)), block trades (Kraus and Stoll (1972) and Bozcuk and Lasfer (2005)), or a more sophisticated daily trades and quarterly holdings mapping procedure proposed by Campbell et al. (2009). However, as suggested by Cready et al. (2014), inferring investor type based on trade size is not reliable. However, Abel Noser data contain institutional trades with type and identities of sample institutions. The main advantage of TAQ compared with Abel Noser is that TAQ data contain all market trades, while Able Noser data only cover trading by a subset of institutional investors.

2.2.5. CAUD and TORQ

The NYSE's Consolidated Equity Audit Trail Data (CAUD) contain detailed information on all orders executed on the exchange, both electronic and manual (those handled by floor brokers). This database includes program trading and index arbitrage trading. One of the fields associated with the buyer and seller of each order, Account Type, specifies whether the order comes from an institutional investor. The Account Type designation of individual investor orders has its origins in the aftermath of October 1987 stock market crash. Kaniel et al. (2008), Kaniel et al. (2012), and Hendershott et al. (2015) contain detailed discussions and analyses of CAUD data.

The trades, orders, reports, and quotes (TORQ) database by the New York Stock Exchange (NYSE) contains detailed information on trades, orders, quotes, and audit reports associated with trades for a size-stratified sample of 144 NYSE-traded equity securities from November 1, 1990, to January 31, 1991. There are approximately fifteen firms randomly selected from within each capitalization decile. The database identifies the party initiating the trade (individual versus institution, for example) and the direction of the trade (buy versus sell). This allows the precise identification of these variables without relying on algorithms intended to infer trade direction and trader identity. See Sias and Starks (1997), Kavajecz (1999), Lee and Radhakrishna (2000), and Chakravarty (2001) for further descriptions of the TORQ database.

The main advantages of Abel Noser data compared to CAUD and TORQ is that Abel Noser covers stocks from all stock exchanges in U.S. and contains identities and types of sample institutional investors, allowing for institution-specific empirical analysis. The disadvantage of Abel Noser database is that it does not cover all trades or traders in the market, while CAUD and TORQ have better coverage in terms of traders since these data are directly from the exchange.

2.2.6. NASDAQ clearing records data and NASDAQ PostData

NASDAQ clearing records based trading data consist of trading by brokerage houses in all NASDAQ-listed firms from January 2, 1997, to December 31, 2002. The data include the date, time, ticker symbol, trade size, and price of each transaction for each stock. These clearing records also include market maker IDs from the settlement process allowing trading volume to be assigned to investment banks. The data can identify the direction of each trade (buy or sell). The data also contain separate principal/agent flags to identify whether the parties are trading for their own account or for a client. The data include both trades reported "to the tape" (tape report) and unreported NASDAQ clearing records (nontape report). The main advantage of Abel Noser data compared to NASDAQ clearing records based trading data is that researchers can separately track trading records of each sample institution in Abel Noser, but not in the latter. The main disadvantage of Abel Noser is its coverage of only a subset of institutional investors and hence market dollar trading volume. According to Griffin et al. (2012) who use NASDAQ clearing records data, the data cover 77.8% of NASDAQ trading volume. Please see Section 3.1.8 below for our results on Abel Noser data's coverage of CRSP trading volume.

NASDAQ PostData contain trade date, ticker symbol, market maker ID and type (including a code for ECNs), number of shares traded, dollar volume of trade, number of trades, and average trade size for all NASDAQ market makers. Each trade is divided into buy, sell, and crossed trades. PostData are attributed to particular market makers. Juergens and Lindsey (2009) use 136 days of the data for 3874 NASDAQ-traded stocks for all quoting market makers. Since NASDAQ PostData is about trading by market makers, it is quite different from the set of investors covered by Abel Noser data, which are mainly either investment managers or plan sponsors.

To summarize, Abel Noser data, though only contain trades for a subset of institutional investors, offer clear advantages over the above long list of alternative data sources. The main advantages include high frequency (at least daily), long and continuous time series, detailed and accurate information on specific transactions such as side (buy versus sell), and type and identities of sample institutions allowing for cross-sectional and institution-specific analyses.

3. Data and results

In this section, we discuss various Abel Noser data issues (Section 3.1) and present results of our empirical analysis on institutional trade sizes and Abel Noser data's coverage of CRSP (Section 3.1.8).

3.1. Data issues

3.1.1. Sample start and end

As mentioned earlier, Abel Noser provides institutional trading data from 1997 to 2015. However, data for the first two years are very small compared to later years and there might be data quality issues as mentioned by Abel Noser during those two years. So we start our sample period in January 1999. Abel Noser provides data on a quarterly basis with a time lag around one year. Sometime in 2017, Abel Noser completely stopped providing data for academic research. So the data ended around 2015. More importantly, Abel Noser removed *clientcode* in the data after September 2011 along with several other variables. Initially, Abel Noser also removed *clienttypecode*. After pleading directly to upper management by the first author of this paper, Abel Noser agreed to add back *clienttypecode*, but not *clientcode*, for data after September 2011.

Clientcode is a key variable that separately identifies trades from different institutions. As discussed earlier, one of the main advantages of Abel Noser data is the ability to separately identify and track trades for different institutions allowing for cross-sectional and institution-specific analyses. Therefore, Abel Noser data after September 2011 are much less useful for academic researchers, especially for most corporate finance, investments, and accounting topics, where separately tracking trades for different institutions enables much more thorough and interesting investigations of relevant research questions. Abel Noser data after September 2011 may still be useful for certain market microstructure and investments studies that only aim to generally examine institutional trades, for example, a study of institutional trading costs that does not rely on knowing something about the trader's characteristics. Though still useful, Abel Noser data after September 2011 are less unique compared to alternative sources for institutional trading such as NYSE CAUD discussed earlier. Table 5 lists sample periods for all 55 publications to date using Abel Noser data. Not surprisingly, sample periods for almost all publications, including forthcoming papers, fall within the sample period we use: January 1999 to September 2011 (12 years and 9 months).

3.1.2. Potential sample biases

Since Abel Noser provides trading cost analysis for clients, presumably, these institutions care more about their trading execution costs in general. Therefore, sample selection biases, if any, should mostly concern microstructure studies. On the other hand, for research questions in corporate finance, investments, and accounting, we cannot think of obvious sample selection biases. Appendices of Puckett and Yan (2011), Anand et al. (2012), and Jame (2018) contain discussions of issues related to potential survivorship and selection biases of Abel Noser data. These papers show that Abel Noser institutions on average do not differ from 13F institutions in stock holdings, return characteristics, and stock trades, although they are larger in size.

3.1.3. Matching with CRSP

Abel Noser data contain three variables that identify each stock: *symbol*, *CUSIP*, and *stockkey*. Both *symbol* and *CUSIP* are as provided by Abel Noser institutional clients. Complicating things, different institutions may have different variations for the same stock's *symbol* and *CUSIP*, e.g., the number of characters in CUSIP may vary and *symbol* may be different for the same stock for different institutions. *Stockkey* is Abel Noser's own stock identifier. However, we have observed that *stockkey* may contain matching errors and may in fact be reused over time for different stocks, though we have observed that the quality of *stockkey* has improved over the years. In order to match with CRSP PERMNO, we therefore recommend first clean up *symbol* and *CUSIP* provided by Abel Noser, and then use them to match with CRSP. This matching process is non-trivial and may involve partial manual matching to ensure matching quality.

3.1.4. Data repetition and lognumber

When Abel Noser receives a batch of trading data from a client, typically monthly, it assigns a unique *lognumber* for each batch of data received. Sometimes clients may send "corrections" for earlier batches of trading data. In other words, Abel Noser data may contain repetitions of trading data for the same client institution over the same time period. To counter this problem, researchers using Abel Noser data will need to remove these data repetitions by making use of the *lognumber*, as it is unique for each batch of client trading data received by Abel Noser.

3.1.5. Intraday time stamps

Abel Noser data contain several intraday time stamps marking the times at different points in the trading process as depicted in Fig. 1: order entry (when the portfolio manager enters the order), order placement (when the institutional trader places the order with an external broker), and order execution (when the trade order is executed). When institutional clients do not provide these time stamps, Abel Noser sets entry and placement times to the market open (9:30 am), and execution times to the market close (4:00 pm). Most academic studies using Abel Noser data so far do not make use of intraday time stamps. One reason is that these time stamps default to market open or close quite frequently, especially during earlier years. For example, Choi et al. (2017) state that "a significant fraction of trades are recorded as occurring at the opening of trading" (footnote 6). Another possible complication that may add further noise in intraday time stamps is the splitting and packaging of trade orders. However, it does not mean that intraday time stamps could not be useful just because they might be noisy.

In a recent paper, Huang et al. (2018a) make an important contribution by carefully analyzing the patterns of intraday time stamps in Abel Noser data. The authors present convincing evidence that these intraday time stamps, though not without limitations and could be noisy, are likely to be authentic and valid. Huang et al. (2018a) first show that there are indeed a large percentage of placement times at the market open and execution times at the market close, though these percentages decline over the sample period significantly. They report that for placement times, the percentage starts at over 80% at the beginning of the sample, but declines to about 40% at the end, and the full sample percentage of trades placed at the market open is 52%. This is consistent with Abel Noser's institutional clients providing intraday time stamps to a greater degree over time. Importantly, Huang et al. (2018a) find that when placement times are not at the market open, the pattern of placement times during the trading day appears to be very realistic and reasonable: high trading volume after the market open, smooth trading volume during the trading day except for small hourly spikes, and increased trading volume as the market close approaches. Their findings suggest that though intraday time stamps in Abel Noser data are frequently at the market open or close, when they are not at the market open or close, they are likely to contain useful information.[2]

3.1.6. Identities of sample institutions

This is an important data issue, since if armed with identities of sample institutions, researchers can then match Abel Noser data with other publicly available data sources such as SEC 13F filing to gain many more variables and additional characteristics of sample institutions. The original data were anonymous, that is, though there is a *clientcode* that separately identifies each sample institution, no institution identities information were provided. However, by cumulating trading by Abel Noser sample institutions for each stock and comparing trading patterns with quarterly portfolio holdings changes from public 13F filings, one may be able to "reverse engineer" and uncover identities of Abel Noser institutions. Hu et al. (2009), an early working paper version of Hu et al. (2018), was the first to implement such a non-trivial algorithm, which was later used in Chemmanur et al. (2010). The Internet Appendix of Hu et al. (2018) describes this matching algorithm in detail. Note that these papers only use the inferred identify information to classify institutions (e.g., transient institutions) without revealing any information on any specific sample institution.

On the other hand, sometime during 2010–2011, Abel Noser data included a *MasterManagerXref* file containing cross-reference identity information for sample institutions. This file was only available during this period and not for data subscribers before or after. The above time period refers to trading data period, and since Abel Noser provides trading data to academic research with a nine to twelve months' time lag, one needs to be a subscriber of Abel Noser data sometime during 2011–2012 to have gained access to this crucial piece of information. With this cross-reference file, the reverse engineering algorithm implemented in Hu et al. (2009, 2018) is no longer necessary, and researchers can match Abel Noser daily trading data with 13F quarterly holdings data, see, e.g., Choi et al. (2017).

However, the revelation of Abel Noser institution identities may indeed be one of the main reasons why Abel Noser removed *clientcode* in the data after September 2011. As discussed previously, since the motivation for Abel Noser to provide data for academic research is not financial gains to begin with, the cost for the firm to stop providing data to academia is small and intangible. In 2017, Abel Noser completely stopped providing data for academic research. As a result, for new researchers to gain access to Abel Noser data, one solution might be to work with researchers who have used the data before. Table 5 contains a list of authors who have published academic studies using Abel Noser data so far.

3.1.7. Clientmgrcode, Clienttdrcode, and Clientbkrcode

For investment managers, *clientmgrcode* identifies funds, fund managers, or separately managed accounts. However, it may change over time for the same fund and is not very reliable, though it has been used in prior studies. For plan sponsors, *clientmgrcode* identifies the investment manager or fund managing plan sponsor clients' assets. Jame (2018) provides a good example of usage of this information to identify hedge fund trades in Abel Noser data. The *clienttdrcode* identifies the trader inside the institutional investor handling the trading transaction (see Fig. 1), and it may be useful for certain microstructure studies. The *clientbkrcode* allows researchers to identify the broker who executed the trading transaction. Anand et al. (2012) and Chemmanur et al. (2015) are good examples that make use of broker information in Abel Noser data.

3.2. Results

Table 1 presents summary statistics of the Abel Noser sample. Following Puckett and Yan (2011), after using *CUSIP*, *symbol*, and *stockkey* provided by Abel Noser to match with CRSP PERMNO, we only keep stocks with CRSP SHRCD equal to either 10 or 11, i.e., U.S. ordinary common shares. We also only keep transactions by either investment managers (*clienttypecode* = 2) or plan sponsors (*clienttypecode* = 1), and remove transactions by brokers (*clienttypecode* = 3) and transactions with missing *clienttypecode*. Not surprisingly, there are more plan sponsors (740) than investment managers (399). As discussed earlier, plan sponsors are the core Abel Noser clients in terms of numbers, and hence there are more plan sponsors than investment managers in the sample. The number of plan sponsors drops fairly steadily over the sample period, with a peak of 344 in 2002 and a low of 138 in 2011. The number of sample investment managers first increases pretty steadily and significantly from 37 in 1999 to a peak of 157 in 2006 and 2007, and then drops steadily but relatively less significantly to 121 in 2011. We do not know the exact reasons for such fluctuations in the number of Abel Noser institutional clients. Our conjecture is that decimalization, and especially the increasing prevalence of algorithm trading where institutions split trades into tiny pieces and use automated programs for trading execution may have resulted in lesser value-added by an external consultant such as Abel Noser. If institutions break up all or most of their trades into very small pieces and spread them throughout the day, it would be very hard for an outside consultant to offer any additional help in improving trading execution strategy and lower trading costs. It also becomes less important for institutions to rely on an outside consultant to

[2] We thank the referee for deepening our understanding of the reliability and usefulness of intraday time stamps in Abel Noser data.

Table 1
Summary statistics.

Year	Investment managers	Plan sponsors	# of trades (millions)	Shares traded (billions)	$ Traded ($billion)	Commissions ($million)	Shares per trade	$ per trade
1999	37	342	5.1	46.2	$2060	$1589	8985	$400,401
2000	43	328	6.5	62.6	$2762	$1941	9671	$426,574
2001	64	334	8.1	89.0	$2699	$2555	10,964	$332,446
2002	81	344	10.5	112.2	$2708	$4488	10,700	$258,287
2003	84	316	10.4	94.3	$2350	$3707	9052	$225,607
2004	117	285	12.7	91.5	$2615	$3444	7199	$205,662
2005	132	244	16.8	98.8	$3059	$3132	5869	$181,774
2006	157	242	28.4	120.0	$3860	$2990	4231	$136,098
2007	157	220	36.0	117.8	$4159	$2647	3269	$115,455
2008	151	182	30.3	138.4	$3960	$2898	4576	$130,915
2009	143	173	24.5	131.8	$2839	$2695	5389	$116,054
2010	139	168	24.9	98.8	$2619	$2140	3971	$105,314
1–9/2011	121	138	18.4	57.7	$1776	$1273	3136	$96,495
Overall	399	740	232.6	1259.2	$37,467	$35,499	6693	$210,083

This table reports summary statistics of Abel Noser data from January 1999 to September 2011. Overall numbers for investment managers and plan sponsors are the total unique numbers of institutions over the whole sample. Following Puckett and Yan (2011), we match Abel Noser data with CRSP daily stock file and only keep those stocks with SHRCD equal to either 10 or 11 (i.e., U.S. ordinary common shares). The dollar value of each trade equals the product of the transaction price and the number of shares bought or sold by the institution. Commissions and shares traded are directly from Abel Noser data and summed across trades executed within each calendar year. The last two columns report the average of shares (dollars) per trade.

measure their trading costs. We further discuss the prevalence of algorithm trading and the resulting dramatic decline in institutional trade size later in this section.

After applying filters mentioned above, Abel Noser data contain 232.6 million transactions, with 1.3 trillion shares and $37.5 trillion traded over the period from January 1999 to September 2011. Sample institutions collectively paid $35.5 billion in brokerage commissions. This translates into 2.8 cents per share and 9.5 basis points per dollar traded. These estimates are in line with prior studies on institutional trading. For example, both using Abel Noser data, Hu (2009) estimates institutional brokerage commissions to be 11 to 12 basis points per dollar traded (Table 2 on page 430) for an earlier and shorter sample period, and Anand et al. (2012) estimate commissions paid by institutions to be also exactly 2.8 cents per share for their sample period of 1999 to 2008 (Table 1 on page 565). These per share or per dollar commission estimates could be biased downwards because brokers may act as market makers thereby profiting from bid-ask spreads and charge zero commission for certain NASDAQ trades. On the other hand, Anand et al. (2012) note that the reduction in spreads that accompanied decimalization in 2001 made the NASDAQ zero commission business model untenable, and institutions began paying commissions on NASDAQ trades (footnote 8 on page 566).

Table 2 reports statistics of Abel Noser data separately for investment managers (Panel A) and plan sponsors (Panel B). Though there are more plan sponsors than investment managers, the quantity and the size of investment managers' trades tend to be larger. Specifically, investment managers account for 191.1 million transactions, with 1.1 trillion shares ($31.5 trillion) traded, and $30.2 billion in brokerage commissions paid; whereas plan sponsors only account for 41.4 million transactions, with 0.2 trillion shares ($6.0 trillion) traded, and $5.3 billion paid for brokerage commissions.

3.2.1. Institutional trade sizes

Fig. 2 plots institutional trade size over time, measured as average shares per trade and average $ per trade, for all sample institutions (Panel A), investment managers (Panel B), and plan sponsors (Panel C), respectively. The data plotted in Fig. 2 are presented in the last two columns in Tables 1 and 2. Across all three panels in Fig. 2, there is a clear declining trend of average institutional trade size over time. For instance, looking at $ per trade in Panel A, average trade size declined from $400 thousand in 1999 to $96 thousand in 2011, a 75% reduction. For investment managers, average trade size declined from $493 thousand in 1999 to $113 thousand in 2011. Similarly for plan sponsors, average trade size dropped from $214 thousand in 1999 to $57 thousand in 2011.

Looking at average trade size alone, though informative, may not tell the whole story, as the distribution of trade size could be highly skewed, with a small fraction of large trades contributing disproportionally to the sample mean. Thus, we further examine percentiles of institutional trade size in Table 3. Some rather striking patterns emerge. First, the distribution of trade size appears to be highly skewed, with average trade size being much higher than median trade sizes. For example, the median $ traded for all institutions over the whole sample is $9222 (Table 3 Panel A), while the corresponding average is $210,083 per trade (Table 1), more than 23 times larger. In fact, the average is much higher than the 75th percentile – $52,602 (Table 3 Panel A).

From 1999 to 2011, the median dollars (shares) traded declined from $58,900 (1600 shares) to $4708 (135 shares) for all institutions, a 92.0% (91.6%) drop, and the median declined from $60,250 (1600 shares) to $9447 (295 shares) for investment managers, a 84.3% (81.6%) drop. The decline is much more dramatic for plan sponsors: from $56,700 (1700 shares) to $194 (5 shares), a 99.7% (99.7%) drop. One possibility is that, since plan sponsors' assets are typically managed by external investment managers, who trade on their behalf, it is possible that these trades were in fact fractions of larger trades executed by external

Table 2
Investment managers versus plan sponsors.

Year	# of trades (millions)	Shares traded (billions)	$ Traded ($billion)	Commissions ($million)	Shares per trade	$ per trade
Panel A. Investment managers						
1999	3.4	36.9	$1694	$1276	10,741	$493,146
2000	4.5	51.6	$2321	$1611	11,462	$515,175
2001	5.7	67.2	$2055	$2084	11,883	$363,441
2002	7.5	89.7	$2158	$3600	11,894	$286,268
2003	8.1	79.5	$1991	$3173	9819	$245,925
2004	10.6	78.5	$2247	$3015	7373	$211,125
2005	15.1	86.5	$2686	$2815	5744	$178,373
2006	26.1	104.6	$3365	$2602	4013	$129,107
2007	32.5	96.7	$3389	$2260	2976	$104,359
2008	26.3	120.6	$3441	$2528	4586	$130,876
2009	20.3	115.5	$2472	$2363	5682	$121,653
2010	17.9	83.5	$2193	$1800	4650	$122,190
1–9/2011	13.1	47.9	$1472	$1069	3667	$112,563
Overall	191.1	1058.6	$31,484	$30,196	7268	$231,861
Panel B. Plan sponsors						
1999	1.7	9.3	$367	$313	5462	$214,291
2000	2.0	11.0	$441	$330	5577	$223,953
2001	2.5	21.8	$644	$471	8855	$261,282
2002	2.9	22.5	$550	$887	7644	$186,650
2003	2.3	14.8	$359	$534	6378	$154,755
2004	2.1	13.1	$368	$429	6304	$177,598
2005	1.8	12.3	$373	$317	6933	$210,686
2006	2.3	15.4	$496	$388	6693	$215,216
2007	3.5	21.1	$769	$387	5961	$217,276
2008	4.0	17.8	$519	$369	4505	$131,173
2009	4.1	16.3	$367	$332	3951	$88,584
2010	6.9	15.3	$426	$339	2209	$61,551
1–9/2011	5.3	9.8	$304	$204	1833	$57,059
Overall	41.4	200.5	$5982	$5302	5562	$169,237

This table reports statistics of transactions in Abel Noser data separately for investment managers (Panel A) and plan sponsors (Panel B) from January 1999 to September 2011.

investment managers on behalf of plan sponsors. However, this cannot fully explain the dramatic declining trend over time for plan sponsors and for the whole sample. Fig. 3 plots percentiles of $ principal traded over time for all sample institutions (Panel A), investment managers (Panel B), and plan sponsors (Panel C), respectively.[3]

One plausible explanation for this dramatic decline in institutional trade size is perhaps the increasing prevalence of algorithm trading as postulated by Chordia et al. (2011): "Algorithmic trading was nonexistent in the early 1990s but was expected to represent about half of the trading volume in 2010" (footnote 49 on page 261). In general, the declining trend of institutional trade size documented here is consistent with and provides direct support for the findings in Chordia et al. (2011), who find that more frequent smaller trades have progressively formed a larger fraction of trading volume over time, and that turnover has increased the most for stocks with the greatest level of institutional holdings.

This significant decline in institutional trade size renders size-based inferences of institutional trades problematic. Pioneered by Cready (1988) and Lee (1992), transaction size-based techniques are widely applied in the academic literature: Cready et al. (2014) identify that "over 30 published papers employing transaction size-based techniques with 10 of them appearing in year 2010 or later (footnote 1 on page 878)." Less than $10,000 is commonly used as a cutoff for small trades done by small (retail) investors. Our evidences show that this cutoff is no longer reliable due to the dramatic decline in institutional trade size over the years. In fact, the median $ traded for our sample institutions is $9222 over the entire sample period, which means more than half of the institutional trades would have been misclassified if one were to apply a $10,000 cutoff. This $10,000 cutoff may be less problematic in earlier years. For example, in 1999, the 25 percentile for all Abel Noser sample trades was $16,929. However, the next twelve years saw a dramatic decline and the 25 percentile in 2011 was only $482. Overall, our findings on institutional trade sizes are consistent and supportive of the findings in Cready et al. (2014).

3.2.2. Abel Noser coverage of CRSP

Abel Noser data's coverage of CRSP volume is an important statistic that authors and referees care about. Puckett and Yan (2011) state that: "On average, this trading activity accounts for approximately <u>8%</u> of the dollar value of CRSP trading volume during the 1999 to 2005 sample period. Assuming that institutional investors, in aggregate, are responsible for 80% of CRSP trading volume, we estimate that ANcerno institutions account for <u>10%</u> of all institutional trading volume" (2nd paragraph on page 606, underlines

[3] The pattern for shares trade is similar, and the figure for shares traded is omitted for brevity.

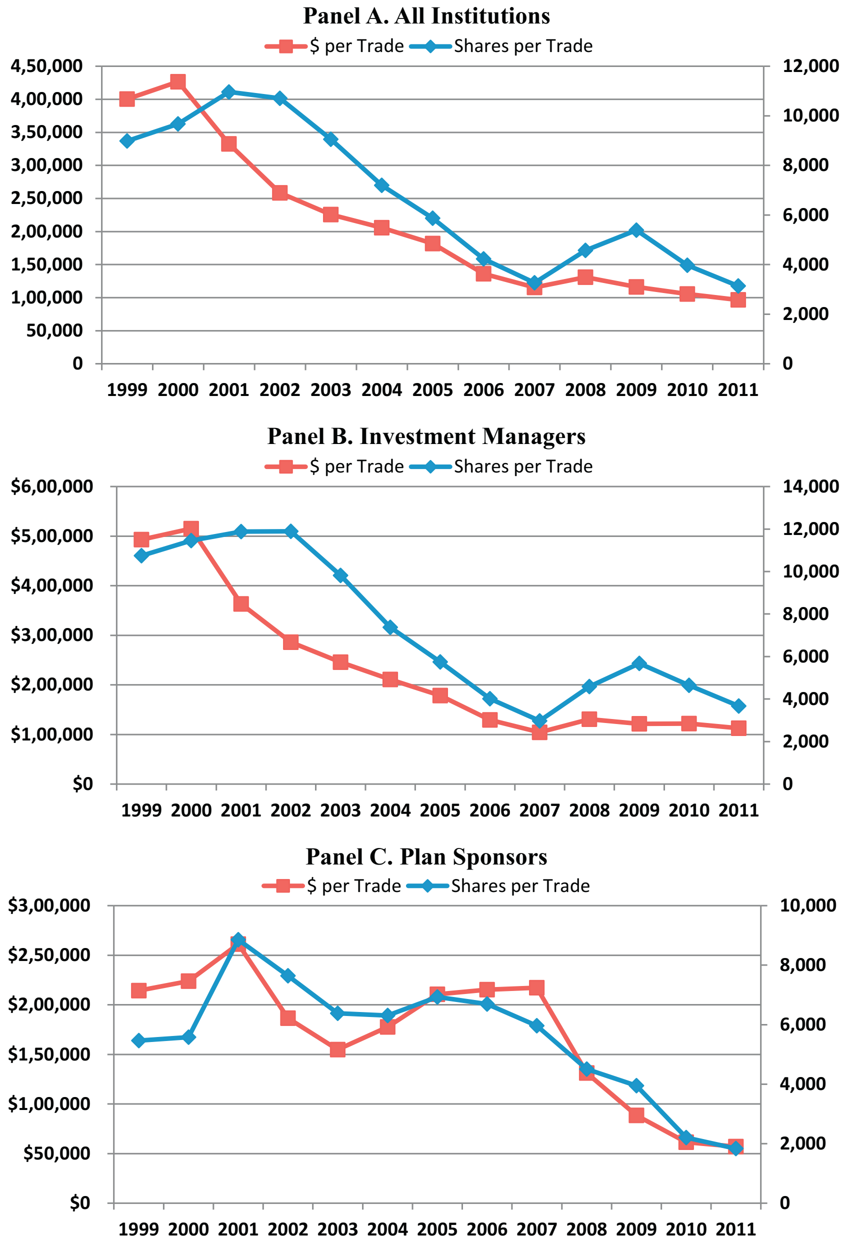


Fig. 2. Average institutional trade sizes over time. This figure plots institutional trade sizes over time, measured as average shares per trade and average $ per trade. Panel A plots statistics for all institutions, while Panels B and C plot results for investment managers and plan sponsors, respectively.

added). This *8%* and (or) the *10%* numbers are widely quoted in subsequent academic literature, such as, Brown et al. (2014), Cready et al. (2014), Ljungqvist and Qian (2016), and Henry and Koski (2017).

Puckett and Yan (2011) also state that: "We believe this estimate represents an approximate lower bound for the size of the ANcerno database" (footnote 9 on page 606). However, when later literature quotes those numbers, the fact that they are *lower bounds* is typically ignored and the numbers are presented as estimates of the size of Abel Noser's coverage. For example, Brown et al. (2014) state that: "According to Puckett and Yan (2011), Ancerno's institutional clients account for 10% of all institutional trading volume. Therefore, the Ancerno data set represents a significant subset of institutional trading and is used by several studies to analyze the trading behavior of institutional investors (page 11);" Cready et al. (2014) state that: "We investigate the reliability of transaction size-based inferences about trader behaviors using a detailed database on institutional transactions from Ancerno Ltd. These investors are all pension or mutual funds and Puckett and Yan (2011) conclude that Ancerno trading accounts for around 10% of institutional trading activity (page 878);" and Ljungqvist and Qian (2016) state that: "Goldstein et al. (2011) and Puckett and Yan

Table 3
Percentiles of institutional trade sizes over time.

	All institutions			Investment managers			Plan sponsors		
Percentile	25%ile	Median	75%ile	25%ile	Median	75%ile	25%ile	Median	75%ile
Panel A. $ Principal traded									
1999	$16,929	$58,900	$216,425	$16,139	$60,250	$247,664	$18,468	$56,700	$174,150
2000	$15,750	$54,200	$209,250	$15,192	$53,630	$230,332	$17,063	$55,362	$175,875
2001	$10,650	$38,363	$148,700	$9910	$37,011	$154,770	$12,494	$41,222	$138,000
2002	$8469	$29,988	$116,469	$8135	$29,127	$117,300	$9380	$32,553	$114,752
2003	$7440	$27,462	$110,279	$7607	$26,851	$110,415	$6700	$29,810	$109,869
2004	$5888	$21,066	$88,738	$5304	$18,740	$80,505	$10,898	$36,788	$127,523
2005	$3051	$12,909	$65,371	$2716	$11,027	$56,794	$11,933	$40,352	$139,089
2006	$1606	$6559	$38,387	$1474	$5649	$31,978	$8916	$35,842	$132,509
2007	$1148	$4160	$24,270	$1143	$3891	$21,144	$1242	$9900	$69,067
2008	$1179	$5747	$34,964	$1427	$6178	$34,969	$282	$2133	$34,927
2009	$1130	$6087	$35,969	$1590	$7074	$38,568	$138	$1262	$23,885
2010	$515	$4470	$30,362	$1692	$7963	$41,254	$80	$257	$3892
1–9/2011	$482	$4708	$31,886	$1928	$9447	$45,080	$80	$194	$2128
Overall	$1710	$9222	$52,602	$1999	$9184	$50,061	$339	$9510	$65,039
Panel B. Shares traded									
1999	500	1600	5600	450	1600	6300	600	1700	4900
2000	500	1500	5500	400	1500	6000	500	1600	4800
2001	400	1400	5200	400	1300	5500	500	1500	4900
2002	400	1300	5000	350	1200	5100	400	1400	4700
2003	300	1055	4400	300	1000	4400	300	1300	4600
2004	200	700	3100	200	600	2800	400	1400	4600
2005	100	400	2100	95	330	1800	400	1400	4670
2006	40	200	1200	36	170	1000	300	1100	4100
2007	25	100	700	25	100	600	46	300	2000
2008	35	200	1271	46	205	1269	9	97	1300
2009	50	275	1600	75	300	1715	4	50	1001
2010	15	150	1092	57	300	1500	2	6	109
1–9/2011	12	135	980	55	295	1394	2	5	68
Overall	50	300	1800	64	300	1700	10	320	2250

This table reports the 25th, 50th, and 75th percentiles of institutional trade sizes over time separately based on dollars traded (Panel A) and shares traded (Panel B) from January 1999 to September 2011. We also report percentiles based on trades by investment managers or plan sponsors separately.

(2011) report that ANcerno institutions account for around 8% of CRSP trading volume and 10% of institutional trading volume (footnote 23 on page 2006)."

In order to arrive at the 8% of CRSP dollar volume and 10% of institutional trading volume estimates, Puckett and Yan (2011) "calculate the ratio of ANcerno trading volume to CRSP trading volume during each day of the sample period. (They) include only stocks with sharecode equal to 10 or 11 in (their) calculation. In addition, (they) divide all ANcerno trading volume by two, since each individual ANcerno client constitutes only one side of a trade" (footnote 9 on page 606, "we" and "our" replaced by "they" and "their", and underlines added). We put an upper bound and four progressively tighter lower bounds on Abel Noser data's coverage of CRSP volume. The lowest lower bound, *% of CRSP_L1*, follows the same methodology in Puckett and Yan (2011), i.e., Abel Noser's total volume (buy + sell) divided by 2. The implicit assumption underlying the lowest lower bound, *% of CRSP_L1*, might be that since Abel Noser data include both institutional buy and sell transactions, the same market transaction might be double counted if both the buyer and the seller for the same transaction are present in Abel Noser data, similar to the logic behind dividing NASDAQ trading volume by two due to its special market maker structure. However, for 2004 and later years, Gao and Ritter (2010) "use a divisor of 1.0, reflecting the fact that there are no longer important differences in the reporting of Nasdaq and NYSE volume." (Appendix B on page 51).

Our other three lower bounds modify the methodology and offer higher (tighter) estimates. *% of CRSP_L2* takes the maximum of buy and sell volume for each stock on a given trading day: Max(buy volume, sell volume), as in Bethel et al., 2009. This is still a very strict lower bound as it implicitly assumes whenever there are double-sided trades in Abel Noser data for the same stock on the same trading day, we recognize them as trading between Abel Noser clients.

For *% of CRSP_L3*, we first round transaction prices to the nearest penny, and then further impose the condition that, in order to be recognized as trades between Abel Noser clients, the rounded transaction price should be equal between buy and sell trades. To implement this method, we aggregate all Abel Noser transactions to PERMNO/tradedate/side/rounded price level, and then take the maximum of buy and sell volume for each unique aggregation of PERMNO/tradedate/rounded price. We round prices to the nearest penny to avoid potential matching errors when Abel Noser transaction prices have many decimal digits (sometimes up to more than 10).

Next, our *% of CRSP_L4* measure is similar to *% of CRSP_L3* except that we do not round prices to the nearest penny, and directly

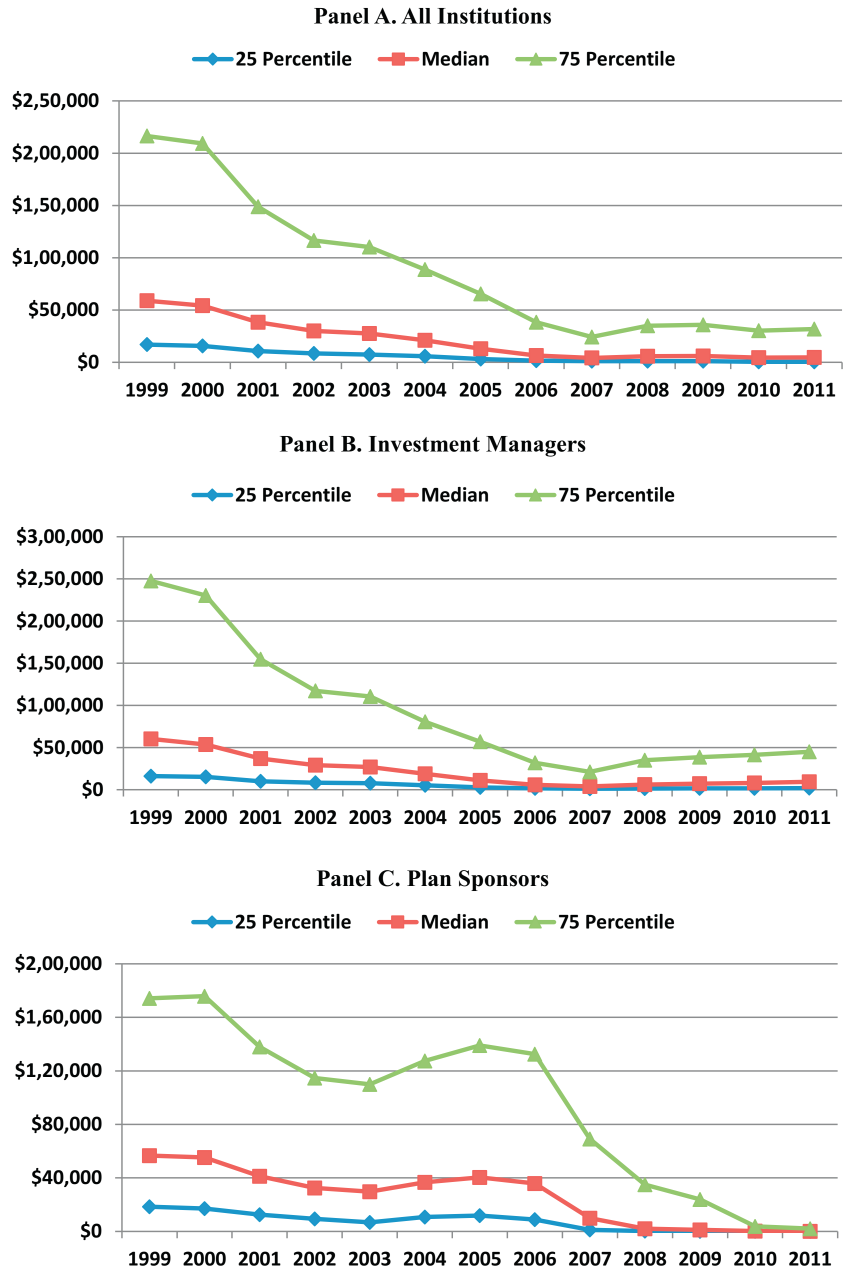


Fig. 3. Percentiles of $ principal traded over time. This figure plots percentiles of $ principal traded over time for all sample institutions (Panel A), and separately for investment managers (Panel B) and plan sponsors (Panel C).

impose the condition that, in order to be recognized as trades between Abel Noser clients, the transaction price should be equal between buy and sell trades. Similarly, in order to implement this method, we sort and aggregate all Abel Noser transactions to PERMNO/tradedate/side/price level, and then take the maximum of buy and sell volume for each unique PERMNO/tradedate/price. Nevertheless, *% of CRSP_L4* is still a lower bound, because even for buy and sell trades in Abel Noser data that have the same price for the same stock on the same day, these transactions could still have been executed with counterparties outside of Abel Noser database, as it accounts for only a fraction of total market transactions.

Finally, *% of CRSP_U* is the upper bound and is calculated as Abel Noser's buy plus sell volume scaled by the corresponding daily volume recorded in the CRSP database (all abovementioned "volume" is the product of the shares traded per transaction and the daily closing price matched from the CRSP daily stock file, in order to eliminate the effect of extreme transaction prices recorded in Abel Noser). For all four lower bounds and the upper bound, the aggregation across different stocks and trading days are done following Puckett and Yan (2011): first we aggregate across stocks and calculate the ratio of Abel Noser-based trading volume (using different

Table 4
Abel Noser data's coverage of CRSP volume.

Year	CRSP # of PERMNOs	Abel Noser # of PERMNOs	# of Trading Days	% of CRSP_U	% of CRSP_L4	% of CRSP_L3	% of CRSP_L2	% of CRSP_L1 (Puckett and Yan)
1999	7617	6222	252	14.8%	14.4%	14.3%	11.8%	7.4%
2000	7234	5968	252	13.6%	13.2%	13.1%	10.8%	6.8%
2001	6524	5119	248	16.7%	16.2%	16.1%	13.3%	8.3%
2002	5811	4746	252	18.6%	18.1%	17.8%	14.2%	9.3%
2003	5405	4772	252	16.8%	16.5%	16.2%	13.1%	8.4%
2004	5199	4953	252	13.4%	13.2%	12.9%	10.5%	6.7%
2005	5142	4820	252	13.1%	12.8%	12.5%	10.3%	6.6%
2006	5068	4730	251	13.8%	13.5%	13.2%	10.7%	6.9%
2007	5056	4770	251	11.0%	10.7%	10.5%	8.5%	5.5%
2008	4746	4375	253	9.4%	9.2%	9.1%	7.2%	4.7%
2009	4464	4263	252	9.2%	9.0%	8.8%	7.0%	4.6%
2010	4272	3970	252	7.6%	7.5%	7.3%	5.9%	3.8%
1–9/2011	4062	3642	189	6.5%	6.4%	6.3%	5.1%	3.2%
1999–2005	9278	8366	1760	15.3%	14.9%	14.7%	12.0%	7.6%
All	10,407	9412	3208	12.6%	12.3%	12.2%	9.9%	6.3%

This table reports the annual summary of Abel Noser dollar trading volume as a percentage of CRSP dollar trading volume from 1999 to September 2011. We adjust NASDAQ volumes according to Gao and Ritter (2010). *% of CRSP_U* is the upper bound which is calculated as Abel Noser's buy plus sell volumes divided by the CRSP daily trading volume. We estimate four progressively tighter lower bounds: *% of CRSP_L1*, *% of CRSP_L2*, *% of CRSP_L3*, and *% of CRSP_L4*. *% of CRSP_L1* follows the method in Puckett and Yan (2011) that divide Abel Noser buy plus sell volumes by 2. *% of CRSP_L2* takes the maximum of buy and sell volumes for each stock on a given trading day (Bethel et al. (2009)). For *% of CRSP_L3*, we first round transaction prices to penny, and then further impose the condition that, in order to be recognized as trades between Abel Noser clients, the rounded transaction price should be equal between buy and sell trades. *% of CRSP_L4* is similar to *% of CRSP_L3* except that we do not round to penny. Volume is calculated as the product of the CRSP daily closing price and the shares traded for the same stock recorded either in Abel Noser or CRSP.

methods for different bounds as described above) to that in CRSP on each trading day, and then the average of these five ratios across trading days within the same calendar year are reported in Table 4.

For 1999–2005, our estimate of *% of CRSP_L1* is 7.6%, similar to the 8% reported in Puckett and Yan (2011) over the same time period. However, *% of CRSP_L2* is significantly higher at 12%, which is simply taking the maximum instead of dividing buy and sell volumes by two (Bethel et al. (2009)). By further imposing the cancel-off restriction that (rounded) prices should be equal, we obtain higher estimates: 14.7% for *% of CRSP_L3* (and 14.9% for *% of CRSP_L4*). The upper bound, *% of CRSP_U*, is 15.3%. These estimates produce a very tight range for Abel Noser data's coverage of CRSP volume for the 1999–2005 sample periods. We therefore conclude that Abel Noser data cover about 15% of CRSP trading volume during 1999–2005. If one were to apply a similar logic as in Puckett and Yan (2011), i.e., assuming that institutional investors are responsible for 80% of CRSP trading volume, then we estimate that Abel Noser institutions account for 19% of all institutional trading volume during 1999–2005.

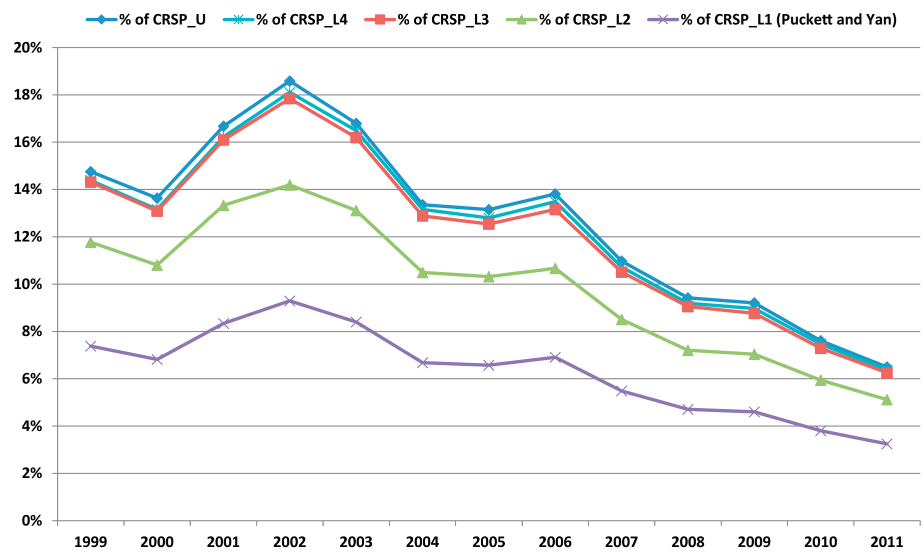


Fig. 4. Abel Noser Data's coverage of CRSP volume over time. This figure plots the four lower and the upper bound estimates of Abel Noser's coverage of CRSP volume over time (1999 – first 9 months of 2011). *% of CRSP_U, % of CRSP_L4*, *% of CRSP_L3*, *% of CRSP_L2*, and *% of CRSP_L1* are defined as in Table 4.

Fig. 4 plots Abel Noser data's coverage of CRSP volume over time. The last two lower bounds: *% of CRSP_L3* and *% of CRSP_L4*, are quite similar and close to the upper bound, *% of CRSP_U*. This is not surprising as Abel Noser institutions taken as a whole, though quite significant, only account for a minority of all market participants. Therefore, for a given Abel Noser transaction, it is much more likely that it traded with a market participant outside of Abel Noser rather than within. Furthermore, it seems that Abel Noser's coverage of CRSP has declined over time, especially in the latter half of the sample. As a result, for the overall sample from January 1999 to September 2011, we estimate that Abel Noser data's coverage of CRSP trading volume to be between 12.3% and 12.6%, or simply about 12%. If one were to assume that institutional investors are approximately responsible for 80% of CRSP trading volume, as in Puckett and Yan (2011), then one can infer that Abel Noser institutions account for 15% of all institutional trading volume over the entire sample period (1999–2011).

4. Publication patterns using Abel Noser data

To identify publications using Abel Noser data, we conduct keyword searches on Google Scholar, using keywords "Abel Noser," "Abel/Noser," or "ANcerno," combined with sources in "Finance," "Financial," "Economics," and "Accounting." We then comb through all search results to identify publications that use Abel Noser data in their analyses, and exclude publications that mention Abel Noser or ANcerno but do not actually use the data. We also check authors' and journals' websites for forthcoming papers. Table 5 lists the 55 publications yielded from this process.[4] We then classify each paper into one of four areas: market microstructure, corporate finance, investments, and accounting. This is not always easy, as some papers span across different areas. The main classification criterion we use is the research question addressed in the paper, and we also use other criteria such as the journal outlet.

The first publication using Abel Noser data was Blume (1993), and there was a fifteen-year gap until the next publication appeared: Hu et al. (2008). The next year, 2009, was a break-out year for Abel Noser data in terms of academic publications, with six papers published using the data. 2009 was also the first year when academic studies using Abel Noser data appeared in top journals: Chemmanur et al. (2009) and Goldstein et al. (2009). The following year, 2010, turned out to be a lean year for Abel Noser data, with only one paper published: Chemmanur et al. (2010), though it was also in a top journal.

Fig. 5 Panel A plots numbers of publications using Abel Noser data by publication year. From 2011 till now, there have been at least three publications using the data every year, with the highest numbers in 2014 and 2017 – nine and eight publications, respectively. It appears that the interests in Abel Noser data have in fact grown in recent years, with many scholars newly adopting the data. This trend is not slowing down, as we have identified two published papers in-print in 2018 so far and eight forthcoming papers using Abel Noser data.

Fig. 5 Panel B plots numbers of publications by area. 23 out of 55 publications are in investments, which also include publications on capital markets and financial intermediation. Since Abel Noser data is a specialized microstructure dataset, it is not surprising that there are 15 publications in market microstructure. What is a bit surprising is that there are only 11 publications in corporate finance, though it is a much larger field than market microstructure. As discussed previously, given the daily frequency, Abel Noser data are especially well-suited for studying the behavior of institutional investors around various corporate events. Since Abel Noser data were first discovered and used by finance scholars, there have been only six accounting publications so far. This may mean that accounting could be the next frontier for Abel Noser data.

Fig. 5 Panel C further separates publications by area into two sub-periods: 1993–2013 versus 2014 and after. Some interesting patterns emerge. The later sub-period, though shorter, has more publications compared to the former, 35 versus 20. Before 2014, half of the 20 publications are in market microstructure, even though it is a small field in finance. However, for 2014 and after, the microstructure field seems to be shrinking – only five out of 35 publications are in microstructure. There are six corporate finance publications after 2014 compared with five before. Therefore, its proportion has in fact declined. On the other hand, though only five out of 20 publications before 2014 are in investments, for 2014 and after, about half of the 35 publications are in investments (including financial intermediation). Accounting has also experienced significant growth, from zero in the earlier period to six publications in the latter.

Fig. 5 Panel D plots numbers of publications by journal. Publications using Abel Noser data have appeared in 22 journals so far and generally in high-quality journals: 20 in "top three" finance journals and four in "top three" accounting journals. If we were to include *Journal of Financial and Quantitative Analysis*, *Management Science*, and *Review of Finance*, then the majority of publications using Abel Noser data appeared in "top" journals (32 out of 55). We do not have a very good explanation for why 14 publications using Abel Noser data appeared in *Review of Financial Studies*. It could be due to pure randomness. One possible reason we could think of is that several of the early publications using Abel Noser data were published in *Review of Financial Studies*: Goldstein et al. (2009), Chemmanur et al. (2010), Anand et al. (2012), and Choi and Sias (2012). Publications using Abel Noser data also appeared in many other high-quality journals, most notably, three in *Journal of Corporate Finance* and five in *Journal Financial Markets*.

5. A survey of academic literature using Abel Noser data

In this section, we survey the growing academic literature, including 55 publications so far (including forthcoming papers) and selected working papers, using Abel Noser data to address various research questions related to institutional trading in market

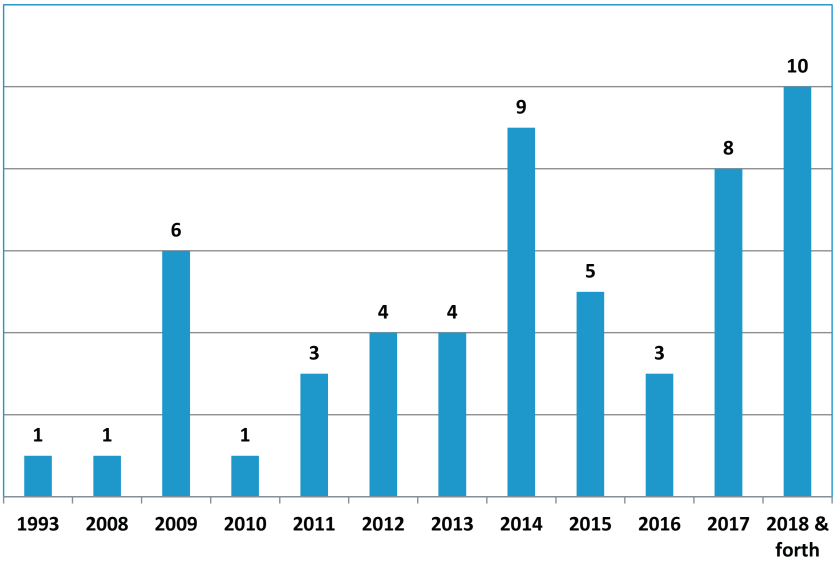
[4] We exclude Kelley and Tetlock (2017), who use Abel Noser data, but only to construct a control proxy for institutional trading in their analysis of retail short selling. All publications in Table 5 directly analyze Abel Noser data.

Table 5
Publications using Abel Noser data.

Authors	Year	Journal	Sample Start	Sample End	Area
Blume	1993	*FAJ*	1/1990	3/1990	Microstructure
Hu, Meng, and Potter	2008	*JBFA*	10/2001	12/2001	Investments
Bethel, Hu, and Wang	2009	*JCF*	1/1999	12/2005	Corporate Finance
Chemmanur, He, and Hu	2009	*JFE*	1/1999	12/2005	Corporate Finance
Goldstein, Ivrine, Kandel, and Wiener	2009	*RFS*	1/1999	12/2003	Microstructure
Hu	2009	*JFM*	10/2001	12/2001	Microstructure
Pagano	2009a, 2009b	*IJMF*	9/2006	11/2007	Microstructure
Pagano	2009a, 2009b	*JoT*	9/2006	11/2007	Microstructure
Chemmanur, Hu, and Huang	2010	*RFS*	1/1999	12/2004	Corporate Finance
Goldstein, Ivrine, and Puckett	2011	*JFQA*	1/1999	12/2005	Corporate Finance
Green and Jame	2011	*JFM*	1/1999	12/2005	Investments
Puckett and Yan	2011	*JF*	1/1999	12/2005	Investments
Anand, Irvine, Puckett, and Venkataraman	2012	*RFS*	1/1999	12/2008	Microstructure
Busse, Green, and Jegadeesh	2012	*JFM*	1/1997	12/2005	Investments
Choi and Sias	2012	*RFS*	1/1999	12/2006	Investments
Edelen and Kadlec	2012	*JFE*	1/1999	6/2009	Microstructure
Anand, Irvine, Puckett, and Venkataraman	2013	*JFE*	1/1999	9/2010	Microstructure
Feinstein, Hu, Marcus, and Ali	2013	*JForE*	1/1999	12/2010	Microstructure
Jain and Wang	2013	*JoT*	1/1998	6/2009	Corporate Finance
Kuvvet	2013	*JoT*	1/1997	12/2009	Microstructure
Agarwal, Gay, and Ling	2014	*RFS*	9/1998	12/2008	Investments
Brogaard, Hendershott, Hunt, and Ysusi	2014	*FR*	4/2010	6/2010	Microstructure
Brown, Wei, and Wermers	2014	*MS*	1/1998	12/2008	Investments
Cready, Kumas, and Subasi	2014	*JAR*	1/2003	12/2010	Accounting
Fang, Peress, and Zheng	2014	*RFS*	1/1997	12/2002	Investments
Goetzmann, Kim, Kumar, and Wang	2014	*RFS*	1/1999	12/2010	Investments
Green, Jame, Markov, and Subasi	2014	*JAE*	1/2004	12/2008	Accounting
Hu, McLean, Pontiff, and Wang	2014	*RFS*	1/1999	12/2010	Investments
Lynch, Puckett, and Yan	2014	*JBF*	1/1999	12/2005	Investments
Ahern and Sosyura	2015	*RFS*	1/2000	12/2011	Corporate Finance
Angel, Harris, and Spatt	2015	*QJF*	1/1999	6/2012	Microstructure
Bernile, Sulaeman, and Wang	2015	*JCF*	1/2000	12/2005	Corporate Finance
Chemmanur, Hu, and Huang	2015	*JFQA*	1/1999	12/2009	Corporate Finance
Hameed, Morck, Shen, and Yeung	2015	*RFS*	1/2001	12/2009	Investments
Akbas, Meschke, and Wintoki	2016	*JAE*	1/2002	12/2011	Accounting
Chemmanur and He	2016	*JCF*	1/1999	12/2004	Corporate Finance
Ljungqvist and Qian	2016	*RFS*	7/2006	12/2010	Investments
Ben-Rephael	2017	*JFI*	1/1999	12/2009	Investments
Ben-Rephael, Da, and Israelsen	2017	*RFS*	2/2010	6/2015	Investments
Chakrabarty, Moulton, and Trzcinka	2017	*JFQA*	1/1997	12/2009	Investments
Cheng, Hameed, Subrahmanyam, and Titman	2017	*JFQA*	1/1999	12/2011	Investments
Chiyachantana, Jain, Jiang, and Sharma	2017	*JFM*	1/2001	12/2012	Microstructure
Henry, Nguyen, and Pham	2017	*JFM*	1/1997	9/2011	Corporate Finance
Henry and Koski	2017	*JF*	1/1999	3/2008	Corporate Finance
Jain, Kuvvet, and Pagano	2017	*IBR*	1/2004	12/2008	Investments
Gantchev and Jotikasthira	2018	*MS*	1/2000	12/2007	Investments
Hu, Ke, and Yu	2018	*JAAF*	1/1999	12/2005	Accounting
Ben-David, Franzoni, and Moussawi	forthcoming	*JF*	1/2000	12/2015	Investments
Ben-Rephael and Israelsen	forthcoming	*RoF*	1/1999	9/2011	Microstructure
Bhattacharya, Cho, and Kim	forthcoming	*TAR*	1/2007	12/2010	Accounting
Busse, Tong, Tong, and Zhang	forthcoming	*RFS*	1/1999	12/2009	Investments
Chen, Da, Huang	forthcoming	*RFS*	6/2006	3/2011	Investments
Huang, Lu, and Wang	forthcoming	*JAAF*	1/2002	12/2010	Accounting
Jame	forthcoming	*MS*	1/1999	12/2010	Investments
Lepone and Wong	forthcoming	*JBFA*	10/2005	12/2006	Microstructure

This table lists 55 publications so far using Abel Noser data, ordered by publication year and alphabetical order within the year. Rows are alternately shaded by publication year. Abbreviations for journals (alphabetical order): *Financial Analyst Journal*-FAJ, *Financial Review*-FR, *International Business Review*-IBR, *International Journal of Managerial Finance*-IJMF, *Journal of Accounting and Economics*-JAE, *Journal of Accounting Research*-JAR, *Journal of Accounting, Auditing and Finance*-JAAF, *Journal of Banking and Finance*-JBF, *Journal of Business Finance and Accounting*-JBFA, *Journal of Corporate Finance*-JCF, *Journal of Finance*-JF, *Journal of Financial and Quantitative Analysis*-JFQA, *Journal of Financial Economics*-JFE, *Journal of Financial Intermediation*-JFI, *Journal of Financial Markets*-JFM, *Journal of Forensic Economics*-JForE, *Journal of Trading*-JoT, *Management Science*-MS, *Quarterly Journal of Finance*-QJF, *Review of Finance*-RoF, *Review of Financial Studies*-RFS, and *The Accounting Review*-TAR.

Panel A. Publications by Year



Panel B. Publications by Area

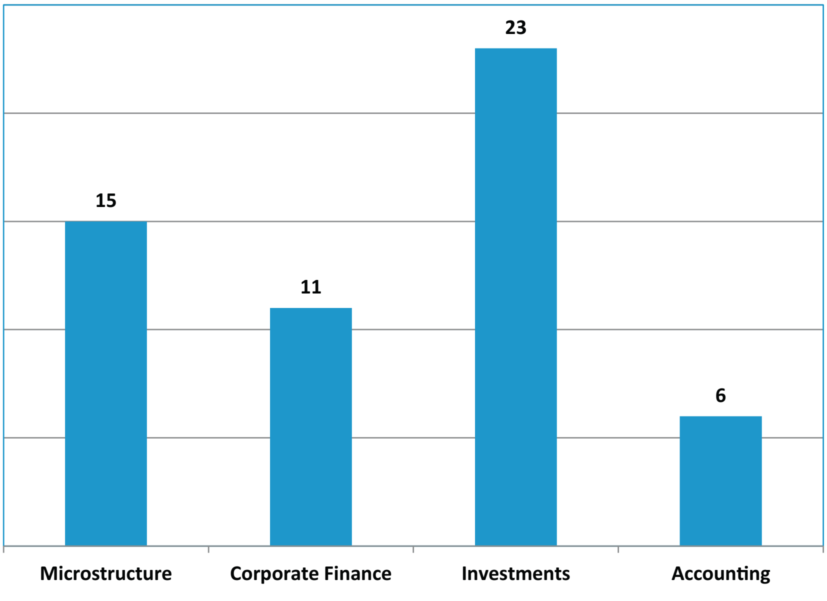
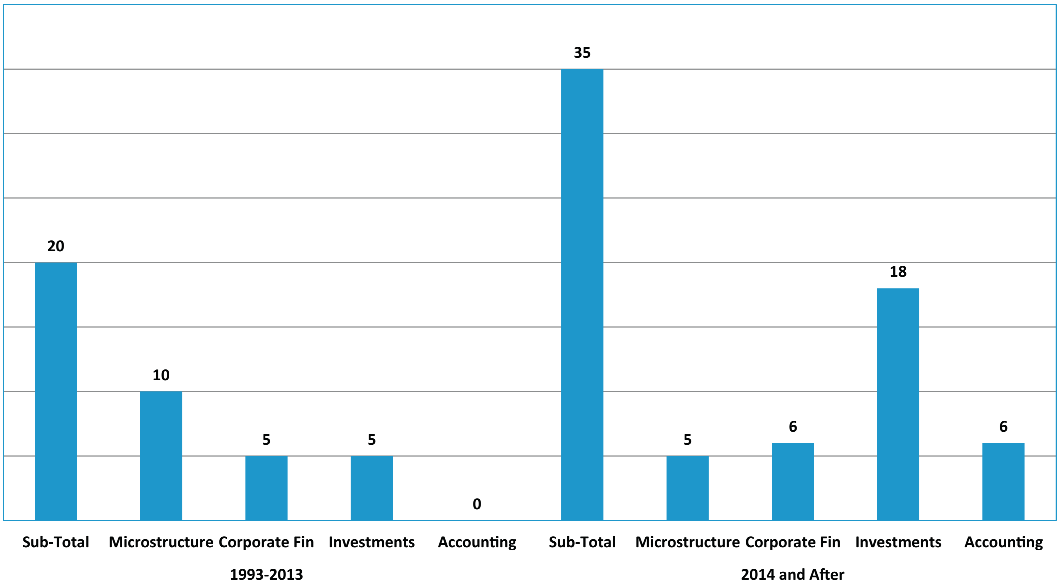


Fig. 5. This plots 55 publications so far using Abel Noser data by the journals they are published in. Abbreviations for journals (alphabetical order): *Financial Analyst Journal*-FAJ, *Financial Review*-FR, *International Business Review*-IBR, *International Journal of Managerial Finance*-IJMF, *Journal of Accounting and Economics*-JAE, *Journal of Accounting Research*-JAR, *Journal of Accounting, Auditing and Finance*-JAAF, *Journal of Banking and Finance*-JBF, *Journal of Business Finance and Accounting*-JBFA, *Journal of Corporate Finance*-JCF, *Journal of Finance*-JF, *Journal of Financial and Quantitative Analysis*-JFQA, *Journal of Financial Economics*-JFE, *Journal of Financial Intermediation*-JFI, *Journal of Financial Markets*-JFM, *Journal of Forensic Economics*-JForE, *Journal of Trading*-JoT, *Management Science*-MS, *Quarterly Journal of Finance*-QJF, *Review of Finance*-RoF, *Review of Financial Studies*-RFS, and *The Accounting Review*-TAR.

Panel C. Publications by Area, Before versus After 2014



Panel D. Publications by Journal

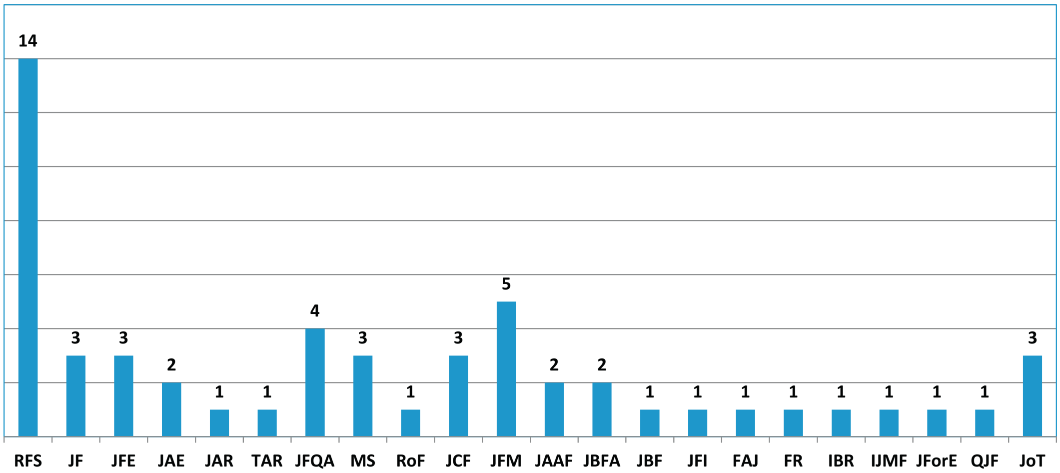


Fig. 5. (*continued*)

microstructure, corporate finance, investments, and accounting, respectively.

5.1. *Institutional trading and market microstructure*

Since Abel Noser data are at transaction-level, it is well-suited for studying microstructure issues. Abel Noser data were first used in Blume (1993), who studies "soft dollars" in brokerage commissions, and its impact on the brokerage industry and institutional trading. Goldstein et al. (2009) find evidences showing that the per-share commission practice offers a convenient way for traditional full-service brokers to charge their clients a pre-negotiated fixed fee for long-term access to their premium services.

Hu (2009) shows that the widely documented buy-sell asymmetry in implicit institutional trading cost is mainly driven by mechanical characteristics of a specific class of measures: pre-trade measures, which is how implementation shortfall and price impact are typically measured. On the other hand, during-trade measures, such as the VWAP (volume-weighted average price) cost, do not suffer from the same problem (see Berkowitz et al. (1988) and Hu (2007) for detailed discussions of VWAP Cost). Henry and

Koski (2017) apply the VWAP Cost measure in Hu (2009) to measure institutional trader skill during ex-dividend periods. In order to alleviate these concerns about traditional pre-trade measures of price impact, Chiyachantana et al. (2017) devise an innovative measure of market-adjusted and risk-adjusted permanent price impact to isolate the price impact related to new fundamental information about the stock and to standardize it across stocks with different risk characteristics. Using this measure and Abel Noser data, the authors find that the sign of the permanent price impact asymmetry between institutional buys versus sells is positive at the initial stage of a price run-up and reverses due to changing constraints with a prolonged price run-up in a stock, supportive of the theoretical predictions in Saar (2001).

Anand et al. (2012) document the presence of persistent trading skill by institutional trading desks, and though some brokers can deliver better executions consistently over time, this trading-desk skill is not limited to broker selection. They conclude that the trade execution process is economically important and can contribute to portfolio performance. In a similar setting, Edelen and Kadlec (2012) model the agency conflict arising from the portfolio manager delegating trade execution to a separated trading desk, and show that optimal trading performance benchmarks often create an incentive for traders to execute orders contrary to concurrent information flow, which leads to delays in the assimilation of information in security prices. Using Abel Noser data, they find evidences consistent with the predictions of their model. Examining the time-series trading activities of Abel Noser clients from 1999 to 2010, Anand et al. (2013) discover that up to 2007 institutions enjoy a consistent decline in trading cost. Yet, after the breaking out of the 2008 financial crisis in October, the average cost of trading increases dramatically and it lasts until the second half of 2009. In addition, they identify a set of liquidity providers among the sample institution during the crisis that withdraw from risky securities significantly and recover slowly until the end of 2009. In a recent paper, Ben-Rephael and Israelsen (2018) find that systematic and persistent differences in execution costs exist across clients for the same management company, and these differences are comparable to the variation across management companies documented in Anand et al. (2012). They also find that clients who receive lower execution costs reward management companies with an increase in dollar trading volume.

Pagano (2009a) finds that brokerage commissions around the world are declining but indirect trading costs remain steady. U.S. leads the way with the steepest decline in trading costs, and the emerging markets of South America is experiencing the fastest growth in trading volume. Pagano (2009b) provides evidences of virtuous and vicious circles between trading costs and volume within a specific geographic region or nation, and that trends in U.S. trading costs and volume play a key role in influencing the costs and trading activity in other parts of the world. Applying the daily news-based index of economic policy uncertainty developed by Baker et al. (2016), Kuvvet (2013) finds institutional investors to be net buyers during abnormal market decreases and net sellers during abnormal market increases. However, the author does not find evidence that institutional investors cause or exacerbate the abnormal market movements.

On the other hand, institutional investors seem to provide liquidity during abnormal days. Brogaard et al. (2014) study whether high-frequency trading (HFT) increases the execution costs of institutional investors. The authors make use of technology upgrades that lower the latency of the London Stock Exchange to obtain variation in the level of HFT over time. They find that following upgrades, the level of HFT increases, but around these shocks to HFT institutional traders' costs remain unchanged, thus finding no clear evidence that HFT impacts institutional execution costs. Angel et al. (2015) show that despite concerns expressed by some practitioners and regulators, various measures of market quality indicate that U.S. markets continue to be very healthy. Trading cost estimates remain low and market depth and execution speeds remain high. Total transaction costs of large block trades indicate that improvements in market quality also have benefited large institutional traders. Lepone and Wong (2018) examine the impact of mandatory International Financial Reporting Standards (IFRS) on the market quality of the Australian Securities Exchange (ASX) 200 constituent stocks. In the era of High Frequency Trading (HFT) and occurrences of 'fleeting' liquidity, the authors provide some evidence that while IFRS may have enhanced 'visible' bid-ask spreads, tangible liquidity for market participants, particularly global institutional investors, has not improved significantly, echoing the findings in Brogaard et al. (2014).

Using both claims data in class action securities cases and Abel Noser data, Feinstein et al. (2013) find that aggregate damages in class action securities cases estimated using public volume data may be understated due to the frequent occurrence of inter-fund trades (or internal crosses). Inter-fund trades are internal crossing trades between funds within the same fund family and are one of the few instances of trading transactions that are *not* reported publicly. Two working papers study inter-fund trades or internal crosses in a general setting. Using Abel Noser data between 1999 and 2010, Chan et al. (2017) estimate total trading cost savings from internal crosses to be \$1.9 billion for institutions in Abel Noser data. The authors also identify potential crosses as market trades that could have been crossed internally absent regulatory or other restrictions. They estimate cost savings from potential crosses to be about \$2.4 billion. Their results suggest that the breadth of funds and accounts in investment management firms can generate economies of scale via this unusual channel. Also using Abel Noser data, Eisele et al. (2017) examine the incentives for mutual funds to trade with sibling funds affiliated with the same group. They find that cross-trades are used either to opportunistically reallocate performance among trading funds or to reduce trading costs where the specific incentive depends on internal monitoring and market conditions.

5.2. *Institutional trading and corporate finance*

One of the main advantages of Abel Noser over other alternative data sources for institutional trading is its daily frequency. This feature makes Abel Noser an ideal source for studying institutional investors' trading behavior around various corporate events. An early example of this line of academic research is Bethel et al. (2009) which examines institutional trading around mergers and shareholder voting outcomes. The authors find that institutions as a whole buy shares and hence voting rights before merger record dates. Trading mentioned here is not related to merger arbitrage proxies or merger announcements trading, and thus not a simple continuation of the latter. Trading and buying before recorded dates are positively related to voting turnout and negatively related to shareholder support of merger proposals. Therefore, there appears to be an active voting rights market for institutional investors around merger events. In a recent working paper also using Abel Noser data, Li and Schwartz-Ziv (2018) examine how shareholder votes and trades are related in a broader context.

Also using Abel Noser data, Chemmanur et al. (2009) explicitly identify institutional SEO (seasoned equity offering) allocations for the first time in the literature. Their empirical results show the sample institution can identify those SEO firms with better long-run performance via obtaining more allocations. Meanwhile, their post-SEO trade orders are in the same direction as their private knowledge: their post-event trading is positively related to the pre-offer trade and allocations, which tend to be larger in SEOs with better performance. Hence, their studies in general support the information production role of institutional investors in the stock market.

In the context of initial public offerings (IPOs), Chemmanur et al. (2010) for the first time provide evidences that institutions play an important role in supporting IPOs in the aftermarket, especially for those with weaker post-issue demand, and they are rewarded by underwriters with more allocations. Moreover, the authors find that by participation in the book-building process, institutions who receive IPO allocations have residual information advantage, which allows them to continuously gain trading profits during months after the IPO, suggesting additional economic rents besides the IPO underpricing. Overall, their results suggest that institutional investors possess significant private information about IPOs, play an important supporting role in the IPO aftermarket, and receive considerable compensation for their participation in IPOs. Goldstein et al. (2011) find evidence that institutions increase round-trip stock trades, increase average commissions per share, and pay unusually high commissions on some trades. These excess commission payments are a particularly effective way for transient investors to receive lucrative IPO allocations. Their results suggest that the underwriter's concern for their long-term client relationships limits the payment-for-IPO practice.

Ahern and Sosyura (2015) study how the incentive of media to publish sensational news affects the accuracy of media coverage amid merger rumors. They find journalists' experiences, specialized education, and industry expertise significantly predict accuracy of media coverage, but investors do not fully account for these predictors of future stock returns. They find that institutional investors covered in Abel Noser are net sellers in the target firm with merge rumors, suggesting that institutions in their sample provide liquidity to individuals who buy the targets upon the merger rumors.

Brennan and Hughes (1991) put forth an information production theory of stock splits based on fixed brokerage commissions per share. Chemmanur et al. (2015) test the theory using Abel Noser data. They find that the immediate trades of the sample institution after a stock split are significantly informative regarding the stock's near future returns. This predictability is stronger among stocks with larger split factors and among institutions that pay more aftermath commissions. Furthermore, the analyst forecast error is decreased among those split stocks for which institutions pay more commission fees. Overall, their paper provides clear evidences that the incentive of brokerage houses to produce more valuable information of a particular stock is directly related to the amount of commissions they can generate from. Examining the role of institutional trading during the option backdating scandal of 2006–2007, Bernile et al. (2015) find that institutional investors behave as informed investors. Institutional investors display negative abnormal trading imbalances in anticipation of firm-specific backdating exposures, with the underlying trades earning positive abnormal short- and long-term profits. When the backdating is likely a more severe issue, the negative abnormal imbalances are even larger in magnitude.

Different from the equity market setting, Jain and Wang (2013) focus on the credit rating change event for the U.S. bond market. In particular, they distinguish two types of events: 'reactive' versus 'proactive', where the former can be anticipated based on recent earnings news, equity analysts' opinion changes, rating watch and etc. Overall, they provide evidences suggesting that institutional investors are informative of the bond downgrade events and their daily profitability surrounding the event is comparable to that in the equity market. Under the setting of corporate spin-offs, where previous empirical studies have shown that firms tend to experience positive market reactions, Chemmanur and He (2016) utilize Abel Noser trading data to examine the role of institutional investors during the event. They find that institutional trades subsequent to the spin-off completion date are informative about the firm's stock and operating performance. Besides, the result shows that the post-event institutional trades tend to be concentrated in one-side (buy or sell) and in either the parent or subsidiary company, which they interpret as evidences of relative information advantages of the institution regarding the sub-division of the conglomerate.

Henry and Koski (2017) examine whether skilled institutions indeed exploit positive abnormal ex-dividend returns, and find that institutions on average concentrate trading around certain ex-dates and are capable of identifying ex-day events with higher trading profits. Dividend capture trades represent 6% of all institutional buy trades but contribute 15% of overall abnormal returns. Institutional dividend capture trading skill is persistent. Institutional ex-day profitability is also strongly related to trade execution skill. Their results suggest that skilled institutions target certain opportunities rather than benefit from the execution skill uniformly over time. Examining another interesting corporate event, dividend cut, Henry et al. (2017) show that up to two quarters before the public announcement of quarterly dividend cut (omission), there are significant net selling activities by institutional traders. They find that 'negative buying' activities are more pronounced among those firms that are more 'opaque,' suggesting that sample institutions have information advantage relative to the market trader.

5.3. Institutional trading and investments

Abel Noser data contain trading records of institutional investors, which makes it suitable for studying research topics in investments and on behaviors of institutional investors as financial intermediaries. How information in the market is disseminated and to what extent investors pay attention to and trade based on the information are important issues in investments. Inspired by and building on the findings of Kacperczyk et al. (2008), Puckett and Yan (2011) directly analyze the performance of interim trades of institutional investors. Their evidences suggest that trading skills documented by previous studies using quarterly portfolio holdings can be biased downwards when interim trades are not accounted for as they document evidence that institutional investors earn significant abnormal returns on their trades within the trading quarter.

Choi and Sias (2012) test whether financial strength information is gradually impounded over time using institutional investor demand as a proxy for revisions in sophisticated investors' expectation. They find that financial strength predicts both future returns and future institutional investor demand. In addition, more sophisticated transient institutions respond to financial strength signals prior to less sophisticated institutions, consistent with the gradual incorporation of information. Hameed et al. (2015) document that analyst follow disproportionally firms whose fundamentals correlate more with those of their industry peers. They find that when analysts revise a bellwether firm's earning forecast, it changes the prices of other firms significantly, while the same does not happen when revisions for firms followed less intensely occurs. They find institutional investors buying up shares in other firms in a bellwether firm's industry when analysts revise the latter's earnings forecasts upward. Ljungqvist and Qian (2016) provide evidence that even small arbitrageurs help make prices efficient by revealing their own information. They show that investors respond strongly to the information, spiking SEC filing views, volatility, order imbalances, realized spreads, turnover, and selling by the longs. The small arbitragers induce target company shareholders to trade on their behalf, as they have limited capital and face server short-sale constraints themselves. When the information is credible, the unconstrained investors sell, thereby accelerating price discovery.

Other studies examine investors' attention through media coverage. Extending Fang and Peress (2009) where they demonstrate that breadth of information dissemination stemming from mass media coverage affects stock returns, Fang et al. (2014) examine the relation between mutual fund trades and mass media coverage of stocks. The authors construct a measure of mutual fund's propensity to buy or sell stocks covered in media, and find funds tend to buy stocks with media coverage more than those without, where this propensity is negatively related to their future performance. However, this result does not extend to fund sells. The authors conclude that their findings suggest limited attention among professional investors. Ben-Rephael et al. (2017) find that institutional attention responds more quickly to major news events, leads retail attention, and facilitates permanent price adjustment. The authors use a direct measure of abnormal institutional investor attention (AIA) by observing news searching and news reading activity for specific stocks on Bloomberg terminals. They also find that the price drifts following both earnings announcements and analyst recommendation changes to be caused by announcements where institutional investors fail to pay sufficient attention.

Interestingly, some researchers have found institutional traders to not exhibit special skills in certain aspects related to investments. Busse et al. (2012) conclude that institutional investors do not particularly exhibit skills when it comes to discerning the quality of recommendations, after examining the performance of buy-side institutional investor trades in conjunction with sell-side brokerage analyst stock recommendations. Behaviorally, buy-side trades follow sell-side analyst recommendations but not the other way around. When recommendation changes, buy-side institutional investors trading in the same direction as the recommendation change earn equally as other institutional investors trading in the opposite direction. Similarly, Chakrabarty et al. (2017) document that the majority of short-term institutional trades lose money. They document over 23% of round-trip trades are held for less than 3 months and these trades on average have a − 3.91% return. The losses are found across all types of stocks, with the worst performance occurring in small stocks, value stocks, and low-momentum stocks. More volatile the market, short-term trades lose even more.

Some allege that institutional investors try to mislead investors by placing trades that inflate performance (portfolio pumping) or distort reported holdings (window dressing). Examining window-dressing behavior of mutual fund managers, Agarwal et al. (2014) find evidence that window dressing is more likely to be done by managers who are less skilled, perform poorly, and tend to incur high portfolio turnover and trade costs, which is followed by worse future performance. On the other hand, Hu et al. (2014) offer depressed institutional selling as a previously unexplored explanation for year-end price inflation and find institutions tend to buy stocks in which they already have large positions, consistent with portfolio pumping. They find no obvious evidence of window dressing in institutions using daily institutional trades. They argue that evidences of window dressing in past studies were inconclusive because they are mainly based on indirect tests with quarterly holdings data instead of daily trading data. Investigating institutions' role as the explanations for the turn-of-the-year (TOY) effect, Lynch et al. (2014) conclude that institutions play a limited role in driving the TOY effect. They find limited evidence that institutional trading impacts TOY returns through window dressing, and little evidence of the tax-loss selling and risk-shifting trading strategies contributing to the TOY returns. In addition, stocks with no institutional trading around the year-end were found to have considerably stronger TOY return patterns than stocks with institutional trading.

Brown et al. (2014) take advantage of the Abel Noser's institutional investor transaction level data setting, and study herding behavior of mutual funds. They document that mutual funds herd into stocks with consensus sell-side analyst upgrades, but herd out of stocks with consensus downgrades. The effect of herding is stronger for downgrades, and among managers with greater career concerns as they are incentivized to follow analyst information given the greater reputational and litigation risk of holding losing stocks. They suggest that career-concerned fund managers herding is influenced by analyst recommendation revisions, and this type of trading destabilize price more as the level of mutual fund ownership of stocks increase. Taking another direction, Hu et al. (2008) examine opinion divergence among professional investment managers and find it to be common. When managers trade in the opposite direction, subsequent returns are low, especially for stocks that are difficult to short, which is consistent with Miller's (1977) hypothesis that, opinion divergence can cause an upward bias in prices when short-sale constraints are present. When managers trade in the same direction, returns are similar regardless of the direction, which the authors concludes is inconsistent with the notion that professional investment managers possess stock picking skills or private information that is of investment value.

Transaction-level data allows for comprehensive research of the relation between investment decisions and liquidity. Examining the trades of index funds and their willingness to accept tracking error in order to trade at more favorable prices, Green and Jame (2011) find index funds purchase stocks beginning with the announcement of composition changes and do not fully establish their positions until weeks after the effective date. They also document that small and mid-cap funds provide liquidity to index funds around additions, and added stocks with a greater proportion of these natural liquidity providers experience lower inclusion returns. Exploring the trading decisions of equity mutual funds during ten periods of extreme market uncertainty, Ben-Rephael (2017) find that mutual funds reduce their holdings of illiquid stocks. He notes that the result is mainly driven by larger withdrawals from funds that hold less liquid stocks only after initial deterioration in market conditions. Cheng et al. (2017) find that stocks that perform poorly in the previous quarter lead to stronger reversals over the subsequent two months, due to changes in the number of active investors in the stock, which influence liquidity provision. Using liquidity supplying institutions identified from Abel Noser data base as one of their proxies for active institutions, they find that active institutions participate less in losing stocks and that the magnitude of monthly return reversals fluctuates with changes in the number of active institutional investors. They argue that the link between past returns and return reversals are partly explained by fluctuations in liquidity provision with past return performance.

Ben-David et al. (2018) argue that ETFs can be a catalyst for short-horizon liquidity traders because of their low trading costs. Thus, liquidity shocks can propagate to the underlying securities through the arbitrage channel, and ETFs may increase the non-fundamental volatility of those securities in their baskets. The authors exploit exogenous changes in index membership and find that stocks with higher ETF ownership display significantly higher volatility. They also find that ETF ownership increases the negative autocorrelation in stock prices and that the increase in volatility appears to introduce undiversifiable risk such that stocks with high ETF ownership earn a significant risk premium.

Jame (2018) shows that hedge funds that engage in short-term contrarian strategies (i.e., liquidity suppliers) earn significantly higher returns on their equity trades and holdings where the superior returns are reflective of greater exposure to a liquidity provision. The liquidity-supplying funds have a tendency to trade against stocks heavily traded by constrained mutual funds but not so much against stocks heavily traded by unconstrained mutual funds. Adding the fact that the outperformance of liquidity-supplying funds is concentrated during periods of low funding liquidity, the author suggests liquidity to be a useful proxy for the fund's financing constraints. Gantchev and Jotikasthira (2018) investigate the liquidity theories of activism, where the activist screens firms based on fundamentals but target firms at a particular time by exploiting institutional liquidity shocks, and demonstrate that firm's probability of becoming an activist target increases with institutional sales. The authors also document that activist purchases closely track institutional sales, such synchronicity is stronger among targets with lower expected monitoring benefits, and that institutional sales accelerate the timing of a campaign at firms already followed by activists.

Busse et al. (2018) propose a new measure of how regularly investors trade. The authors find that institutional investors who trade regularly outperform and such outperformance is persistent for at least a year. Among those who trade most regularly, larger funds perform relatively worse due to higher trading costs associated with larger trades. Institutions that regularly trade generate superior performance by behaving as contrarians and trading more aggressively on information. In contrast, the authors find no relation between their measure of trading regularity and outperformance among index funds. Chen et al. (2018) examine net arbitrage trading (NAT) measured by the difference between quarterly abnormal hedge fund holdings and abnormal short interest. The authors find that NAT strongly predicts stock returns in the cross section, and abnormal returns are only present among stocks experiencing large NAT across well-known stock anomalies. The authors confirm their findings using daily Abel Noser institutional trading data.

Several studies have examined institutional trading with riveting topics such as weather-based mood or corruption. Using survey, weather, and Abel Noser data, Goetzmann et al. (2014) show that weather-based mood have real impact on perceptions of mispricing and trading decisions of institutional investors. They find that cloudier days increase perceived overpricing in individual stocks and the Dow Jones Industrial Index, increasing selling propensities of institutions. Jain et al. (2017) find corruption to have a significant effect on a nation's financial markets through negative impact on foreign portfolio investment (FPI). They find that while highly transparent nations attract the most foreign investment, very corrupt countries actually attract more foreign investment than moderately corrupt countries, exhibiting a nonlinear and reverse J-shape relationship between effects of corruption on FPI.

Finally, Huang et al. (2018a) examine institutional trading surrounding corporate news by combining a comprehensive database of news releases with Abel Noser institutional trades. In order to more precisely identify the ability of institutional investors to predict or quickly interpret news, the authors form "news clusters" of related news that occur in succession. They find that institutions mainly trade on the tone of news directly after the earliest news release in a cluster. As discussed earlier, Huang et al. (2018a) make an important contribution by carefully analyzing and making full use of intraday time stamps in Abel Noser data.

5.4. Institutional trading and accounting

Several recent studies in accounting have also used Abel Noser data. In the past, considerable body of research had explored how investors process information by categorizing investor size (individual vs institutional) based on trading sizes building on early works by Cready (1988) and Lee (1992). Cready et al. (2014) use Abel Noser data to examine the reliability of transaction size-based inference about trader behaviors and find that institutions are heavily involved in small transaction activities. Furthermore, they document that institutions substantially increase their order size in announcement periods, presumably in response to earnings news, and trade in unsophisticated manners contrary to prior evidence that large traders trade in the direction of the analyst forecast error and ignore the random walk error. Despite the fact that sample constitutes sizable portion of overall market trading activities, they find no evidence of announcement period trading affecting the magnitude of post-earnings-announcement drift.

Green et al. (2014) examine the determinants and consequences of broker-hosted investor conferences. The authors use Abel Noser data to measure trade commission share across brokers, and document a significant increase in the host's commission share for conference stocks in weeks 0 and + 1. Moreover, the commission share increase for a conference stock conditional on an informative disclosure being made is approximately twice as large as the increase for a conference stock conditional on an uninformative disclosure. The variation in commissions around conferences suggests that investors reward brokers for providing management access, and that these rewards are greater when access to management results in a greater transfer of value-relevant information.

Investigating whether the directors' professional connections can affect the likelihood of information transfer to sophisticated traders, Akbas et al. (2016) find that financial institutions tend to be better informed when trading stocks of firms with more connected board members. For firms with large director networks, the difference in risk-adjusted returns between most traded stocks and least traded stocks by sophisticated investors is between 4%~7.2% when compared to those of firms with less connected directors. Overall, the evidence shows that connectedness of corporate board affects the information set of short sellers, options traders, and institutional investors.

Motivated by conflicting views on transient institutions' degree of sophistication and managerial allegation that transient institutions sell firm's shares whenever there is a small shortfall of reported earnings, Hu et al. (2018) find economically significant abnormal selling by transient institutions in response to small negative earnings surprises. While transient institutions' selling in response to small negative earnings surprises is associated with contemporaneous stock price declines, there is no reversal of stock prices subsequent to transient institutions' trading, indicating that the sale is not an overreaction. This suggests that transient institutions are sophisticated traders and can correctly interpret small negative earnings surprises.

Bhattacharya et al. (2018) investigate how XBRL adoption affects smaller institutions' access to financial statement information relative to their larger ones. The authors examine three aspects of trading responsiveness: abnormal trading volume, response speed to 10-K information, and decision to trade immediately following the 10-K filing. With regard to all three aspects of trading responsiveness, they find that small institutions' responsiveness to 10-K news increases significantly more relative to the change experienced by large institutions from the pre- to post-XBRL periods. They further document that small institutions' stock picking skills in the 10-K filing period increase more compared to those of large institutions following the regulation. Overall, their results suggest that the informational playing field between small and large institutions has become more even following the SEC's XBRL mandate.

Huang et al. (2018b) examine option backdating announcements and the information advantage of institutional investors. The authors use a two-step research design where they first identify fund-firm pairs with heavily sold shares before public revelation of option backdating investigations. In the second stage, the authors focus on trading at other times and find that these funds are more likely to make correct trades before the earnings announcements of their paired firms, and they trade more actively and perform better on these paired firms in general. Their results support the notion that institutions possess information advantage on their paired firms, but this advantage deteriorates following the backdating revelation.

6. Conclusion and directions for future research

In this paper, we analyze institutional trading using Abel Noser data. We provide background information and suggestions for cleaning and using the data, and discuss advantages and disadvantages of Abel Noser compared to various other potential data sources for institutional trading. We document two simple facts: 1) institutional trade sizes decline dramatically over time, rendering trade size-based inferences of institutional trades problematic; 2) we estimate that Abel Noser data cover 12% of CRSP volume over the 1999–9/2011 sample period and 15% for 1999–2005, significantly higher than the previously estimated 8% in Puckett and Yan (2011) for 1999–2005, a widely quoted number in the literature.

We then turn to survey the growing academic literature, including 55 publications thus far, using Abel Noser data to address various research questions related to institutional trading in market microstructure, corporate finance, investments, and accounting. We also analyze publication patterns and show how the availability of a specialized microstructure dataset propagates across different areas in finance and other disciplines such as accounting.

One potentially fruitful direction for future research is to combine Abel Noser institutional trading data with other novel data sources so as to answer interesting research questions. Some good examples of publication along this line include: Bethel et al. (2009), which combines Abel Noser data with shareholder voting data around mergers and acquisitions; Goetzmann et al. (2014), which combines Abel Noser data with institutional investor survey data and weather data; Ahern and Sosyura (2015), which combines Abel Noser data with merger rumors data; and Ben-Rephael et al. (2017), which combines Abel Noser data with news searching and news reading activity on Bloomberg terminals. In terms of working papers, some good examples are: Alldredge and Puckett (2017), which combines Abel Noser data with supply-chain data; Bhattacharya et al. (2017), which combines Abel Noser data with credit ratings issued by EJR – an investor-paid rating agency; Choi et al. (2017), which combines Abel Noser institutional trading data with public institutional holdings data to identify opening and closing of short trades by hedge funds; and Huang et al. (2018a), which combines Abel Noser data with a comprehensive database of news releases for all U.S. firms during 2000 to 2010.

In terms of areas with most opportunities for future research using Abel Noser data, we believe the trend of finance studies using Abel Noser data will remain strong in the foreseeable future, as evidenced by the ten 2018 or forthcoming publications. Within finance, corporate finance appears to be the most under-researched area given that it is a large field with rich literature and a large number of interesting corporate events and contexts. Only 11 out of the 49 finance publications so far are in corporate finance, even less than 15 in microstructure. Out of the 55 publications using Abel Noser data to date, 49 are in finance and only six are in accounting. Since empirical research in accounting and finance (especially corporate finance) are closely related and sometimes even overlap, we believe corporate finance and accounting possess the biggest potential with the most promising opportunities for future scholarly work using Abel Noser data.

Acknowledgement

We thank the editor (Stuart L. Gillan) and the referee for their detailed comments that helped us improve the paper significantly. For helpful discussions, we thank Amber Anand, Eli Bartov, Ekkehart Boehmer, Thomas Chemmanur, Agnes Cheng, Zhi Da, Ilia Dichev, Fangjian Fu, Tyler Henry, Dashan Huang, Pankaj Jain, Michael Jung, Byoung Kang, Bin Ke, Jennifer Koski, Weikai Li, JC Lin, Andy Puckett, Zhen Shi, Wanshan Song, Chester Spatt, Jun Tu, John Wei, Russ Wermers, Franco Wong, Yong Yu, and participants at AsianFA 2018 Conference in Tokyo, 2018 JAAF Conference in Jeju and seminars at Chinese University of Hong Kong Shenzhen, Hong Kong Polytechnic University, and Singapore Management University. Hu thanks Ani Chitaley, Peter Lert, Mark Matthews, and other former colleagues at Fidelity Management & Research Company's Global Equity Trading group for introducing him into the world of institutional trading, and their instrumental help in obtaining Abel Noser data for academic research. Hu also thanks Abel Noser, especially Gene Noser and Jamie Noser, for generously providing their institutional trading data, and Siu Chang, Allison Keane, Judy Maiorca, Bruce Stewart, John Thomas, and Marsha Tracer at or formerly at Abel Noser for excellent technical assistance and answering numerous data questions over the years. Hu acknowledges support from the National Natural Science Foundation of China (NSFC No. 71471130). All remaining errors or omissions are our own.

References

Agarwal, Vikas, Jiang, Wei, Tang, Yuehua, Yang, Baozhong, 2013. Uncovering hedge fund skill from the portfolio holdings they hide. J. Financ. 68, 739–783.
Agarwal, Vikas, Gay, Gerald D., Ling, Leng, 2014. Window dressing in mutual funds. Rev. Financ. Stud. 27, 3133–3170.
Agarwal, Vikas, Mullally, Kevin A., Tang, Yuehua, Yang, Baozhong, 2015. Mandatory portfolio disclosure, stock liquidity, and mutual fund performance. J. Financ. 70, 2733–2776.
Ahern, Kenneth R., Sosyura, Denis, 2015. Rumor has it: sensationalism in financial media. Rev. Financ. Stud. 28, 2050–2093.
Akbas, Ferhat, Meschke, Felix, Wintoki, M. Babajide, 2016. Director networks and informed traders. J. Account. Econ. 62, 1–23.
Alldredge, Dallin M., Puckett, Andy, 2017. The Performance of Institutional Investor Trades across the Supply Chain. (Working paper).
Anand, Amber, Irvine, Paul, Puckett, Andy, Venkataraman, Kumar, 2012. Performance of institutional trading desks: an analysis of persistence in trading costs. Rev. Financ. Stud. 25, 557–598.
Anand, Amber, Irvine, Paul, Puckett, Andy, Venkataraman, Kumar, 2013. Institutional trading and stock resiliency: evidence from the 2007–2009 financial crisis. J. Financ. Econ. 108, 773–797.
Angel, James J., Harris, Lawrence E., Spatt, Chester S., 2015. Equity trading in the 21st century: an update. Q. J. Financ. 5, 1–39.
Aragon, George O., Hertzel, Michael G., Shi, Zhen, 2013. Why do hedge funds avoid disclosure? Evidence from confidential 13F filings. J. Financ. Quant. Anal. 48, 1499–1518.
Baker, Scott R., Bloom, Nicholas, Davis, Steven J., 2016. Measuring economic policy uncertainty. Q. J. Econ. 131, 1593–1636.
Ben-David, Itzhak, Franzoni, Francesco, Moussawi, Rabih, 2018. Do ETFs increase volatility? J. Financ (forthcoming).
Ben-Rephael, Azi, 2017. Flight-to-liquidity, market uncertainty, and the actions of mutual fund investors. J. Financ. Intermed. 31, 30–44.
Ben-Rephael, Azi, Israelsen, Ryan D., 2018. Are some clients more equal than others? An analysis of asset management companies' execution costs. Rev. Financ (forthcoming).
Ben-Rephael, Azi, Da, Zhi, Israelsen, Ryan D., 2017. It depends on where you search: institutional investor attention and underreaction to news. Rev. Financ. Stud. 30, 3009–3047.
Berkowitz, Stephen A., Logue, Dennis E., Noser, Eugene A., 1988. The total costs of transacting on the NYSE. J. Financ. 43, 97–112.
Bernile, Gennaro, Sulaeman, Johan, Wang, Qin, 2015. Institutional trading during a wave of corporate scandals: "Perfect payday"? J. Corp. Finan. 34, 191–209.
Bethel, Jennifer, Gillan, Stuart L., 2002. The impact of the institutional and regulatory environment on shareholder voting. Financ. Manag. 31, 29–54.
Bethel, Jennifer, Hu, Gang, Wang, Qinghai, 2009. The market for shareholder voting rights around mergers and acquisitions: evidence from institutional daily trading and voting. J. Corp. Finan. 15 (2009), 129–145.
Bhattacharya, Utpal, Wei, Kelsey D., Xia, Han, 2017. Follow the Money: Investor Trading around Investor-Paid Credit Rating Changes. (Working paper).
Bhattacharya, Nilabhra, Cho, Young Jun, Kim, Jae B., 2018. Leveling the playing field between large and small institutions: evidence from the SEC's XBRL mandate. Account. Rev (forthcoming).
Blume, Marshall E., 1993. Soft dollars and the brokerage industry. Financ. Anal. J. 49, 36–44.
Bozcuk, Aslihan, Lasfer, M. Ameziane, 2005. The information content of institutional trades on the London Stock Exchange. J. Financ. Quant. Anal. 40, 621–644.
Brennan, Michael J., Hughes, Patricia J., 1991. Stock prices and the supply of information. J. Financ. 46, 1665–1691.
Brogaard, Jonathan, Hendershott, Terrence, Hunt, Stefan, Ysusi, Carla, 2014. High frequency trading and the execution costs of institutional investors. Financ. Rev. 49, 345–369.
Brown, Nerissa C., Wei, Kelsey D., Wermers, Russ, 2014. Analyst recommendations, mutual fund herding, and overreaction in stock prices. Manag. Sci. 60, 1–20.
Busse, Jeffrey A., Clifton Green, T., Jegadeesh, Narasimhan, 2012. Buy-side trades and sell-side recommendations: interactions and information content. J. Financ. Mark. 15, 207–232.
Busse, Jeffrey A., Tong, Lin, Tong, Qing, Zhang, Zhe, 2018. Trading regularity and fund performance. Rev. Financ. Stud (forthcoming).
Campbell, John, Schwartz, Allie, Ramadorai, Tarun, 2009. Caught on tape: institutional trading, stock returns, and earnings announcements. J. Financ. Econ. 92, 66–91.
Chakrabarty, Bidisha, Moulton, Pamela C., Trzcinka, Charles, 2017. The performance of short-term institutional trades. J. Financ. Quant. Anal. 52, 1403–1428.
Chakravarty, Sugato, 2001. Stealth-trading: which traders' trades move stock prices? J. Financ. Econ. 61, 289–307.
Chan, Louis K.C., Lakonishok, Josef, 1993. Institutional trades and intraday stock price behavior. J. Financ. Econ. 33, 173–199.
Chan, Louis K.C., Lakonishok, Josef, 1995. The behavior of stock prices around institutional trades. J. Financ. 50, 1147–1174.
Chan, Yue-Cheong, Conrad, Jennifer, Hu, Gang, Wahal, Sunil, 2017. Trading by Crossing, Working Paper, Hong Kong PolyU, UNC Chapel Hill, and ASU.
Chemmanur, Thomas J., He, Shan, 2016. Institutional trading, information production, and corporate spin-offs. J. Corp. Finan. 38, 54–76.
Chemmanur, Thomas J., He, Shan, Hu, Gang, 2009. The role of institutional investors in seasoned equity offerings. J. Financ. Econ. 94, 384–411.
Chemmanur, Thomas J., Hu, Gang, Huang, Jiekun, 2010. The role of institutional investors in initial public offerings. Rev. Financ. Stud. 23, 4496–4540.
Chemmanur, Thomas J., Hu, Gang, Huang, Jiekun, 2015. Institutional investors and the information production theory of stock splits. J. Financ. Quant. Anal. 50, 413–445.
Chen, Yong, Da, Zhi, Huang, Dayong, 2018. Arbitrage trading: the long and the short of it. Rev. Financ. Stud (forthcoming).
Cheng, Si, Hameed, Allaudeen, Subrahmanyam, Avanidhar, Titman, Sheridan, 2017. Short-term reversals: the effects of past returns and institutional exits. J. Financ. Quant. Anal. 52, 143–173.
Chiyachantana, Chiraphol, Jain, Pankaj K., Jiang, Christine, Sharma, Vivek, 2017. Permanent price impact asymmetry of trades with institutional constraints. J. Financ. Mark. 36, 1–16.
Choi, Nicole Y., Sias, Richard W., 2012. Why does financial strength forecast stock returns? Evidence from subsequent demand by institutional investors. Rev. Financ. Stud. 25, 1550–1587.

Choi, Jaewon, Park, Ji Min, Pearson, Neil D., Sandy, Shastri, 2017. A First Glimpse into the Short Side of Hedge Funds. (Working paper, UIUC).
Chordia, Tarun, Roll, Richard, Subrahmanyam, Avanidhar, 2011. Recent trends in trading activity and market quality. J. Financ. Econ. 101, 243–263.
Conrad, Jennifer S., Johnson, Kevin M., Wahal, Sunil, 2001. Institutional trading and soft dollars. J. Financ. 56, 397–416.
Conrad, Jennifer S., Johnson, Kevin M., Wahal, Sunil, 2003. Institutional trading and alternative trading systems. J. Financ. Econ. 70, 99–134.
Cready, William, 1988. Information value and investor wealth: the case of earnings announcements. J. Account. Res. 26, 1–27.
Cready, William, Kumas, Abdullah, Subasi, Musa, 2014. Are trade size based inferences about traders reliable? Evidence from institutional earnings related trading. J. Account. Res. 52, 877–909.
Daniel, Kent, Grinblatt, Mark, Titman, Sheridan, Wermers, Russ, 1997. Measuring mutual fund performance with characteristics-based benchmarks. J. Financ. 52, 1035–1058.
Edelen, Roger M., Kadlec, Gregory B., 2012. Delegated trading and the speed of adjustment in security prices. J. Financ. Econ. 103, 294–307.
Eisele, Alexander, Nefedova, Tamara, Parise, Gianpaolo, Peijnenburg, Kim, 2017. Trading out of Sight: An Analysis of Cross-Trading in Mutual Fund Families. Working paper. Université Paris-Dauphine and HEC, Paris.
Fang, Lily H., Peress, Joel, 2009. Media coverage and the cross-section of stock returns. J. Financ. 64, 2023–2052.
Fang, Lily H., Peress, Joel, Zheng, Lu, 2014. Does media coverage of stocks affect mutual funds' trading and performance? Rev. Financ. Stud. 27, 3441–3466.
Feinstein, Steven, Hu, Gang, Marcus, Mark, Ali, Zann, 2013. Underestimation of securities fraud aggregate damages due to inter-fund trades. J. Forensic Econ. 24, 161–173.
Gantchev, Nickolay, Jotikasthira, Chotibhak, 2018. Institutional trading and hedge fund activism. Manag. Sci. 64, 2930–2950.
Gao, Xiaohui, Ritter, Jay R., 2010. The marketing of seasoned equity offerings. J. Financ. Econ. 97, 33–52.
Gillan, Stuart L., Starks, Laura T., 2000. Corporate governance proposals and shareholder activism: the role of institutional investors. J. Financ. Econ. 57, 275–305.
Goetzmann, William N., Kim, Dasol, Kumar, Alok, Wang, Qin, 2014. Weather-induced mood, institutional investors, and stock returns. Rev. Financ. Stud. 28, 73–111.
Goldstein, Michael A., Irvine, Paul, Kandel, Eugene, Wiener, Zvi, 2009. Brokerage commissions and institutional trading patterns. Rev. Financ. Stud. 22, 5175–5212.
Goldstein, Michael A., Irvine, Paul, Puckett, Andy, 2011. Purchasing IPOs with commissions. J. Financ. Quant. Anal. 46, 1193–1225.
Green, T. Clifton, Jame, Russell, 2011. Strategic trading by index funds and liquidity provision around S&P 500 index additions. J. Financ. Mark. 14, 605–624.
Green, T. Clifton, Jame, Russell, Markov, Stanimir, Subasi, Musa, 2014. Broker-hosted investor conferences. J. Account. Econ. 58, 142–166.
Griffin, John M., Shu, Tao, Topaloglu, Selim, 2012. Examining the dark side of financial markets: do institutions trade on information from investment bank connections? Rev. Financ. Stud. 25, 2155–2188.
Hameed, Allaudeen, Morck, Randall, Shen, Jianfeng, Yeung, Bernard, 2015. Information, analysts, and stock return comovement. Rev. Financ. Stud. 28, 3153–3187.
Hendershott, Terrence, Livdan, Dmitry, Schürhoff, Norman, 2015. Are institutions informed about news? J. Financ. Econ. 117, 249–287.
Henry, Tyler R., Koski, Jennifer L., 2017. Ex-dividend profitability and institutional trading skill. J. Financ. 72, 461–494.
Henry, Darren, Nguyen, Lily, Pham, Viet Hung, 2017. Institutional trading before dividend reduction announcements. J. Financ. Mark. 36, 40–55.
Hu, Gang, 2005. Essays in Corporate Finance and Institutional Trading. Ph.D. Dissertation. Boston College.
Hu, Gang, 2007. VWAP cost excluding own trades. J. Trading 2, 30–34.
Hu, Gang, 2009. Measures of implicit trading costs and buy-sell asymmetry. J. Financ. Mark. 12, 418–437.
Hu, Gang, Ginger Meng, J., Potter, Mark, 2008. Opinion divergence among professional investment managers. J. Bus. Financ. Account. 35, 679–703.
Hu, Gang, Ke, Bin, Yu, Yong, 2009. Do Transient Institutions Overreact to Small Negative Earnings Surprises? Working paper. https://ssrn.com/abstract=1495384.
Hu, Gang, McLean, R. David, Pontiff, Jeffrey, Wang, Qinghai, 2014. The year-end trading activities of institutional investors: evidence from daily trades. Rev. Financ. Stud. 27, 1593–1614.
Hu, Gang, Ke, Bin, Yu, Yong, 2018. Can transient institutions correctly interpret small negative earnings surprises in the absence of access to management's private information? J. Account. Audit. Financ. 33, 3–33.
Huang, Alan Guoming, Tan, Hongping, Wermers, Russ, 2018a. Institutional Trading around Corporate News: Evidence from Textual Analysis. Working paper. University of Waterloo, McGill University, and University of Maryland.
Huang, Wenli, Lu, Hai, Wang, Xiaolu, 2018b. Option backdating announcement and information advantage of institutional investors. J. Account. Audit. Financ (forthcoming).
Irvine, Paul, Lipson, Marc, Puckett, Andy, 2007. Tipping. Rev. Financ. Stud. 20, 741–768.
Jain, Pankaj, Wang, Qin, 2013. Credit rating changes and institutional trading. J. Trading 8, 38–47.
Jain, Pankaj K., Kuvvet, Emre, Pagano, Michael S., 2017. Corruption's impact on foreign portfolio investment. Int. Bus. Rev. 26, 23–35.
Jame, Russell, 2018. Liquidity provision and the cross section of hedge fund returns. Manag. Sci (forthcoming).
Juergens, Jennifer L., Lindsey, Laura, 2009. Getting out early: an analysis of market making activity at the recommending analyst's firm. J. Financ. 64, 2327–2359.
Kacperczyk, Marcin, Sialm, Clemens, Zheng, Lu, 2008. Unobserved actions of mutual funds. Rev. Financ. Stud. 21, 2379–2416.
Kaniel, Ron, Saar, Gideon, Titman, Sheridan, 2008. Individual investor trading and stock returns. J. Financ. 63, 273–310.
Kaniel, Ron, Liu, Shuming, Saar, Gideon, Titman, Sheridan, 2012. Individual investor trading and return patterns around earnings announcements. J. Financ. 67, 639–680.
Kavajecz, Kenneth A., 1999. A Specialist's quoted depth and the limit order book. J. Financ. 54, 747–771.
Keim, Donald B., Madhavan, Ananth, 1995. Anatomy of the trading process: empirical evidence on the behavior of institutional traders. J. Financ. Econ. 37, 371–398.
Keim, Donald B., Madhavan, Ananth, 1997. Transaction costs and investment style: an inter-exchange analysis of institutional equity trades. J. Financ. Econ. 46, 265–292.
Kelley, Eric K., Tetlock, Paul C., 2017. Retail short selling and stock prices. Rev. Financ. Stud. 30, 801–834.
Kraus, Alan, Stoll, Hans R., 1972. Price Impacts of block trading on the New York Stock Exchange. J. Financ. 27, 569–588.
Kuvvet, Emre, 2013. The impact of political uncertainty and abnormal market conditions on institutional trading behavior. J. Trading 8, 15–22.
Lee, Charles M.C., 1992. Earnings news and small traders: an intraday analysis. J. Account. Econ. 15, 265–302.
Lee, Charles M.C., Radhakrishna, Balkrishna, 2000. Inferring investor behavior: evidence from TORQ data. J. Financ. Mark. 3, 83–111.
Lee, Charles M.C., Ready, Mark, 1991. Inferring trade direction from intraday data. J. Financ. 46, 733–746.
Lepone, Andrew, Wong, Jin Boon, 2018. The impact of mandatory IFRS reporting on institutional trading costs: evidence from Australia. J. Bus. Financ. Account (forthcoming).
Li, Sophia Zhengzi, Schwartz-Ziv, Miriam, 2018. How Are Shareholder Votes and Trades Related? Rutgers University and Michigan State University, Working paper.
Ljungqvist, Alexander, Qian, Wenlan, 2016. How constraining are limits to arbitrage? Rev. Financ. Stud. 29, 1975–2028.
Lynch, Andrew, Puckett, Andy, Yan, Xuemin Sterling, 2014. Institutions and the turn-of-the-year effect: evidence from actual institutional trades. J. Bank. Financ. 49, 56–68.
Miller, E.M., 1977. Risk, uncertainty and divergence of opinion. J. Financ. 32, 1151–1168.
Pagano, Michael S., 2009a. International market structure: global problems and micro solutions. Int. J. Manag. Financ. 5, 5–15.
Pagano, Michael S., 2009b. Which factors influence trading costs in global equity markets? J. Trading 4, 7–15.
Puckett, Andy, Yan, Xuemin Sterling, 2011. The interim trading skills of institutional investors. J. Financ. 66, 601–633.
Saar, Gideon, 2001. Price impact asymmetry of block trades: an institutional trading explanation. Rev. Financ. Stud. 14, 1153–1181.
Schwarz, Christopher G., Potter, Mark E., 2016. Revisiting mutual fund portfolio disclosure. Rev. Financ. Stud. 29, 3519–3544.
Sias, Richard W., Starks, Laura T., 1997. J. Financ. 55, 1543–1562 Institutions and individuals at the turn-of-the-year.
Wermers, Russ, 1999. Mutual fund herding and the impact on stock prices. J. Financ. 54, 581–622.
Zheng, Lu, 1999. Is money smart? A study of mutual fund investors' fund selection ability. J. Financ. 54, 901–933.