Short-Term Trading Skill: An Analysis of Investor Heterogeneity and Execution Quality*

Mehmet Sağlam†
Lindner College of Business
University of Cincinnati
email: mehmet.saglam@uc.edu

Ciamac C. Moallemi
Graduate School of Business
Columbia University
email: ciamac@gsb.columbia.edu

Michael G. Sotiropoulos
Deutsche Bank
email: michael.sotiropoulos@db.com

This Version: September 2016

Abstract

We examine short-horizon return predictability using a novel proprietary dataset of institutional traders with known identities. We estimate investor-specific short-term trading skill and find that there is pronounced heterogeneity in predicting short-term returns among institutional investors. Incorporating short-term predictive ability, our model explains much higher fraction of variation in asset returns. Ignoring the heterogeneity in short-term trading skill has major implications in modeling price impact. We analyze the differences between trading characteristics of skilled and unskilled investors. A simple trading strategy exploiting our skill estimates yields statistically significant abnormal returns supporting the skill-based interpretation.

Keywords: Short-term Trading Skill, Price Impact, Execution Costs
JEL Classification: G11, G14, G24.

*We are grateful for helpful comments from Robert Battalio, Alex Borisov, Zhi Da, Larry Glosten, Hui Guo, Brian Hatch, Gerry Tsoukalas, Kumar Venkataraman, Haoxiang Zhu and seminar participants at Columbia University, 2012 INFORMS conference, Princeton University and University of Cincinnati. Sağlam acknowledges support from Deming Doctoral Fellowship. Moallemi acknowledges support from NSF Grant CMMI-1235023.

†Corresponding author: Mehmet Sağlam, University of Cincinnati, 2925 Campus Green Drive, Cincinnati, OH, 45221-0195, Phone: (513) 556-9108; Fax: (513) 556-0979; mail: mehmet.saglam@uc.edu

Short-Term Trading Skill: An Analysis of Investor Heterogeneity and Execution Quality

Abstract

We examine short-horizon return predictability using a novel proprietary dataset of institutional traders with known identities. We estimate investor-specific short-term trading skill and find that there is pronounced heterogeneity in predicting short-term returns among institutional investors. Incorporating short-term predictive ability, our model explains much higher fraction of variation in asset returns. Ignoring the heterogeneity in short-term trading skill has major implications in modeling price impact. We analyze the differences between trading characteristics of skilled and unskilled investors. A simple trading strategy exploiting our skill estimates yields statistically significant abnormal returns supporting the skill-based interpretation.

Keywords: Short-term Trading Skill, Price Impact, Execution Costs
JEL Classification: G12, G14, G24.

1. Introduction

There is ample evidence that excess stock returns are predictable at various horizons by macroeconomic and firm-level characteristics such as dividend-price and book-to-market ratios, short-term rates, aggregate volatility, and lagged returns. Indeed, there is ongoing active research in uncovering predictive variables or proposing trading strategies generating abnormal returns against standard asset-pricing models. This literature is motivated by the fact that investors can exploit predictable returns when making portfolio decisions dynamically. For example, Johannes et al. (2014) find statistically and economically significant benefits for investors using models of return predictability.

In addition to the documented sources of return predictability, some investors may have private information about the fundamental value of the asset. It is unlikely that this informational advantage is due to having access to non-public material information as a corporate insider but, rather, some investors may just be more skilled in processing short-term information flows to identify under- or over-valued stocks. For example, using a large set of institutional trading data, Yan and Zhang (2009) show that trades of short-term institutional investors are positively correlated with future stock returns. Similarly, Diether et al. (2009) examine daily short-sale trading activity and argue that short-sellers can detect when the asset prices deviate from their fundamental value.

These examples show that, exploiting particular return-predictive signals, some investors will be able to forecast short-term price movements. The literature on the evidence of informed trading by institutional investors mostly focuses on these types of skilled investors. However, not all investors are motivated by short-term goals. Investors will also undertake certain trading strategies that are idiosyncratically dependent on their own investment objectives, style, and horizon, and may end up employing trading strategies that are at odds with short-term return predicting signals. As an extreme case, consider a fund manager with a value-style investment view and a long-horizon performance benchmark. Asness et al. (2013) document that value and momentum signals are negatively correlated and have different time horizons, hence a value investor may systematically trade in a way that is opposite to short-term predictable returns generated through momentum effects.

In this paper, we analyze a large data set of intraday institutional trading data with masked investor identities, and decompose the performance of trades into two components: (i) an investor-specific "trading skill" component which captures the timing of the decision to trade relative to favorable or unfavorable short-term price movements; and (ii) a market impact or price impact component, which measures investor-independent execution costs. We document the presence of *skilled* short-term investors in our data set. They are identified as the ones that systematically buy (resp., sell) an asset during a period when the asset return is positive (resp., negative). In addition, we find strong evidence for the presence of *unskilled* investors in our data. These systematically decide to buy (resp., sell) the asset during a trading interval when the asset return is negative (resp., positive) on average.[1] This heterogeneity in short-term trading skill has important implications for measures of execution costs, such as the implementation shortfall, introduced by Perold (1988). Since skilled traders correctly predict short-run future returns, the cost of their trades appear high when compared to the trades of a benchmark noise trader. Similarly, because unskilled traders make trading decisions that are systematically opposite to short-term returns, the execution cost of their trades appear low when compared to a noise trader. As a result, measured execution costs may not be an unbiased estimate of the true cost of trading, which has been the crucial measure of market quality assessment in the literature. For example, a number of earlier studies (e.g., Huang and Stoll (1996), Bessembinder and Kaufman (1997a), Bessembinder and Kaufman (1997b)) utilize execution costs to compare execution quality differences between NYSE and NASDAQ. Similarly, in order to improve the transparency on market quality, the Securities and Exchange Commission (SEC) adopted Rule 605 on November 15, 2000, which requires market centers to make monthly public disclosure of certain execution costs. Given this regulatory emphasis, execution costs have also been a popular comparison metric with various changes in market structure. Brogaard et al. (2014) and Tong (2015) examine the impact of high-frequency trading on the executions costs of institutional investors. Similarly, Korajczyk and Murphy (2015) and Van Kervel and Menkveld (2015) use execution costs to study the interaction between high-frequency liquidity provision and large order institutional executions. All of these studies motivate the significance of obtaining accurate measures of execution quality for the use of brokers and policy-makers.

There is also evidence from both theoretical and empirical literature that heterogeneity in trading skill may affect the choice of selecting different venues for trading needs. Zhu (2014) and

[1] We use the terms, *skilled* and *unskilled*, to highlight the timing ability of the investor as the prices move in the same direction with his trading in a permanent basis.

Iyer et al. (2015) argue that informed traders strategically choose the lit markets for their execution needs, whereas dark pools are relatively more attractive to uninformed traders. Consequently, a naive execution cost analysis that does not take this into account may systematically suggest that (assuming all else is equal) dark pools have better execution quality. Similarly, there is strong evidence that short-term information may affect the choice of limit orders versus market orders, or the choice of a high-rebate or low-rebate trading venue (e.g., Kaniel and Liu (2006), Maglaras et al. (2012), Collin-Dufresne and Fos (2015)). Thus, an execution cost analysis across trading venues controlling for all other effects but not investor heterogeneity should similarly be systematically biased.

In this paper, we are interested in jointly estimating an investor-dependent short-term trading skill and an investor-independent measure of execution costs. To our knowledge, the effects of heterogeneous short-term predictive ability have been largely ignored in prior execution cost studies. In general, these effects are much more difficult to model *ex ante*, since short-term predictions cannot be observed directly. In a typical algorithmic trading situation, where an investor executes a large order through an algorithm provided on an agency basis by a broker, the investor rarely communicates their short-term price views directly to the executing broker. Instead, investors might implicitly express their alpha view by choosing the asset, direction, and time of the trade, and by adjusting the parameters of the broker's trading algorithm. By not accounting for short-term trading skill, any subsequent transaction cost analysis may misestimate the price impact associated with the investor's trades.

For this purpose, we propose a model to attribute the asset returns observed during the execution of a large order between the short-term predictive skill of the investor and the price impact of the resulting trades. Specifically, we consider short-term trading skill as a characteristic of the investor. Besides the usual price impact factors such as the relative size of the order, speed and volatility, the model introduces the investor's short-term predictive ability in the form of risk-adjusted performance metric as in the typical usage with Sharpe ratio. We do not impose any *a priori* grouping of investors into categories that might explain their predictive ability, such as institutional investors, quantitative funds, or retail investors. The risk-adjusted measure of short-term trading performance allows our model to capture the dependence of future price movements on the mere desire of an investor to trade a specific asset at a specific point in time.

We estimate our proposed model on a unique and proprietary historical data set consisting of a large sample of intraday equity execution data along with masked investor identifiers, obtained from a large broker who provides algorithmic trading services. We analyze our estimation results for robustness, and our contributions can be summarized as follows:

1. There is strong evidence for investor heterogeneity in short-term trading skill. We find that approximately one third of the investors are systematically skilled or unskilled relative to the rest. In other words, ability to predict short-term price changes may be a significant motivation for many investors in our sample to trade a specific asset at a specific point in time. A falsification test on bootstrapped samples provides further evidence that the numbers of skilled and unskilled investors are abnormally high.

2. Short-term trading skill significantly increases the power of the model in explaining the variation of returns relative to arrival price. In fact, including investor specific skill variables improves the R^2 of the model relative to a model that only considers the price impact of orders by an order of magnitude, from 0.5% to 10%. In other words, the identity of an investor who wishes to trade is highly predictive of future price movements relative to considering only the orders the investor places. Moreover, ignoring investor identity results in systematic misestimation of the price impact of trades. Our results are robust to alternative model specifications and can actually predict out-of-sample returns with skill estimates.

3. We analyze the trading characteristics of skilled and unskilled investors and find that skilled traders differ significantly from unskilled investors by trading (relatively) larger orders and trading more in lit markets. Unskilled traders tend to sell recent winners and buy recent losers over the past month. Both types of investors are not able to time favorable liquidity conditions. Finally, out-of-sample execution costs have major statistical dependence on our skill estimates as well. Expected execution cost difference between short-term skilled and unskilled traders is 25 bps which is economically substantial. This finding reinforces the potential problem of using execution costs as a standalone metric of trading costs by ignoring heterogeneous short-term trading skill.

4. We construct a simple out-of-sample trading strategy based on skill estimates and find that

this trading strategy generates significant abnormal returns when benchmarked against the Carhart (1997) four-factor model. In other words, our classification of skilled and unskilled traders is consistent with their ability to predict future returns in the short-term.

In current literature, there has been little evidence of the cross-sectional structure of short-term trading skill across a universe of institutional investors. Our paper proposes a methodology to identify such investor behavior. We demonstrate that short-term predictive ability is very heterogeneous among an institutional investor base. This is consistent with theoretical agent-based microstructure models where information asymmetry provides a major motivation to trade (see e.g., Easley et al. (2002)). Moreover, while the literature focuses on informed and uninformed investors, our results reveal the presence of another type that systematically places orders in the opposite direction of short-term future returns.

From a policymaker perspective, our results illustrate that mere comparison of execution costs cannot be a standalone measure of execution quality. Venues populated either with skilled or unskilled traders may have misestimated measures of execution quality if the heterogeneity in short-term trading skill is ignored. Consequently, these biased estimates may not lead to an optimal policy recommendation.

From a practical perspective, moreover, our results illustrate that incorporating short-term trading skill is important in the estimation of execution costs and, in particular, of price impact. Ignoring skill heterogeneity results in models that have both much lower predictive ability and systematically biased estimates of price impact, which may often be conflated with skill. Eliminating this bias may result in improved decision-making throughout the trading process. In the pre-trade phase, for example, more accurate transaction cost estimates will result in better portfolio construction. During trade execution, accounting for short-term trading skill will allow brokers to tailor their trading algorithms on an investor-by-investor basis and achieve better execution results. Finally, our predictive variables for short-term trading skill can be utilized in the absence of investor identities or limited execution data availability so that execution costs will be estimated more accurately in the pre-trade phase.

The rest of the paper is organized as follows: in Section 1.1, we present a brief literature review. In Section 3, we set up the underlying statistical model. Section 4 describes our experimental study, while Section 5 contains our model estimation and analysis. Sections 6 examines the robustness of our results and provide strong evidence for our interpretations with investor heterogeneity in short-term trading skill. Section 7 discusses the differences in trading styles of skilled and unskilled investors. Section 8 presents empirical evidence that the skill estimates can be utilized to generate a profitable long-short trading strategy. Finally, we conclude in Section 9.

1.1. Literature Review

Our paper is related to two main strands of the literature: studying skill in institutional trading and estimating the price impact of trading activity.

A large literature on institutional trading activity addresses the question whether institutional investors are informed. Gompers and Metrick (2001) find that there is positive relationship between institutional ownership and future stock returns. Yan and Zhang (2009) argue that this relationship is driven by short-horizon institutions. Using a more high-frequency data, Puckett and Yan (2011) find that institutional investors are skilled even after accounting for trading costs. In a more recent study using news analytics, Hendershott et al. (2015) find that institutional investors are informed and their trading direction can predict the sentiment of the future news. There are also a number of studies that document skill in the general context of fund management (e.g., Cohen et al. (2005), Kacperczyk et al. (2005), Mamaysky et al. (2008)). On the other hand, Anand et al. (2012) document that institutional trading desks have persistent trading costs – institutions that have low trading costs continue to have low trading costs over time. Our paper is related to this literature but focuses on studying the heterogeneity in short-term trading skill. In terms of documenting unskilled short-term investors, our results also resemble the underperformance of fund managers after accounting for management fees as in Wermers (2000).

The relationship between trading activity and asset prices in financial markets has been an important question in the economic microstructure literature for several decades. The theoretical origins of price impact arise from the presence of informed traders as, for example, in the celebrated models of Kyle (1985) or Glosten and Milgrom (1985). As a result, a line of literature has emerged focusing on the empirical analysis of the impact of trades on prices, motivated by the economic question of understanding the role of information asymmetry in markets. This work is nicely summarized by Hasbrouck (2007) and it is still actively pursued, see Easley et al. (2012).

More recently, however, with the rise of electronic and algorithmic trading, a new line of literature has emerged. Motivated by the concerns of practitioners, this literature focuses on the decision problem faced by an investor seeking to algorithmically spread his trades out over time, in order to minimize execution costs. A key ingredient in such algorithmic trading is the estimation of the effect of a sequence of "child" orders placed by an algorithm executing an investor's "parent" order on the asset price across future time horizons. The most notable early works here are those of Bertsimas and Lo (1998) and Almgren et al. (2005). More broadly, Bouchaud et al. (2008) provides a summary of theoretical and empirical results on models which predict the impact of trades on prices, bid-ask spreads, and other market dynamics over time. They theorize that much of these dynamics can be explained by the presence of algorithmic traders strategically spreading their orders across time.

A closely related question is the estimation of overall transaction costs for large block trades (e.g., Keim and Madhavan (1996), Almgren (2008)). These cost functions play an important role in portfolio optimization and other pre-trade analysis. Obizhaeva (2009) estimates such trading cost functions using a data set of large portfolio transitions. Kyle and Obizhaeva (2014) provide a theoretical model that seeks to explain the cross-sectional variation of trading costs across a universe of stocks. Hendershott et al. (2013) propose an approach for measuring the temporary component of the total trading cost of a large execution. Our work extends this line of inquiry by explicitly including investor identity as a predictive factor of order execution costs.

Finally, our paper is related to the analysis of transaction costs with respect to different investment strategies. Using equity executions from 21 institutional traders, Keim and Madhavan (1997) find that total trading costs for a technical-style investment strategy is higher compared with value-style investment strategy. Intuitively, they relate this finding to the differences in aggressiveness as value investors trade patiently via worked orders. Similarly, using transactions of 37 money managers, Chan and Lakonishok (1995) find that growth-oriented strategies incur higher transaction costs due to differences in their demands for immediacy. These papers mainly study the variation in execution skill across institutional invetors. Controlling for differences in trading schedules or demands for immediacy, our paper complements these studies by focusing on a fundamentally different theme. In our model, heterogeneity across investors does not stem from better trading schedules as this is ultimately controlled by the algorithm of the broker, but from differences in skill levels of timing ability. Investor's information set at the time of his trading decision determines the execution strategy which remains unchanged throughout the execution. In this context, the trading skill emerges due to the particular timing of the trade decision by anticipating favorable short-term price changes.

2. Theoretical Background

In this section, we briefly review the relevant theoretical models that guide our empirical work. Specifically, we summarize potential theories regarding the short-term price movements around large order executions.

In a survey of market microstructure, Biais et al. (2005) summarize two main competing theories for price formation which is also applicable for executions of large orders: inventory (liquidity) and information (adverse selection) paradigm. In the inventory paradigm (see e.g., Ho and Stoll (1981)), uninformed investors trade with risk-averse market makers who can control the order flow by changing their quotes. For example, when the uninformed investor buys (sells) a large number of shares for liquidity needs, market makers are forced to a net short (long) position deviating from their preferred inventory positions. In this case, they raise (lower) their bid and ask quotes in order to bring their inventory back to their preferred position which leads to increase (decrease) in mid-quote prices. In the information-based trading paradigm (see e.g., Glosten and Milgrom (1985) and Kyle (1985)), an investor trades a large order due to his private information about the fundamental value of the asset and thus the market maker accounts for the information content of the order and set his quotes accordingly. In this framework, prices are formed according to the expectations of the value of the asset conditioned on the realized order flow and consequently buy (sell) orders imply higher (lower) valuation and increase (decrease) equilibrium prices. Another difference between these theories is their different implication on the transitory or permanence of the the price impact. Liquidity-based trading causes temporary price impact whereas information-based trading moves the prices permanently.

The literature mostly focuses on two types of traders, informed and liquidity. However, information based-trading models can be generalized to accommodate investors with differing beliefs. For example, Easley et al. (2002) propose an information-based theoretical model in which there is investor heterogeneity with regards to their beliefs of expected returns and correctness of these priors. Therefore, it is also possible to observe unskilled investors who are systematically deciding to buy (sell) the asset during a trading interval when the asset return is negative (positive) on average due to heterogeneous beliefs. Their trades can be profitable in the long-term but they can be subject to short-term losses.

Another reason for unskilled trading can emerge from behavioral biases in decision-making. Based on findings of prospect theory, Shefrin and Statman (1985) document the investor's tendency to sell winners too early and ride losers too long. They refer to this behavior as "the disposition effect." Empirical evidence suggests that this behavior is more pronounced for individual investors. This theory suggests that a liquidity trader prone to disposition effect will sub-optimally sell (buy) a stock if it has recently realized large positive (negative) returns and as the asset value continues to increase (decrease), the liquidity trader will be subject to short-term losses.

These theories suggest that there may be investor-specific short-term trading skill originating from information-based trading (correct beliefs) or insensitivity to behavioral biases. Furthermore, considering the inventory paradigm, there is also investor-independent measure of price impact that serves as market maker's compensation for inventory risk. Thus, it is an empirical question to determine the most important drivers of short-term price movements during a large execution among these potential factors of price impact.

3. The Model

We consider a population of J investors sending a total of N orders to an executing broker. The mapping $i \xrightarrow{c} j$ identifies order i, $i = 1 \dots N$ as belonging to investor $j = c(i)$, with $j = 1 \dots J$. Each order is for a quantity of Q_i shares of an asset, with $Q_i > 0$ ($Q_i < 0$) for buy (sell) orders, respectively. Each order also has an execution duration of T_i, measured as a fraction of the trading day. We define the participation rate $\rho_i \triangleq |Q_i|/V_i$, where V_i is the total market volume traded within the interval T_i. The arrival price $P_{i,0}$ is the last traded price prior to the order's arrival and the terminal price is the last execution price P_{i,T_i}. In our model, given order i, we consider the expected return of the asset over the execution interval T_i, that is, $\log(P_{i,T_i}/P_{i,0})$. We posit that this return is driven by two predictable effects. The first effect is *short-term trading skill.* Some investors will be able to predict short-term asset returns using models of return predictability or correct beliefs. Thus, we expect this impact to be persistent in the short-term. Conditional on the arrival of a buy (resp., sell) order i, we expect a return in the asset price of $\alpha_{c(i)}\sigma_i\sqrt{T_i}$ over the execution interval T_i. Here, the coefficient $\alpha_{c(i)}$ represents the short-term predictive ability of investor $c(i)$ and σ_i is the daily volatility of the mid-quote of the asset price, typically estimated as an average of daily volatilities over the prior month. $\alpha_{c(i)}$ can be positive, zero, or negative which can be interpreted as skilled, unidentified or unskilled respectively. Note that this predictive ability is parameterized in a risk-adjusted fashion, i.e. we assume that each particular investor has a constant short-term Sharpe ratio or information ratio over all of the trades.[2]

The second effect is *price impact*, or, the direct effect of the trades placed on behalf of the investor. The price impact for order i is given by $\lambda\sigma_i\sqrt{T_i}h(\rho_i)$. Here, λ is a (broker-specific) price impact coefficient. We normalize the price impact component with the volatility of the asset during execution horizon captured by $\sigma_i\sqrt{T_i}$ so that we represent the impact as a fraction of the typical movement of the stock return. Transaction cost models sometimes include a bid-offer spread term to incorporate stock-specific liquidity costs. However, since we are modeling *returns* over the execution horizon as opposed to the average cost, we did not include spreads in our baseline specification. In Section 6.3, we include an additional spread component in the price impact specification and our findings remain largely unchanged.[3] Our price-impact assumption is consistent with the literature (e.g., Almgren et al. (2005)). On the theoretical front, Keim and Madhavan (1996) derive that price impact is a concave function of the trade size. Similarly, Chacko et al. (2008) also find that expected price impact is proportional to the volatility and empirically validates this claim.

The price impact function $h(\cdot)$ captures the effect of the participation rate (or trading speed)

[2]Here, we ignore the role of a benchmark risk-free return in the definition of a Sharpe ratio, i.e. we do not consider excess returns. This is reasonable since the risk-free return is effectively zero over the intraday time horizons of interest.

Further, note that we do not scale Sharpe ratio with the square root of the investment horizon (as typically done with longer horizons in asset management). In Section 6.2, we explore an alternative specification where we define $\alpha_{c(i)}$ to be a daily Sharpe ratio that is scaled by the square root of the length of the execution horizon, and find that both models provide very similar findings.

[3]In Section 6.4, we also consider stock fixed effects and obtain very similar findings due to our high-liquid stocks in the data.

on price. The idea here is that orders executed with a higher participation rate will have a larger price impact. In order to illustrate the robustness of our results with respect to our price impact formulation, we will consider two explicit forms for the price impact function: a linear price impact function, i.e.,

$$h(\rho_i) \triangleq \rho_i,$$

or a square root price impact function, i.e.,

$$h(\rho_i) \triangleq \rho_i^{1/2}.$$

The choice of sublinear price impact has been extensively studied both theoretically and empirically. There is a long line of literature supporting the choice for a square root price impact law. For example, Chacko et al. (2008) provides empirical evidence that the expected price impact is proportional to the square root of the quantity traded. Using a large sample of US equity trades, Almgren et al. (2005) also estimate the exponent to be very close to 0.5. This exponent is also consistent with the well-known Barra model for market impact costs outlined in Torre and Ferrari (1998).

Putting everything together, we assume that the sign-adjusted log-return of an order relative to the arrival price and over the execution horizon can be expressed as an additive model of the form

$$\operatorname{sgn}(Q_i) \log\left(\frac{P_{i,T_i}}{P_{i,0}}\right) = \alpha_{c(i)}\sigma_i\sqrt{T_i} + \lambda\sigma_i\sqrt{T_i}h(\rho_i) + \epsilon_i\,, \tag{1}$$

with ϵ_i having a mean of zero and variance of ν_i^2.

Explicit forms of the impact function $h(\cdot)$ fully specify the model as a linear regression of the risk-normalized interval return against short-term trading skill and broker impact. For example, with a linear price impact function we obtain

$$\operatorname{sgn}(Q_i) \log\left(\frac{P_{i,T_i}}{P_{i,0}}\right) = \beta_0 + \sigma_i\sqrt{T_i}\sum_{j=1}^{J}\mathbb{I}_{\{c(i)=j\}}\alpha_{c(i)} + \lambda\sigma_i\sqrt{T_i}\rho_i + \epsilon_i\,, \tag{2}$$

with $\mathbb{I}$ the indicator function. Likewise, a square-root impact function leads to the model

$$\operatorname{sgn}(Q_i) \log\left(\frac{P_{i,T_i}}{P_{i,0}}\right) = \beta_0 + \sigma_i \sqrt{T_i} \sum_{j=1}^{J} \mathbb{I}_{\{c(i)=j\}} \alpha_{c(i)} + \lambda \sigma_i \sqrt{T_i} \rho_i^{1/2} + \epsilon_i \,. \tag{3}$$

When fitted on a historical sample of short-term execution returns, the above models (2) and (3) identify the short-term trading of each investor j, along with the impact coefficient.[4] We proceed by presenting our data set and model estimation results in Section 5.

4. Data

For our empirical study we use a novel proprietary execution data from the historical order databases of a large investment bank ("The Bank") which is one of the top five electronic trading brokers in the US by market share. The orders originate from a diverse pool of investors, such as institutional portfolio managers, quantitative investment funds, internal trading desks and retail customers. Our data set consists of two widely used algorithmic execution strategies, the volume weighted average price (VWAP) and the percentage of volume (POV). These algorithms collectively constitute roughly 75% of all execution strategies employed by The Bank. The VWAP algorithm aims to achieve an average execution price that is as close as possible to the volume weighted average price over the execution horizon. The main objective of the POV algorithm is to have constant participation rate in the market within the execution interval. VWAP and POV have relatively small discretion on opportunistically speeding up or slowing down the execution so the aggressiveness of the execution is mainly controlled by the investor choosing a particular urgency level in the pre-trade phase. With VWAP and POV algorithms, we eliminate any potential broker-specific effects such as the usage of trading signals and market events that drive more opportunistic algorithms. Furthermore, we also avoid any biases that can occur with endogenous selection of the algorithm itself. For example, if we were to have more sophisticated algorithms in our dataset,

[4] The above models could also be expressed in terms of the arrival slippage $\log\left(\bar{P}_i/P_{i,0}\right)$ instead of the interval return $\log\left(P_{i,T_i}/P_{i,0}\right)$, where $\bar{P}_i$ is the average execution price of the i-th order. This would lead to an approximate rescaling of the coefficients $\alpha_{c(i)}$ and λ by a factor of $1/2$. The execution algorithms considered here trade at constant participation rate. Therefore, the execution price $\bar{P}_i$ is close to the realized interval VWAP, and for a price path with a constant drift, the VWAP return is half of the interval return.

one might argue that skilled traders may actually be just better in choosing algorithms. However, with VWAP or POV, an investor can only have superior short-term trading skill by starting the execution at a particular time.

This proprietary data set provides a rich set of attributes. For each order i we have access to the following: investor identity tag,[5] $c(i)$, ticker of the traded stock, order size, Q_i, order side (buy/sell), $\operatorname{sgn}(Q_i)$, execution duration, T_i, participation rate, ρ_i, average volatility of the stock over the last 20 trading days, σ_i, the percentage return over the execution interval, $P_{i,T_i}/P_{i,0} - 1$. These data allows us to fully estimate the model of Section 3. In addition, our data include the daily, average (i.e., over the last 20 trading days[6]) and interval (i.e., during the execution horizon) proportional bid-ask spread, mid-quote volatility and traded volume for each stock.

We use a restricted subset of the execution data, defined by the following selection criteria:

- The trading period is from January 2011 to June 2012, inclusive.
- The asset universe consists of the S&P 500 stocks. We focus on highly liquid stocks to focus on the differences on short-term predictive ability as a result of following certain set of strategies. For this set of stocks, it is hard to have an investor trading on an insider information.
- Orders come from active investors only: an investor is considered active if he has at least 100 and at most 500 orders within the period of study. This is to prevent any specific investor from having major influence on our results.
- All orders have been fully filled without intermediate replacements or cancellations.
- The execution duration is greater than 5 minutes but no longer than a full trading day, with no participation in opening/closing auctions. We exclude executions that last less than 5 minutes to avoid any short term effects from market orders and auctions.

Using the above criteria, our final sample consists of 63,379 executions coming from 30,438 buy and 32,941 sell orders.[7] The trading algorithms used are 41,339 VWAP and 22,040 POV. The orders

[5]Investors are identified by numerical aliases to protect anonymity.

[6]Throughout the analysis we will refer to this measure as the average over the past-month for conciseness as there are 21 trading days in each month on average.

[7]On average, each parent-order has approximately 120 child-order executions so the total number of trades are roughly 10 million.

came from a set of 293 active investors, with 216.3 orders per investor and 168 orders per trading day on average. The highest number of executions on a single stock is 454 which corresponds to 0.71% of all executions. Table I provides additional summary statistics for our data sample.

[Insert Table I here]

The average percentage return realized during the execution interval is 0.3 bps. We observe that bid-ask spread and volatility lay in a tight range. More than half of the executions have a bid-ask spread between 2 bps and 5 bps and a mid-quote annualized volatility between 15% and 27%. The mean duration of the executions is a little less than 2.5 hours. Finally, we have a wide range of participation rates across executions with an average (median) of 6.44% (1.59%).

5. Model Estimation Results

We analyze the execution data using our full model with two price impact specifications, linear price impact as in (2) and square root price impact as in (3). We also fit a reduced model to the data by ignoring trading skill terms, i.e., $\alpha_j = 0$,

$$\operatorname{sgn}(Q_i) \log\left(\frac{P_{i,T_i}}{P_{i,0}}\right) = \beta_0 + \lambda^{\mathsf{base}} \sigma_i \sqrt{T_i} \rho_i^{\gamma} + \epsilon_i \,, \tag{4}$$

where $\gamma = 1$ for the linear model and $\gamma = \frac{1}{2}$ for the square-root model. We use the superscripts, base, to emphasize the difference between the reduced and full models. We are concerned with heteroscedasticity, contemporaneous correlation across stocks, and auto-correlation within each stock and adjust our standard errors by clustering on calendar day and stock throughout the analysis as suggested by Petersen (2009).

In this section, we first present the differences between the reduced and the full model with respect to the price impact coefficient and adjusted R^2. Then, we illustrate the cross-sectional distribution of short-term trading skill across investors. Finally, we use bootstrap analyses to examine the statistical significance of the biased price impact coefficient in the reduced model and the presence of short-term trading skill.

5.1. Price Impact Coefficients and R^2

Table II summarizes the regression results from the full and the reduced model.

[Insert Table II here]

This table suggests a number of interesting observations. First, consider the price impact parameter, λ. In all cases, the estimate of λ is statistically significant. The square-root model in the absence of trading skill corresponds to the well-known Barra market impact model as outlined in Torre and Ferrari (1998). Here, our estimate of λ is of order unity — this is consistent with the prior literature. Yet, we observe that without accounting for investor's short-term trading skill level, price impact by itself has a very low explanatory power with the maximum adjusted R^2 of 0.52% from both models. The linear price impact model has relatively better fit than the square root price impact model, even though the difference is negligible. However, the inclusion of the short-term trading skill term substantially increases the goodness of fit, leading to an adjusted R^2 of approximately 10%. This significant difference illustrates that the variation in short-term returns can be explained much better when the systematic short-term trading skill of the investor is acknowledged.[8]

Moreover, if we ignore the predictive abilities of the investors, we observe that price impact is misestimated. If the price impact is linear in participation rate, then accounting for alpha view of the investors reduces the price impact coefficient by approximately 20%. This is also observed for the square-root model, but to a lesser degree. This result suggests that in our sample short-term skilled trader activity is relatively higher when compared with unskilled trader activity. The standard errors do not allow us to conclude that the difference between $\hat{\lambda}$ and $\hat{\lambda}^{\mathsf{base}}$ is statistically different. For this reason, we use a bootstrapping analysis in Section 5.3 to formally test for the statistical significance of the bias introduced in the absence of skill terms.

In summary, we observe that accounting for investor heterogeneity in short-term trading skill explains much higher variation of asset returns during an execution. We observe that the usual practice of ignoring investor specific view may introduce a systematic bias in price impact estimates. We further investigate this implication in the context of execution costs in Section 7.3, and find

[8]To highlight the economic significance of this increase, it is worth to note that 500 stock dummies move the adjusted R^2 by mere 0.1% in our analysis in Section 6.4.

that the standard measure of execution cost, implementation shortfall, also depends on short-term predictive ability with statistical significance.

5.2. Heterogeneity in Short-Term Trading Skill

We observe that there is significant investor heterogeneity in predicting short-term returns. We label an investor as *skilled* (resp., *unskilled*) if his short-term trading skill estimate is positive (resp., negative) and is statistically significant under the 10% level. We label the remaining investors as *unidentified*. Table II reports that with the linear price impact model, 49 out of 293 investors are *skilled*, while 48 investors are *unskilled* and 196 investors are *unidentified*. With the square root price impact model, the alpha estimates slightly drop as the square-root model puts more weight on the price impact unilaterally. Figure 3 illustrates this drop in Sharpe ratio estimates graphically. In this case, we obtain that 36 investors are *skilled* whereas 63 are *unskilled*. In both models, approximately one third of our investor universe are *skilled* or *unskilled*.

Evaluating short-term trading skill for a large number of investors can be viewed as a multiple hypothesis testing problem. By random chance, some investors may appear to have significant α coefficients. Consequently, the number of *skilled* or *unskilled* investors may not be statistically significant in our dataset as we do not know the true distribution of these statistics under the null hypothesis. Therefore, we employ a bootstrap analysis in Section 5.3 to formally test the statistical significance of the number of *skilled* or *unskilled* investors.

[Insert Figure 1, Figure 2, and Figure 3 here]

We now discuss the magnitudes of the estimated short-term predictive abilities. Figure 1 shows histograms of investor skill estimates both for the linear and square root price impact model specifications, including statistically insignificant estimates. Figure 2 shows the histograms of the alpha estimates which are statistically significant. We observe that estimates which are small in absolute value are likely to be insignificant. We report the investor short-term trading skill estimates as annualized Sharpe ratios. In the linear model, the range of the estimated Sharpe ratios is between -27.7 and 14.3 with the sample mean (median) of -0.59 (-0.40). We find that the distribution of skill estimates have negative skew and large kurtosis suggesting that the distribution of the skill estimates are asymmetric and non-normal. The Sharpe ratio estimates arising from our models are much larger than typical Sharpe ratios observed in the traditional asset management industry. However, current empirical literature reports similar Sharpe ratio estimates over short, intraday investment horizons (e.g., for high-frequency traders). For example, Clark-Joseph (2013) estimates that annualized Sharpe ratios of high-frequency traders are in the neighborhood of 10 to 11. Baron et al. (2013) report that the average high-frequency trader Sharpe ratio in their data set is 9.2.

Our findings suggest that, at the individual investor level, there is substantial variation in short-term predictive ability. Roughly, one third of the investors have statistically significant Sharpe ratio estimates. On the other hand, we observe that half of these investors make systematically wrong bets in the short-term. Note that these statistics do not imply that investors in this group are losing money in the long-run. These investors are possibly long-term investors who are not exploiting the short-term predictability of asset returns. We further investigate the source of unskilled trading in Section 7.4 and Section 7.5 with regards to disposition effect and long-term objectives, respectively.

5.3. Bootstrap Analyses

In this section, we run two bootstrap analyses to formally assess whether price impact coefficients estimated from the full and reduced model are statistically different and whether the number of *skilled* or *unskilled* investors are statistically significant and hence the heterogeneity in short-term trading skill exists.

[Insert Table III here]

In the first bootstrap test, we randomly sample our data and construct $1,000$ datasets each with $10,000$ executions. For each dataset, we estimate the full and the reduced models and store the price impact coefficients, $\hat{\lambda}$ and $\hat{\lambda}^{\mathsf{base}}$, for both linear and square-root price impact specifications. Table III reports the mean difference of $\hat{\lambda} - \hat{\lambda}^{\mathsf{base}}$ and its corresponding standard error. We find that the difference between $\hat{\lambda}$ and $\hat{\lambda}^{\mathsf{base}}$ is highly statistically significant in both price-impact specifications.

[Insert Table IV here]

In the second bootstrap analysis, we create $10,000$ different samples of our execution dataset by randomly permuting the investor identifiers across executions. Each investor has still the same number of assigned executions and the total number of executions remain the same but each investor is now assigned a random selection of executions. This bootstrap procedure allows us to generate the empirical distribution of the number of *skilled* and *unskilled* traders under the null hypothesis that investor identifiers are unrelated to log-returns during the execution horizon.

For each randomly sampled dataset, we estimate our full model with skill terms and compute the number of the number of *skilled* or *unskilled* investors. We then derive the empirical distribution for the desired parameters: the numbers of *skilled* and *unskilled* investors. Table IV illustrates the p-values of our original estimates of the numbers of *skilled* and *unskilled* investors with respect to the empirical distribution. In both linear and square-root models, we find strong evidence that the estimated number of *skilled* or *unskilled* investors is indeed abnormally high suggesting that investor heterogeneity is present with statistical significance. Finally, this analysis also shows that, the estimated price impact coefficient, $\hat{\lambda}$, is also statistically different than what we would obtain under the null hypothesis that investor identities do not matter.

6. Robustness Tests

In this section, we assess the robustness of our results in six ways. First, in order to control for the possibility of over-fitting, we assess the robustness of our model predictions on out-of-sample data. Using our estimated model in-sample, we predict out-of-sample short-term returns over an execution. Second, we explore a different alpha specification that scales with the square root of the execution horizon. Third, we consider an alternative price impact specification with an additional spread component. Fourth, we add stock fixed effects to control for the cross-sectional variation across stocks. Fifth, we test the robustness of our findings when execution horizon is largely the same across orders and is not in the subset of slowest or fastest executions. Sixth, we consider an alternative specification that incorporates the market return over the execution time horizon.

6.1. Out-of-Sample Predictions

Our model specifications in equations (2) and (3) contain a number of parameters, namely, one for each investor. In order to eliminate the possibility of over-fitting, in this section, we consider a cross-validation experiment that illustrates the ability of our model to predict out-of-sample execution returns.

First, we divide the data into two parts: in-sample data and out-of-sample data. We perform this by randomly allocating half of each investor's executions into the in-sample data set and the remaining ones into the out-of-sample data set. We then estimate the model parameters by running the regressions specified in equations (2) and (3) using only the in-sample data.

[Insert Table V and Figure 4 here]

Table V illustrates the regression results for the in-sample data set. The estimated regression coefficients for price impact are very similar to those obtained using all the data. For example, using the linear price impact specification, the price impact estimate, $\hat{\lambda}$, is 1.79 whereas using the complete data, the estimate is 1.81. Similarly, we observe that investor skill estimates are also very stable. Figure 4 compares skill estimates between the in-sample and the complete data sets. In both price impact models, these are very close to each other, implying the robustness of the estimates. Formally, we find that the correlation between the skill estimates is 77% in both price impact models.

[Insert Table VI here]

Using the skill and price impact estimates obtained from the in-sample data, we can test whether our model can explain out-of-sample execution returns. Table VI illustrates the root mean squared prediction error (rMSPE) and R^2 estimates both in-sample and out-of-sample. We observe that in-sample and out-of-sample mean-squared errors are very close to each other. Similarly, we obtain an out-of-sample R^2 of more than 8.1% in both price impact models suggesting that our regression model does not suffer from over-fitting. Both of these findings emphasize that our model has out-of-sample predictive power.

6.2. Robustness in Alpha Specification

Sharpe ratio is typically defined over a reference time horizon, and is often scaled with the square root of the investment horizon when it is projected across different horizons. In our model, however, we assumed that the Sharpe ratio is held constant, independent of the execution horizon. We can also consider the alternative model. In this specification, investor j has an expected return of

$\sigma_i \alpha_j T_i$ when he trades ith stock during the trading horizon, T_i. Consequently, our skill estimation models are given by

$$\operatorname{sgn}(Q_i) \log\left(\frac{P_{i,T_i}}{P_{i,0}}\right) = \sigma_i T_i \sum_{j=1}^{J} \mathbb{I}_{\{c(i)=j\}} \alpha_{c(i)} + \lambda \sigma_i \sqrt{T_i} \rho_i^{\gamma} + \epsilon_i\,, \tag{5}$$

where $\gamma = 1$ for the linear model and $\gamma = \frac{1}{2}$ for the square-root model.

[Insert Table VII here]

Table VII illustrates the regression results for the model presented in Equation 5. The estimated regression coefficients for price impact are very similar to those obtained in the original model. For example, using the linear price impact specification, the price impact estimate, $\hat{\lambda}$, is 1.875 whereas in our original specification, the estimate is 1.811. Similarly, we find that the sets of *skilled* and *unskilled* investors from both models are very similar. For example, we find that the exact same set of 40 (resp., 46) investors are identified as *skilled* (resp. *unskilled*) in both models. Collectively, this common group nearly constitutes 90% of the original set of the identified traders suggesting that our results are robust to the choice of alpha specification.

6.3. Spread Component in Price Impact

We can generalize our price impact term by including a spread component. Due to a liquidity premium, the price impact during an execution may be higher for a less liquid stock, keeping all else equal. To test the robustness of our results with respect to a spread component, we explore an alternative specification with an additional independent variable controlling for the spread. If the model with the spread component explains much higher variation without the skill terms, accounting for short-term trading skill may lose its attractiveness. Formally, we have the following skill estimation models:

$$\operatorname{sgn}(Q_i) \log\left(\frac{P_{i,T_i}}{P_{i,0}}\right) = \sigma_i \sqrt{T_i} \sum_{j=1}^{J} \mathbb{I}_{\{c(i)=j\}} \alpha_{c(i)} + \lambda \sigma_i \sqrt{T_i} \rho_i^{\gamma} + \delta S_i + \epsilon_i\,, \tag{6}$$

where S_i denotes the time-weighted average bid-offer spread over the course of the execution, $\gamma = 1$ for the linear model and $\gamma = \frac{1}{2}$ for the square-root model.

[Insert Table VIII here]

Table VIII summarizes the results of these regressions along with the reduced models where we drop the skill terms. We observe that our conclusions with the original model remain unchanged. Spread parameter is only significant (at the 10% level) in the models augmented with short-term trading skill. Inclusion of the spread component do not change the R^2 in the models with and without short-term trading skill. The universe of *skilled* and *unskilled* traders and the price impact coefficient also stay largely unchanged. These results also illustrate that skilled traders are not particularly trading illiquid stocks which could have been alternative explanation in the absence of this analysis.

[Insert Table IX here]

We also provide robustness check for our bootstrap analysis that compares regression coefficients in the presence and absence of skill terms. We again randomly sample our data and construct $1,000$ datasets each with $10,000$ executions. For each dataset, we estimate the full model in Equation 6 and the reduced model in which we ignore the skill terms. We store the regression coefficients, $\hat{\lambda}$, $\hat{\lambda}^{\mathsf{base}}$, $\hat{\delta}$, and $\hat{\delta}^{\mathsf{base}}$ using the linear price impact model. Table IX reports the mean differences and their corresponding standard errors. We find that both coefficients are statistically different when compared between the full and reduced models. This finding implies that additional explanatory variables of price impact may also be biased when trading skill terms are ignored.

6.4. Fixed Stock Effects

We can also generalize our specification by incorporating stock fixed effects. If certain characteristics of stocks are highly correlated with price movements, and investors are consistently trading these particular stocks, skill could be spuriously assigned to investors. To mitigate this potential concern, we re-run our regression with stock dummies:

$$\operatorname{sgn}(Q_i)\log\left(\frac{P_{i,T_i}}{P_{i,0}}\right) = \sigma_i\sqrt{T_i}\sum_{j=1}^{J}\mathbb{I}_{\{c(i)=j\}}\alpha_{c(i)} + \lambda\sigma_i\sqrt{T_i}\rho_i^{\gamma} + + \sum_{k=1}^{S}\gamma_k\mathbb{I}_{\{s(i)=k\}} + \epsilon_i\,, \tag{7}$$

where we use the mapping $i \xrightarrow{s} k$ to identify the executed stock k, $k = 1\ldots S$.

[Insert Table X here]

Table X summarizes our findings. We observe that in both models, the numbers of skilled and unskilled investors are largely unaffected along with the price impact coefficients. Furthermore, we find that including 500 stock dummies moves the adjusted R^2 by mere 0.1%. This negligible increase contrasts with the substantial explanatory power of 293 investor skill coefficients which resulted in 10% increase. Overall, these findings do not provide any evidence that skilled investors are trading particular set of illiquid stocks.

6.5. Choice of Execution Horizon

The investor has the potential to control the execution horizon by adjusting the urgency parameters of the trading algorithms. Investors having short-term horizon may be more inclined to choose faster executions compared with long-horizon investors who are satisfied with complete one-day executions. In other words, the choice of execution horizon may be endogenous.

Summary statistics in Table II also show that fifty percent of the executions are either less than 15 minutes or larger than 325 minutes. To address the concern that the choice of execution horizon may drive our findings, we consider a subset of executions for which execution horizons are very similar and do not fall into the fastest and slowest category. For this purpose, we construct another sample data where execution horizon is between 156 and 234 minutes (between 40% and 60% of a full trading day). We obtain 1651 executions submitted by 39 distinct investors.

[Insert Table XI here]

Table XI illustrates the regression results using this subset of data. Our earlier conclusions regarding increased explanatory power with investor identity and heterogeneity in short-term trading skill also emerge with this substantially different data set. Without accounting for short-term trading skill, R^2 values are very close to zero whereas incorporating skill terms lead to adjusted R^2 values of 16% in both price impact specifications which is again an order of magnitude difference. We also observe pronounced heterogeneity in short-term trading skill with roughly 50% of the investors being either *skilled* or *unskilled*. Finally, we note that in both models, price impact coefficients differ significantly when short-term trading skill is taken into account. Our findings suggest that earlier findings from the original data set remain unchanged qualitatively when using the restricted data set, in which execution horizon is no longer a strategic choice and is taken to be roughly one half of a trading day.

6.6. Accounting for Market Returns

Our model uses raw execution returns as dependent variables as outlined in equations (2) and (3). Given that active fund managers are evaluated against market-driven benchmarks, we can adjust our model specifications with benchmarked returns. For this reason, we explore an alternative specification in a one factor asset-pricing model dealing with abnormal returns, expressed as the difference between the execution return of a single asset and the market return. Consequently, this adjusted model illustrates that short-term trading skill can also be quantified in a benchmark setting.

In the following regressions, we check whether such an abnormal return specification results in different findings with respect to significance of short-term trading skill and price impact. Formally, we have the following specification:

$$\operatorname{sgn}(Q_i)\log\left(\frac{P_{i,T_i}}{P_{i,0}}\right) - r_i = \sigma_i\sqrt{T_i}\sum_{j=1}^{J}\mathbb{I}_{\{c(i)=j\}}\alpha_{c(i)}^{\mathsf{mkt}} + \lambda^{\mathsf{mkt}}\sigma_i\sqrt{T_i}\rho_i^{\gamma} + \epsilon_i, \tag{8}$$

where r_i denotes the log-return of the S&P 500 index (excluding dividends) over the date of the execution, scaled proportionally with the length of the execution duration, $\gamma = 1$ for the linear model and $\gamma = \frac{1}{2}$ for the square-root model. We use the superscripts, mkt, to emphasize the difference between these and original regression models.

[Insert Table XII here]

Table XII summarizes the results of these regressions. When returns are benchmarked against the market return, we observe that the number of skilled (unskilled) traders decreases (increases) when compared with our original model. This result is also consistent with the traditional performance measurement studies in the mutual fund literature as in Wermers (2000). We observe that accounting for market returns does also increase the goodness of fit slightly with an adjusted R^2 of 11.8%. Our main result from this analysis is the sustained heterogeneity in the short-term trading skill. The results support our earlier findings regarding the significance of short-term trading skill and price impact.

7. How do Skilled and Unskilled Investors Trade?

In this section, we would like to uncover the differences in trading styles of skilled and unskilled investors. In each analysis, we do not claim causality but we are interested in identifying associations in out-of-sample data. We first identify *skilled* and *unskilled* investors by estimating our full model using the in-sample data. Consistent with our definition in Section 5, we label an investor as *skilled* (resp., *unskilled*) if his short-term trading skill estimate is positive (resp., negative) and is statistically significant under 10% level. We then label the executions in the out-of-sample data with dummy variables to account for *skilled* and *unskilled* investors. IsSkilled_i (resp., IsUnskilled_i) takes a value of 1 if the ith execution is sent by a *skilled* (resp., an *unskilled*) investor. Thus, our empirical design allows to cleanly evaluate the trading differences between these two types of investors.

7.1. Timing Liquidity

There are a few theoretical studies that show that investors who are informed about the (long-term) fundamental value of the asset may choose to trade during high episodes of liquidity (see e.g., Admati and Pfleiderer (1988) and Collin-Dufresne and Fos (Forthcoming)). We test whether investors with short-term trading skill differ from unskilled investors in this respect. Using different liquidity measures, L_i, we can run the following regression model with control variables and stock fixed effects to formally test whether out-of-sample liquidity measures have statistical dependence on the type of the investor:

$$L_i = c_0 + \beta_s IsSkilled_i + \beta_u IsUnskilled_i + \sum_j c_j \, ControlVariables_j + \sum_{k=1}^{S} \gamma_k \mathbb{I}_{\{s(i)=k\}} + \epsilon_i. \quad (9)$$

Our liquidity measures include proportional spread, logarithm of share volume, turnover, and Amihud Illiquidity measure (*ILLIQ*) realized during the execution period. Higher values of spread and *ILLIQ* indicate lower liquidity whereas higher share volume and turnover tend to be correlated with higher liquidity.

[Insert Table XIII here]

Table XIII summarizes the estimated models. Overall, we do not observe a clear pattern between favorable liquidity conditions and skill level. All of the coefficients of the skill dummies are insignificant at 5% level. The signs of the coefficients are also mixed. With respect to share volume and turnover, the coefficient on the *skilled* investors indicates higher liquidity whereas for the remaining spread and *ILLIQ*, the sign of the coefficient signals lower liquidity. The coefficient on the *unskilled* investors matches with that of *skilled* investors except in the case of *ILLIQ*. Moreover, the linear hypothesis, $H_0 : \beta_s = \beta_u$, testing for the formal difference between the trading styles of *skilled* and *unskilled* investors returns insignificant test statistic in each liquidity measure. Overall, we do not find strong evidence of liquidity timing that is correlated with skill level.

7.2. Execution in the Dark Pools

As we mentioned earlier, skilled investors may also differ from their unskilled counterparts with regards to participation in dark pools. As illustrated in Zhu (2014), execution risk in the dark pools disincentivizes informed traders to send their orders to dark pools. This uncertainty in execution emerges as there may be additional informed orders accumulating at the same side of the market and compete for execution with the investor's order.

Our dataset allows us to verify this theoretical conjecture cleanly as the investors may opt out of dark venue executions in the pre-trade phase by marking a check box. We have information about the venue of child-order executions for roughly 20% of the out-of-sample data. Using this data set, we can compute the ratio of executed shares in the dark pools. We let DP_i be the percentage of the shares traded in the dark pools for the ith execution. We fit the following regression model with stock dummies and control variables to formally test whether the allocation to dark pools has statistical dependence on the skill level of the investor in out-of-sample executions:

$$DP_i = c_0 + \beta_s IsSkilled_i + \beta_u IsUnskilled_i + \sum_j c_j \, ControlVariables_j + \sum_{k=1}^{S} \gamma_k \mathbb{I}_{\{s(i)=k\}} + \epsilon_i. \qquad (10)$$

We consider different sets of execution-level control variables consisting of participation rate, interval bid-offer spread, interval mid-quote volatility, and interval turnover.

[Insert Table XIV here]

Table XIV summarizes the regression results. In each column, we observe that the coefficient on *skilled* investors is negative and significant at 5% level. We observe that the dark pool usage of unskilled investors is not statistically significant. With respect to skilled investors, unskilled investors use dark pools more often and this additional allocation is roughly 5%. Considering that the unconditional average of dark venue usage is at approximately 10%, this difference is economically significant. Our Wald tests indicate that this difference is also statistically significant at 5% when all of the control variables are included. Overall, our findings are consistent with the theory presented in Zhu (2014). This result also provides empirical support to earlier conjecture that market venues may significantly differ in terms of the proportion of *skilled* and *unskilled* investors.

7.3. Out-of-Sample Execution Costs

Our main analysis illustrated that the price impact coefficients are biased if investor's short-term trading skill is ignored. Our model implies that traditional measures of execution cost will also suffer from the same bias in the presence of systematic short-term predictive ability. Controlling for execution characteristics, our model predicts that execution costs of *skilled* (resp., *unskilled*) short-term investors will be higher (resp., lower).

In order to explore this hypothesis, we use implementation shortfall (IS) as a measure of execution cost as introduced by Perold (1988). IS is the widely preferred measure of trading cost for institutions and has been frequently employed in the literature to proxy institutional trading cost. It is computed as the normalized difference between the average execution price and the price of the asset prior to the start of the execution. Formally, the IS of ith execution in our data is given by

$$\mathrm{IS}_i = \mathrm{sgn}\left(Q_i\right) \frac{P_i^{\mathrm{avg}} - P_{i,0}}{P_{i,0}}, \tag{11}$$

where P_i^{avg} is the volume-weighted execution price of the parent order. We fit the following regression model with stock fixed-effects and control variables to formally test whether out-of-sample execution costs have statistical dependence on the type of the investor:

$$IS_i = c_0 + \beta_s IsSkilled_i + \beta_u IsUnskilled_i + \sum_j c_j ControlVariables_j + \sum_{k=1}^{S} \gamma_k \mathbb{I}_{\{s(i)=k\}} + \epsilon_i. \qquad (12)$$

We consider execution-level control variables including participation rate, interval bid-offer spread, interval mid-quote volatility, execution duration, and average bid-offer spread and mid-quote volatility over the past month.

[Insert Table XV here]

Table XV reports the estimated coefficients of the models with different sets of control variables. In each column, we observe that the coefficient on skilled (unskilled) investors is positive (negative) and they are both highly significant. The cost difference between *skilled* and *unskilled* investors is approximately 25 bps which is both statistically and economically significant as verified by Wald tests.

Combined with our analysis in Section 7.2, these findings have important implications for comparing execution quality between market venues. For example, comparing price impact coefficients of a dark pool and a lit exchange on a standalone basis cannot be conclusive, as dark pools seem to be utilized more by *unskilled* investors. Our model can be employed to correct for this bias by accounting for the heterogeneity in short-term trading skill.

7.4. Disposition Effect

There are several empirical studies (see Barber and Odean (2011) for a survey) that document investors' exposure to disposition effect, i.e., they have tendency to sell winners too early and ride losers too long. The disposition effect may be one channel of unskilled trading as our main analysis suggests that the stock price tends to fall (increase) while *unskilled* investors are selling (buying) it.[9] Thus, we need to test whether unskilled traders buy (sell) an asset that has realized large positive (negative) returns due to this behavioral bias.

Let R_i be the cumulative return of the asset over the previous month before the execution. If the *unskilled* investors in our dataset are exposed to disposition effect, we would expect that they

[9]The presence of retail investors in our dataset can strengthen this behavioral bias.

would sell (buy) the asset, when R_i is positive (negative). We fit the following regression model with stock fixed-effects and control variables to formally test this hypothesis:

$$-\text{sgn}\left(Q_i\right) R_i = c_0 + \beta_s IsSkilled_i + \beta_u IsUnskilled_i + \sum_j c_j\, ControlVariables_j + \sum_{k=1}^{S} \gamma_k \mathbb{I}_{\{s(i)=k\}} + \epsilon_i, \tag{13}$$

where the dependent variable measures the investor's holding return before the execution. We consider average volatility and turnover over the past month as our control variables.

[Insert Table XVI here]

Table XVI reports the estimated coefficients of the model. We observe that the group dummy of *unskilled* investors has statistically significant positive coefficients at 1% level suggesting that higher holding period is correlated with their execution decisions. We find that skilled investors do not make trading decisions based on prior month returns. Furthermore, we formally test for the differences between β_s and β_u using a Wald test. In all cases, we reject the null hypothesis suggesting that *unskilled* investors differ significantly from *skilled* ones in selling (buying) winning (losing) stocks.

7.5. Long-Term Objective

Our estimation results illustrated that *unskilled* investors are making unprofitable trading decisions in the short-term. However, this does not mean that they are losing money consistently. Another explanation for the presence of *unskilled* investors may be due to their long-term objectives. In this section, we will compare the long-term performance of *skilled* and *unskilled* investors.

Let LR_i be the cumulative return of the investor's execution over the next year in basis points using the average price of the execution as the initial cost. We fit the following regression model with stock-dummies to formally test whether out-of-sample long-term performance have statistical dependence on the skill level of the investor:

$$LR_i = c_0 + \beta_s IsSkilled_i + \beta_u IsUnskilled_i + \sum_j c_j\, ControlVariables_j + \sum_{k=1}^{S} \gamma_k \mathbb{I}_{\{s(i)=k\}} + \epsilon_i. \tag{14}$$

We use the volatility of the asset and prior yearly (momentum signal) and monthly (short-term reversal signal) stock returns as control variables.

[Insert Table XVII here]

Table XVII reports the regression results. We observe that the dummy for the set of unskilled investors is positive and statistically significant at 10% level in each specification. These coefficients are also larger than those estimated for *skilled* investors suggesting that their trades tend to perform better in the long-term. However, Wald tests do not provide statistical significance for this difference.

7.6. Trade Size

In a theoretical model, Easley and O'hara (1987) show that informed investors can choose to trade larger amount of shares. Thus, in this section, we study the relationship between trading skill and order size decision. We can fit the following regression to formally test whether out-of-sample order sizes have statistical dependence on the short-term trading skill level of the investor:

$$\log(Q_i) = c_0 + \beta_s IsSkilled_i + \beta_u IsUnskilled_i + \sum_j c_j \, ControlVariables_j + \sum_{k=1}^{S} \gamma_k \mathbb{I}_{\{s(i)=k\}} + \epsilon_i, \quad (15)$$

where we use stock dummies and execution-level control variables that are in the information set of the investor when he decides on the number of shares to be executed. These can include prior stock and market returns, and average share volume and turnover over the past month.

[Insert Table XVIII here]

Table XVIII reports the estimated coefficients of the model. In five different specifications, we observe that the dummy for the set of *skilled* investors has positive coefficient but it is not statistically significant at 5% level. On the other hand, the dummy for *unskilled* investors has statistically significant negative coefficients. Furthermore, we formally test for the difference between β_s and β_u, $H_0 : \beta_s = \beta_u$, using a Wald test. The results show a clear rejection pattern of the hypotheses highlighting that skilled investors choose to execute larger asset quantities than their unskilled counterparts.

8. Does α Measure Skill?

Our estimation methodology allowed us to identify skilled and unskilled traders qualified by a short-term Sharpe ratio over the interval of execution. If skilled trading arises from superior predictive ability (based on short-term information), as is our interpretation, then skilled trading should also be predictive of *future* returns beyond the execution horizon. As outlined in Section 2, the price change after a large order execution will be permanent when the trading motive is information-based. The construction of a profitable trading strategy based on the short-term informational asymmetry will be consistent with this theory, as it implies that the price impact is permanent as opposed to transitory effects observed in liquidity-based trading.

In this section, we propose a long-short portfolio strategy that exploits the classification of traders based on skill. Our interpretation based on trading skill would be supported if this trading strategy were to generate significant abnormal returns. This conjecture is based on the theoretical evidence summarized in Section 2, arguing that information-based trading induces permanent price impact.

8.1. Data

We use the estimated short-term predictive ability of each trader from our model estimation in Section 5. In order to test our trading strategy, we complement the original data set by extracting the next day returns of the executed stock from the CRSP database. We download daily four factor data based on portfolio returns of market risk premium, HML, SMB and UMD (momentum) and the risk-free rate from Ken French's webpage.[10]

8.2. Construction of Trading Strategy

We construct a simple trading strategy according to the sign of the estimated short-term trading skill for each trader. We will restrict our universe of traders to *skilled* and *unskilled* based on our earlier definition in Section 5. The main idea of the strategy is to follow the trades of the skilled investors and perform trades in the opposite direction of what unskilled investors trade.

[10] See Ken French's data library at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

This strategy is motivated by the hypothesis that skill should be persistent in the short-run.

We rebalance our portfolio at the end of each trading day between January 3rd, 2011 and June 29th, 2012. Specifically, we construct our long-short portfolio specifically as follows:

- On trading day t, there are L_t distinct stocks that skilled traders have bought and unskilled traders have sold. Then, at the beginning of $(t+1)$th trading day, our long portfolio will put equal (positive) weight to each of these securities with a maximum possible weight of w_{max}. Formally, each security will have a weight of $\min\left(\frac{1}{L_t}, w_{\mathsf{max}}\right)$.

- On trading day t, there are S_t distinct stocks that skilled traders have sold and unskilled traders have bought. Then, at the beginning of $(t+1)$th trading day, our short portfolio will put equal (negative) weight to each of these securities with a minimum possible weight of $-w_{\mathsf{max}}$. Formally, each security will have a weight of $\max\left(-\frac{1}{S_t}, -w_{\mathsf{max}}\right)$.

- The remaining portfolio is invested in the risk-free rate. Formally, at the beginning of $(t+1)$th trading day, the weight in the risk-free security is

$$w_{\mathsf{rf},t+1} = 1 - L_t \min\left(\frac{1}{L_t}, w_{\mathsf{max}}\right) - S_t \max\left(-\frac{1}{S_t}, -w_{\mathsf{max}}\right).$$

We set w_{max} to 5% as a base calibration and also provide sensitive analysis for this parameter.

8.3. Results

Using this long-short trading strategy, we obtain the time-series of returns for our strategy resulting from 377 trading days. The average daily return from our trading strategy is 6.91 bps corresponding to 17.4% of annualized return. The standard deviation of the daily return resulting from our strategy is 71.23 bps corresponding to annualized volatility of 11.3%. The resulting annualized Sharpe ratio is 1.94. In order to ensure that these raw return statistics are abnormal, we regress our strategy returns against the four-factor model of Carhart (1997).

[Insert Table XIX here]

Table XIX shows that our long-short portfolio returns have statistically significant Jensen's alpha. In annualized terms, employing our strategy earns an excess return of 18.8% which is also economically significant. These abnormal return statistics provide further supporting evidence of our classification of traders with regards to short-term predictive ability. Note that this also illustrates that the trades of the skilled and unskilled investors do predict the returns of the next day, which is actually a stronger result.

Table XIX also illustrates that our results are robust to the choice of w_{max}. We use three other choices of w_{max} with 2.5%, 10% and 15% and in each scenario we obtain statistically significant excess returns for our strategy. The annualized excess returns are between 13.9% and 20.3% and Sharpe ratios are between 1.54 and 1.99.

We observe that the excess return of our strategy is very high compared to the reported returns of the standard anomalies studied in the empirical asset pricing literature. This is expected as this long-short portfolio construction is not implementable with public data sources. Thus, the main key takeaway from this analysis is its supporting evidence on skill-based trading at the investor level. Our profitable trading strategy implies that the realized execution returns are persistent in the short-term.

9. Conclusion

Statistical models for short-term returns observed during the execution of a large order typically have low explanatory power. It is difficult to separate price impact due to demand for liquidity from predicting price changes due to trading skill. Consequently, trading cost models are estimated on large order samples, where the effect of short-term predictive ability is expected to cancel out.

In this paper, we first propose a model to explain the variation in short-term returns with the short-term trading skill of the investors along with a parametric modeling of price impact. Motivated by the performance metrics for the fund management industry, we measure trading skill in a risk-adjusted way using short-term Sharpe ratios. We estimated our model on a large sample of executions with masked investor identities and our results show that incorporating short-term predictive ability offers drastic improvements in explaining the variation in security returns over an execution horizon. We also observe that ignoring short-term trading skill may lead to biased price impact estimates which are economically large.

The estimated trading skill is specific to each investor in the sample, and can be used to classify investors according to the success of their predictive ability. We find that in addition to the presence of skilled investors, a significant portion of the investor universe is unskilled in the sense that their trades are in the opposite direction of future short-term price movements. This cross-sectional variation implies a pronounced heterogeneity in short-term trading skill among institutional investors. Bootstrapping analyses rigorously illustrate that the numbers of skilled and unskilled investors are abnormally high and the price impact coefficients are biased when heterogeneity in short-term trading skill is ignored. Furthermore, this classification is robust and has predictive power about the future trading performance of an investor. In order to test for persistence in the short-term, we propose a trading strategy exploiting the signs of estimated skill coefficients and find that this strategy generates economically substantial abnormal returns even against a benchmark Fama-French four-factor model.

We analyze the trading characteristics of skilled and unskilled investors and find that skilled traders differ significantly from unskilled investors by trading (relatively) larger orders and trading more in lit markets. Unskilled traders tend to sell recent winners and buy recent losers over the past month. Both types of investors are not able to time favorable liquidity conditions. Finally, out-of-sample execution costs have major statistical dependence on our skill estimates as well. Expected execution cost difference between short-term skilled and unskilled traders is 25 bps which is economically substantial. These findings have important policy implications. Our results illustrate that mere comparison of execution costs cannot be a standalone measure of execution quality. We find that dark pools are avoided by skilled investors, which may consequently bias their measures of execution quality. In the presence of investor heterogeneity, these biased estimates may not ultimately lead to an optimal policy recommendation. Moreover, our findings have several other practical applications. For example, an agency broker can use the historically estimated trader skill to advise individual investors on the choice of algorithm and parameter settings. Tracking the predictive ability of an investor through time gives a measure of trading efficiency, that would be of interest to this investor. Such measures can complement the traditional transaction cost analysis (TCA) that brokers typically provide.

References

Anat R Admati and Paul Pfleiderer. A theory of intraday patterns: Volume and price variability. *Review of Financial studies*, 1(1):3–40, 1988.

R. Almgren, C. Thum, E. Hauptmann, and H. Li. Direct estimation of equity market impact. *Risk*, 18(7):58–62, 2005.

Robert Almgren. Execution costs. *Encyclopedia of quantitative finance*, 2008.

Amber Anand, Paul Irvine, Andy Puckett, and Kumar Venkataraman. Performance of institutional trading desks: An analysis of persistence in trading costs. *Review of Financial Studies*, 25(2): 557–598, 2012.

Clifford S Asness, Tobias J Moskowitz, and Lasse Heje Pedersen. Value and momentum everywhere. *The Journal of Finance*, 68(3):929–985, 2013.

Brad M Barber and Terrance Odean. The behavior of individual investors. *Available at SSRN 1872211*, 2011.

M. Baron, J. Brogaard, and A Kirilenko. The trading profits of high frequency traders. Working paper, 2013.

Dimitris Bertsimas and Andrew W Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50, 1998.

Hendrik Bessembinder and Herbert M Kaufman. A comparison of trade execution costs for nyse and nasdaq-listed stocks. *Journal of Financial and Quantitative Analysis*, 32(03):287–310, 1997a.

Hendrik Bessembinder and Herbert M Kaufman. A cross-exchange comparison of execution costs and information flow for nyse-listed stocks. *Journal of Financial Economics*, 46(3):293–319, 1997b.

Bruno Biais, Larry Glosten, and Chester Spatt. Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets*, 8(2):217–264, 2005.

J.-P. Bouchaud, J. D. Farmer, and F. Lillo. How markets slowly digest changes in supply and demand. In *Handbook of Financial Markets: Dynamics and Evolution*, pages 57–156. Elsevier: Academic Press, 2008.

Jonathan Brogaard, Terrence Hendershott, Stefan Hunt, and Carla Ysusi. High-frequency trading and the execution costs of institutional investors. *Financial Review*, 49(2):345–369, 2014.

Mark M Carhart. On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82, 1997.

George C. Chacko, Jakub W. Jurek, and Erik Stafford. The price of immediacy. *The Journal of Finance*, 63(3):1253–1290, 2008.

Louis KC Chan and Josef Lakonishok. The behavior of stock prices around institutional trades. *The Journal of Finance*, 50(4):1147–1174, 1995.

Adam D Clark-Joseph. Exploratory trading. *Unpublished working paper). Harvard University, Cambridge MA*, 2013.

Randolph B Cohen, Joshua D Coval, and L'uboš Pástor. Judging fund managers by the company they keep. *The Journal of Finance*, 60(3):1057–1096, 2005.

Pierre Collin-Dufresne and Vyacheslav Fos. Do prices reveal the presence of informed trading? *Journal of Finance*, 70(4):1555–1582, 2015.

Pierre Collin-Dufresne and Vyacheslav Fos. Insider trading, stochastic liquidity and equilibrium prices. *Econometrica*, Forthcoming.

Karl B Diether, Kuan-Hui Lee, and Ingrid M Werner. Short-sale strategies and return predictability. *Review of Financial Studies*, 22(2):575–607, 2009.

D. Easley, M. López de Prado, and M. O'Hara. Flow toxicity and liquidity in a high frequency world. *Review of Financial Studies*, 25(5):1457–1493, 2012.

David Easley and Maureen O'hara. Price, trade size, and information in securities markets. *Journal of Financial economics*, 19(1):69–90, 1987.

David Easley, Soeren Hvidkjaer, and Maureen O'Hara. Is information risk a determinant of asset returns? *The Journal of Finance*, 57(5):2185–2221, 2002. ISSN 1540-6261. doi: 10.1111/1540-6261.00493. URL http://dx.doi.org/10.1111/1540-6261.00493.

L. R. Glosten and P. R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100, 1985.

Paul A Gompers and Andrew Metrick. Institutional investors and equity prices. *Quarterly Journal of Economics*, pages 229–259, 2001.

J. Hasbrouck. *Empirical Market Microstructure.* Oxford University Press, 2007.

Terrence Hendershott, C Jones, and Albert J Menkveld. Implementation shortfall with transitory price effects. *High Frequency Trading; New Realities for Trades, Markets and Regulators, Easley, D., M. Lopez de Prado, and M. O?Hara (editors), Risk Books (London: 2013)*, 2013.

Terrence Hendershott, Dmitry Livdan, and Norman Schürhoff. Are institutions informed about news? *Journal of Financial Economics*, 117(2):249–287, 2015.

Thomas Ho and Hans R Stoll. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial economics*, 9(1):47–73, 1981.

Roger D. Huang and Hans R. Stoll. Dealer versus auction markets: A paired comparison of execution costs on NASDAQ and the NYSE. *Journal of Financial Economics*, 41(3):313 – 357, 1996.

K. Iyer, R. Johari, and C. C. Moallemi. Welfare analysis of dark pools. SSRN Working paper, 2015.

Michael Johannes, Arthur Korteweg, and Nicholas Polson. Sequential learning, predictability, and optimal portfolio returns. *The Journal of Finance*, 69(2):611–644, 2014.

Marcin Kacperczyk, Clemens Sialm, and Lu Zheng. On the industry concentration of actively managed equity mutual funds. *The Journal of Finance*, 60(4):1983–2011, 2005.

Ron Kaniel and Hong Liu. So what orders do informed traders use?*. *The Journal of Business*, 79 (4):1867–1913, 2006.

Donald B Keim and Ananth Madhavan. Transactions costs and investment style: An inter-exchange analysis of institutional equity trades. *Journal of Financial Economics*, 46(3):265–292, 1997.

Donald Bruce Keim and Ananth Madhavan. The upstairs market for large-block transactions: Analysis and measurement of price effects. *Review of Financial Studies*, 9(1):1–36, 1996.

Robert A Korajczyk and Dermot Murphy. High frequency market making to large institutional trades. *Available at SSRN 2567016*, 2015.

A. P. Kyle and A. Obizhaeva. Market microstructure invariants: Theory and empirical tests. SSRN Working paper, 2014.

A. S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, November 1985.

C. Maglaras, C. C. Moallemi, and H. Zheng. Queueing dynamics and state space collapse in fragmented limit order book markets. Working paper, 2012.

Harry Mamaysky, Matthew Spiegel, and Hong Zhang. Estimating the dynamics of mutual fund alphas and betas. *Review of Financial Studies*, 21(1):233–264, 2008.

A. Obizhaeva. Liquidity estimates and selection bias. Working paper, 2009.

Andre F Perold. The implementation shortfall: Paper versus reality. *The Journal of Portfolio Management*, 14(3):4–9, 1988.

Mitchell A Petersen. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of financial studies*, 22(1):435–480, 2009.

Andy Puckett and Xuemin Sterling Yan. The interim trading skills of institutional investors. *The Journal of Finance*, 66(2):601–633, 2011.

Hersh Shefrin and Meir Statman. The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of finance*, 40(3):777–790, 1985.

Lin Tong. A blessing or a curse? the impact of high frequency trading on institutional investors. *SSRN Working Paper*, 2015.

N Torre and Mark J Ferrari. The market impact model. *Horizons, The Barra Newsletter*, 165, 1998.

Vincent Van Kervel and Albert J Menkveld. High-frequency trading around large institutional orders. *Available at SSRN 2619686*, 2015.

Russ Wermers. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *The Journal of Finance*, 55(4):1655–1703, 2000.

Xuemin (Sterling) Yan and Zhe Zhang. Institutional investors and equity returns: Are short-term institutions better informed? *Review of Financial Studies*, 22(2):893–924, 2009.

Haoxiang Zhu. Do dark pools harm price discovery? *Review of Financial Studies*, 27(3):747–789, 2014.

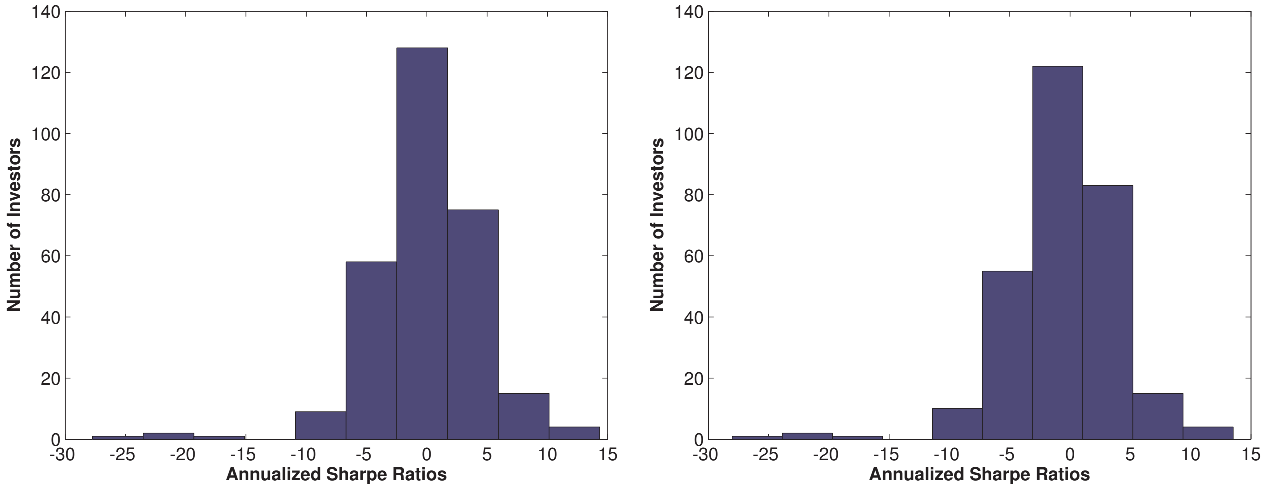

Figure 1: Histogram of all investor skill estimates expressed as annualized Sharpe ratios, when the price impact is proportional to the participation rate (left) and when the price impact is proportional to the square root of participation rate (right).

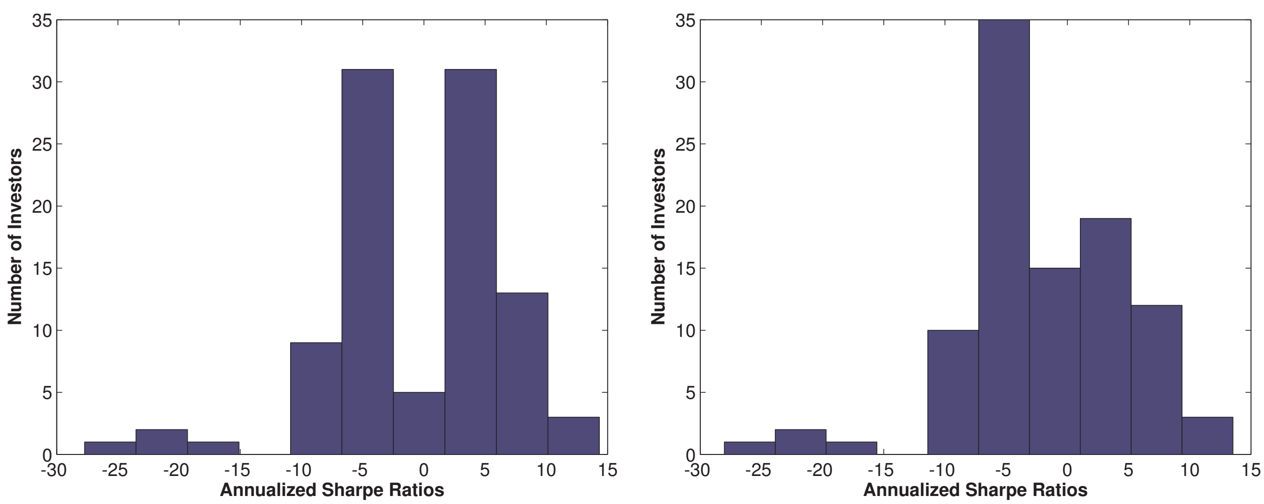

Figure 2: Histogram of statistically significant investor skill estimates, expressed as annualized Sharpe ratios, when price impact is proportional to the participation rate (left) and when the price impact is proportional to the square root of the participation rate (right).

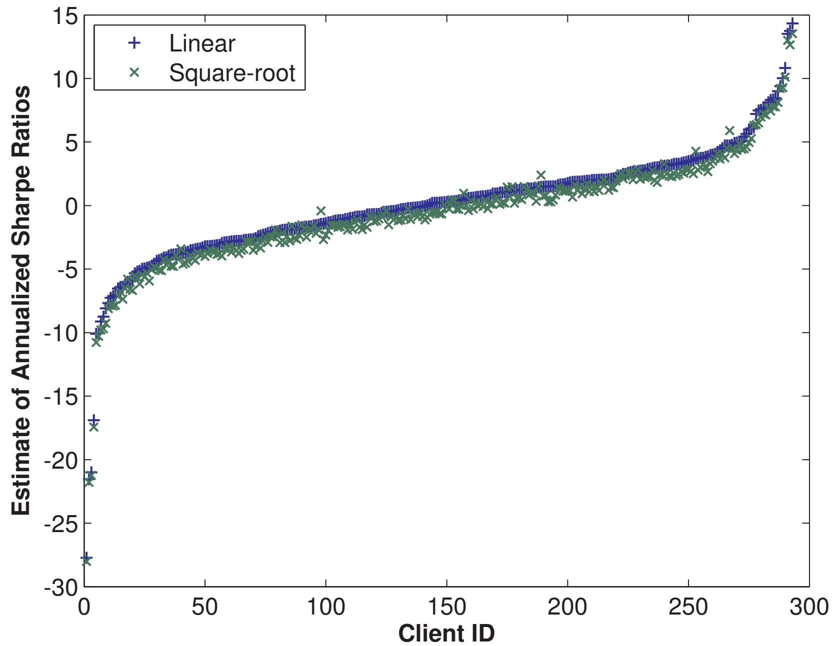

Figure 3: The difference in trading skill estimates between the linear and square root price impact models in Equations (2) and (3).

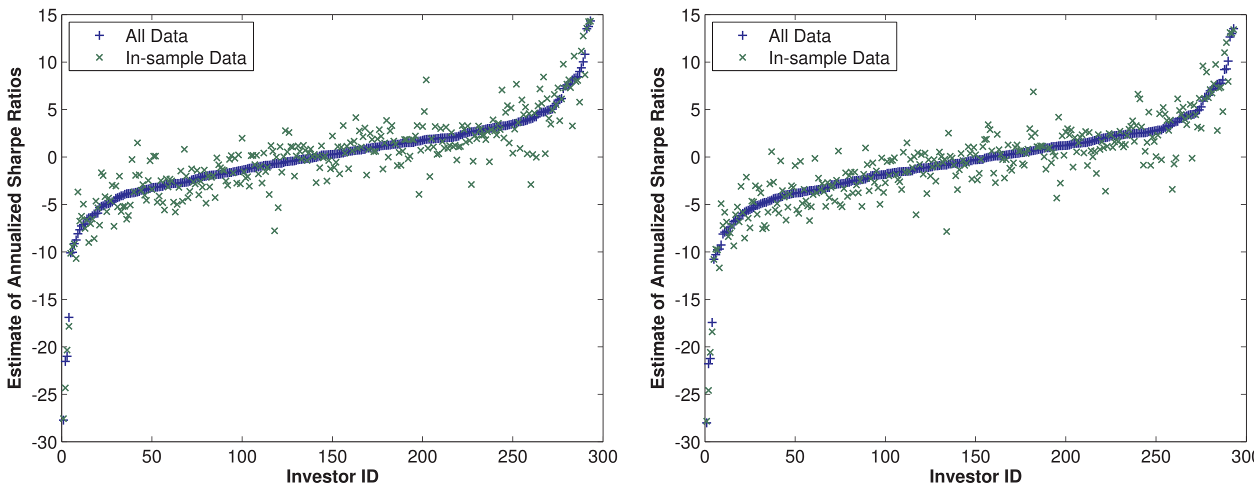

Figure 4: As a robustness check, we compare trading skill estimates computed from the complete data set and a randomly constructed in-sample data set in which for every investor only random half of his executions are considered. We use two price impact specifications for our comparisons: the linear model (left) and the square-root model (right).

Table I: Summary statistics for the main attributes in our execution data. The bid ask spread is normalized using the mid-quote price. Average daily volatilities are computed using the previous 20 trading days before the execution date. The duration of a full trading day in U.S. equity markets is 390 minutes.

Statistic	Interval Return (%)	Bid Ask Spread (bps)	Average Daily Volatility (%)	Execution Duration (mins)	Participation Rate (%)	Percentage of Daily Volume (%)
Min.	-18.102	0.68	0.128	5.00	0.0001	<0.01
1st Qu.	-0.307	2.45	0.980	15.23	0.17	0.06
Median	0.000	3.35	1.267	59.82	1.59	0.26
Mean	0.003	4.16	1.417	148.13	6.44	0.64
3rd Qu.	0.333	4.87	1.685	324.50	10.50	0.72
Max.	10.793	54.76	25.740	390.00	100.00	28.25

Table II: Regression results for two price impact models with and without trading skill terms. Two price impact specifications are estimated, linear and square root. We label an investor as *skilled* (resp., *unskilled*) if his short-term trading skill estimate is positive (resp., negative) and is statistically significant under 10% level. In each column, we report estimated coefficients and their standard errors, adjusted by clustering on calendar day and stock as suggested by Petersen (2009).

	Linear		Square root	
Trading Skill?	Yes	No	Yes	No
Intercept (bps)	1.26	-1.34	1.51	-3.27
	(1.24)	(2.42)	(1.22)	(2.69)
λ	1.811^{***}	2.277^{***}	0.744^{***}	0.837^{***}
	(0.291)	(0.363)	(0.173)	(0.154)
Number of *skilled* investors	49	N/A	35	N/A
Number of *unskilled* investors	48	N/A	63	N/A
R^2	10.5%	0.5%	10.5%	0.4%
Adj. R^2	10.1%	0.5%	10.0%	0.4%

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table III: Bootstrapping results for the difference in price impact coefficients by constructing $1,000$ random datasets each with $10,000$ executions. In each column, we report the mean difference between the price impact coefficients and its corresponding standard errors.

	Linear	Square root
$\hat{\lambda} - \hat{\lambda}^{\mathsf{base}}$	-0.467^{***}	-0.096^{***}
	(0.012)	(0.007)
N	$1,000$	$1,000$

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table IV: Results of the falsification test based on a bootstrapping analysis via shuffling investor identifiers across executions. We randomly construct $10,000$ different samples of our execution dataset by permuting the investor ids. In each column, we report the estimated coefficients from the original models and their corresponding p-values in square brackets. Empirical distribution for the parameters are obtained under the null hypothesis that investor identifiers are unrelated to log-returns realized during each execution horizon.

	Linear	Square root
Number of *skilled* investors	49^{***}	35^{***}
	[< 0.001]	[< 0.001]
Number of *unskilled* investors	48^{***}	63^{**}
	[0.001]	[0.016]
$\hat{\lambda}$	1.811^{***}	0.744^{***}
	[< 0.001]	[< 0.001]
N	$10,000$	$10,000$

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table V: Regression results for two price impact models with an in-sample data set constructed using random half of each investor's executions. We label an investor as *skilled* (resp., *unskilled*) if his short-term trading skill estimate is positive (resp., negative) and is statistically significant under 10% level. We estimate two price impact models with and without predictive ability of the investor. In each column, we report estimated coefficients and their standard errors, adjusted by clustering on calendar day and stock as suggested by Petersen (2009).

	Linear		Square root	
Trading Skill?	Yes	No	Yes	No
Intercept (bps)	1.41	-0.76	1.55	-2.93
	(1.54)	(2.41)	(1.53)	(2.62)
λ	1.793^{***}	2.160^{***}	0.767^{***}	0.838^{***}
	(0.310)	(0.366)	(0.179)	(0.156)
Number of *skilled* investors	42	N/A	32	N/A
Number of *unskilled* investors	45	N/A	53	N/A
R^2	11.4%	0.5%	11.4%	0.5%
Adj. R^2	10.6%	0.5%	10.6%	0.5%

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

	rMSPE		R^2	
Model	In-Sample	Out-of-Sample	In-Sample	Out-of-Sample
Linear	93.19	95.06	10.6%	8.2%
Square root	93.19	95.07	10.6%	8.1%

Table VI: This table reports root mean squared prediction errors (rMSPE) between in-sample and out-of-sample execution returns and in-sample and out-of-sample R^2. Predicted execution returns use the skill and price impact coefficients estimated from in-sample data set. rMSPE values are reported in basis points.

Table VII: Regression results for two price impact models using constant daily Sharpe ratio for each investor. We only consider the presence of trading skill terms. We label an investor as *skilled* (resp., *unskilled*) if his short-term trading skill estimate is positive (resp., negative) and is statistically significant under 10% level. In each column, we report estimated coefficients and their standard errors, adjusted by clustering on calendar day and stock as suggested by Petersen (2009).

	Linear	Square root
Intercept (bps)	1.48	1.08
	(1.03)	(1.20)
λ^{mkt}	1.875***	0.759**
	(0.243)	(0.140)
Number of *skilled* investors	45	35
Number of *unskilled* investors	61	71
R^2	10.9%	10.9%
Adj. R^2	10.5%	10.5%

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table VIII: Regression results for two price impact models with an additional spread component in the price impact model. We label an investor as *skilled* (resp., *unskilled*) if his short-term trading skill estimate is positive (resp., negative) and is statistically significant under 10% level. We estimate two price impact models with and without predictive ability of the investor. In each column, we report estimated coefficients and their standard errors, adjusted by clustering on calendar day and stock as suggested by Petersen (2009).

	Linear		Square root	
Trading Skill?	Yes	No	Yes	No
Intercept (bps)	0.41	−1.12	0.63	−2.78
	(1.31)	(1.80)	(1.29)	(2.03)
λ	1.758^{***}	2.284^{***}	0.721^{***}	0.846^{***}
	(0.310)	(0.366)	(0.179)	(0.176)
δ (bps)	0.304^{*}	−0.055	0.316^{*}	−0.135
	(0.179)	(0.433)	(0.179)	(0.180)
Number of *skilled* investors	46	N/A	32	N/A
Number of *unskilled* investors	48	N/A	63	N/A
R^2	10.5%	0.5%	10.5%	0.4%
Adj. R^2	10.1%	0.5%	10.0%	0.4%

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table IX: Bootstrapping results for the difference in coefficients of the linear price impact model when skill terms are included or excluded. Superscript base refers to the exclusion case. We construct $1,000$ random datasets each with $10,000$ executions. In each column, we report the mean difference between the price impact coefficients and its corresponding standard errors.

	Linear
$\hat{\lambda} - \hat{\lambda}^{\mathsf{base}}$	-0.530^{***}
	(0.012)
$\hat{\delta} - \hat{\delta}^{\mathsf{base}}$	0.362^{***}
	(0.008)
N	$1,000$

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table X: Regression results for two price impact models with stock fixed-effects in the price impact model. We label an investor as *skilled* (resp., *unskilled*) if his short-term trading skill estimate is positive (resp., negative) and is statistically significant under 10% level. We estimate two price impact models with and without predictive ability of the investor. In each column, we report estimated coefficients and their standard errors, adjusted by clustering on calendar day and stock as suggested by Petersen (2009).

	Linear		Square root	
Trading Skill?	Yes	No	Yes	No
Intercept (bps)	-7.60	-10.51	-7.65^{***}	-10.27^{***}
	(7.12)	(7.41)	(0.60)	(1.70)
λ	1.75^{***}	2.183^{***}	0.733^{***}	0.812^{***}
	(0.178)	(0.135)	(0.194)	(0.191)
Number of *skilled* investors	46	N/A	31	N/A
Number of *unskilled* investors	50	N/A	60	N/A
R^2	11.2%	1.4%	11.2%	1.3%
Adj. R^2	10.1%	0.6%	10.1%	0.5%

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table XI: Regression results for two price impact models for a subset of executions that lasted between 156 and 234 minutes (between 40% and 60% of a full trading day). We label an investor as *skilled* (resp., *unskilled*) if his short-term trading skill estimate is positive (resp., negative) and is statistically significant under 10% level. We estimate two price impact models with and without predictive ability of the investor. In each column, we report estimated coefficients and their standard errors, adjusted by clustering on calendar day and stock as suggested by Petersen (2009).

	Linear		Square root	
Trading Skill?	Yes	No	Yes	No
Intercept (bps)	-2.29	-3.79	-3.28	-4.70
	(8.77)	(6.28)	(8.65)	(3.29)
λ	2.726^{**}	2.455^{*}	1.063^{*}	0.464
	(1.243)	(1.411)	(0.554)	(0.649)
Number of executions	1651	1651	1651	1651
Number of investors	39	39	39	39
Number of *skilled* investors	9	N/A	8	N/A
Number of *unskilled* investors	9	N/A	10	N/A
R^2	17.9%	0.3%	18.0%	0.1%
Adj. R^2	15.8%	0.2%	15.9%	0.1%

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table XII: Regression results for two price impact models using abnormal (excess) returns. We only consider the presence of trading skill terms. We label an investor as *skilled* (resp., *unskilled*) if his short-term trading skill estimate is positive (resp., negative) and is statistically significant under 10% level. In each column, we report estimated coefficients and their standard errors, adjusted by clustering on calendar day and stock as suggested by Petersen (2009).

	Linear	Square root
Intercept (bps)	1.89	2.46
	(2.49)	(2.44)
λ^{mkt}	1.609***	0.555**
	(0.406)	(0.265)
Number of *skilled* investors	33	25
Number of *unskilled* investors	55	61
R^2	12.2%	12.2%
Adj. R^2	11.8%	11.8%

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table XIII: The dependent variables are liquidity measures based on proportional spread, logarithm of share volume, turnover, and Amihud Illiquidity measure (*ILLIQ*) realized during the execution period. We regress these measures on our skill dummies and execution level control variables including volatility, average spread and logarithm of average volume over the past month, absolute values of asset and market return over the past week. Formally, we run the model with stock dummies specified in Equation 9. We use the out-of-sample data constructed for robustness checks in Section 6. Using the estimation from in-sample data, IsSkilled_i (resp., IsUnskilled_i) takes a value of 1 if the ith execution is sent by a *skilled* (resp., an *unskilled*) investor (as defined in Section 5). Standard errors are given in parentheses and are adjusted by clustering on calendar day as suggested by Petersen (2009).

	Dependent variable: Liquidity Measures			
	Spread	Log Volume	Turnover	ILLIQ
IsSkilled	0.03	0.09	0.22	0.01
	(0.04)	(0.12)	(0.44)	(0.01)
IsUnskilled	0.004	0.12	0.49*	−0.02
	(0.04)	(0.10)	(0.29)	(0.01)
Volatility	105.89***	−4.47	20.54	6.05***
	(3.69)	(3.16)	(17.24)	(0.75)
Average Past Spread	0.10	0.01	−0.02	−0.001
	(0.08)	(0.004)	(0.02)	(0.01)
Log Average Past Volume	0.59***	0.53***	1.85***	−0.06***
	(0.15)	(0.07)	(0.25)	(0.02)
|Prior Market Return|	−3.94*	8.77*	45.74*	−0.56
	(2.30)	(4.92)	(23.85)	(0.53)
|Prior Week Return|	−0.83	0.93	9.07***	−0.11
	(0.51)	(0.59)	(3.45)	(0.10)
Observations	31,670	31,670	31,670	31,670
R^2	0.75	0.41	0.20	0.27
Adjusted R^2	0.75	0.40	0.19	0.26
Wald Test: $H_0 : \beta_s = \beta_u$				
Chi-squared statistic	0.23	0.04	0.29	2.1
p-value	[0.63]	[0.84]	[0.59]	[0.15]

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table XIV: We regress fraction of shares executed in the dark pool on our skill dummies and execution level control variables including participation rate, bid-offer spread, mid-quote volatility, and turnover during the execution horizon. Formally, we run the model with stock dummies specified in Equation 10. We use the out-of-sample data constructed for robustness checks in Section 6. Using the estimation from in-sample data, IsSkilled_i (resp., IsUnskilled_i) takes a value of 1 if the ith execution is sent by a *skilled* (resp., an *unskilled*) investor (as defined in Section 5). Standard errors are given in parentheses and are adjusted by clustering on calendar day as suggested by Petersen (2009).

	Dependent variable:				
	DP				
	(1)	(2)	(3)	(4)	(5)
IsSkilled	-0.04^{**}	-0.04^{**}	-0.03^{**}	-0.04^{**}	-0.03^{**}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
IsUnskilled	0.01	0.01	0.01	0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Participation Rate		-0.37^{***}	-0.38^{***}	-0.40^{***}	-0.30^{***}
		(0.08)	(0.08)	(0.08)	(0.07)
Spread			-0.003	0.0001	0.001
			(0.002)	(0.003)	(0.003)
Volatility				-1.44^{*}	-1.90^{***}
				(0.78)	(0.72)
Turnover					0.002^{***}
					(0.001)
Observations	6,633	6,633	6,633	6,633	6,633
R^2	0.10	0.11	0.11	0.11	0.12
Adjusted R^2	0.03	0.04	0.04	0.04	0.05
Wald Test: $H_0 : \beta_s = \beta_u$					
Chi-squared statistic	3.4	3.6	3.6	4.0	4.0
p-value	[0.067]	[0.059]	[0.058]	[0.046]	[0.046]

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table XV: We regress implementation shortfall on our skill dummies and execution level control variables including including participation rate, bid-offer spread, mid-quote volatility, and turnover during the execution horizon. Formally, we run the model with stock dummies specified in Equation 11. We use the out-of-sample data constructed for robustness checks in Section 6. Using the estimation from in-sample data, IsSkilled_i (resp., IsUnskilled_i) takes a value of 1 if the ith execution is sent by a *skilled* (resp., an *unskilled*) investor (as defined in Section 5). Standard errors are given in parentheses and are adjusted by clustering on calendar day as suggested by Petersen (2009).

	Dependent variable:				
	Implementation Shortfall				
	(1)	(2)	(3)	(4)	(5)
IsSkilled	11.24***	11.43***	11.42***	11.42***	11.43***
	(2.53)	(2.62)	(2.63)	(2.62)	(2.71)
IsUnskilled	−14.16***	−13.49***	−13.49***	−13.49***	−13.53***
	(3.71)	(3.69)	(3.69)	(3.68)	(3.62)
Participation Rate		35.52***	35.10***	35.12***	23.52***
		(8.66)	(8.90)	(8.95)	(6.82)
Spread			0.44	0.53	0.42
			(0.40)	(0.38)	(0.36)
Volatility				−35.69	7.09
				(205.66)	(183.15)
Turnover					−0.43
					(0.27)
Observations	31,690	31,690	31,690	31,690	31,689
R^2	0.03	0.03	0.03	0.03	0.03
Adjusted R^2	0.01	0.01	0.01	0.01	0.01
Wald Test: $H_0 : \beta_s = \beta_u$					
Chi-squared statistic	35.5	33.6	33.5	33.6	33.5
p-value	[<0.0001]	[<0.0001]	[<0.0001]	[<0.0001]	[<0.0001]

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table XVI: We regress out-of-sample holding returns realized over the past month on our skill dummies and execution level control variables including average volatility and turnover over the past month. Formally, we run the model with stock dummies specified in Equation 13. We use the out-of-sample data constructed for robustness checks in Section 6. Using the estimation from in-sample data, IsSkilled_i (resp., IsUnskilled_i) takes a value of 1 if the ith execution is sent by a *skilled* (resp., an *unskilled*) investor (as defined in Section 5). Standard errors are given in parentheses and are adjusted by clustering on calendar day as suggested by Petersen (2009).

	Dependent variable:		
	Past Month Holding returns		
	(1)	(2)	(3)
IsSkilled	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)
IsUnskilled	0.01***	0.01***	0.01***
	(0.004)	(0.004)	(0.004)
Average Volatility		0.22	0.22
		(0.35)	(0.35)
Average Turnover			0.0001
			(0.0004)
Observations	31,640	31,640	31,640
R^2	0.03	0.03	0.03
Adjusted R^2	0.01	0.01	0.01
Wald Test: $H_0 : \beta_s = \beta_u$			
Chi-squared statistic	4.5	4.5	4.5
p-value	[0.034]	[0.034]	[0.034]

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table XVII: We regress out-of-sample one-year future returns on our skill dummies and execution level control variables including average volatility and turnover over the past month. Formally, we run the model with stock dummies specified in Equation 14. We use the out-of-sample data constructed for robustness checks in Section 6. Using the estimation from in-sample data, IsSkilled_i (resp., IsUnskilled_i) takes a value of 1 if the ith execution is sent by a *skilled* (resp., an *unskilled*) investor (as defined in Section 5). Standard errors are given in parentheses and are adjusted by clustering on calendar day as suggested by Petersen (2009).

	Dependent variable:		
	Future Annual Holding Return (bps)		
	(1)	(2)	(3)
IsSkilled	95.94	100.82*	101.87*
	(59.38)	(58.52)	(58.50)
IsUnskilled	146.45*	151.15*	152.46*
	(89.01)	(87.84)	(87.48)
Average Volatility (bps)		1.08	0.93
		(0.78)	(0.68)
Prior Yearly Return			155.22
			(169.50)
Prior Monthly Return			−464.41
			(286.37)
Observations	31,663	31,663	31,663
R^2	0.02	0.02	0.02
Adjusted R^2	0.001	0.001	0.002
Wald Test: $H_0 : \beta_s = \beta_u$			
Chi-squared statistic	0.24	0.25	0.26
p-value	[0.62]	[0.62]	[0.61]

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table XVIII: We regress logarithm of order size on our skill dummies and execution level control variables including logarithm of average daily volume (over the past month), absolute values of prior weekly asset and market return and average turnover (over the past month). Formally, we run the model with stock dummies specified in Equation 15. We use the out-of-sample data constructed for robustness checks in Section 6. Using the estimation from in-sample data, IsSkilled_i (resp., IsUnskilled_i) takes a value of 1 if the ith execution is sent by a *skilled* (resp., an *unskilled*) investor (as defined in Section 5). Standard errors are given in parentheses and are adjusted by clustering on calendar day as suggested by Petersen (2009).

	Dependent variable:				
	$\log(Q_i)$				
	(1)	(2)	(3)	(4)	(5)
IsSkilled	0.16	0.16	0.17*	0.16	0.16
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
IsUnskilled	−0.30**	−0.30**	−0.30**	−0.30**	−0.30**
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
Log Average Past Volume		0.28***	0.22***	0.26***	0.39***
		(0.09)	(0.09)	(0.08)	(0.09)
|Prior Week Return|			2.12***	2.44***	2.48***
			(0.64)	(0.68)	(0.68)
|Prior Market Return|				−4.39	−4.46
				(3.53)	(3.53)
Average Turnover					−0.01***
					(0.004)
Observations	31,690	31,686	31,670	31,670	31,670
R^2	0.13	0.13	0.13	0.13	0.13
Adjusted R^2	0.12	0.12	0.12	0.12	0.12
Wald Test: $H_0 : \beta_s = \beta_u$					
Chi-squared statistic	7.9	8.0	8.0	8.0	8.0
p-value	[0.0049]	[0.0046]	[0.0046]	[0.0047]	[0.0047]

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table XIX: Regression results for long-short portfolio returns against the four factor model due to Carhart (1997). In each column, we report estimated coefficients and their standard errors, calculated using heteroskedasticity and auto-correlation consistent standard errors.

	(1) $w_{\mathsf{max}} = 2.5\%$	(2) $w_{\mathsf{max}} = 5\%$	(3) $w_{\mathsf{max}} = 10\%$	(4) $w_{\mathsf{max}} = 15\%$
Intercept (bps)	5.50^{*}	7.48^{**}	8.07^{**}	6.59^{*}
	(3.05)	(3.34)	(3.51)	(3.51)
Mkt-RF	-0.024	-0.044	-0.033	-0.024
	(0.059)	(0.059)	(0.037)	(0.049)
SMB	-0.090	-0.075	-0.025	-0.051
	(0.073)	(0.075)	(0.065)	(0.070)
HML	-0.080	-0.089	-0.037	0.032
	(0.091)	(0.104)	(0.115)	(0.119)
UMD	0.006	0.021	0.104^{*}	0.114^{*}
	(0.061)	(0.066)	(0.055)	(0.066)
Sharpe Ratio	1.61	1.94	1.99	1.54
N	377	377	377	377
R^2	0.02	0.03	0.03	0.02
Adjusted R^2	0.01	0.02	0.02	0.01

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$