The Non-Linear Market Impact of Large Trades: Evidence from Buy-Side Order Flow

We perform an empirical study of a set of large institutional orders executed in the U.S. equity market. Our results validate the hidden order arbitrage theory proposed by Farmer et al. (2013) of the market impact of large institutional orders. We find that large trades are drawn from a distribution with tail exponent of roughly 3/2 and that market impact approximately increases as the square root of trade duration. We examine price reversion after the completion of a trade, finding that permanent impact is also a square root function of trade duration and that its ratio to the total impact observed at the last fill is roughly 2/3. Additionally, we confirm empirically that the post-trade price reverts to a level consistent with a fair pricing condition of Farmer et al. (2013). We study the relaxation dynamics of market impact and find that impact decay is a multi-regime process, approximated by a power law in the first few minutes after order completion and subsequently by exponential decay.

Keywords: market impact, permanent impact, order size, price reversion and liquidity

JEL Classifications: G10, G12, G17, G19

Nataliya Bershova, Ph.D.,
VP, Quantitative Trader, AllianceBernstein LP,
Email: nataliya.bershova@alliancebernstein.com

Dmitry Rakhlin, Ph.D.,
SVP, Global Head of Quantitative Trading, AllianceBernstein LP,
Email: dmitry.rakhlin@alliancebernstein.com

July, 2013

INTRODUCTION

Market impact estimation and modeling has been a central topic of market microstructure research because of its conceptual and practical importance. Practitioners use impact models as a pre-trade tool to estimate the expected transaction cost of an order and to optimize execution strategy. On a post-trade basis, impact models serve as performance benchmarks to evaluate trading results. More broadly, market impact reflects the balance of supply and demand, which is of central importance for economics.

The market microstructure literature has empirically studied diverse types of market impact models. One strand of research examines the market impact of a single transaction. These studies identify such impact as a concave function of the transaction volume, although the reported shapes of market impact vary. For example, Lillo et al. (2003) show that the price impact of a single transaction is a power-law function of the transaction volume with the exponent in the range of 0.2 to 0.5, depending on the size of the volume. Potters et al. (2003) find that a logarithmic function provides a better fit to the trading from the Paris Bourse and NASDAQ. Hopman (2003) examined the market impact of individual orders executed at the Paris Bourse and showed that the impact is a power-law function with an average exponent of 0.4. The author claims that the exponent is sensitive to the time it takes to execute the order.

Another group of studies concentrates on the price impact of aggregate transactions, i.e., the sum of signed transaction volumes over a given number of trades or time intervals. Gabaix et al. (2003) used U.S. and European transaction data from the London and Paris stock markets over 15-minute intervals to study the price impact as a function of the transaction volume. The authors found the power-law function with an exponent of 0.5 to be a good fit with the data. Plerou et al. (2002) examined NYSE data on time scales ranging from 5 to 195 minutes and showed that the price impact fits the hyperbolic tangent of the traded volume. Although all these studies confirm the concavity of the impact, they suggest different functional forms. The differences in results are likely attributable to the different markets and time scales that are used in the studies.

Our study belongs to a series of price impact studies focusing on large institutional orders that are often referred to as hidden orders or “metaorders” (Bouchaud et al. (2009)). They are “hidden” in the sense that buy-side traders keep the true order size secret to minimize information leakage. Rather than being traded as a block, these orders are typically split into small slices and executed incrementally. The study of the impact of metaorders requires a different approach compared to individual or aggregate orders. The reason is that metaorders come from the same agent and generate strong correlations in order flow through a series of incremental executions (Lillo et al. (2004)). In turn, highly correlated order flow results in a heavy tailed distribution of trading volume.

This study was originally motivated by Farmer et al. (most recent version is FGLW 2013), who have proposed a theory for the market impact of hidden orders that the authors refer to as metaorders. FGLW theory is based on a traditional view in finance that market impact reflects information. Namely, the information on future stock prices maps into the distribution of order sizes, which in turn dictates the shape of market impact function. The authors assume that a transaction price is a martingale and they further assume fair pricing condition, which stipulates that the post-trade price after metaorder completion is equal to the average transaction price, and with these two assumptions they derive a relationship between the functional forms of two observable characteristics: order size distribution and the shape of market impact. This allows for the direct empirical tests that we discuss in this article.

The existing empirical literature on the market impact of metaorders is limited (Torre (1997), Almgren et al. (2003), Moro et al. (2009), Toth et al. (2011)) because of the difficulty of obtaining access to the data that would identify the ultimate trading accounts. Moro et al. (2009) statistically reconstructed hidden orders from the book of market transactions from the exchanges, using information about market member codes. The advantage of our study is that it uses a proprietary set of large trades that clearly identifies large orders. As a proxy of metaorder size distribution in the market, we use the distribution of proprietary large orders executed by AllianceBernstein's buy-side trading desk for approximately 50 distinct institutional funds. We show a Pareto law with the exponent $\beta \approx 1.56$ approximates our distribution. We empirically examine the shape of the total market impact (the slippage of the last fill price relative to the arrival price, defined as "peak" impact in the FGLW 2013) during order execution and find that the well-known square root function has the best fit, thus validating the FGLW prediction that, in case of a Pareto distribution, the market impact is a power function of order size with exponent $\beta - 1$.

We observe that the fair pricing condition holds across different order size groups and market conditions, suggesting that the condition may be enforced by the collective behavior of market participants. If both the realized and the post-trade prices can be measured directly, trade counterparties may only perceive it fair if "the current price of the asset, as priced by the market, is equal to the price that was paid [to acquire it]" (quotation from FGLW). This empirical observation immediately suggests the concavity of the permanent impact, as it is broadly accepted that the realized impact is a strongly concave function of trade size[1].

[1] R. Almgren, C. Thum, E. Hauptmann, and H. Li, Direct Estimation of Equity Market Impact, Risk, 18 (2005), p. 57; B. Toth, Y. Lemperiere, C. Deremble, J. De Lataillade, J. Kockelkoren, and J.-P. Bouchaud, Anomalous price impact and the critical nature of liquidity in financial markets, Physical Review X, 1 (2011).

Indeed, one of the most important findings of this study is that permanent impact is a strongly concave function of order size and is best described by a square root function. This observation contradicts existing conventional wisdom that permanent impact must be linear.

Another important finding is the nonlinear dependence of post-trade reversion time on order size. We address the question of the temporal dynamics of impact after the order is completed. To our knowledge, this is the first empirical study that examines the dynamics of market liquidity around large order arrivals, i.e., how quickly liquidity is being replenished after it has been removed during execution of an order. We study the relaxation dynamics of temporary impact, which consists of two distinct phases: fast initial power decay and slow exponential relaxation. We discuss the nature of such phenomena, as well as a coexistence of the non-linear permanent impact and the power/exponential decay of temporary impact.

This paper is organized as follows. In Section 1, we briefly summarize the major predictions of FGLW's theory. In Section 2, we present the descriptive statistics of the data and discuss the order size distribution. Section 3 discusses functional forms for the dependence of the total and permanent impact on trade size. We validate the relation between the total and permanent impact and provide testing of the fair pricing condition. In Section 4, we examine a functional form of impact decay and formulate the relation between order size and post-trade reversion time. Section 5 concludes the paper.

1. REVIEW OF THE FGLW THEORY: MARKET IMPACT OF HIDDEN ORDERS

In this section, we give a brief description of the FGLW (2013). The theory considers market ecology consisting of the three agent types: informed traders, liquidity (day) traders and market makers. Informed traders are presumed to be rational, long-term institutional investors who act on a common information signal α drawn from exogenously given distributions $p(\alpha)$ by generating buy or sell orders. All informed traders receive the same α signal, i.e., they are homogeneously informed. The day traders generate order flow based on their private information signal η_t at the beginning of each period t and submit a market buy or sell order, where η_t is the zero-mean IID noise process. Both informed traders' and day traders' orders are bundled together and traded incrementally over time via equally sized lots by taking liquidity supplied by competitive and profit maximizing market makers. The size distribution p_N of bundled metaorders maps into the distribution of the information signal $p(\alpha)$. Whereas the true size of the metaorders remains hidden, the size distribution p_N is known. The power of FGLW theory is that the impact is predicted in terms of p_N, which is directly measurable with the proper data, which we have. Therefore, FGLW theory adopts the traditional view in finance that impact reflects information, which translates into the distribution of order sizes.

There are two major assumptions. First, the theory imposes a martingale condition for a market price during the lifetime of a metaorder. Thus, the transaction price $\tilde{S}_t$ $(t = 1,2,...N)$, with t being successive transaction time for a metaorder of size $\kappa \cdot N$ split into N equal slices of size κ, is a martingale:

$$P_t(\tilde{S}_{t+1} - \tilde{S}_t) + (1 - P_t)(S_{t+1} - \tilde{S}_t) = 0 \qquad (1)$$

where P_t is the probability that the order continues beyond t executions, $(1 - P_t)$ is the probability that the order stops. $\tilde{S}_t$ and $\tilde{S}_{t+1}$ are the prices before and after the execution of the *(t+1)*-th slice and S_{t+1} is the post-trade price assuming the metaorder was fully executed in *t* slices. Thus, the price S_{t+1} reflects the permanent price impact. Eq. (1) may be rewritten as

$$\frac{\tilde{\Re}_t}{\Re_t} = \frac{1-P_t}{P_t}, \tag{2}$$

where $\tilde{\Re}_t = \tilde{S}_{t+1} - \tilde{S}_t$ and $\Re_t = \tilde{S}_t - S_{t+1}$ are incremental price responses to continuation and completion of a metaorder. The likelihood that a metaorder will persist depends on the distribution p_N and the number of executions *t* that it has already experienced. In particular, if p_N has tails heavier than an exponential, P_t will increase with time and thus, according to Eq. (2), incremental price response $\tilde{\Re}_t$ will decrease with time (and also with executed size, since size is equal to $\kappa \cdot t$) resulting in concave market impact.

The second FGLW assumption, the fair pricing condition, states that the average execution price is equal to the post-trade price, implying that the implementation shortfall is equal to the permanent price impact:

$$\frac{1}{N}\sum_{i=1}^{N}\tilde{S}_i - S_{N+1} = 0 \tag{3}$$

The martingale and fair pricing condition together form a system of equations that can be solved to express price dynamics in terms of probabilities P_t. The shape of market impact is then fully explained by the distribution of order size. In a special case of Pareto distribution of order size $p_N \sim \frac{1}{\xi(\beta)}\frac{1}{N^{\beta+1}}$, where $\xi(\beta)$ is the Riemann zeta function, FGLW theory makes the following predictions:

A. In the limit of large t, the total market impact (the slippage of the last price S_N relative to the arrival price, the authors also call it "peak" market impact) increases asymptotically as

$$\mathcal{I}_t \sim t^{\beta-1} \text{for } \beta \neq 1,$$

$$\mathcal{I}_t \sim \log(t+1) \text{for } \beta = 1$$

B. For $\beta \neq 1$, the permanent impact behaves asymptotically for large time t as

$$\mathrm{I}_t \sim \frac{1}{\beta} t^{\beta-1}$$

C. Combining A and B, the ratio of the permanent impact to the total impact is equal to

$$\frac{\mathrm{I}_t}{\mathcal{I}_t} \sim \frac{1}{\beta}$$

In Section 4, we will test predictions **A-C**, as well as the price fairness condition (3).

2. DATA

We examine single-day market orders executed by AllianceBernstein's buy-side trading desk from January 2009 to June 2011 in the U.S. equity market. For our sample, we select orders with a fill rate higher than 80%, trade duration longer than 10 minutes, participation rate in the range of 5 to 50% and order size in the range of 0.1 to 30% of the median daily volume (MDV). We remove trades in illiquid names with MDV less than 300,000 shares. We exclude orders arriving before 10 A.M. to avoid the effects of accentuated volatility in the morning. We also require the completion of trading by 3 P.M. so as to capture the post-trade price reversion. To further control for available liquidity, we removed trades with a trading rate that is outside of 0.01 to 0.1% MDV per minute: namely the trades that are either too aggressive or too passive were excluded (Figure 1). In our context MDV per minute is a proxy for available liquidity. The final sample contains 12,500 metaorders. To study the post-trade reversion over an extended period of time (we have looked up to 2 hours after the last fill), we removed orders with extreme slippages versus several post-trade benchmarks such as available VWAP and 10% PWP (participation weighted price) as explained in Appendix 1.

2.1. GENERAL STATISTICS

Table 1 presents the descriptive statistics of the trades. The sample is dominated by mid- and large-cap names with an average median daily volume of 3 million shares. Order sizes range from 0.1 to 21.8 % MDV. The data suggest that on average, it takes 42 minutes to complete a trade at an average participation rate of 16%. The average trading rate is 0.036%MDV per minute or 14%MDV per day; trading rate range of [0.01-0.1] MDV/minute is enforced for all orders in sample. Figure 1 shows the order size plotted against the trade duration of an order and the OLS results. Our MDV per minute filtering procedure lessens the notorious bias of underestimating the market impact of large trades that were (1) initiated and executed because unusually high liquidity was available or (2) traded too slowly because of price limits. It also results in linear dependence of average trade size on trade duration, a setup that is consistent with the FGLW model. The average trade participation rate is fairly stable across order durations: it is found within 15% to 19% range.

2.2 ORDER SIZE AND DURATION DISTRIBUTION

Figures 2 and 3 show the distributions of order size and trade duration that indicate power-law decay. We fit 60 different functional forms to the empirical distributions and rank them based on the Kolmogorov-Smirnov goodness of fit test. For both size and trade duration, Pareto distribution is ranked in top 30% of functional forms (18^{th} and 14^{th} out of 60) with the estimated exponents of 1.56 and 1.64.

We have looked at other distributions, such as lognormal and Inverse Gaussian (2P/3P) which were ranked above Pareto and noticed that extreme tails explain most of difference in the goodness of fit. In the middle the probability decay for both lognormal and Inverse Gaussian (2P/3P) is very close to that of Pareto: $\sim order\ size^{-2.5}$. This finding is further confirmed by the probability density plot (Figure 4), which is drawn in the log-log scale and shows good linear fit in the 1 to 20% MDV range with the slope ~2.48 (for probability density function, which corresponds to $\beta \approx 1.48$ for cumulative density function). Figures 2 and 3 include P-P plots that reveal more information about the goodness of fit in the middle and at the tails of empirical distributions.

We are not concerned with an accurate fit of the left tail of empirical distribution for small orders as filters described in the previous chapter may have distorted it significantly. Namely, the filters have removed a significant number of small orders that were executed either too passively or too aggressively (the cone in Figure 1 has minimal width for small orders). FGLW theory is based on the premise that market makers are able to observe large institutional orders via the autocorrelations of order imbalances or similar metrics. From this perspective, truncating very small orders makes logical sense; small orders do not create strong autocorrelations. We are more interested in the characteristics of large orders, and with this in mind we focus on examining the asymptotic distribution of the upper tail, e.g., the extreme quantile of order size. We ran a goodness of fit test based on the exponential tail (ET) estimator proposed by Diebolt et al. (2002).

We cannot reject the null that lognormal distribution gives an adequate approximation for the upper tail (0.05 quantile) of order size distribution at the 5% confidence level. This quantile corresponds to the threshold of 7% MDV in our sample. Additionally, we ran the

GPD (Diebolt et al. (2007)) test that approximates the survival distribution function by a General Pareto Distribution above 0.05 quantile and check goodness of fit for a Pareto distribution. We rejected Pareto distribution as an adequate extrapolation of the upper tail. Table 2 reports the results of the ET and GPD tests.

There may be different explanations for this result. We should keep in mind that there is a strong censorship in act and a cone shape of the trading rate data, which implies larger uncertainty in trade size for long durations. In fact, the data set includes only those trades that were completed by the end of the day. Therefore, on average, one should expect that trades submitted at midday have a shorter length than those submitted in the morning. This finding potentially may result in the underestimation of the frequency of long trades. Table 3 presents the summary statistics of time duration, conditional on the hour of the day in which the order was routed, to give an idea of how the submission time affects trade duration.

Another factor that may explain the relative scarcity of large trades is trade optionality. Portfolio managers tend to cancel costly orders rather than chasing a price at any cost. This optionality results in a faster decaying order size distribution and limits realized market impact of large trades. Section 3 indicates that a logarithmic model better describes market impact for large orders. This finding contradicts the prediction of the FGLW model for lognormal distribution in the limit of extremely large order sizes: while for smaller orders the lognormal distribution of order size has asymptotic behavior of Zipf's distribution, in case of very large orders, it would imply a convex market impact. Theoretically, this case would be possible for extremely large orders when trades have exhausted natural liquidity providers (i.e., those that can absorb some inventory without hurting their risk profile) and need to reward liquidity providers for taking additional risk. Hostile takeover would be an example of such trade, but this scenario takes us beyond the framework of the FGLW model, which assumes elastic liquidity supply.

Yet another explanation may come from the fact that we analyze the trades from a single asset manager. Vaglica et al (2008) shows that the metaorder size distribution for individual managers is lognormal, but once aggregated across different managers, a power law better fits the order size distribution. With this in mind we analyzed aggregated U.S. trading volumes by symbol and side executed by a diverse set of buy-side institutions that were provided to us in highly aggregated form by a third party TCA vendor. Appendix 2 gives a description of the data and a summary of the results. Our analysis of aggregated third-party data validates the results obtained solely for AllianceBernstein's trades. Specifically, we show that Pareto, with an exponent of roughly 1.5, closely describes the order size distribution for the largest portion of the data for large- and mid-cap trades, except the tails[2]. For the upper tail, presenting the top 20% of trades, the ET and GPD tests suggest rejecting Pareto distribution in favor of lognormal. Finding similar implications for the order size distribution forms in two different data sets provides a robustness check. See Appendix 2 for more details.

In summary, the order size distribution of metaorders in both proprietary and aggregated data sets is not inconsistent with a Pareto distribution with tail exponent of 3/2 for the bulk of the trades in both samples except extremely large order sizes.

3. IMPACT

It is well documented that large orders exhibit a price impact that decreases over time after order completion (Ben-Rephael et al. (2011)). However, the impact does not vanish completely. Thus, the market impact has two components: temporary impact that is caused by

[2] We were unable to establish statistically reliable results for small cap because of significant noise in the stitched data.

market frictions (noise caused by price discovery and market microstructure effects, such as a bid-ask bounce) and permanent impact that is caused by the change in the market's perception of the security as a result of a large trade. The total market impact is the sum of temporary and permanent components and for an order of duration T it is calculated as the return from the midpoint price S_{t_0} at trade arrival to the midpoint price S_{t_0+T} at the last fill:

$$\mathcal{I} = \varepsilon \frac{S_{t_0+T} - S_{t_0}}{S_{t_0}} \tag{4}$$

where ε the trade sign, $\varepsilon = 1$ for buys and $\varepsilon = -1$ for sells. We use the notations found in the FGLW paper, where the authors call this return the "peak" impact, implying that on average the market impact is at maximum at the last fill. The permanent price impact is defined as

$$\mathrm{I} = \varepsilon \frac{S_{t_0+T+T_d(T)} - S_{t_0}}{S_{t_0}} \tag{5}$$

where $S_{t_0+T+T_d(T)}$ is the midpoint price at T_d after the order completion. Almgren et al. (2005) use a fixed 30-minute interval after the completion of the trade to compute the permanent impact. As we will see, one of the results of this paper is that there is a monotonically increasing relation between the trade duration and the time T_d (T) needed to dissipate the transient price impact. Finally, the realized impact of an order executed at price $\overline{S}$ is

$$J = \varepsilon \frac{\overline{S} - S_{t_0}}{S_{t_0}} \tag{6}$$

3.1. TOTAL IMPACT

In this section, we investigate the dependence of the total price impact $\mathcal{I}$ on the trade duration. We consider three functional forms: linear, power-law and logarithmic:

$$\mathcal{I} = \beta \sigma T + \eta$$

$$\mathcal{I} = \beta\sigma T^{\gamma} + \eta \tag{7}$$

$$\mathcal{I} = \beta\sigma \ln(T) + \eta$$

where η is a noise term, and σ is a daily stock volatility calculated as the 30-day trailing average of Parkinson volatility. We estimate β and γ with a nonlinear least square regression against trade duration T. We compare the Schwartz Information Criteria (BIC) to select the best model, as is typically done in model selection. According to these figures, the best model is a power law with an exponent of roughly 0.5. Table 4 shows the estimated parameters from the three models. Figure 5 displays the market impact.

The empirical finding of square root impact as a function of time is consistent with a Pareto order size distribution with an exponent β of 1.56 according to the FGLW. However, it is worth noting that the difference in the goodness of fit between a square root and a logarithm is small. Interestingly, Gerig et al. (2011) argue that the type of market impact function depends on how a liquidity provider discerns transactions. If a market maker can associate each transaction with its corresponding parent order, she can predict future order flow by calculating the conditional probability that the order continues given prior information on transactions for this order. The authors show that in this case, market impact is a logarithmic function because there is more predictability in order flow that results in a lower price response for later transactions. By contrast, market impact is a power-law of time if liquidity providers use only past transactions to predict the future. In this case, the price impact is higher because order flow predictability is diminished. Gerig et al. (2011) argue that in reality, it is highly unlikely for market makers to precisely associate a transaction with its parent order. Therefore, it is plausible that the market impact form is somewhere in-between the two scenarios.

To provide further evidence of different functional forms for different order sizes, we formed 10 buckets of approximately equal sample size (~1000 orders per bucket) based on trade duration. Figure 6 shows the average total market impact against the average duration time for each bucket. Each point on this chart is given an equal weight – we purposely neglect the specifics of order size distribution and the fact that market impact is much noisier for larger orders (we do not weight each point by the inverse of its standard deviation). The chart demonstrates that smaller orders are best approximated by a square root, whereas larger orders are best described by a logarithmic function. This result makes intuitive sense: if FGLW fair pricing condition holds, a power-law impact implies increasing post-trade price reversion for larger orders - a situation that is not observed in the market. A logarithmic impact model, on the other hand, results in the average traded price $J = \int_0^T \beta \ln(t)dt = \beta(\ln(T) - 1)$, implying that temporary impact does not grow with the order size: $I - J = \beta \ln(T) - J = \beta = const$ and that the post-trade reversion expressed as a fraction of market impact, decreases with the order size. This finding is supportive of the explanation for the changing impact function based on the order flow predictability identified by Gerig et al. (2011). Large orders that have a long sequence of successive transactions create more predictable order flow that translates into a smaller marginal price response to subsequent transactions and a flatter impact curve compared to small orders.

3.2. POST-TRADE PRICE REVERSION

Calculating the post-trade price reversion and the permanent impact of an order is more complicated because the price profile becomes much noisier as we are moving further away from the order arrival. Instead of using a fixed time interval after the completion of the trade to compute the permanent impact, we take into account the dependence of the reversion time on trade duration.

First we construct a series of 5-minute volume-weighted average prices from the time of the last fill to market close. For example, for the price at 5 minutes after the last execution, we take the 5-minute VWAP after the last fill: $VWAP_{t_0+T+5}$. Similarly, the price at 10 minutes from the last execution is the VWAP between 5 and 10 minutes after the last fill. In this manner, we construct a post-trade reversion series that is defined by price changes from the midpoint price at the time of last fill to X minute VWAP vs. midpoint at arrival of an order:

$$R_X = \varepsilon \frac{VWAP_{t_0+T+X} - S_{t_0+T}}{S_{t_0}}, \quad R_0 \equiv 0 \tag{8}$$

We avoid taking closing prices as post-trade reversion prices by excluding orders that were still trading during the last 60 minutes of the trading day. To ensure that we are capturing the entire reversion, we started by calculating the price profile over an even longer time window, while being mindful about retaining the majority of orders. Table 5 reports flat averages for the reversions, total market impact and implementation shortfall up to 2 hours after an order has been completed. While all orders have at least a 60-minute reversion profile, only 77% of the orders have a 120-minute post-trade profile. On average, we infer that the maximum price reversion over the entire sample is achieved by 50 minutes after the order arrival and comprises 1/3 of the total price impact. Figure 7 displays this result. Additionally, Figure 7 shows the dollar-weighted average impact, suggesting that it takes longer for large orders to achieve price reversion.

To obtain improved granularity on price reversion time, we bucket trades based on execution time. We form 3 buckets of approximately equal size: 3,585 orders with durations of less than 20 minutes; 3,527 trades with durations between 20 and 45 minutes; and 3,054 orders with durations of more than 45 minutes. Reversion profiles for all three buckets are shown in Figure 8. For orders of < 20 minutes, the side-adjusted price increases from the post-trade level after approximately 60 minutes. This phenomenon is explained by multiple small orders in the same symbol and side traded throughout the day for related funds; in choosing our filters, we only made sure that new orders in the same symbol are not placed within a 60-minute window from the last fill. For each bucket, we run *t*-tests to find the first 5-minute interval where the price reversion is statistically indistinguishable from the reversion in all consecutive 5-minute bins within the 60-minute window. *T*-tests indicate that for the first bucket (<20 minutes), reversion dissipates in the 15 minutes after the last execution. For the second and third buckets, we find this time to be equal to 30 minutes and 40 minutes, respectively. To reduce noise in the post-trade price profile, we define the permanent price for orders in each bucket as the VWAP from the time of the first indistinguishable peak of price reversion to the end of the 60-minute window. For example, for orders with durations of less than 20 minutes, the permanent impact price is the VWAP from 15 minutes to 60 minutes after the last execution.

3.3. PERMANENT IMPACT: A NON-LINEAR CONCAVE FUNCTIONAL FORM

We calculate the permanent impact as the return from the price at trade submission to the permanent impact price which is defined in the previous section. We performed the regression analysis of Equations 7 of the permanent impact against the time duration for all trades. Table 6 shows the estimated parameters and BIC from the three models. A square root model provides a better fit to the data. This result is in line with the predictions from FGLW (2013) that a Pareto distribution with an exponent that is roughly near 1.5 determines the square root dependence of the permanent impact on execution time.

Our finding is in contrast to the linear permanent impact assumption prevalent in the literature. For example, Huberman and Stanzl (2004) argue that permanent impact must be linear in trade quantity to preclude quasi-arbitrage. The authors are careful to specify that linearity only applies when liquidity is constant. However, observed concave shape of total market impact implies that available liquidity increases with traded size (price impact per share decreases). It is feasible that large orders, which create larger temporary price dislocation, attract broader spectrum of market participants willing to bet for larger reward.

Almgren et al. (2005) also confirm empirically that the permanent impact is linear. We believe that the linear functional form of permanent impact obtained empirically by Almgren et al. (2005) is a result of measuring the post-trade permanent price at a fixed time, i.e., 30 minutes after the last fill for all orders regardless of their size. This potentially introduces a bias in the estimates of permanent impact because large orders may not exhibit the dissipation of temporary impact within 30 minutes of the last fill. Moreover, in cases when the post-trade time is after the market close, Almgren et al. (2005) roll over to the next morning. This is another source of noise. In addition, the empirical results for the permanent impact are influenced by the composition of orders in a sample. If a test sample is dominated by small orders, both the total and permanent market impacts deviate from the square root toward a less concave shape, and it becomes more difficult to distinguish between a linear and non-linear form for the permanent impact. This result is likely to occur if an empirical study is conducted based on small broker placements instead of large institutional orders.

3.4. RELATION BETWEEN PERMANENT AND TOTAL IMPACT

We are interested in investigating the relation between the total and permanent impact, i.e., we seek to quantify what fraction of the total impact persists after the trade is completed. We compute the ratio between the means of the permanent impact and of the total impact, i.e.

$$\eta = \frac{E(\mathcal{I} | T)}{E(\mathrm{I} | T)} \tag{9}$$

The estimates from Table 4 and Table 6 suggest that η=0.66 +/- 0.09. We observe that the value of η is consistently below 1. The ratio of the means is consistent with the theoretical value predicted by FGLW (2013). Specifically, a Pareto distribution with an estimated exponent of approximately 1.5 implies that the ratio of the permanent impact to the total impact roughly equals 2/3.

3.5. THE FAIR PRICING CONDITION: DIRECT TEST

According to the assumptions of FGLW, the post-trade reversion price (i.e., the arrival price plus the permanent impact) is equal to the average price. In other words, the expected realized impact J is equal to the expected permanent impact I.

We directly test this hypothesis. Specifically, we check whether the difference between the realized and the permanent impact has a zero mean by using a *t*-test. The results of the test are shown in Table 7. We cannot reject the null hypothesis of no statistically significant differences in the means for the realized and permanent impacts for the pool of our orders.

Alternatively, we can show that if the expected market impact is a power-law of trade duration, i.e., $\mathcal{I} = \beta T^{\gamma}$, $\gamma \in (0,1]$, then the expected realized impact J is equal to $\frac{1}{T}\int_0^T \beta t^{\gamma}\, dt = \frac{1}{1+\gamma}\beta T^{\gamma}$. In our case, $\gamma = 0.5$ implies that $J = 2/3\mathcal{I} \longrightarrow J = \mathrm{I}$.

We empirically established that on average, the execution price is equal to the post-trade price after reversion. This is a fairly straightforward test that can be conducted by directly measuring two observable quantities, at least for single-day orders: the average execution price and the average post-trade price. To ensure an accurate estimation of the latter, it is important to consider a reasonably large post-trade time window, while avoiding market after-hours or rolling to the next day. We checked that the fair pricing condition is a robust empirical fact that holds in different market conditions (periods) and across different order and stock characteristics such as trade aggressiveness, volatility and market capitalization.

Although the fair pricing condition is an empirically observed phenomenon, it has not been adequately addressed in the market microstructure literature. Our empirical findings strongly suggest that further studies of this striking phenomenon are warranted.

In conclusion, we wish to stress that concave permanent market impact is an inescapable consequence of the observed concavity of total market impact together with our empirical confirmation of the fair pricing condition.

4. DEVIATIONS FROM THE FGLW MODEL: RELAXATION DYNAMICS OF MARKET IMPACT

Although the FGLW model predicts both total and permanent impact levels, it does not explicitly address impact relaxation dynamics. In fact, a strict application of martingale condition would imply instantaneous post-trade reversion. Prior studies on market impact indicate two commonly assumed functional forms for impact decay: power-law (Bouchaud et al. (2009)), Gatheral (2010) and exponential (Obizhaeva and Wang (2005)). On the other hand, the relaxation dynamics of various physical processes observed in materials science, electronics, chemistry, etc. frequently combines both regimes over different time scales. According to Karagiannis et al. (2007), finding two different regimes of impact dissipation seems to be a broad empirical phenomenon found in multiple studies. A fast power law dissipation of system perturbations that transition over time to a slower exponential regime is typical. The reason behind complex relaxation dynamics is the coexistence of diverse mechanisms manifesting themselves over different time scales. Our approach in fitting the data to the relaxation process highlights the importance of separating diverse regimes.

4.1. ESTIMATION PROCEDURE

We begin empirical tests for a functional form of market impact relaxation by studying price reversion $R_{i,n}$ series as it was defined in Section 3.2, Eq. 8, where i is an order identifier, and n is time elapsed since the order's last fill, measured in 1-minute bin. To reduce noise in the post-trade reversion data, we normalize $R_{i,n}$ by the expected total market impact $\mathrm{E}[\mathcal{I}_i] \sim 0.0225 * \sigma_i \sqrt{T_i}$ obtained from the regression (Table 4). Thus, all impact relaxation trajectories start from the same point ($R_{i,0} \equiv 0$) and have the same magnitude.

Figure 10 displays the flat average of market impact relaxation for 4 buckets of orders. Initially the dataset was split into 6 buckets, which were equally spaced by order duration on the logarithmic scale: 11-17, 18-29, 30-49, 50-78, 79-127, >=128 minutes. Our analysis did not uncover statistically significant differences in relaxation dynamics between buckets 3 and 4 and between buckets 5 and 6; therefore, we combined the buckets.

As we saw in the Section 3, for all order durations the price relaxes to approximately 2/3 of the total market impact (and thus Fig. 10 shows stabilization around -0.33 levels). Figure 11 depicts log-log and log-linear charts of the average impact relaxation for all 4 buckets. To draw these charts, we compute $\ln(1+\overline{R}_{i,n} / (\overline{\mathcal{I}}_i - \overline{\mathrm{I}}_i))$, where $\overline{R}_{i,n}$ is average impact relaxation series across all orders in a bucket, $\overline{\mathcal{I}}_i$ is the average total market impact from Eq. 4 and $\overline{\mathrm{I}}_i$ is the average permanent impact, defined by Eq. 5. The last formula is equivalent to $\ln((VWAP_{t_0+T+n} - S_{t_0+T+T_d(T)})/(S_{t_0+T} - S_{t_0+T+T_d(T)}))$ with $VWAP_{t_0+T} \equiv S_{t_0+T}$, $VWAP_{t_0+T+\infty} \approx S_{t_0+T+T_d(T)}$

The exact values of decay exponents are sensitive to the level of permanent impact (e.g. the choice of $T_d(T)$ for each bucket). However, we noticed that for all buckets in sufficiently wide range of assumed $T_d(T)$, the impact relaxation series is better described by power-law decay in the first 10 minutes after order completion. After that period, exponential impact decay provides a better fit to the data over the rest of the post-trade period until the temporary impact is fully dissipated. Figure 11 demonstrates the change in the regime for specific normalized levels $(\overline{P}_i - \bar{I}_i) = 0.3, 0.3, 0.38, 0.35$ found for $T_d(T)$=12,28,41,62 minutes.

Fast power-law market impact decay in the first few minutes after order completion can be explained by an aggregate trading activity of market makers and arbitrageurs. Stock price is most dislocated immediately after order completion, suggesting profitable arbitrage opportunities that attract competitive arbitrageurs and market makers and put significant pressure on price, resulting in rapid price correction. Once the price moves closer to its permanent level, arbitrageurs flip the side of the trade to keep flat positions, e.g. that they start buying/selling the stock that they were selling/buying earlier. The arrival of contra liquidity slows the price impact decay, which is revealed in the slower exponential decay. We found the magnitude of power-law decay is about a half of bid-ask spread for two smallest buckets k=1,2 which have order durations under 30 minutes and about full bid-ask spread (~5bps) for buckets k=3,4. For k=1,2 buckets most of the temporary impact is dissipated through power-law regime: 100% for k=1 and 72% for k=2. For larger orders, these numbers are 55% and 32%, respectively. Thus, prior findings by Bouchaud et al. (2004) on the power-law decay of the impact are consistent with our results for relatively small orders with durations of up to 30 minutes, where the observed impact decay is dominated by a power-law regime.

For the purpose of quantifying exponential decay for the buckets with duration less than 30 minutes, the residual temporary impact after initial power decay is smaller than the minimum tick size and thus bid-ask bounce make it difficult to determine the parameters of the decay. For larger buckets k=3,4 and for large multi-day orders discussed in Appendix 2, exponential impact dominates and therefore we need a method that can consistently determine both the level of permanent impact and the value of exponent in the exponential decay. We suggest estimating the parameters of an Ornstein-Uhlenbeck (O-U) process

$$dR_{i,n} = -\theta(R_{i,n} - \mu)dt + \tilde{\sigma}_i dB_{i,n}, \quad R_{i,n}(n=0)=0, \quad R_{i,n}(n=N_k) \to \mu \tag{10}$$

for $n \in [10, N_k]$ is the time in minutes by which price reverts to its permanent level in each bucket k=1,2,3,4. $R_{i,n}$ is impact reversion series normalized by E[$\mathcal{I}_i$] $\sim 0.0225 * \sigma_i \sqrt{T_i}$, μ is the expected permanent level, $\theta > 0$ is a speed of impact reversion, $\tilde{\sigma}_i = \sigma_i / E[\mathcal{I}_i] = a/\sqrt{T_i}$ is normalized volatility and $B_{i,n}$ is a Brownian motion. Integrating equation (10) produces

$$R_{i+1,n} = e^{-\theta} R_{i,n} + (1 - e^{-\theta})\mu + Z_i, \text{ where } Z_i \sim N(0, 1 - e^{2\theta}/2\theta T_i) \tag{11}$$

It remains for us to determine the appropriate relaxation time N_k after which the price impact no longer decays. We observe in our data (Fig. 10) that once the price impact reaches the permanent level, it stabilizes around that level, showing small fluctuations due to bid-ask bounce. Consequently, estimates of the O-U parameter μ should not depend on the choice of sub-window within this impact stabilization zone $n > N_k$.

Thus, we run O-U for each bucket with a sliding window $\Delta = 5\,\text{min}$, starting from the end of the post-trade time interval (60 minutes for the first bucket and 75 min for the other buckets) and moving toward the beginning of the price reversion with a 1-minute step. Figure 12 shows the values for O-U parameter μ for each bucket. A critical time N_k for each bucket is chosen as the time when μ reaches its first local minima, e.g., when monotonic decay ends. Critical values of N_k are found to be 12, 28, 41 and 62 minutes for k=1,2,3,4, respectively. These numbers are consistent with the charts in Figure 10, suggesting that it takes longer for larger orders to achieve a relaxation of temporary impact. We also examined other criteria, such as finding the points where O-U coefficients θ or μ permanently shift away from the values observed in the stabilization zone, e.g., $\| \theta_{l<k} - \bar{\theta}_{m>k} \| > \delta_1$, $\| \mu_{l<k} - \bar{\mu}_{m>k} \| > \delta_2$, where $\delta_{1,2}$ are 95% confidence intervals in the stabilization zone. This last method generates of lower estimates of N_k, while μ are still hovering around -0.33.

Having determined the critical times N_k, we then run O-U for each order duration bucket in the time window $[10, N_k]$. Table 8 summarizes the results: O-U estimates of the permanent impact levels $\mu = -0.3, -0.3, -0.38, -0.35$ closely agree with the methodology described in Section 3.2; however O-U method also provides the estimates for the duration and the intensity of the decay. For the two large buckets 30-78 minutes and >79 minutes the estimated exponents in the exponential decay stage appear to be statistically different.

We run diagnostic tests for the standardized residuals to validate the model specification of price reversion by the O-U process. We find that the standardized residuals for each bucket are closely described as i.i.d normal ~ N (0, ∑), except fat tails - the empirical phenomenon in financial time series; 98% of the standardized residuals are bounded within [-2,2] range. We do not find significant autocorrelation patterns in the residual time series: the first order autocorrelation is the largest at 0.1.

Four buckets are not enough to confidently estimate the relationship between order size *T* and characteristic decay time. We notice that the exponent in the power-law decay correlates well with the exponent in the market impact function. Therefore, it was not surprising to find that for buckets where power-law decay dominates, the characteristic decay time has the same magnitude as the order size *T*. The last bucket exhibits mostly exponential decay, and at 62 minutes, the decay time is faster than the average order size in the bucket (126 minutes). This observation was independently confirmed for multi-day orders in Appendix 2: for large mid- and large-cap orders with duration of 2 to 8 days, post-trade reversion took only 1 day. We find that the market impact for such orders is better described by a logarithmic function. Notably, a fair pricing condition holds and thus percent of impact decay decreases with order size. As expected in case of logarithmic impact, the magnitude of price reversion does not increase with order size, suggesting that market impact decay time may not increase as well. This phenomenon can be explained by the interaction with a broader spectrum of liquidity providers attracted by the significant price dislocations that these large orders have caused.

5. CONCLUSION

In this paper, we have presented an empirical study of the price impact of hidden market orders in the U.S. market. Our findings provide supportive evidence for the theory proposed by FGLW (2013) showing that order size distribution fairly closely defines the shape of market impact during execution. The major empirical finding of the paper is a confirmation of the fair pricing condition and thus nonlinearity of permanent impact, as in the FGLW theory; the fair pricing condition imposes permanent impact to be on average a square root of trade duration because the realized impact is known to be a square root of order duration/trade size. In our tests, we observed robustness of the fair pricing condition across different market conditions and stock characteristics. While we present empirical evidence of nonlinear market impact, we are not aware of any theoretical study that can explain this empirical phenomenon from the perspective of no price manipulation. We anticipate further research on nonlinear market impact models that can adequately capture the realistic features of financial markets.

Although in aggregate we find that the order size distribution is described by Pareto with an exponent of roughly 1.5, implying a square root of market impact, we have found different impact curves for different order size groups, suggesting that market impact is a multi-regime process. In particular, we find that the upper tail of order size distribution in our sample is better approximated by a lognormal distribution that has asymptotic behavior of Zipf's law. While theoretically lognormal order size distribution would imply convex impact function for very large orders, we do not confirm it possibly because of trade optionality. We showed that impact decay is a multi-regime process, approximated by a power law in the first few minutes after order completion and exponential decay thereafter.

In conclusion, FGLW (2013) theory determines the market impact of hidden orders during execution solely from the (assumed known) distribution of order size. While our empirical results seem to largely validate this approach, the question remains as to whether the full complexity of order flow can be effectively described by a single characteristic such as distribution of order size. Market impact is a result of price discovery that reflects supply and demand, i.e., the provision and removal of liquidity. Therefore, this paper suggests a venue for the further research that can potentially extend the theoretical foundation of the hidden order arbitrage theory to include other variables driving supply-demand dynamics, which ultimately defines the shape of market impact function.

REFERENCES:

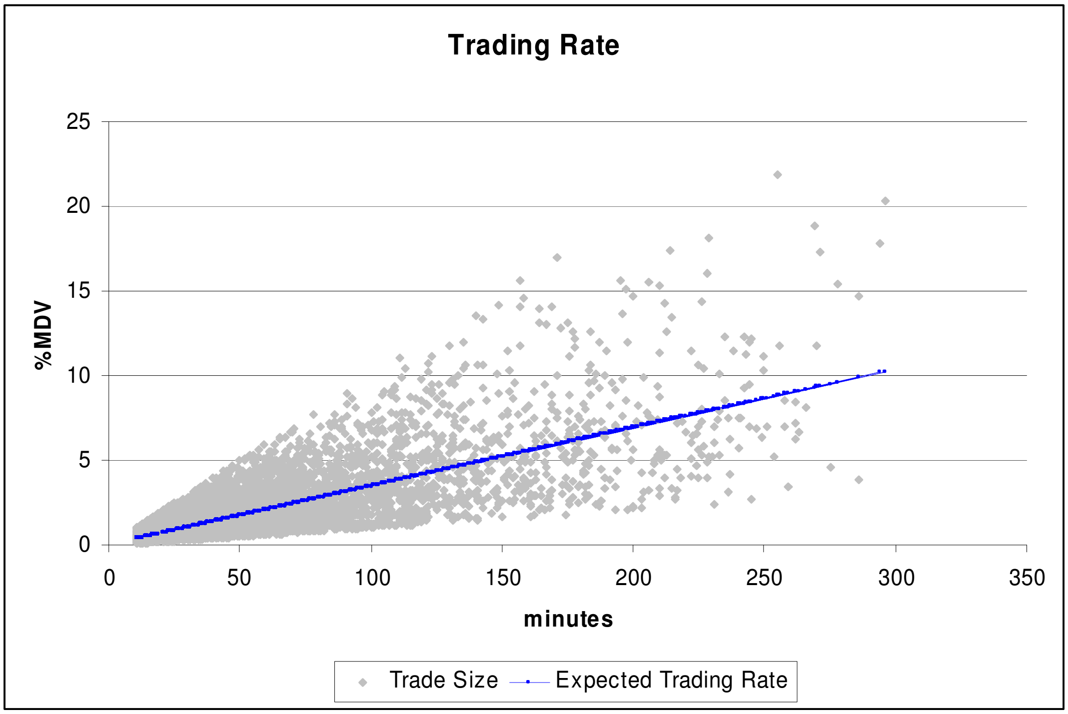
1. Almgren R. and N. Chriss, Optimal execution of portfolio transactions, Journal of Risk, 2000, 3:5-39.
2. Almgren R., Optimal executions with non-linear impact functions and trading-enhanced risk. Applied Mathematical Finance, 2003, 10, 1-18.
3. Almgren R., C. Thum, H. Hauptmann and H. Li, Equity market impact. Risk, 2005.
4. Ben-Rephael A, S. Kandel and A.Wohl, The price pressure of aggregate mutual fund flows. Journal of Financial and Quantitative Analysis, 2011, Vol. 46, pp.585-603.
5. Bouchaud J., Y. Gefen, M. Potters and M. Wyart, Fluctuations and response in financial markets: the subtle nature of "random" price changes. Quantitative Finance, 2004, Vol. 4, pp. 176-190.
6. Bouchaud J., D. Farmer and F.Lillo, How markets slowly digest changes in supply and demand. Handbook of Financial Markets: Dynamics and Evolution, 2009.
7. Diebolt J., M. Garrido and S. Gerard, Asymptotic normality of the ET method for extreme quantile estimation. Application to the ET Test, INRIA Research Report, 2002.
8. Diebolt J., M. Garrido and S. Gerard, A goodness-of–fit test for the Distribution Tail, Topics of Extreme Value Distributions, Chapter 5, pp. 95-109, Nova Scotia, 2007.
9. Gabaix X., P. Gopikrishnan, V. Plerou and E. Stanley, A theory of power-law distributions in financial market fluctuations. Nature 2003, 423, 267-270.
10. Gatheral J., No-dynamic arbitrage and market impact. Quantitative Finance 2010, Vol. 10, pp.749-759.
11. Gerig A., J.D. Farmer and F. Lillo, How prices respond to worked orders. Working paper, 2011.
12. Hopman C., Essays on the relation between stock price movements and orders. Ph. D. thesis, MIT, 2003.
13. Huberman G. and W. Stanzl, Optimal liquidity trading. Review of Finance, 2005, Vol. 9. Issue 2, 165-200.
14. Farmer D., Gerig A., Lillo F. and H. Waelbroeck, Information arbitrage and the rational pricing of market impact. Working paper, 2009.
15. Farmer D., Gerig A., Lillo F. and H. Waelbroeck, How efficiency shapes market impact, http://arxiv.org/abs/1102.5457, 2013.
16. Farmer D., Gillemot L, Lillo F., Mike S. and A. Sen, What really causes large price changes? Quantitative Finance, 2004, 4.
17. Karagiannis T., Le Boudec J-Y. and M. Voinovic. Power law and exponential decay of inter contact times between mobile devices", Microsoft Research, Proceedings of ACM Mobicom 2007.
18. Klößner S., M. Becker and R. Friedmann, Modeling and measuring intraday overreaction of stock prices, Journal of Banking and Finance, 2012, Vol.36, p.1152-1163.
19. Lillo F., D. Framer and R. Mantegna, Master curve for price impact function. Nature, 2003, 421, pp.129-130.
20. Lillo F. and D. Farmer, On the origin of power law-tails in price fluctuations. Quantitative Finance, 314, 7-10, 2004.

21. Moro E., J.Vicente, L.Moyano, A. Gerig, D.Farmer, G.Vaglica and F.Lillo, Market impact and trading profile of hidden orders in stock markets. Physical Review, 2009, Vol. 80.
22. Obizhaeva A. and J. Wang, Optimal trading strategy and supply-demand dynamics. AFA Boston Meeting Paper, 2005.
23. Plerou V., Gopikrishnan P., X. Gabaix, H. Eu. Stanley, Quantifying stock-price response to demand fluctuations. Physical Review, 2002, Vol. 66.
24. Potters M. and J. Bouchaud, More statistical properties of order books and price impact. Physica, 2003, 324, 133.
25. Torre N., BARRA market impact model handbook. BARRA Inc., Berkeley, 1997.
26. Toth B., Y. Lemperiere, C. Deremble, J de Lataillade, J Kockelkoren and J. Bouchaud, Anomalous price impact and the critical nature of liquidity in financial markets. Physical Review X 1(2), 021006, 2011
27. Vaglica G., F. Lillo, E. Moro and R. Mantegna, Scaling laws of strategic behavior and size heterogeneity in agent dynamics. Physical Review 2008, Vol.7.

Table 1. Descriptive statistics of the trade sample (10,166 trades)

Variables	**Duration (min)**	**Order Size (% MDV)**	**Participation Rate (%)**	**Spread (bps)**	**MDV (mln. shares)**	**Daily High to Low Volatility (%)**
Mean	42	1.4	16	5.1	6	2.7
Median	27	0.8	14	4.1	3	2.4
Stddev	40	1.8	8	3.5	11	1.5

Figure 1. We control for available liquidity by imposing minimum and maximum trading rate (%MDV/minute) constraint on our sample of trades. Min=0.01%/minute, Max=0.1%/minute. This measure lessens the notorious bias of underestimating the market impact of large trades that were (1) initiated and executed because unusually high liquidity was available or (2) traded too slowly because of price limits.



The trading rate constraint results in constant average trading rate across all order durations.

$S = const + \beta T + \eta$, S is trade size (%MDV), T is trade duration (minutes), η is error term; $R^2 = 62\%$

Model	**Mean**	**StdErr**	**T-stat**	**95% CI**	
const	-0.096	0.016	-5.91	-0.128	-0.064
β	0.036	0.0003	128.91	-0.035	-0.036

Average participation rate in our sample has very weak dependence on order duration. For durations [10, 209] minutes which cover 99% of our trades, the average participation rate stays within narrow [15, 19]% range.

$PartRate = const + \beta T + \eta$, $PartRate$ in %, T is trade duration (minutes), η is error term; $R^2 = 1\%$

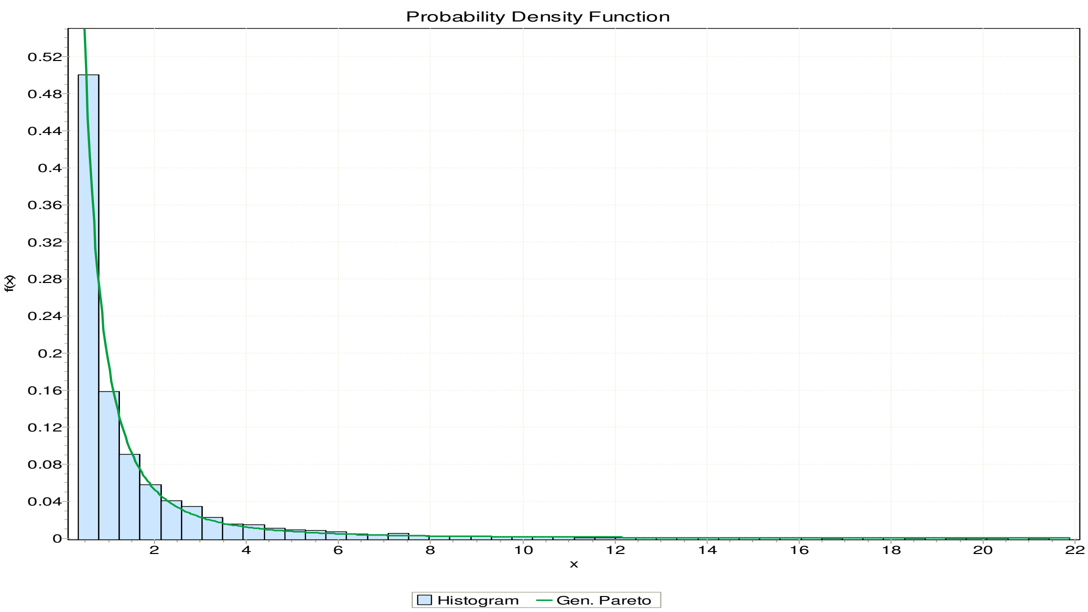
Model	**Mean, %Volume/min**	**StdErr**	**T-stat**	**95% CI**	
const	14.659	0.118	124.62	14.4	14.9
β	0.021	0.002	10.57	0.017	0.025

Figure 2. Distribution of order sizes: Cumulative Distribution Function, Generalized Pareto:

$$F_{(\xi,\mu,\sigma)}(x) = 1 - (1 + \frac{\xi(x-\mu)}{\sigma})^{\frac{-1}{\xi}} \; for \xi \neq 0,$$

$$F_{(\xi,\mu,\sigma)}(x) = 1 - \exp(-\frac{x-\mu}{\sigma}) \; for \xi = 0.$$

The shape parameter $\xi = 0.64$, suggesting $\beta \sim 1/\xi = 1.56$.



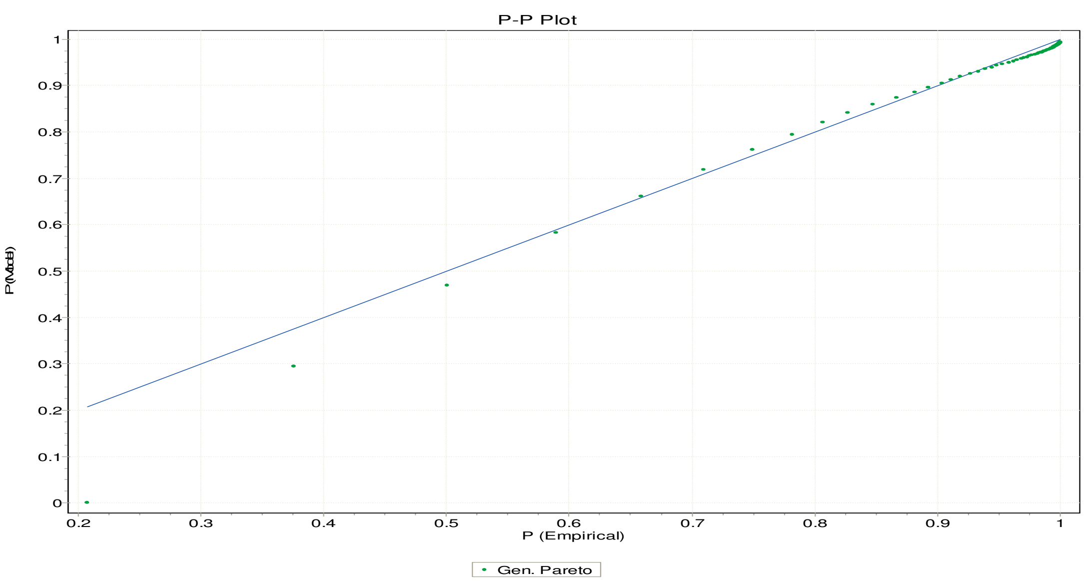
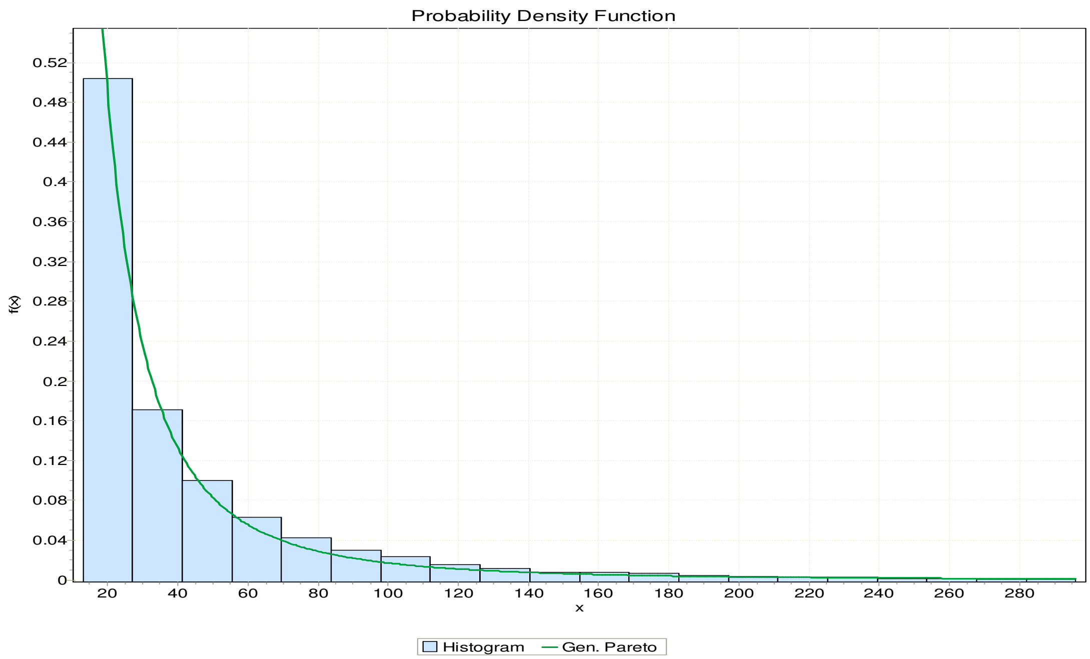


Figure 3. Distribution of trade durations: Cumulative Distribution Function, Generalized Pareto Distribution: $\xi = 0.61$, $\beta \sim 1/\xi = 1.64$.



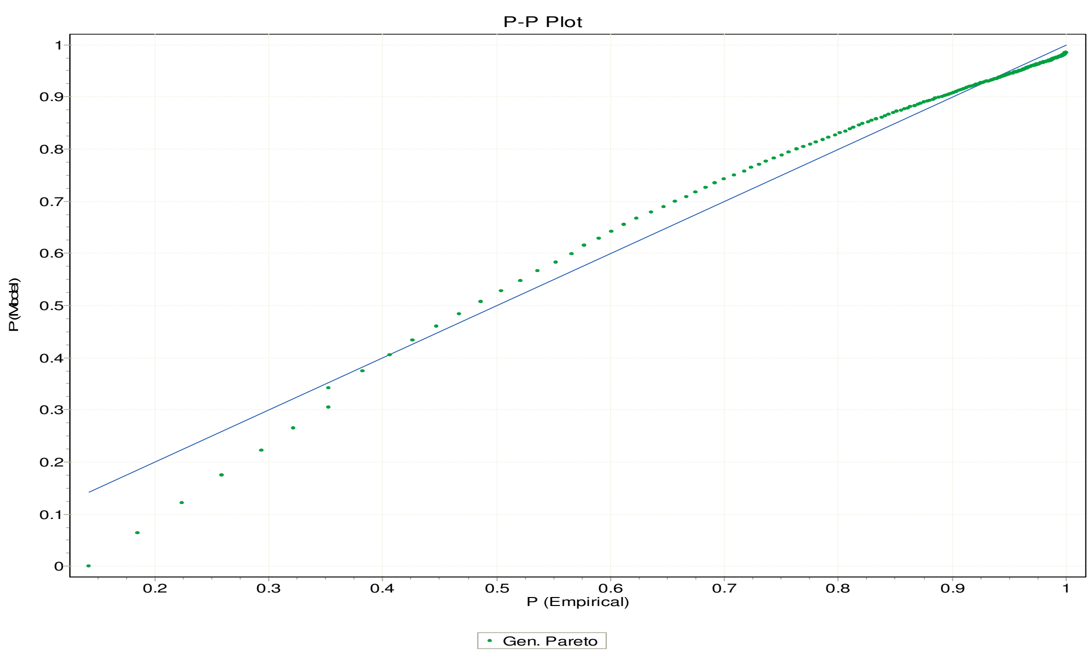
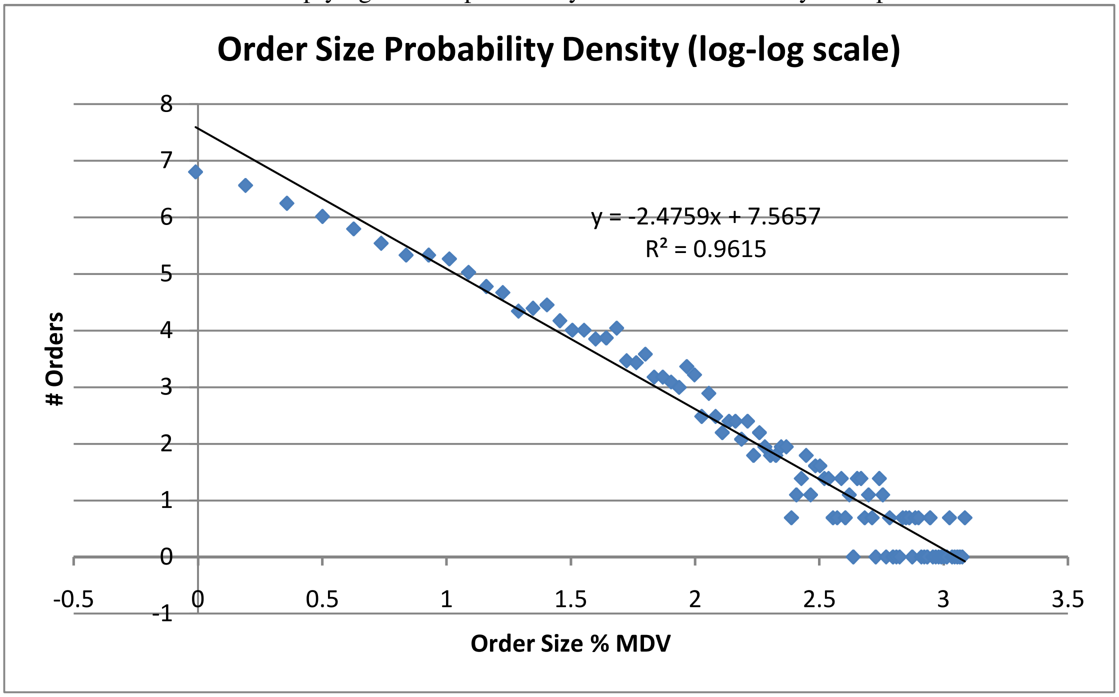


Figure 4. Log-log plot of number of orders versus order size expressed as %MDV. The table below shows statistics for linear fit model implying that the probability of order size S decays as a power function of S.



$R^2 = 96\%$

Model	Mean	StdErr.	T-stat	95% CI
Coefficient	-2.4759	0.0511	-48.44	-2.5773 -2.3744
Intercept	7.5657	0.1198	63.11	7.3277 7.8037

Table 2. This table summarizes the results of ET and GPD tests for 0.05 quantile of the distribution at the 5% confidence level.

TEST	**ET**
Choice Model:	Lognormal
Test Level:	0.05
# of excesses	220
# bootstrap sample	1000
Quantile p of order 1-p	0.05
Test Results:	**Accepted**
Estimated Quantile ET	4.6819
Estimated q	4.3478
qET-qParam	0.334
Inf CI	4.3357
Sup CI	5.3

TEST	**GPD**
Choice Model:	Pareto
Test Level:	0.05
# of excesses	220
# bootstrap sample	1000
Quantile p of order 1-p	0.05
Test Results:	**Rejected**
Estimated Quantile GDP	7.0236
Estimated q	46.1553
qET-qParam	39.1317
Inf CI	149.2318
Sup CI	276.055

Table 3. Distribution of order sizes by time of the day. Removing orders that come late in the day shortens average order duration the in afternoon and reduces number of large orders, introduces a bias in the order size distribution.

Hour	#Observations	Mean Duration (min)	Median Duration (min)	StdErr Duration (min)
10:00 AM	1347	58	60	31
11:00 AM	2008	53	32	50
12:00 AM	2614	43	30	36
1:00 PM	2500	35	26	25
2:00 PM	1697	26	21	17

Table 4. Estimated parameters for three functional models (Eq. (7)) for the total market impact as a function of trade duration

Model	Mean β	StdErr β	T-stat β	Mean γ	StdErr γ	T-stat γ	BIC	R-square, %
Linear	.0020	.00003	53.56				89741.92	24
Power	.0225	.0018	12.26	.4713	.0196	23.94	89139.7	29
Log	.0367	.0006	59.63				89234.16	28

Figure 5. Total market impact is the best approximated by a square root function: $I = \beta\sigma T^{\gamma} + \eta$ where γ=0.4713.

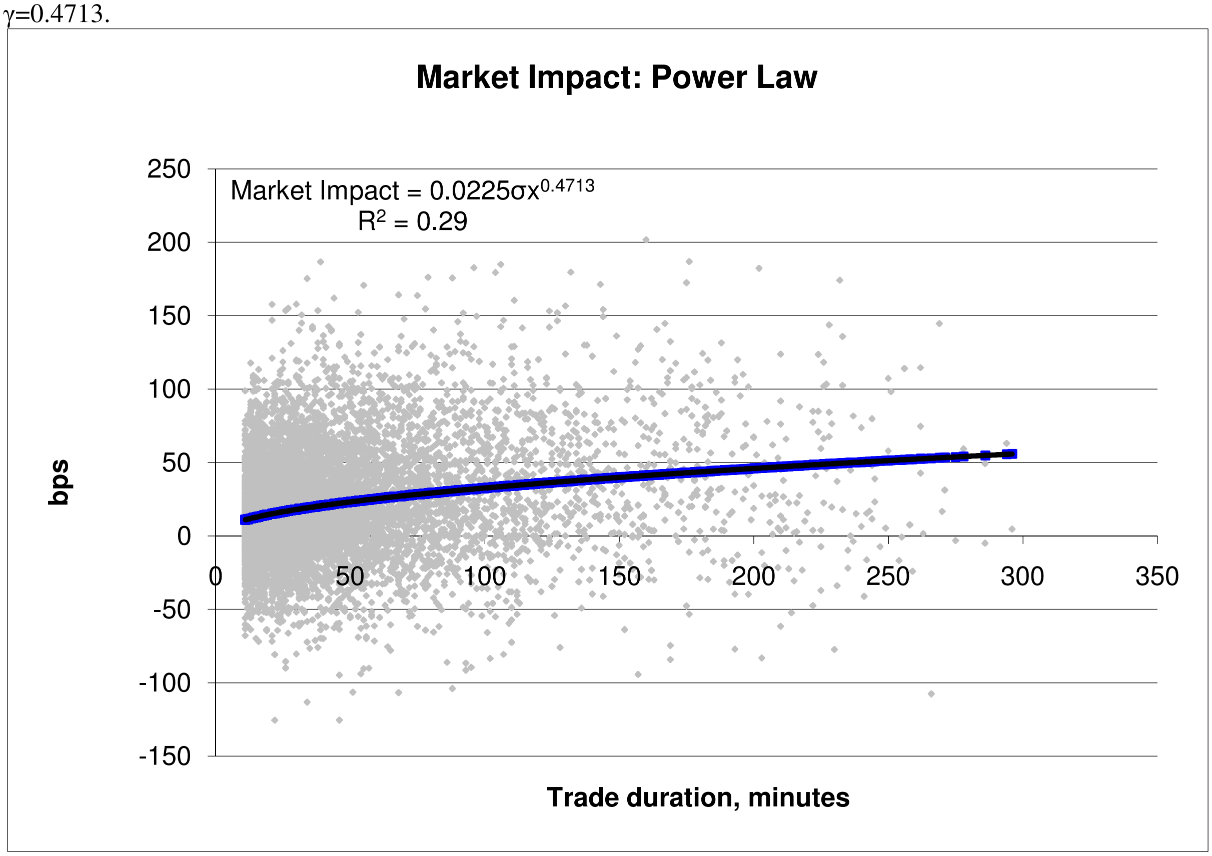


Table 5. The table displays flat averages for total market total impact, realized impact and reversion. All estimates are based on the filtered set of data (see in Appendix 1).

Variable	# Obs.	Mean (bps)	StdErr. (bps)	95% CI (bps)
$\mathcal{I}$	10166	19.4718	.3420	18.8014 20.1423
J	10166	13.9800	.2432	13.5032 14.4569
R_5	10166	-1.3110	.0549	-1.4187 -1.2033
R_{10}	10166	-3.1270	.1159	-3.3543 -2.8997
R_{15}	10166	-4.0950	.1648	-4.4183 -3.7718
R_{20}	10166	-4.5503	.1998	-4.9420 -4.1585
R_{25}	10166	-5.0957	.2260	-5.5389 -4.6525
R_{30}	10166	-5.5480	.2446	-6.0276 -5.0683
R_{35}	10166	-5.9104	.2644	-6.4289 -5.3920
R_{40}	10166	-6.0698	.2826	-6.6239 -5.5158
R_{45}	10166	-6.2335	.2992	-6.8201 -5.6468
R_{50}	10166	-6.3581	.3155	-6.9766 -5.7396
R_{55}	10166	-6.3319	.3342	-6.9870 -5.6767
R_{60}	10166	-6.2346	.3548	-6.9303 -5.5389
R_{65}	10113	-6.1679	.3802	-6.9134 -5.4224
R_{70}	9901	-6.0348	.4099	-6.8383 -5.2313
R_{75}	9695	-6.0371	.4350	-6.8898 -5.1843
R_{80}	9506	-5.9282	.4579	-6.8258 -5.0306
R_{85}	9306	-5.8022	.4826	-6.7482 -4.8561
R_{90}	9090	-5.8198	.5027	-6.8053 -4.8343
R_{95}	8895	-5.8318	.5204	-6.8519 -4.8116
R_{100}	8699	-5.8180	.5445	-6.8855 -4.7506
R_{105}	8523	-5.8807	.5666	-6.9915 -4.7698
R_{110}	8317	-5.6356	.5875	-6.7873 -4.4839
R_{115}	8121	-5.6250	.6089	-6.8187 -4.4313
R_{120}	7918	-5.4236	.6321	-6.6628 -4.1845

Figure 6. This figure displays market impact and its 95% confidence intervals for 10 buckets of roughly equal size.

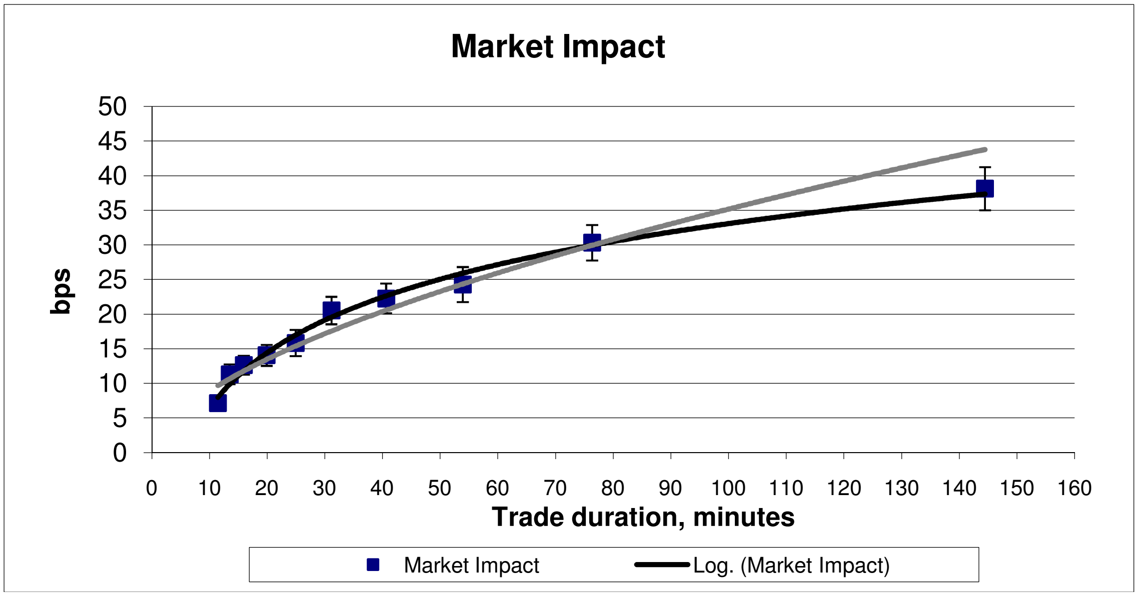
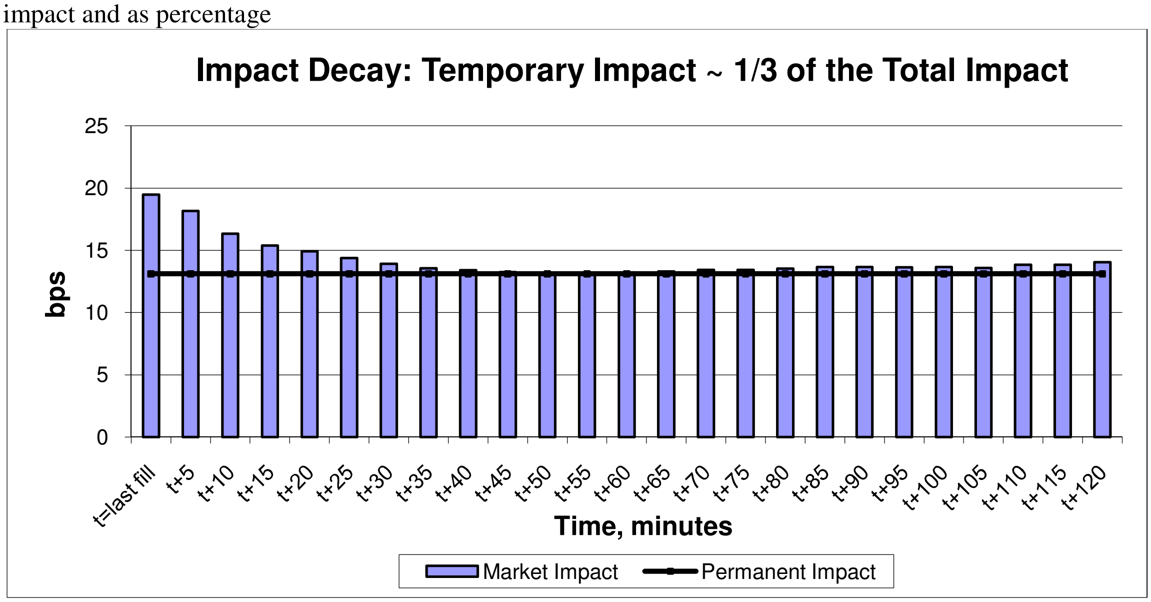
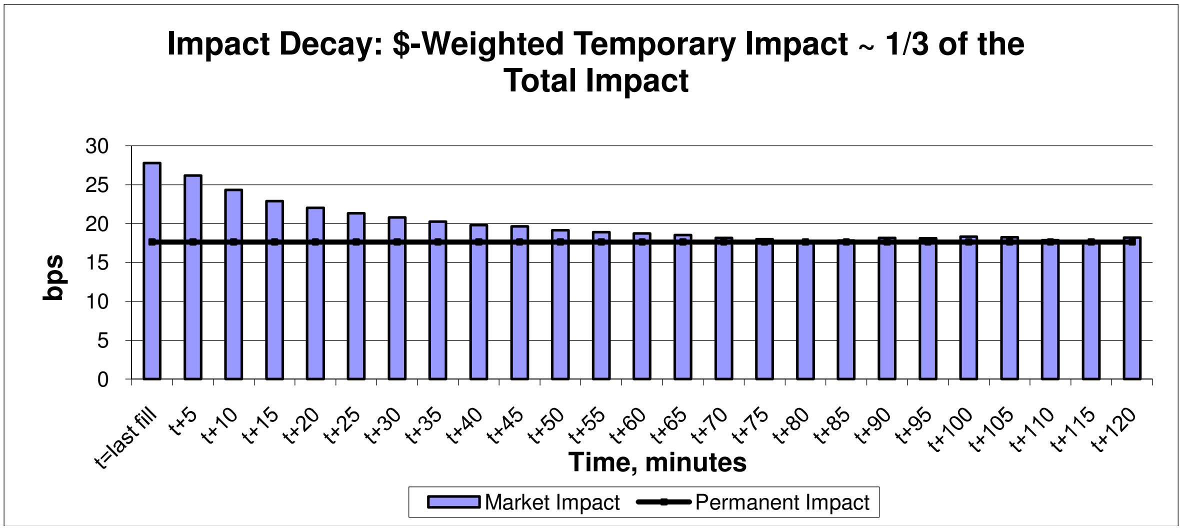
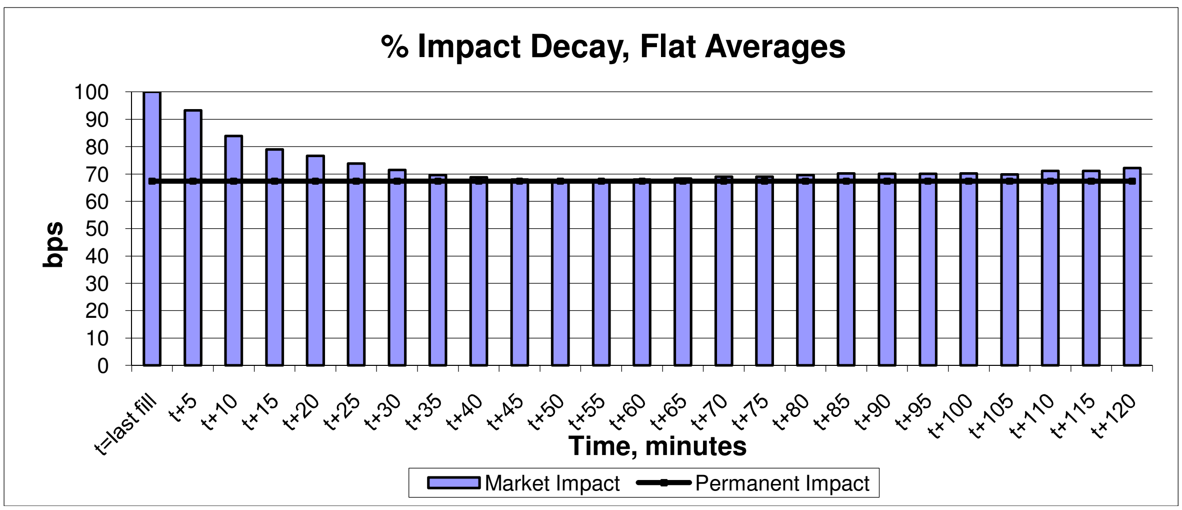


Figure 7. This table shows flat and $-weighted average temporary impact decay as proportion to total market impact and as percentage




Impact Decay: $-Weighted Temporary Impact ~ 1/3 of the Total Impact
bps
30
25
20
15
10
5
0
t=last fill
t+5
t+10
t+15
t+20
t+25
t+30
t+35
t+40
t+45
t+50
t+55
t+60
t+65
t+70
t+75
t+80
t+85
t+90
t+95
t+100
t+105
t+110
t+115
t+120
Time, minutes
Market Impact
Permanent Impact


% Impact Decay, Flat Averages
bps
100
90
80
70
60
50
40
30
20
10
0
t=last fill
t+5
t+10
t+15
t+20
t+25
t+30
t+35
t+40
t+45
t+50
t+55
t+60
t+65
t+70
t+75
t+80
t+85
t+90
t+95
t+100
t+105
t+110
t+115
t+120
Time, minutes
Market Impact
Permanent Impact

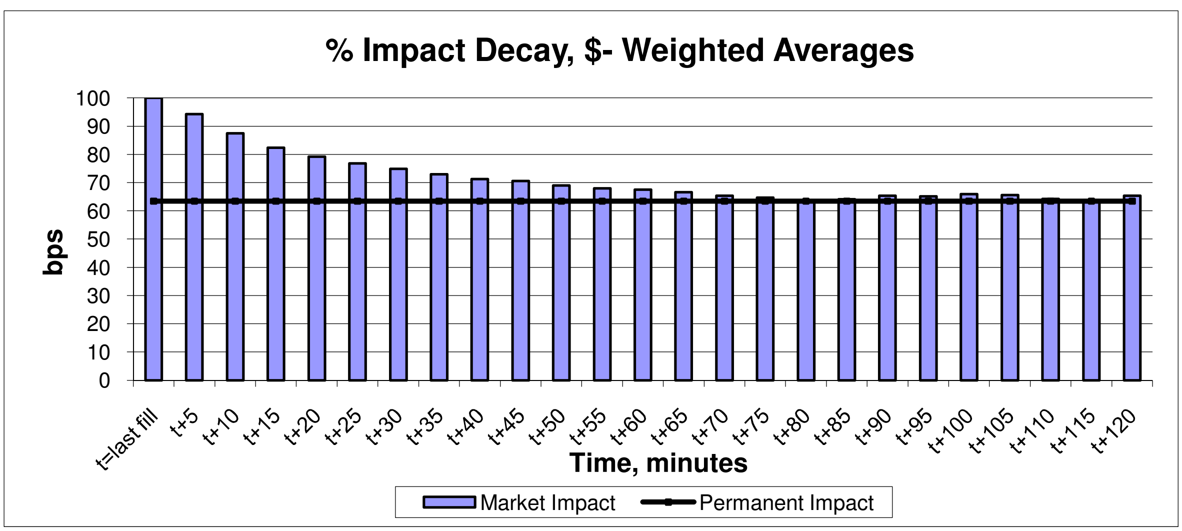
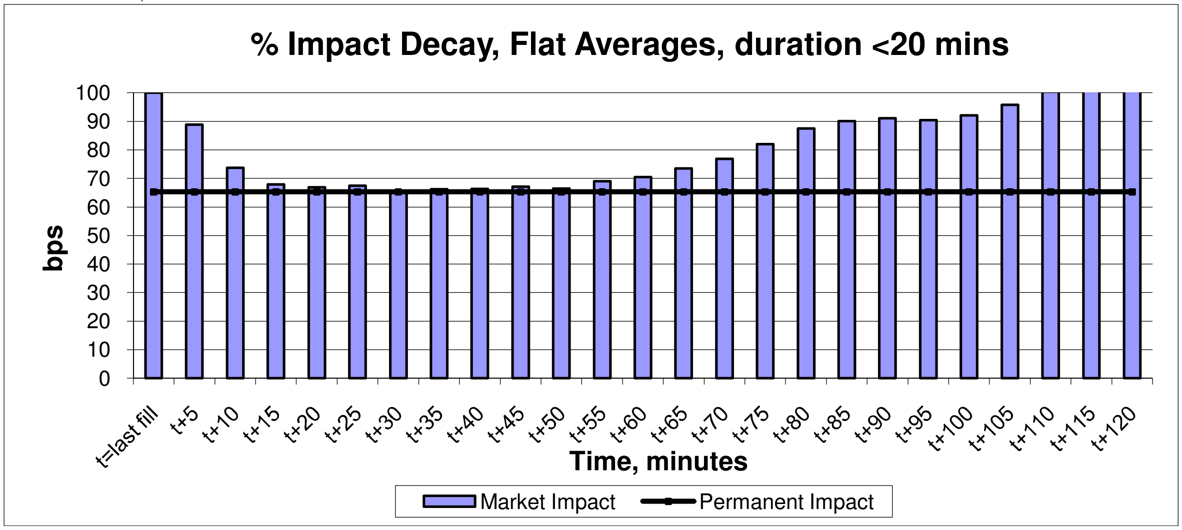
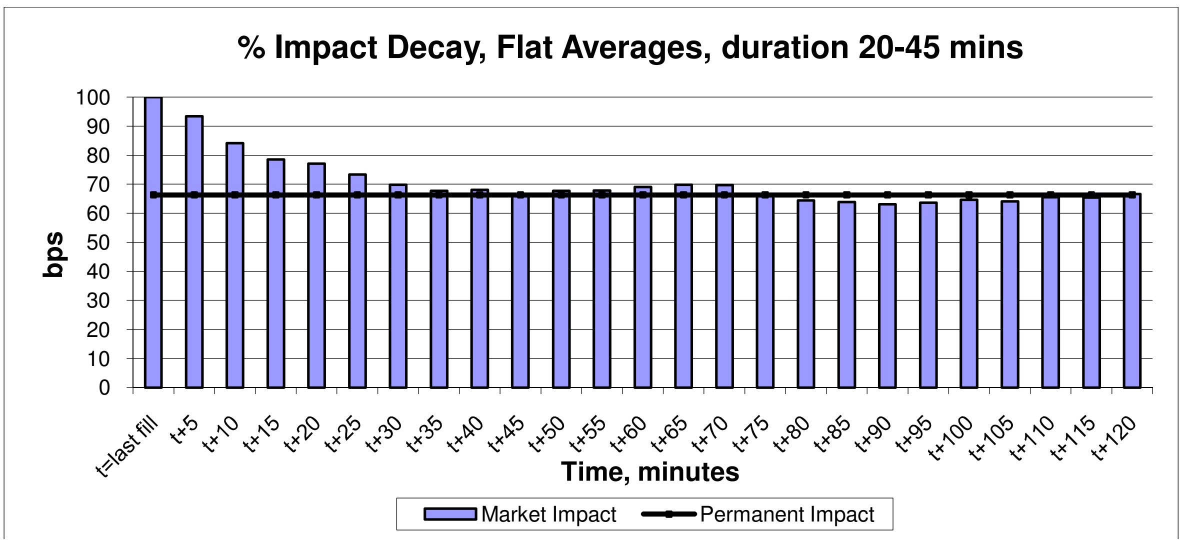

% Impact Decay, $- Weighted Averages
bps
100
90
80
70
60
50
40
30
20
10
0
t=last fill
t+5
t+10
t+15
t+20
t+25
t+30
t+35
t+40
t+45
t+50
t+55
t+60
t+65
t+70
t+75
t+80
t+85
t+90
t+95
t+100
t+105
t+110
t+115
t+120
Time, minutes
Market Impact
Permanent Impact

Figure 8. This figure displays flat average percentage of temporary impact decay for 3 buckets by order duration: <20 minutes , 20-45 minutes and >45 minutes.





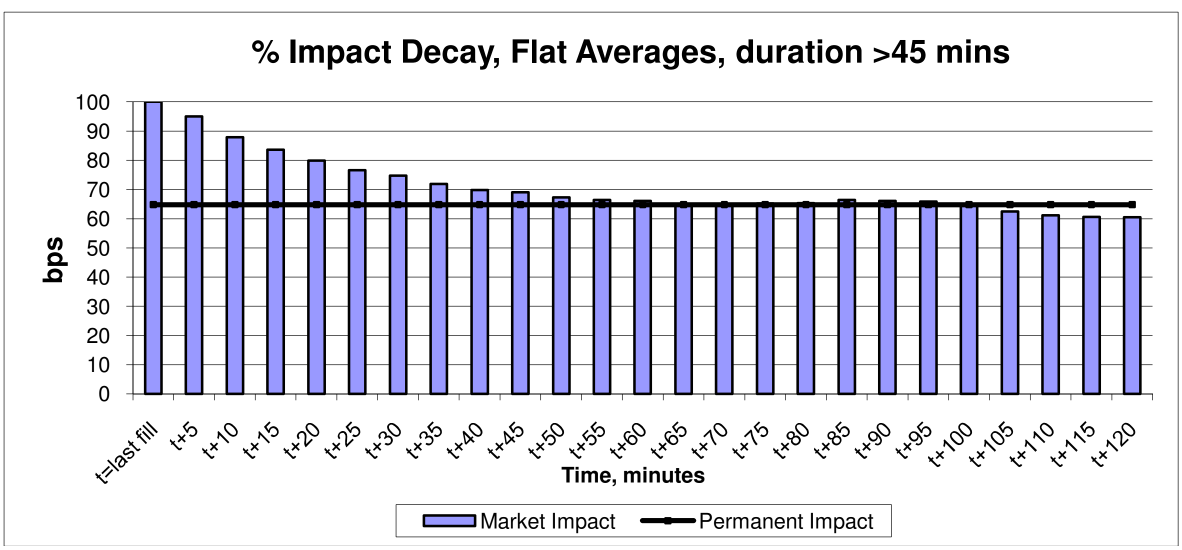


Table 6. This table shows estimated parameters for three functional models (Eq. (7)) for the permanent price impact as a function of trade duration

Model	Mean β	StdErr β	T-stat β	Mean γ	StdErr γ	T-stat γ	BIC	R-square, %
Linear	.0009	.00003	31.75				102000.1	9
Power	.0105	.0015	6.94	.4722	.0351	13.43	101806.5	11
Log	.0170	.0004	34.56				101832.1	10

Figure 9. This figure displays permanent and total market impact.

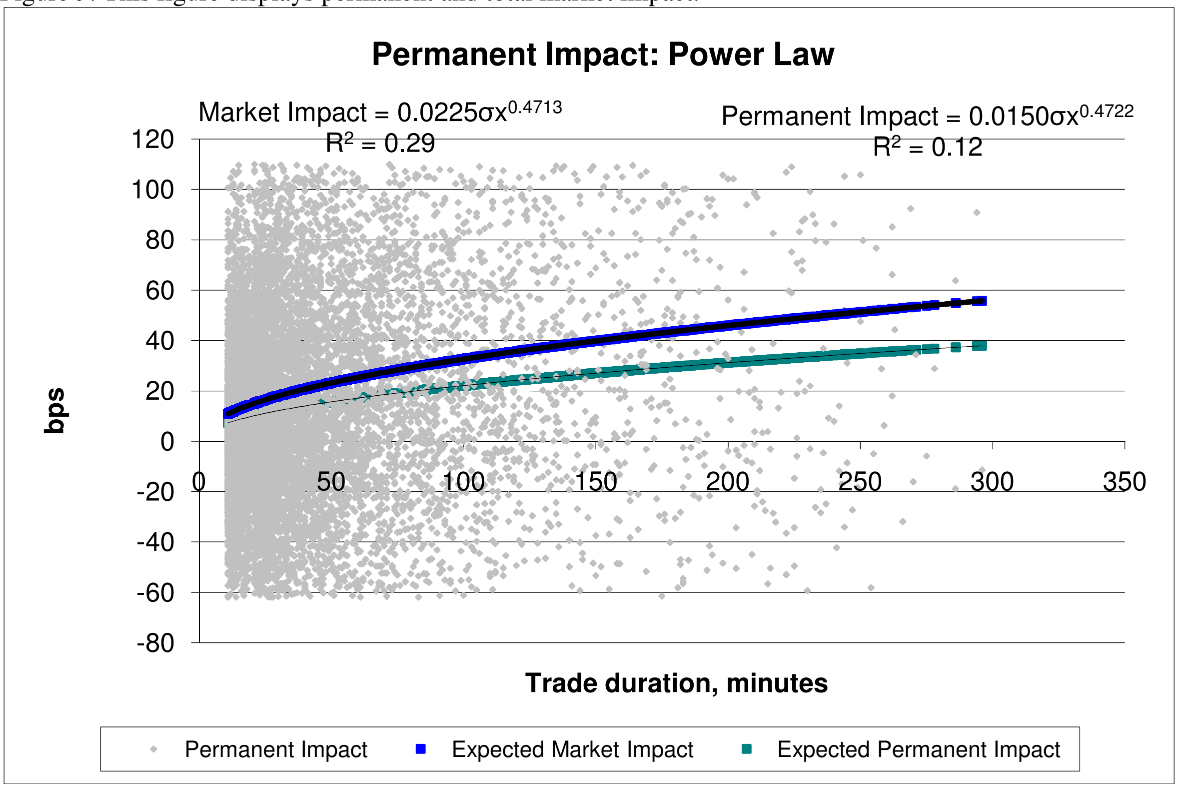


Table 7. This table shows the mean, standard deviation and 95% confidence level intervals for realized and permanent price impacts.

Variable	Mean	StdErr.	95% CI	
J	13.9800	.2432	13.5032	14.4569
I	13.1677	.3588	12.4642	13.8712
$(J-\mathrm{I})$	.8123	.4335	-.0374	1.6621

Figure 10. Minute-by-minute flat average profile of post-trade impact decay along with 95% confidence intervals. The market impacts series R_n (defined in Eq. 8) are normalized by expected total price impact

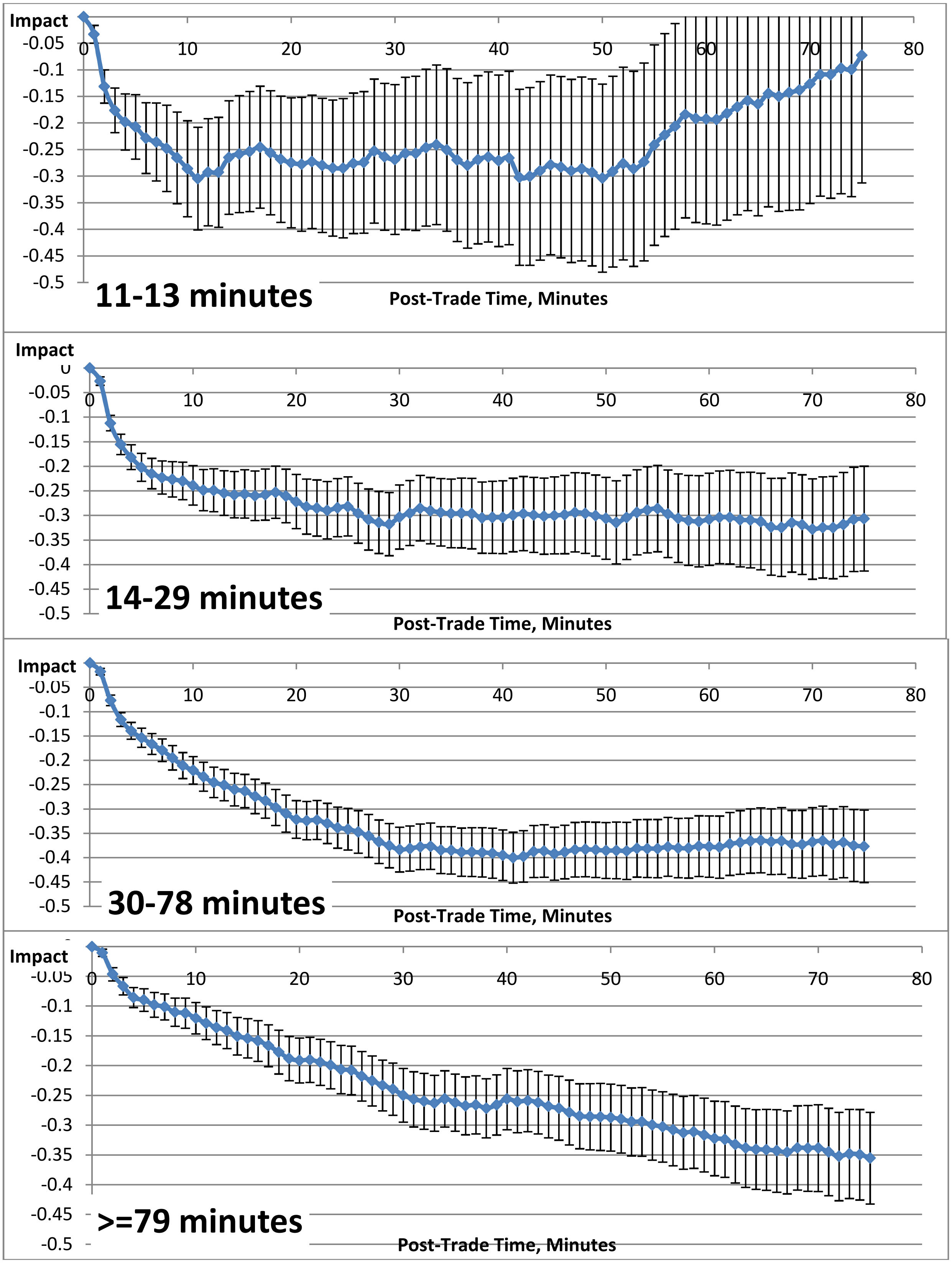
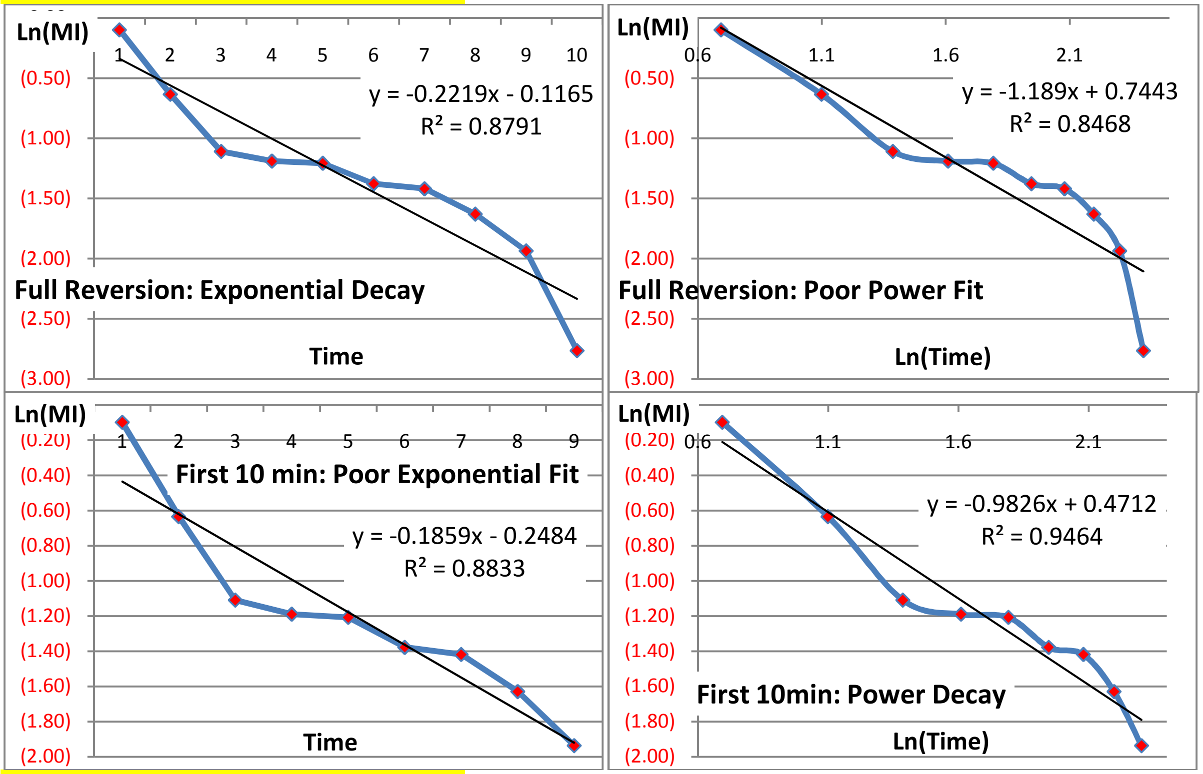
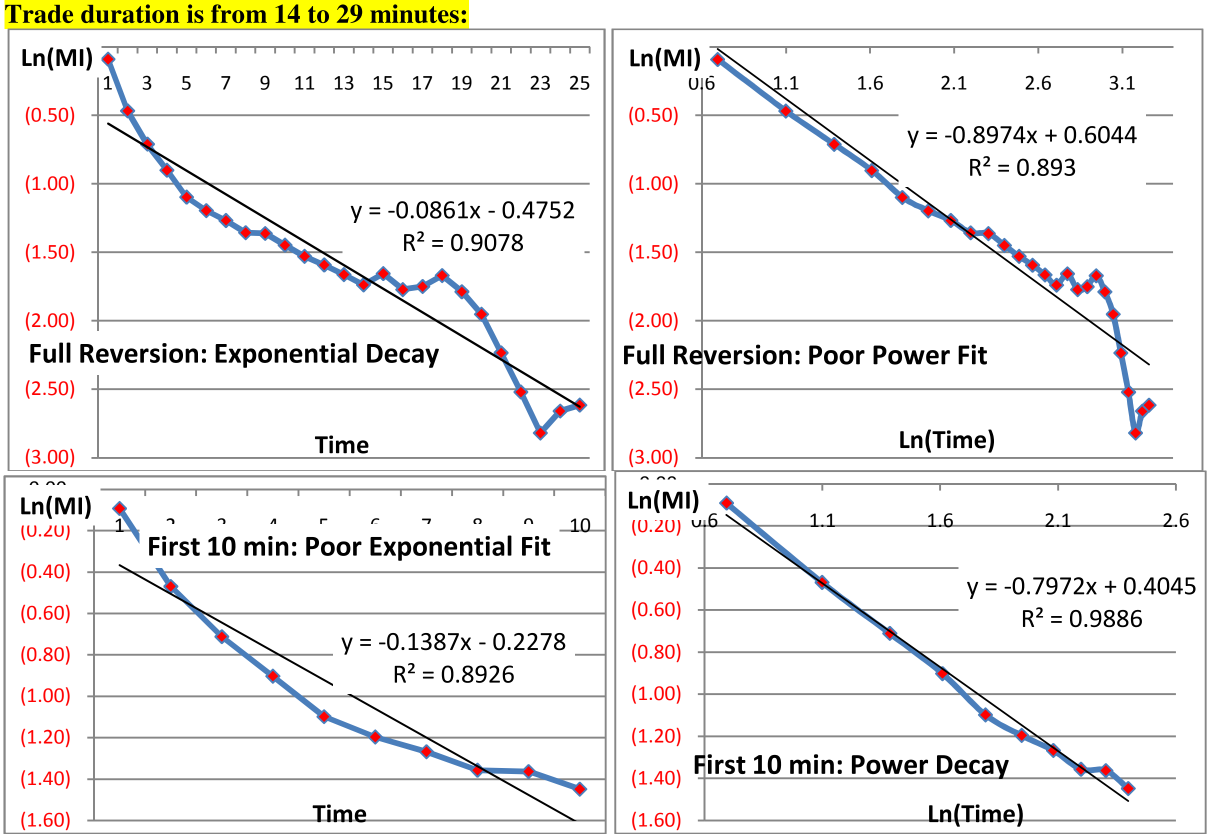


Figure 11: Temporary impact decay $\left(1+\bar{R}_{i,n}/(\bar{\mathcal{I}}_i-\bar{\mathrm{I}}_i)\right)$ in log-log and log-linear axes for different durations.

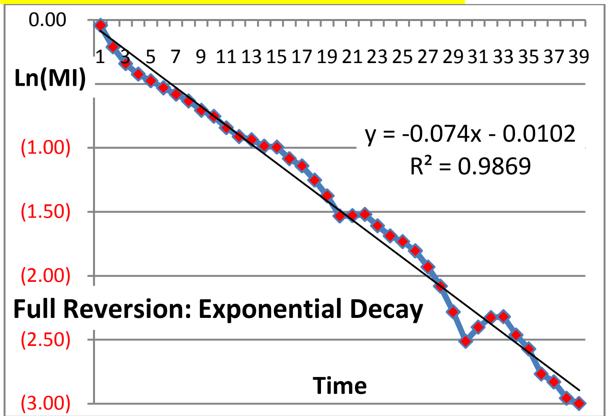
Trade duration is from 11 to 13 minutes:

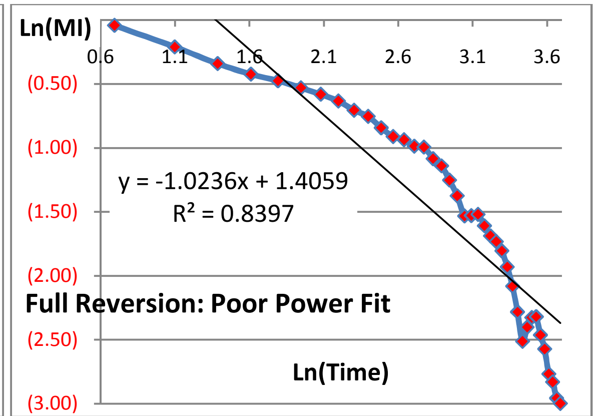


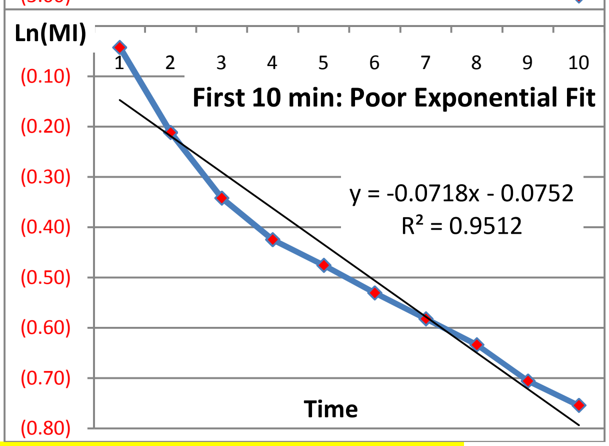
Trade duration is from 14 to 29 minutes:

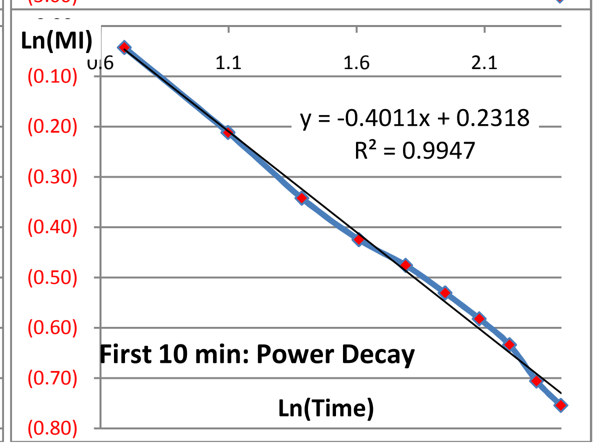


Trade duration is from 30 to 78 minutes:

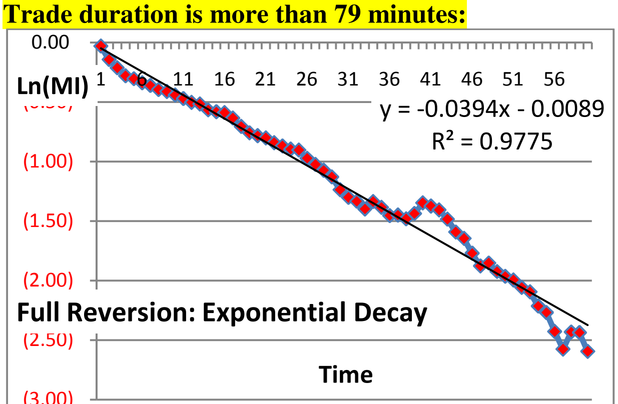


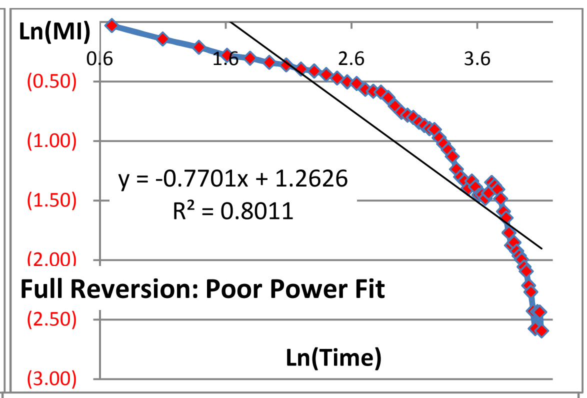


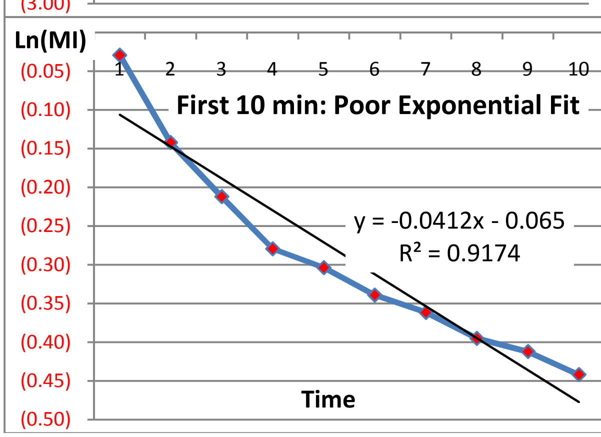




Trade duration is more than 79 minutes:







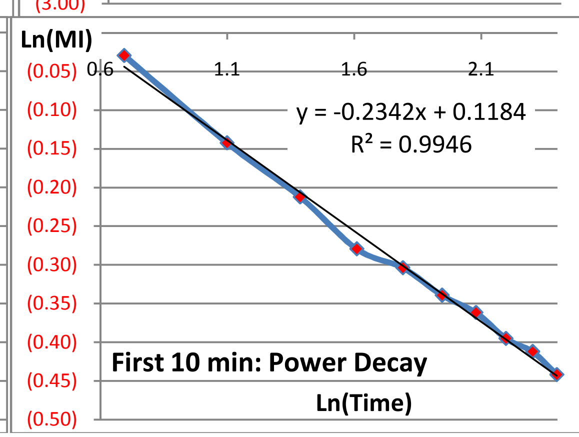


Figure 12. This figure shows adjusted values of permanent level μ as defined by Ornstein-Uhlenbeck process fit on 5-minutes sliding window [t, t+5]. The adjustments are equal to the average permanent impact levels over "stability zone" when price has already reversed: bucket 11-13 is adjusted by 0.34, bucket 14-29 by 0.33, bucket 30-78 by 0.38 and bucket >=79 by 0.33.

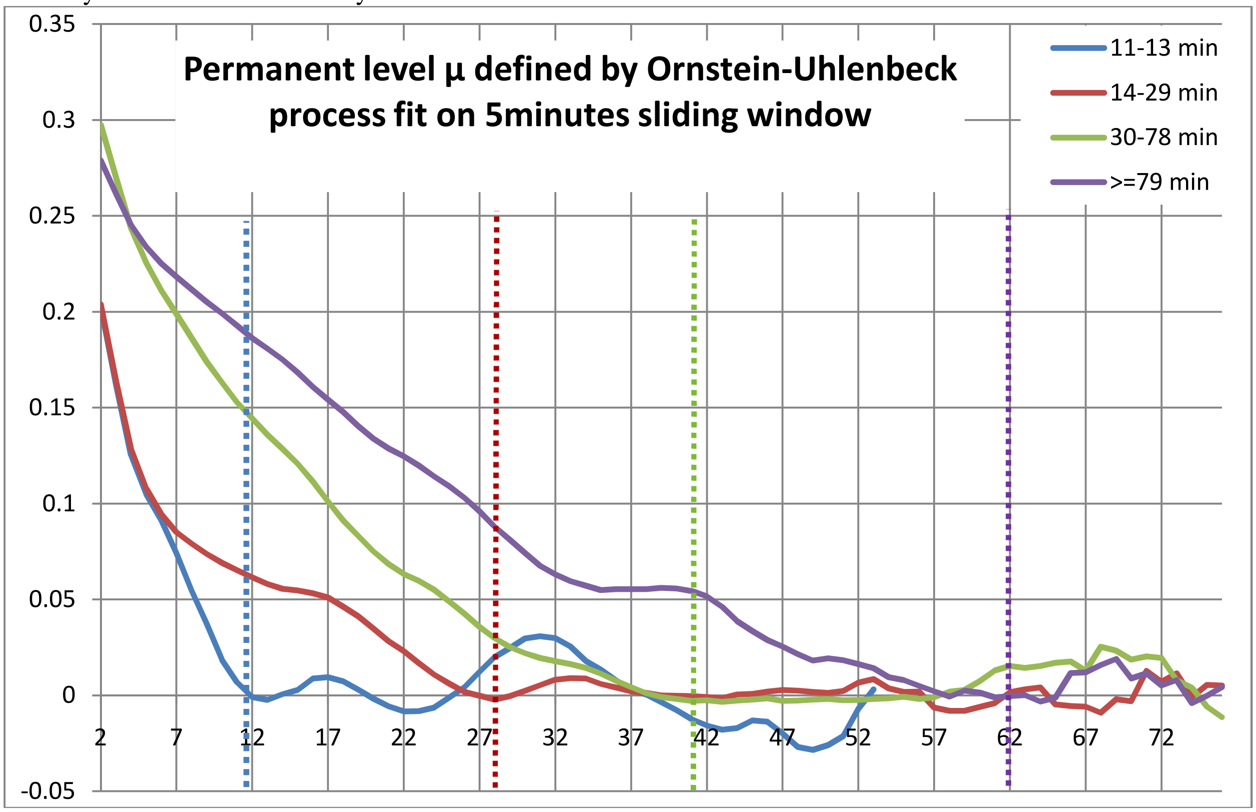


Table 8. This table reports summary characteristics of temporary impact relaxation process. The table displays mean estimates of the permanent impact and its level μ obtained from inspection of Figure 11 and from O-U process. The table reports estimates of exponent θ from log and log-log impact models in the power-law and exponential decay stages: along with the O-U estimate.

Variable	Bucket1 11-13 min	Bucket 14-29 min	Bucket 30-78 min	Bucket >=79 min
#Observations	1178	3359	2841	1099
Average duration, min.	12	20	47	126
Bid-Ask Spread, bps	5	5	5	5
Power-Law impact decay time, min.	10	10	10	10
Overall impact decay time, min	12	28	41	62
Total mkt Impact $\mathcal{I}$, bps	8	14	24	36
Permanent Impact I, bps	6	10	15	23
μ	-0.30	-0.30	-0.38	-0.35
Magnitude power-law decay, bps	3	3	5	4
%Power-law decay	100	72	55	32
Power-law decay/ $\mathcal{I}$	-0.30	-0.24	-0.22	-0.12
θ, power-law decay	0.97 (0.06)	0.87 (0.02)	0.46 (0.02)	0.22 (0.01)
θ, exponential decay	0.22 (0.02)	0.08 (0.01)	0.07 (0.01)	0.04 (0.01)
O-U θ	0.21 (0.008)	0.08 (0.002)	0.06 (0.001)	0.04 (0.001)
O-U μ	-0.27 (0.046)	-0.27 (0.031)	-0.34 (0.024)	-0.25 (0.033)

APPENDIX 1. POST-TRADE REVERSION FILTERS

The original sample size is 12,501 orders. To filter out outliers, we removed market impact and shortfall values from the top 1% and bottom 1% of their distributions, leaving realized impact (IS) within [-88,187]bps range and total market impact within [-137,288]bps range. To examine post-trade price reversion, we discarded trades with extreme post-trade price movements. Specifically, we dropped orders that exhibit slippages in the top 5% of AVWAP distribution (orders with >110 bps slippage to AVWAP or "available VWAP", which is the volume-weighted average price between the time of order arrival and market close). Most of such orders are sells. One of the possible explanations for this finding is a psychological bias, i.e., overreaction to bad news. One strand of literature, for example (Klößner et al. 2012), provides evidence of intraday stock price overreaction to bad news for the majority of stocks in the S&P500 and the XETRA DAX. Sell orders that we discarded because of large slippages to AVWAP tend to be completed soon after arrival to avoid possible adverse price movement in the presence of bad news. Apparently, in cases when a stock price rebounds after a massive selloff, such orders incur large slippages to AVWAP. Additionally, we removed trades with significant slippages to the 10% participation weighted price (top 5% trades, slippage vs. PWP10 > 46 bps). Our average participation rate for trades is 15%, and the average duration is 40 minutes. This means the PWP10% benchmark period on average spans only 20 minutes beyond the order completion time, and slippage >46 bps over such a short interval is likely event-driven. Thus, we discard trades that incur adverse selection cost and therefore do not have a clear reversion price profile.

Finally, to reduce noise in the post-trade profile, we discarded trades for which price reversion during the first hour after a trade lies outside of its 2-standard-deviation range. This removes the top and bottom 5% of permanent impact values. Overall, to study post-trade reversion profiles, we removed 20% of the data.

APPENDIX 2. AGGREGATED ORDER SIZE DISTRIBUTION FOR INSTITUTIONAL MANAGERS

Understandably, the distribution of order sizes generated by a single asset manager may differ from the order size distribution collectively generated by many asset managers. Market makers, on the other hand, may choose to track orders coming from a single agent, or more plausibly, may trace overall market imbalances created by all buy-side participants (market makers themselves do not create any lasting order imbalance). With this in mind, we analyzed the order size distribution and the market impact of aggregated trading volumes by symbol and side, executed by a diverse set of buy-side institutions and provided to us in highly aggregated form by a third party TCA vendor. First, all individual clients' daily trades for the 2010-2011 period were aggregated by the vendor using allocation data to obtain daily imbalances (the difference between shares bought and sold). The initial set included symbol-date pairs with absolute order imbalances >5% MDV (~40% pairs). If the aggregated order imbalance for that symbol exceeded 5% MDV over multiple consecutive days, the aggregated daily buy orders and sell orders were stitched together to form multi-day orders. The size of stitched trades was calculated as the total executed quality normalized by the MDV at the first day of the trade. Additionally, net of contra order size, e.g. order size net of total trading volume executed in aggregate by buy-side clients on the opposite side of the market was reported. Finally, we received the output file with aggregated order size information separately for each market cap category.

We asked the TCA vendor to winsorize the data by removing 1% tails in the participation rate, trade rate, IS, market impact, order size and duration – about 7% of orders – which left us with 11,114 observations in large-cap stocks. For mid-cap stocks, we removed 5% of outliers by trade rate, 14% observations in total. This result is not surprising, given the more irregular, opportunistic nature of trading in less liquid names. Our total sample size in mid-cap after filtering was 25,548 orders. Finally, small-cap names required more thorough filtration; however, because we did not have access to the underlying data (it was in the hands of the TCA vendor), we decided not to include small caps in this study.

Tables and figures below report very similar results for the aggregated dataset that we have observed for proprietary AllianceBernstein trades: First, Pareto distribution approximates well the order size distribution everywhere except the largest quintile (this quintile corresponds to sizes 49 % MDV and 60% MDV for large- and mid-caps or 36% and 48% MDV net of contra side). Second, total market impact is best fit by a square root. Third, the realized impact is equal to 2/3 of total market impact and it is not statistically different from the permanent market impact (with even the largest multiday orders showing full relaxation of temporary market impact by the close of the next day after the order completion).

Large-cap stocks:

Table A1. General characteristics

Variable	Mean	Stderr	95% CI
Order size (%MDV)	57.2	0.4	56.4 58.0
Net size (net of contra)	41.5	0.3	40.9 42.1
Duration, days	2.8	0.0	2.8 2.9
Realized Impact, IS	74.8	2.3	70.3 79.3
Total Market Impact	104.6	2.3	99.8 109.5
MDV	6,402,134	118,136	6,170,566 6,633,701
Participation rate	16.0	0.1	15.9 16.1
Trade rate	19.5	0.1	19.4 19.7

Table A2. Fitting market impact gives a slight edge to the power-law model (adjusted R^2=15.45% vs. R^2=14.91% for logarithmic impact), although the BIC coefficient is marginally higher for the logarithmic model.

Model	Mean β	StdErr β	Mean γ	StdErr γ	BIC
Power	12.6776	1.8415	0.5342	0.0334	155599.9
Log	27.9754	0.6331			155662.8

Figure A1. Pareto fits most data in the middle of the distribution, while deviating significantly at the tails. Note that our resolution is 1 day, and we require a minimum market imbalance 5% MDV, which removes a significant fraction of smaller orders below 20% MDV. On the chart below ln(20)=3.

$R^2 = 93\%$

Model	Mean	StdErr.	T-stat	95% CI
Coefficient	-2.5756	0.0563	-45.73	-2.6868 -2.4642
Intercept	15.6404	0.2822	55.40	15.0826 16.1982

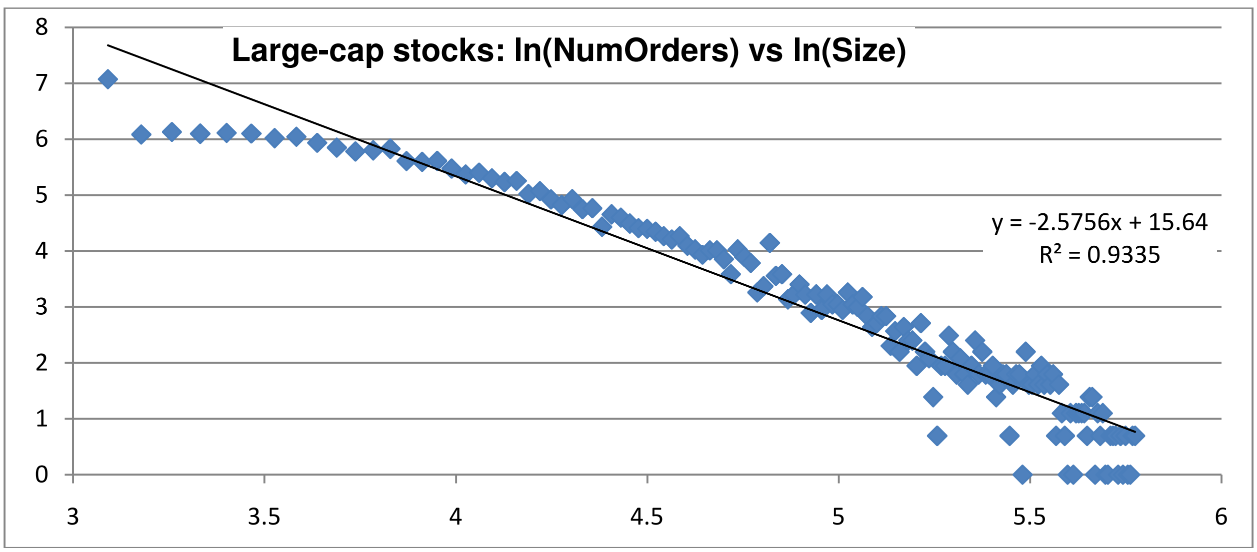


Table A3. The tail test accepts lognormal distribution and rejects Pareto. For this set of data, the tail that fits to lognormal distribution belongs to the upper 20% of data. In contrast, only the upper 5% of AB data fits to lognormal distribution. We attribute this discrepancy to the difference in how the two datasets were created and the filters applied.

TEST	**ET**
Choice Model:	Lognormal
Test Level:	0.05
# of excesses	2000
# bootstrap sample	2000
Quantile p of order 1-p	0.2
Test Results:	**Accepted**
Estimated Quantile ET	77.93936
Estimated q	78.08188
qET-qParam	-0.14252
Inf CI	75.89612
Sup CI	80.00184

TEST	**GPD**
Choice Model:	Pareto
Test Level:	0.05
# of excesses	2000
# bootstrap sample	2000
Quantile p of order 1-p	0.2
Test Results:	**Rejected**
Estimated Quantile GDP	93
Estimated q	104.87995
qET-qParam	-11.88
Inf CI	0.80933
Sup CI	4.23374

Table A4. Fitting market impact for the upper tail (the upper 20% of data) gives an edge to the logarithmic model in terms of BIC.

Model	**Mean β**	**StdErr β**	**Mean γ**	**StdErr γ**	**BIC**
Power	27.1421	8.1930	0.3596	0.0657	67578.29
Log	30.6907	0.9329			67574.21

Mid-cap stocks:

Table A5. General characteristics

Variable	**Mean**	**Stderr**	**95% CI**
Order size (%MDV)	67.1	0.3	66.4 67.7
Net size (net of contra)	51.6	0.3	51.1 52.1
Duration, days	3.0	0.0	3.0 3.0
Realized Impact, IS	66.9	1.9	63.2 70.6
Total Market Impact	92.2	2.1	88.1 96.3
MDV	1,977,944	19,772	1,939,189 2,016,699
Participation rate	17.7	0.0	17.6 17.7
Trade rate	21.5	0.1	21.3 21.6

Table A6. Fitting market impact gives a slight edge to the power-law model (adjusted R^2=7.7% vs. R^2=7.6% for logarithmic impact); BIC coefficient is marginally higher for the logarithmic model.

Model	**Mean β**	**StdErr β**	**Mean γ**	**StdErr γ**	**BIC**
Power	15.7384	2.1787	0.4337	0.0312	369277.4
Log	23.6285	0.5164			369306.5

Figure A2. Pareto fits most data in the middle of the distribution and deviates at the tails. Removing the tail Ln(Size)<5.5, e.g., Size<240% MDV produces an exponent very close to 2.5.

$R^2 = 93\%$

Model	Mean	StdErr.	T-stat	95% CI
Coefficient	-2.7408	0.0483	56.67	-2.8361 -2.6455
Intercept	17.3895	0.2598	66.93	16.8775 17.9015

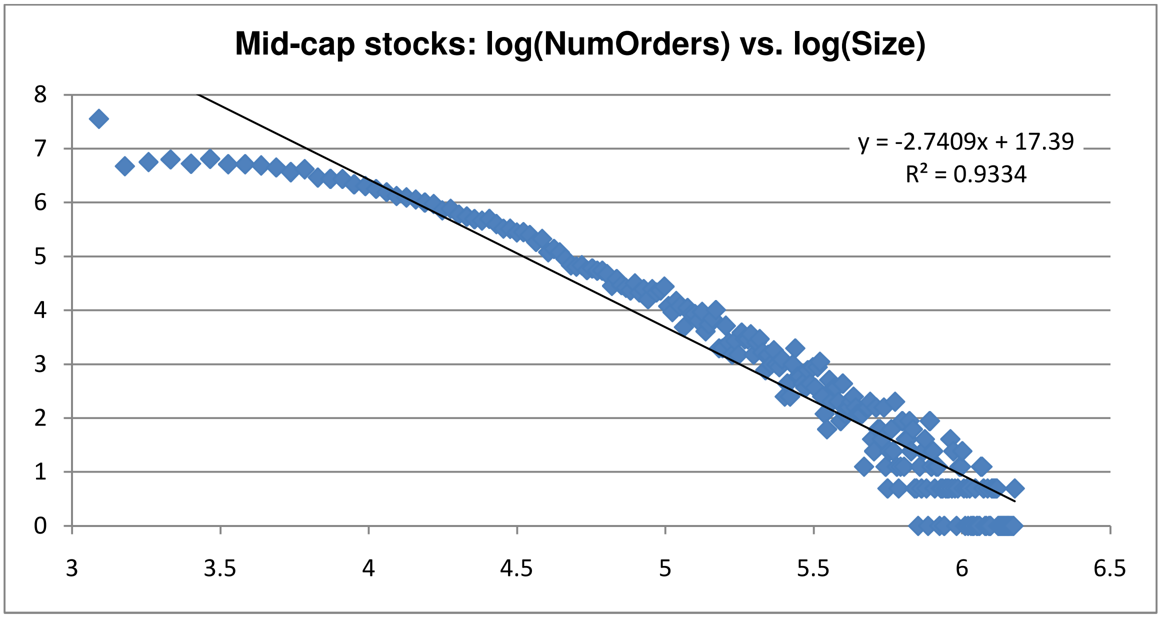


Table A7. The tail test accepts lognormal distribution and rejects Pareto. For this set of data, the tail that fits to lognormal distribution belongs to the upper 20% of data.

TEST	ET
Choice Model:	Lognormal
Test Level:	0.05
# of excesses	5000
# bootstrap sample	2000
Quantile p of order 1-p	0.2
Test Results:	**Accepted**
Estimated Quantile ET	91.74172
Estimated q	91.082
qET-qParam	-0.24028
Inf CI	90.1333
Sup CI	93.35102

TEST	GPD
Choice Model:	Pareto
Test Level:	0.05
# of excesses	5000
# bootstrap sample	2000
Quantile p of order 1-p	0.2
Test Results:	**Rejected**
Estimated Quantile GDP	93
Estimated q	104.87995
qET-qParam	-11.88
Inf CI	0.78243
Sup CI	4.19497

Table A8. Fitting market impact for the upper tail (the upper 20% of data) gives an edge to the logarithmic model in terms of BIC.

Model	Mean β	StdErr β	Mean γ	StdErr γ	BIC
Power	31.2430	11.4461	0.2734	0.0743	130952.2
Log	24.4286	0.8382			130943.8