Round Numbers and Security Returns[*]

Edward Johnson
Visiting Scholar
Haas School of Business
UC Berkeley
egj@stanfordalumni.org

Nicole Bastian Johnson
Haas School of Business
UC Berkeley
njohnson@haas.berkeley.edu

Devin Shanthikumar
Harvard Business School
Harvard University
dshanthikumar@hbs.edu

First Draft: March 2007
This Draft: August 2008

Abstract

We document consistent return patterns related to round numbers. We examine returns following previous day closing prices that are just above or below round number benchmarks. We find, for one-digit, two-digit and three-digit levels, that returns following closing prices just above a round number benchmark are significantly higher than returns following prices just below. For example, returns following "9-ending" prices, which are just below round numbers, such as $25.49, are significantly lower than returns following "1-ending" prices, such as $25.51, which are just above. Our results hold when controlling for bid/ask bounce, and are robust for a wide collection of subsamples based on year, firm size, trading volume, exchange and institutional ownership. While the magnitude of return difference varies depending on the type of round number or the subsample, the magnitude generally amounts to between 5 and 20 basis points per day (roughly 15% to 75% annualized). We explore alternate explanations for the pattern.

[*] We would like to thank Terrance Odean and seminar participants at Rice University, the University of Vienna, the Third Annual Early Career Women in Finance Mini-Conference, and State Street Global Advisors for helpful discussions and feedback.

In this paper we document systematic differences in daily stock returns in U.S. markets based on the proximity of the previous day's closing price to round numbers. Specifically, we find that returns are, on average, higher following prices that lie just above a round number, and lower following prices that lie just below a round number. This generates a consistent pattern of return differences based on the last digits of the closing price. For example, returns following a closing share price that ends in .01 tend to be higher than following a price that ends in .99.

Stock prices tends to cluster on round number (i.e. numbers that end in one or more zeros, or, to a lesser extent, those that end in 5 or 25) and stock price clustering is stronger for the "rounder" prices (such as $25.00), than for prices that are "less round" (such as $26.25). Likewise, the return differentials we find are strongest around the "roundest" numbers. While there are a number of studies that document the clustering of orders and prices at round numbers, our study is the first that we know of to document differences in returns. But while the return patterns we document appear to be related to clustering, they are do not appear to be a simple consequence of clustering.

Our investigation begins by documenting that average U.S. daily stock returns differ based on the last digit of the previous closing price. Figure 1 demonstrates this result. Strikingly, we find that similar (and even stronger) patterns are observed at other scale levels (i.e. when partitioning by the 'dime' or 'dollar' digit instead of the 'penny' digit).

Exploiting the move to decimalization in US exchanges in 2001, we test whether returns differ systematically based on the last one, two or three digits of closing price, and in particular whether those digits place price just above or just below a round number. For example, is the average daily return following a closing price just above a round number, such as $24.11 or 25.01, different from the average return following a closing price just below a round number, such as $24.09 or 24.99? Our results indicate that returns are significantly related to closing prices, with higher returns following closing prices just above round numbers and lower returns following closing prices just below round numbers.

To get a sense of the magnitude of the returns differences we find, let us focus on returns following closing prices just above or below an even dollar amount. Our estimates show that returns following prices ending in 01 through 09 cents are, on average, 12.9 basis points higher than those ending in 91 through 99 cents. This corresponds to an annualized rate of over 38% per year. It is unlikely that these return differences alone could be used to form a profitable trading strategy given the almost-daily rebalancing required, but they may have a significantly positive impact on returns if exploited for optimizing the timing of order execution.

The patterns in returns that we document are robust across a variety of cross-sectional partitions of the data, including partitions based on price levels, liquidity, and stock exchange. When the data is partitioned by size, the return patterns are weaker but still statistically significant for the largest quintile of stocks. Further, when the data is partitioned by year, we find that the phenomenon has weakened somewhat since 2001, but is still significant in 2006, despite large changes in market microstructure. In particular, during our sample period there were very large decreases in clustering and in quoted spreads, increases in automated trading, and changes in commission structures, yet the returns patterns have remained robust.

We test several market microstructure-based explanations for the return patterns that we document and find that they are insufficient to explain our findings. In particular, we use intraday data to demonstrate that the excess returns associated with closing prices near round numbers are realized throughout the following day and are not simply an artifact of irregularities in closing prices, overnight trading, or the opening auction.

We also test several behavioral explanations for our results. While we find some evidence that retail investors trade decisions' are affected by whether stock prices are just above or below round numbers, we are unable to find evidence that our results are driven by traders using technical analysis or uninformed traders. We are also unable to find a mechanism that would lead directly from limit order or stock-price clustering to the pattern of returns that we find.

The remainder of the paper is organized as follows. Section 2 provides a brief review of previous literature on price clustering and on psychological effects related to round numbers. Section 3 describes the data employed and the sample selection procedure. Section 4 describes the main empirical findings about the patterns of returns related to round number prices. Section 5 deals with some potential explanations for these patterns, and generally finds them unable to explain the patterns we describe. Section 6 concludes.

II. Background and Prior Literature

Generally, we would expect markets to be efficient with respect to the proximity of closing prices to round numbers. But a large finance literature on price clustering in financial markets and work on round numbers as cognitive reference points in the cognitive psychology literature lead us to ask whether it's possible that some small inefficiency related to round numbers and their effects on consumer (investor) decision making could be present. Prior literature has not examined this question, perhaps partially be due to the lack of a mapping between numbers that individuals see in most settings, typically in decimal form, and the practice of quoting stock prices in eighths or sixteenths of dollars on U.S. exchanges[1]. We exploit the move to decimalization by U.S. exchanges in 2001 to explore this question.

In the finance literature, researchers have documented price clustering on round numbers in many settings, including U.S. equity markets (Osborne (1962), Nierderhoffer (1965 and 1966), Harris (1991), Christie and Schult (1994), Grossman et. al. (1997), Cooney, VanNess and VanNess (2003)). Previous to decimalization, prices were most

[1] One exception is Bagnoli, Park and Watts (2006), who find similar results in a related working paper. Their analysis is restricted to raw, overnight returns during 2002 and only briefly explores possible explanations for the results.

likely to fall on round dollar amounts, second most likely to fall on half-dollar amounts, then quarter-dollar amounts. Price clustering has continued to be a feature of U.S. stock markets since decimalization was completed in 2001, with transaction prices tending to cluster at zero-ending prices and, to a lesser extent, on five-ending prices. There is even stronger clustering at the two-digit level, with strong clustering at prices ending in .00, .50, and to a lesser extent, .25. Price clustering also occurs in foreign equity markets (Grossman et. al. (1997), Aitken et. al. (1996)), Israeli IPO auctions (Kanhel, Sarig and Wohl (2001)), US SEOs (Corwin (2003)), currency markets (Goodhart and Curcio (1990), Osler (2003)) and a variety of other settings.

Research in cognitive psychology has shown that people tend to give too much weight in decision-making to information that is easily retrieved from memory (Tversky and Kahneman (1973)) and that round numbers are more easily recalled than other numbers[2] (Schindler and Wiman (1989)). Rosch (1975) shows that zero-ending numbers tend to be cognitive reference points in the sense that experimental subjects tend to evaluate the magnitude of a number differently when it is in close proximity to a round number than when it is not. For example, subjects were more likely to agree with the statement "996 is almost 1000" than with the statement "1023 is almost 1027", even though the difference between the numbers in each statement is the same. The effects of cognitive reference points have been examined in various literatures.

A significant literature in marketing has shown that consumers behave as if the demand curve is discontinuous at prices ending in certain digits, such as 9 and 99.[3] Researchers have found evidence for several possible explanations for this phenomenon, including a 'truncation effect' in which consumers tend to put more emphasis on the left-hand digits of numbers when comparing two prices, a 'perceived gain effect' in which 9-ending (1-ending) prices are perceived as a discount (premium) relative to zero-ending prices, or an 'image effect,' resulting from consumers having been conditioned to view nine-ending prices as a signal of a sale or bargain and zero-ending prices as a signal of

---

[2] See Schindler and Kirby (1997) for a discussion of the literature on cognitive accessibility.

[3] See Gedenk and Sattler (1999) and Bizer and Schindler (2005) for reviews of the empirical literature on the effect of nine-ending prices on consumer behavior.

quality (Stiving and Winer (1997), Schindler and Kirby (1997), Bizer and Schindler (2005)). Note that these effects do not have to be mutually exclusive and Stiving and Winer find evidence of all three effects in their analysis of grocery store scanner data.

In the accounting literature, Herrmann and Thomas (2005) explore the effects of round numbers on decision making and show that financial analysts tend to round earnings forecasts and that the market reaction to earnings surprises seems to be based on the rounded forecasts. Additional studies find evidence that a disproportionate number of earnings reports (both total earnings and earnings per share) include round numbers[4]. Among the explanations that have been proposed for this finding is that managers may believe that investors are subject to truncation, perceived-gain or image effects in their evaluations of earnings numbers.

Our paper also relates to a large literature in behavioral finance on the affects of non-standard preferences, psychological biases and cognitive limitations on prices and returns (see Barberis and Thaler (2003) and Shleifer (2000) for surveys of this literature). In particular, Baker and Wurgler (2006) and Kumar and Lee (2006) examine retail investor sentiment and find that investor psychology can have a significant impact on the returns of small firms with high arbitrage costs and less sophisticated investors. However, the reliance on cognitive reference points such as round numbers in decision making seems to be more ingrained than other types of cognitive biases,[5] so it is possible that even the most sophisticated investors will unconsciously rely on these reference points whenever making subjective judgments. In addition, because the return patterns we examine are short-run and a given firm can fall both into "above" and "below" round number categories, the patterns we find are unlikely to be related to systematic risk. For both these reasons, round-number reference points provide an excellent setting to test the potential impact of investor psychology on even the largest and most liquid stocks.

[4] See Carlsaw (1988), Thomas (1989), and Das and Zhang (2003), for example.

[5] See Op de Beeck, Wagemans and Vogels (2003) for a discussion of the literature in psychology, neuroscience and biology on the use of "prototypical" reference points. Op de Beeck, Wagemans and Vogels (2003) show that the use of such reference points occurs even in monkeys, very similarly to humans.

III. Data

Our primary data sample is limited to common stock for US firms trading on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX) or NASDAQ since decimalization. On June 8, 2000, the Securities and Exchange Commission (SEC) issued an order requiring all U.S. exchanges to move to decimal pricing by April 9, 2001.[6] The change was completed on time (Unger (2001)); however some newspapers did not immediately begin quoting prices in pennies. For example, The Wall Street Journal began showing stock prices in dollars and cents as of April 30, 2001 (Wall Street Journal (2001)). Based on the switching dates of exchange quoting practices, and allowing some time for the lag in newspaper's reporting practices, we begin our sample period on May 1, 2001. The sample extends from May 2001 through the end of 2006.

We collect closing prices and quotes, daily returns, and stock characteristics from the CRSP database over this time period. We then remove stocks that have closing prices that were below $5 per share on any day of the previous calendar month.

In calculating returns, we exclude stock-days for which CRSP has imputed a closing price based on the closing bid and ask quotes, which occurs when no trade-based closing price is available. We also drop firm-days on which dividends are distributed, as well as firm-days on which a stock split, merger, or other such event occurs that would make the returns different from those calculated simply on the basis of share prices, because we might expect returns on these days to be less sensitive to the previous quoted closing.

For most of the analysis in the paper, we do not use the simple return based on closing prices. Instead, we analyze returns calculated by using the midpoint of the closing bid and ask quotes, as reported in CRSP. As we explain in detail in the Appendix, we do this in order to remove the effect of the bid-ask bounce that might bias

[6] See Securities and Exchange Act Release No. 42914, June 8, 2008

our results. In order to avoid the influence of bad data or outliers when using this midpoint, we drop firm-day observations for which either the bid is zero, the ask is zero, the bid exceeds the ask, or the midpoint is more than 10% away from the reported closing price. This eliminates return observations for which the midpoint is unavailable or appears unreliable for either the previous day or current day. Eliminating these observations shrinks the sample by about 10%, leaving us with about 3.3 million firm-day observations, on 4,967 unique securities.

For the last digit(s) of closing price, we use the closing price reported by CRSP, which is the price of the last trade during regular trading hours, rounded to the nearest cent. This closing price is also likely to be the price quoted by common sources such as newspapers and online outlets. For example, the Wall Street Journal states "Wall Street Journal stock tables reflect composite regular trading as of 4 p.m.", Yahoo! Finance reports the last trade price in large bold font, and only updates the trade price until the closing trade at or before 4 p.m., and NYSE, Nasdaq and AMEX all use the same method on their own websites – reporting the last trade price as the price, and updating that price until market close. In addition, some of our analysis described in later sections is based on intra-day prices and quotes collected from the New York Stock Exchange's Trades and Quotations (TAQ) database.

To examine retail investor trading, we use a dataset of brokerage account information for 78,000 households at a large discount brokerage firm. The dataset, provided by Terrance Odean, covers the six years from 1991 through 1996 and contains all trades made in each account, for all securities traded through the discount brokerage firm, including bonds, stocks, and mutual funds. We restrict our analysis to common stock[7]. The database also includes monthly portfolio holdings and household demographic and location data; however we do not use these data for our analysis.

Finally, we use institutional ownership data from Thompson Financial's CDA/Spectrum database. We also use data which classifies each institution into one of

[7] We thank Terrance Odean for providing the brokerage data.

three categories: ``transient", ``dedicated" and ``quasi-indexer", following Bushee (1998).[8] We provide more detail on these data and measures in Section V.

IV. Primary Return Results

In this section we document return patterns following closing prices, conditional on the last digits of the closing price. We first present graphically the major patterns of return differences that arise on the one, two, and three-digit levels. We then report results from regression models that allow us to assess the statistical significance of these differences, and to assess how strongly they hold over time and for various sub-categories of stocks. Next we assess how these return patterns evolve over time. Finally, we discuss the robustness of our findings.

Figure 1a displays the basic pattern of returns conditional on the last digit (or "penny digit") of the prior trading day's closing price. The mean daily size-adjusted excess return, calculated as the midpoint-based return minus the return on the firm's size quintile benchmark portfolio based on NYSE size cutoffs, is shown in percentage form on the y-axis, and the final digit (i.e. the "penny" digit) of the previous day's closing price is shown on the x-axis. The error bars show 95% confidence intervals for the mean daily returns (i.e. t-statistic=1.96). The basic pattern is one in which mean returns appear to be high following a last digit of 1, with returns generally declining for higher digits, except for a "bump" up for last digits at or above 5. The error bars show that this pattern is statistically significant.

Figures 1b and 1c are similar to Figure 1a except that they show how returns vary with the second to last digit (i.e. the "dime digit") and third to last digit (i.e. the last digit before the decimal point, or "dollar digit"). The pattern that appears here is quite similar to that for the "penny" digit; returns tend to decline as the digits become larger, with the exception of a bump up around digit 5. The pattern of differences is strongest for the "dime digit" level, but is statistically significant for all three levels.

---

[8] We thank Brian Bushee for providing the institution classification data.

The returns shown in the figures are calculated using the midpoint of the bid-ask spread, as described in the previous section. If we instead use returns based on closing transaction prices, as reported in CRSP, we get a pattern similar to Figure 1, but with even larger differences. A portion of the differential returns that we find when using raw returns seems to be due to differential bid-ask bounce related to clustering patterns in limit order submission. By calculating returns using the midpoint of the bid-ask spread, we eliminate this effect. See the appendix for additional details about the differential bid-ask bounce effect. The data used to construct these figures is shown in Table I, along with the corresponding patterns for raw (non-midpoint-based) returns. When we examine patterns at the "dime" and "dollar" digit levels, we find little difference in the pattern between raw and midpoint-based return results. Nevertheless, except where specified, we use midpoint-based returns for the remainder of the paper in order to eliminate any confounding effects due to differential bid-ask bounce. Note, however, that while the returns are based on the midpoints, the conditioning "final digits" on the x-axis are based on the reported closing transaction price, since this is the price investors would typically see.

Figure 2 shows the pattern of returns conditional on the last two digits of the previous closing price. This combines the information in the "penny digit" and "dime digit" graphs (Figures 1a and 1b) into one figure. The scatterplot is noisier than the previous graphs, but the pattern of declining returns as the last two digits increase is still apparent. The figure also shows fitted regression lines, fitted separately for prices ending in 01-49 and prices ending in 51-99.

The patterns shown in these figures suggest how the return differences we find are related to clustering points, since stock prices tend to cluster on prices ending in 0, and to a lesser extent, those ending on 5. The overall pattern of returns in the figures can be described as follows: returns are higher for prices just above a major clustering point, and they tend to decline until they are lowest just below the next major clustering point, with a smaller "bump up" along the way at minor clustering points. This pattern appears to repeat itself at all levels, whether the clustering point is at the 10-cent level (i.e. \$22.30), the dollar level (i.e. \$22.00), or the 10-dollar level (i.e. \$20.00).

Figure 3 shows returns conditional on the previous closing price for the subsample of stocks with share prices between \$10 and \$60. In this figure, the points on the scatter-plot are mean returns for prices grouped into 10-cent intervals, and the solid lines are fitted regressions for the points within each 10-dollar interval. This figure again shows a clear pattern of higher returns following prices in the region just above every major cluster point at 10-dollar intervals (at prices ending in 0.00), and lower returns following prices just below these points.

Sawtooth Model

The results described above suggest that there is a general pattern in returns related to round numbers: returns tend to be greater after closing prices just above round numbers, and lower just below round numbers. In particular, Figures 1, 2 and 3 suggest there is a jagged "sawtooth" pattern in returns that repeats at different levels of rounding. To examine these results further, we estimate a regression model that allows us to capture this pattern on multiple levels of roundness all at once, using all the clustering points identified by the data.

The model simply estimates daily excess returns as a piecewise-linear function of previous closing price. Specifically, the model allows (but does not require) a discrete "jump" up (or down) at each category of round numbers where prices are found to cluster. These "classes" of cluster points are those ending in round numbers at the 5-, 10-, 25-, 50-, 100-, 500-, and 1000-cent increments, respectively. These categories are defined to be mutually exclusive; for example, the "50" class counts prices ending in 50 cents, ignoring whole dollar prices that are already included in the 100, 500, and 1000 classes. The regression model is specified as follows:

$$\text{Return} = \alpha + \beta_5 \text{N5} + \beta_{10}\text{N10} + \cdots + \beta_{1000}\text{N1000} + \gamma_5\text{C5} + \gamma_{10}\text{C10} + \cdots + \gamma_{1000}\text{C1000} + \delta_1 \text{LP} + \delta_2 \text{log-LP} + \varepsilon. \quad (1)$$

The "N" variables and the "C" variables are simple discrete functions of the last closing price, denoted LP. The Nk variables contain a count of the number of "class k" cluster points at or below price LP. Thus, each time a round number is passed, the sawtooth function can jump up (or down) by an amount determined by its "class." The Ck variables are indicator variables that are equal to one if LP is a "type k" cluster point - they allow the fitted return at the cluster points themselves to be more or less than the full "jump" amount; in practice, we find that the returns at these points are in between the returns for prices above and below. The last closing price, LP, is included to capture the (possible) trend between the jumps at the cluster points, as well as any overall trend which might appear in the data. The logarithm of price is included to capture possible non-linearity in this trend.

Table II shows the results from estimating the sawtooth model on stocks with prices between 10 and 80 dollars. The first two columns show results from estimating the model without additional controls, while the second two columns show results including controls for the prior day's market returns, stock-specific returns, stock-specific trading volume, relative to the stock's trading volume in the prior month, and stock-specific closing spread, relative to the stock's spread in the prior month. While not shown, we have also estimated the regression with additional controls, such as the stock's prior week return, and estimated standard errors which are robust to heteroskedasticity. The results are extremely similar both in magnitude and significance. As seen in Table II, the results are strikingly consistent with the earlier description of the pattern. The coefficients on the "Nk" variables are all positive and strongly statistically significant, indicating that the model predicts an upward "jump" in returns around each cluster point. The "Ck" variables are all negative, indicating that the returns following the prices at the cluster points themselves are lower than the returns at prices just above the cluster points. In fact the magnitude of the Ck coefficients are generally about half the size of the corresponding Nk coefficients, indicating that the returns at the cluster points fall about midway between the abnormally high returns just above, and the abnormally low returns just below. The estimates of the return "jump" for each count variable are as high or slightly higher once the additional control variables are included, and remain statistically significant. We discuss these controls and additional robustness checks in more detail in subsequent sections.

The right-most panel of Table II, columns 5-6, presents the amount of excess clustering observed in the data at each "class" of round number. The size of the "jumps" fitted by the model correlate closely with the magnitude of excess clustering at each category of round number. For the "rounder" numbers, both the amount of clustering and the size of the fitted jump are substantially larger. Figure 4 depicts the return pattern graphically, showing the fitted values from the model for the range of prices between 15 and 25. The fractal-like sawtooth shape of the return pattern is clearly visible.

It is important to note that our model does not constrain the data to have the predicted sawtooth shape at all. If returns were random, we would expect some of the coefficients to be positive and some to be negative, with few statistically significant. The results of our model provide powerful evidence that there is a pattern of differential returns around round numbers, with the strength of the effect related to the "roundness" of the number, similar to the way in which clustering is more pronounced around "rounder" stock prices. Furthermore, the size of the effect for the round numbers at the one-, five- and ten-dollar levels is surprisingly large, with around a two-tenths of a percentage point difference in one-day returns.

Results in Subsamples

In this section we examine how the strength of the patterns described above varies in magnitude and statistical significance over time and over various categories of stocks. To simplify this task, we focus on two measures that represent the most salient features of the previously-described pattern of returns: On the one digit level, we compare returns following prices ending in "1" with those following prices ending in "9." At the two digit level, we compare returns following prices ending with digits 01-09, with those following prices ending in 91-99. (The choice to group digits 01-09 and 91-99 together is somewhat arbitrary. In fact, we find even larger differences if, for example, we only compare the returns of prices ending in 01 to those ending in 99. However, grouping allows more precise estimation, since there are more observations in each group.)

The comparison is done by estimating a simple regression model of the form

$$\text{Return}_{i,t} = \alpha + \beta_1 \text{ LD1}_{i,t\text{-}1} + \beta_2 \text{ LD9}_{i,t\text{-}1} + \varepsilon_{it}, \qquad (2)$$

where $\text{Return}_{i,t}$ is the midpoint-based daily percent return for stock i, $\text{LD1}_{i,t\text{-}1}$ is an indicator variable that equals one if the last digit of the closing price at time *t-1* is 1, and $\text{LD9}_{i,t\text{-}1}$ is an indicator variable that equals one if the last digit of the closing price at time *t-1* is 9. Thus the estimated values of the coefficients β_1 and β_2 show how high or low the returns are for prices ending in 1 or 9, respectively, relative to the returns for the other digits (0 and 2-7). The difference between β_1 and β_2 gives the difference between the returns following a 1-ending price and a 9-ending price. A similar regression is used at the two digit level:

$$\text{Return}_{i,t} = \alpha + \beta_1 \text{ LD2LOW}_{i,t\text{-}1} + \beta_2 \text{ LD2HI}_{i,t\text{-}1} + \varepsilon_{it}, \qquad (3)$$

where $\text{LD2LOW}_{i,t\text{-}1,}$ and $\text{LD2HI}_{i,t\text{-}1}$ are indicator variables that equal one if the last two digits of the closing price at time *t-1* are between 01 and 09, or 91 and 99, respectively.

Table III shows results from estimating Equations (2) and (3). (Alternatively, we have estimated all regressions throughout the paper using standard errors which are robust to heteroskedasticity, and results are similar. We do not discuss them in detail.) The columns in the table under the heading "one-digit regression" show the estimated values of β_1 and β_2, along with the t-statistic from the test of the hypothesis that $\beta_1 = \beta_2$, (or, in other words, that the returns for prices ending in 1 or 9 are the same). The results for the full sample, shown in the first row, show that daily returns following prices ending in 1 are estimated to be 2.8 basis points higher relative to mean returns from the other digits. Likewise, the returns following prices ending in 9 are estimated to be 4.1 basis points lower than those for other digits, for a total estimated 1-9 difference in returns of 6.9 basis points. The hypothesis that $\beta_1 = \beta_2$ can be rejected with a very high degree of confidence (t-statistic = 9.82), showing that this pattern is a highly statistically significant feature of the data. On the two-digit level, the differences are even larger and more significant: returns after prices ending in 01-09 are 5.0 basis points above other digits, while those after prices ending in 91-99 are 7.9 basis points lower, for a total difference of 12.9 basis points.

The rest of Table III shows the results of estimating Equations (2) and (3) on firm-days broken up into various subsamples by year, firm size, share price, volume, and exchange, in order to see if the pattern holds in general or if it is driven by a subset of stocks. It is immediately apparent that the basic pattern shows up at both the one- and two-digit levels across a wide variety of subsamples; almost every estimate for β_1 is positive, every estimate of β_2 is negative, and the t-statistics for the difference between the two are statistically significant in all but one case, usually at very high levels. However, we can see that the magnitude of the difference appears to be larger for some categories of stocks than for others.

Looking first to the variation in the return effect over time, the magnitude of the return differences at the one-digit level appears to be fairly stable over time, ranging from a low of 4.4 basis points in 2003 to a high of 8.8 basis points in 2004. At the two-digit level, the differences range from 9.1 in 2006 to 17.5 in 2001; it appears that there may have been some decline in magnitude over time, but the differences remain large. The stability of the pattern over time is important, because it shows that the pattern has continued to hold in up and down markets and through large changes in market structure.

The next lines in the table show results broken down by firm size. The firms are divided into quintiles based on the quintiles of market capitalization on the NYSE. (The large number of smaller firms on NASDAQ explains why the quintiles are not of equal sizes.) Here there is a clear pattern that the return differences are largest for small and mid-sized firms, although the differences are apparent and significant even for the largest firms.

When the results are sub-divided by price, the return difference in both the one-digit and two-digit regression is largest for the lowest-priced stocks (price≤$20) but is still pronounced for the two higher price categories. Similarly, when the results are broken down by previous day trading volume quintiles, the lowest trading volume group shows the largest return difference for both regressions. The return difference decreases over the volume quintiles, but are still statistically significant for all but the largest quintile in the 1-digit regression.

Finally, we separate the sample into two stock exchange groups – NYSE/AMEX and NASDAQ. In the one-digit regression, we find that the return difference effect is slightly more pronounced for NYSE/AMEX stocks, while in the two-digit regression, the return difference is higher for NASDAQ stocks. The return differences for both groups are statistically significant for both exchanges.

V. Exploring Possible Explanations

In the previous section, we documented an unusual pattern of stock returns related to the proximity of closing prices to round numbers and showed that the result was robust in various cross-sectional subsamples. In this section, we explore several potential explanations for our results. We first explore explanations related to market microstructure and then explanations related to investor psychology. Overall, we are unable to provide an explanation that fully explains our results.

Market Microstructure Artifact

The most immediate explanation for any short-run return pattern is that it is due to market microstructure considerations. However our tests suggest this is unlikely for the pattern we document. In our previous tests, we have already controlled for some microstructure effects, such as the bid-ask bounce effect (see Appendix A) and shown that the patterns we find are robust in subsamples of more liquid firms, such as large firms and firms with high trading volume (see Table III). Additionally, the result is robust to a large degree of variation in market structure. As shown in Table III, the results are strong for both NYSE/AMEX firms and NASDAQ firms, although the exchanges differ significantly in their structures (for example NYSE has a physical trading floor with a single market maker for each security, while NASDAQ is a computerized system with multiple market makers for each stock). In addition, there has been a large degree of change in market microstructure over the period of our sample. Automated execution has become much more prevalent, there have been large changes in commission structures, new order types such as pegged limit orders have been introduced, and various ECNs and alternative trading systems have risen and fallen and risen again.

In the subsections below, we explore market microstructure explanations more thoroughly by examining intra-day returns to remove the effects of irregularities in closing prices, overnight trading, and the opening auction, and we test whether our results are related to unusual trading activity. We do not find any evidence that the market microstructure characteristics we examine are driving the patterns we document.

Characteristics of Closing Price, Overnight Trading, and the Opening Auction

We first examine whether the return differences we document could be an artifact of irregularities in closing prices, overnight trading, or the opening auction by examining whether the return differences are concentrated in overnight returns or whether they persist into the next day. To answer this question, we examine returns between 11AM and closing on the current day, conditional on the last digit(s) of the previous-day reported closing price.[9] We find a similar (albeit weaker) pattern in these returns, showing that the "round numbers" effect lasts beyond the trading at or immediately following the opening of the market.

---

[9] The choice of 11AM is somewhat arbitrary, but we find similar results using 10AM or 12PM as the starting point. The data for 11AM price is taken from the NYSE TAQ database. We match firm-days using the ticker symbol, and record the price for the last transaction that occurred at or before 11AM. We drop firm-days that have no recorded transactions during the 5 minutes leading up to 11AM, and we also drop a very few observations that look like they are likely to be errors, because they imply a very large price swing that is almost completely reversed by day's end.

Table IV shows the regression results for the 11AM-to-close return for the entire sample and for sub-samples, analogous to Table III. The returns in this table are based on transaction prices, rather than midpoints, at both 11AM and at the close. This should not create a problem with bid-ask bounce, because there were many transactions between the previous close and 11AM that would wipe out the effect of any bounce effect correlated with the last digit of the previous day's closing stock price. In order to adjust for size effects, the regressions used to generate the table included a term for the log of market capitalization for the firm, and a term for the daily value-weighted market return.

The first line in the table shows the estimates for the whole sample. Once again, the estimate of β_1 is positive (although small), and the estimate for β_2 is negative. We again reject the hypothesis that the coefficients are equal with a high degree of confidence (t=3.05). For the two digit regression, the estimated differences are still larger and more statistically significant than for the one-digit regression.

Turning to the rest of the table, we again see evidence that the pattern holds across sub-samples. Because the differences in returns between 11AM and close are smaller than the full day returns, the tests have less statistical power. Nevertheless, we find that in the majority of the sub-samples, the coefficients have the expected sign. Moreover, the estimated difference (β_1 - β_2) has the expected positive sign in all sub-samples with the exception of 2003 in the single-digit regressions, and 2002 and the 4th size quintile in the two-digit regressions, and for each of these exceptions the difference is small and statistically insignificant. For each of the other sub-samples we find as expected that $\beta_1 > \beta_2$, and in many this difference is statistically significant.

Finally, we examine returns over a longer time window and find that the differential returns associated with closing prices above or below round numbers persist for several days. Figure 5 shows the average returns for each day from Day -4 to Day +3, conditional on the closing price on Day -1, for four groups of stocks, based on the last two digits of the closing price on Day -1. The groups are those with prices ending in 01-19, 20-39, 60-79, and 80-99. Unreported statistical tests show that the differences between returns for the 01-19 group and the 80-99 group remain significant three days after the conditioning closing price, reinforcing the results above to show that the return differences persist over time, rather than occurring in the trades immediately after the previous close. See the *Technical Analysis* section below for more discussion of Figure 5.

<u>*Unusual Trading Activity*</u>

Next, we test whether our results are related to unusual trading activity. To do so, we partition the sample in three additional ways. To test whether the results are robust to relative trading volume, we designate each firm-day as "high volume" or "low-volume" depending on whether the trading volume for the day exceeds the median share volume for the same firm in the previous calendar month. The results, reported in Table V, show that our previously documented pattern of returns looks substantially the same for high and low volume firm-days, for both Day *t-1* (previous day) and Day *t* (same day). These results are interesting because volume is strongly predictive of returns in our sample; overall, returns on high volume days are 16 basis points higher than on low volume days, and volume is correlated across consecutive days. Furthermore, volume is related to clustering, with more clustering occurring on high volume days[10]. Nevertheless, the pattern and even the magnitude of return differences was largely the same between the low and high volume groups.

The next section of the table compares the return differences based on whether the excess return on the stock on Day *t-1* (the day on which the conditioning stock price was recorded) was negative or positive. As before the return variable used is midpoint-based, size-adjusted excess returns. Once again the direction and magnitude of the return differences do not differ greatly between groups. Using the sawtooth model to estimate the pattern also yields very similar results following high vs. low return days, as does estimating the sawtooth model including control variables for these characteristics, as shown in Table II.

[10] Clustering is more common for low-volume stocks. However, for a specific stock, clustering is actually stronger on high-volume days than on low-volume days.

In additional, un-tabulated tests, we also find the basic pattern described above to be robust to the inclusion of a variety of other control variables, including previous market return, market volume, bid-ask spread, and price.

Behavioral Explanations and Considerations

In this section, we explore several behavioral issues that could be related to our results. First, using brokerage data from the early 1990's, we show that retail investors appear to trade differently when prices are just below a round number than when they are just above a round number. While we aren't able to show directly how this finding would affect returns, it does provide evidence that behavioral considerations may play an important role in explaining our results. We also construct tests to determine whether trading by small traders, technical trading, or order clustering on round numbers is driving our results, but are not able explain our results using any of these explanations.

Retail Investor Trading

In this section we examine whether trading by retail investors is consistent with the pattern of returns we document in Section IV. If the returns pattern is driven by active trading, and particularly if it is driven more strongly by less sophisticated investors, we would expect to see stronger retail investor buying after closing prices above round numbers and stronger retail investor selling after closing prices below round numbers.

We analyze retail investor trading using a dataset of brokerage accounts for 78,000 households at a large discount brokerage firm, covering the six years from 1991 through 1996. This time period is before the exchanges moved to quoting prices in decimals, and so we may not expect as strong of a pattern as in the post-decimalization period, however this data can still be informative about the returns pattern. The dataset contains each trade made within the brokerage accounts, for all securities. It includes the number of shares purchased or sold, the price at which they were traded, as well as additional investor-level information. Barber and Odean (2000, 2001 and 2002) describe the brokerage account data in detail. We restrict our analysis to common stock for U.S. firms, as in the main analysis.

We create measures, for each stock-day, of how strongly retail investor trading is weighted towards buying or selling. For each firm-day, we define trade imbalance as

$$TradeIMB_{x,t} = \frac{buys_{x,t} - sells_{x,t}}{buys_{x,t} + sells_{x,t}}, \tag{4}$$

where $buys_{x,t}$ represents the number of buy trades for firm x on day t made by the retail investors in the sample, and $sells_{x,t}$ represents the number of sell trades. We define *DollarIMB* analogously, using the total dollar value bought and sold by investors in the sample. Finally, we construct an abnormal trade imbalance measure following Malmendier and Shanthikumar (2007), as

$$IMB_{x,t}^{abnormal} = \frac{IMB_{x,t} - E\left(IMB_{x,year(t)}\right)}{\sqrt{Var(IMB_{x,year(t)})}}, \tag{5}$$

where $E(IMB_{x,year(t)})$ and $Var(IMB_{x,year(t)})$ are the mean and variance of the given imbalance measure, calculated for the firm-year. This normalization takes the place of firm-year fixed effects, and controls for differences in the variance of trade across stocks. We then examine how these trade imbalance measures (based on number of trades and dollars traded) vary with respect to prior-day closing price.

Over the years 1991 through 1996, prices on the NYSE were primarily quoted in 1/8 dollar increments, though some more heavily traded stocks were quoted in 1/16 increments. Because of this, we cannot use the same type of "digit" analysis as used in Section IV. Instead, we define analogous variables to measure whether the previous day's closing price was above or below a round number, specifically above or below a whole dollar value. We estimate the following regression, which is analogous to Equations (2) and (3),

$$TradeVariable_{i,t} = \alpha + \beta_1\, FracLow_{i,t-1} + \beta_2\, FracHigh_{i,t-1} + \varepsilon_{it}, \qquad (6)$$

where $FracLow_{i,t-1}$ is an indicator variable that equals one if the fractional part of the closing price on Day *t-1* is 1/8 or 1/16, $FracHigh_{i,t-1}$ is an indicator variable that equals one if the fractional part of the closing price at time *t-1* is 7/8 or 15/16, and $TradeVariable_{i,t}$ is one of the four trade imbalance variables defined using Equations (4) and (5). In addition, because trading might be affected by concurrent returns, (i.e. investors may buy more strongly on days with positive returns simply in response to the returns), we estimate the regression including a control for current-day returns.

Table VI presents results. Returns for the full sample, shown in the first row, follow a similar pattern as in the post-decimalization period, though the difference in returns is smaller in magnitude during the earlier period. Returns are 0.009 higher following closing prices which end in 1/16 or 1/8, just above a round dollar number, while they are 0.027 lower for prices ending in 15/16 or 7/8, just below a round dollar number. This difference of 0.036 compares to a difference of 0.129 in the post-decimalization period when comparing prices ending in 01-09 with prices ending in 91-99.

The next four rows show results for the four trade imbalance variables defined above. In each of the eight specifications, trade imbalance differs significantly above and below round numbers, in the same direction as the returns pattern. An examination of trade at each of the eight common closing prices shows that buying is stronger for 0, ¼, ½ and ¾ prices. Because of this, the estimated coefficients are negative for both the 1/16-1/8 prices and the 7/8-15/16 prices. However, the trade reaction is significantly more negative for prices just below a round number, i.e. ending in 7/8 or 15/16, than for prices just above.

Although the results in this section should be interpreted with some caution, since they come from a different sample period than our main results, they do provide some indication that retail investors' trade activity seems to be related to whether the price is immediately above or below a round number. Further, the net trade imbalance in the market must be zero, so that another set of investors may be trading in the opposite manner as the retail investors in this sample. To explore these findings further, we next explore whether small traders are driving our results by examining the relationship between institutional trading and our returns pattern.

Naïve Small Traders

If the pattern we document is driven by investors' use of round number "cognitive reference points" it may be that only novice individual investors are subject to this cognitive pattern, while sophisticated institutional investors are not. Given prior work on financial market anomalies[11], it seems quite likely that the pattern we document should be strongest among firms with low institutional ownership and weakest or non-existent in firms with high institutional ownership.

Using institutional ownership data from Thompson Financial's CDA/Spectrum database,[12] we form quintiles based on institutional ownership, and test for the differences in returns following closing prices just above and just below one-digit and two-digit round numbers.[13] Results are displayed in Table VII. As can be seen in the "mean percentage" column, there is a high degree of variation in institutional ownership levels. However, differences in returns between prices ending in "1" and "9" or "01-09" and "91-99" are similar in magnitude and significant for each of the five groups.

---

[11] See Barberis and Thaler (2003) and Shleifer (2000) for surveys of this literature and Baker and Wurgler (2006) and Kumar and Lee (2006) for two specific examples.

[12] All institutional investment managers managing more than $100 million in equity must report their holdings quarterly, for all holdings of at least 10,000 shares or $200,000 in market value, in form 13f. We use the 13f data to calculate quarterly measures of institutional ownership. We then match these measures to our stock price and return data for the following quarter. Not all institutions report quarterly. For those that do not, we use an average of the holdings in the prior and following quarters as an estimate of the intermediate quarters' holdings.

[13] Quintiles are formed over the full sample; however forming quintiles for each sample day separately yields similar results.

Since some institutional investors are not active traders and are therefore unlikely to impact short-run returns, we use the classification scheme in Bushee (1998) to form a second set of quintiles based on the proportion of actively traded shares that are owned by institutions.[14] Again, the results are displayed in Table VII. While the magnitude of the difference between returns on prices ending in "1" and "9" is low for the stocks with the lowest levels of "Institutional ownership of actively traded shares," all other values are similar across the groups and there are no clear trends. Once again, the differences are statistically significant for each group.

Overall, we find no clear relationship between the return pattern we document and levels of institutional ownership or our proxy for relative institutional trading activity. This is consistent with our prior that, because the use of cognitive reference points is potentially a more ingrained bias, it is likely that even the most sophisticated investors will unconsciously rely on these reference points when making subjective judgments.

Technical Analysis

Technical analysis is the practice of trading stocks based on price patterns (rather than company fundamentals or valuation alone). Technical traders often refer to a stock's "resistance" and "support" levels, which can be defined in various ways but the general idea is that a firm's stock price tends to move around within a range of prices bounded by a resistance level on the top and a support level on the bottom. Therefore, if the price increases (decreases) to just under (over) a resistance (support) level, it is likely to go back down (up). If, however, the price breaks upwards (downwards) through a resistance (support) level, then technical analysts expect that it will continue to go up (down), as the stock now has a new resistance (support) level. Thus technical traders may be more likely

[14] Following Bushee (1998), we classify institutions as "transient" if they are active traders, "dedicated" who "buy and hold" for a longer period, and "quasi-indexer" if they hold portfolios similar to major indices. In general, we would expect that "dedicated" and "quasi-indexer" investors are not active traders, so the remaining shares not owned by these two groups is a proxy for the number of actively traded shares. We calculate the percentage of actively traded shares owned (and traded) by institutions as:

$$\frac{\%Shares_{transient}}{1-(\%Shares_{Dedicated}+\%Shares_{Quasi-indexer})}.$$

to buy a stock whose price is just above a support level or a resistance level and would be more likely to sell a stock whose price is just below either of these levels.

Although there are no technical reasons why these resistance and support levels should be round numbers, anecdotal evidence indicates that they often are. In this case, technical traders would be more likely to sell when the stock price is just below a round number and more likely to buy when the stock price is just above a round number. It is difficult to test explicitly whether this type of technical trading drives our results because different technical analysts use different time periods and methodologies to construct their estimates of resistance and support levels and trade in firms with varying characteristics, such as size. However, in the tests that we do construct below we are unable to find evidence that technical trading is driving our results.

To examine whether our results appear to be related to technical trading activities, we examine a longer return pattern by looking at the pattern of returns over several days before and after the day the conditioning "last digits price" was recorded. We return to Figure 5, which shows the average returns on each day from Day -4 to Day +3, conditional on the closing price on Day -1, for four groups of stocks based on the last two digits of the closing price on Day -1.

The large "spikes" on day 0 represent the differential returns already extensively documented above. As expected, they are unusually high for the "01-19" group and much lower for the "80-99" group. The figure also shows that the odd patterns of returns are not limited to the day after the closing price. Most notably, there is a smaller pattern of opposite sign on Day -1, the day that the conditioning price is recorded. It would seem that the large differences on day "0" are at least in part a reversal of the opposite differences on Day -1, and to a lesser extent on days before Day -1.

To test more for technical trading effects more explicitly, We regress daily returns on indicator variables that indicate whether the previous day's closing price was just above or below a round number and whether the closing price is a new high or low over the previous two days. If technical trading as described above is driving our results, we would expect positive returns for a new high or new low that is just above a round number and negative returns for a new high or new low that is just below a round number, so we also include interaction terms. Table VIII shows the results. While the coefficients on 'new high' and some of the interaction terms are statistically significant, the coefficients on the indicators for prices just above/below round numbers remain large and significant. In unreported regressions, we repeat these tests requiring that the price be a new high or low over longer time periods, such as twenty days, and also test for returns differences following three days of consecutive positive or negative returns. We still find no evidence that our returns patterns are driven by technical trading activity.

Clustering

Finally, we explore more carefully the relationship between clustering and our pattern of returns. It is well documented that stock prices and limit orders cluster on round numbers, but none of the prior explanations for clustering have had explicit returns predictions, and to our knowledge, no analytical models of clustering have returns predictions. To explore the potential link between clustering and stock returns, we compare changes in clustering over time to changes in our returns pattern over time. Additionally, we simulate trading with clustered limit orders to determine whether there is some mechanism that leads from clustered limit orders to our pattern of returns.

The clustering of stock prices on round numbers has decreased over time. Table IX shows that in 2006, the frequency of each of the cluster points listed had decreased to roughly half of its frequency in 2001. Given that the "expected" frequency of each of these round numbers is non-zero (roughly 10% for the 1-digit round numbers and roughly 1% for the 2-digit round numbers), the level of clustering has decreased by roughly 65-75%. In comparison, the returns difference at the one-digit level has shown no consistent decrease over the years and remains at roughly the same levels as in 2001 and the returns difference at the two-digit level has decreased by less than half. These results provide evidence that our returns patterns are not driven entirely by price or limit order clustering.

In our next approach to understanding the link between clustering and stock prices, we generate two simulations to test whether order rounding would generate differential returns based on the last digit of the closing stock price, further varying parameters within each model. In the first model, we simulate an opening auction in which overnight orders cluster on round numbers. We ran the test using various assumptions about the order distribution, but did not find next-day differential returns associated with closing prices above or below round numbers.

Our second simulation model is based on the notion that our results might be driven by stale limit orders clustered on round numbers that are not cancelled when new information arrives. This might cause a stock price that is adjusting to new information to get "stuck" at round number cluster points until the stale orders are filled. For example, a stock price ending in one could move up four cents before getting "stuck" at a five-ending round number or move down one cent before encountering resistance at a zero-ending round number. This phenomenon could then result in positive average returns for stocks with closing prices just above a round number, since the stock price can move up more freely than it can move down. A stock price just below a round number would experience the opposite effect – it could move down more freely than it could move up, leading to lower average returns. We attempted to operationalize this notion of price "stickiness" in simulations using various asymmetric distributional assumptions about order arrivals, but we did not find differential returns based on closing prices above or below round numbers in any of these variations.

VI. Conclusion

We document a robust pattern of returns in which one-day returns following a closing price just above a round number are persistently higher than one-day returns following a closing price just below a round number. We find that the pattern holds for many subsamples, including the largest and most liquid stocks and does not appear to be driven by common market microstructure characteristics, such as the bid-ask bounce effect or characteristics of closing prices, overnight trading, or the opening auction.

We test several behavioral explanations for our results and find no evidence that our results are driven exclusively by technical trading, naïve small traders or order clustering. We do find some evidence that retail investors trade differently when the stock price is just above a round number rather than just below, but we are not able to connect this finding directly to the pattern of returns that we document. While psychology research has documented that individuals tend to use round numbers as cognitive reference points and that this affects their perceptions and reactions in a variety of settings, and while prior literature has documented various ways in which individual investors trade sub-optimally, it would be somewhat surprising to see that something as simple as whether a price ends in a 9 or a 1 affects returns for a large part of the market. On the other hand, the use of cognitive reference points is one of the more deeply ingrained and unconscious psychological tendencies, and so it may be more persistent even in individuals with financial training and education, unless explicitly addressed. Further work is needed in this area.

References

Aitken, Michael, Philip Brown, Christine Buckland, H. Y. Izan, and Terry Walter, "Price Clustering on the Australian Stock Exchange," *Pacific Basin Finance Journal* 4 (1996): 297-314.

Baker, Malcolm, and Jeffrey Wurgler, "Investor Sentiment and the Cross-Section of Stock Returns," *The Journal of Finance* 61:4, (2006), 1645-1680.

Barber, Brad M., and Terrance Odean. "Boys will be Boys: Gender, Overconfidence, and Common Stock Investment." *Quarterly Journal of Economics* 116, (2001), 261-292.

Barber, Brad M., and Terrance Odean. "Online Investors: Do the Slow Die First?" *Review of Financial Studies*, 15, (2002), 455-489.

Barber, Brad M., and Terrance Odean. "Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors." *Journal of Finance*, 55, (2000), 773-806.

Barberis, Nicholas, and Richard Thaler, "A Survey of Behavioral Finance," in *Handbook of the Economics of Finance,* (2003), Elsevier Science.

Bizer, George Y. and Robert M. Schindler, "Direct Evidence of Ending-Digit Drop-Off in Price Information Processing," *Psychology & Marketing* 22:10, (2005) 771-783.

Bushee, Brian, "The Influence of Institutional Investors on Myopic R&D Investment Behavior," *The Accounting Review* 73:3, (1998) 305-333.

Carslaw, Charles A.P.N., "Anomalies in Income Numbers: Evidence of Goal Oriented Behavior," *The Accounting Review* 63:2, (1988), 321-327.

Christie, William G. and Paul H. Schultz, "Why Do NASDAQ Market Makers Avoid Odd-Eighth Quotes?", *The Journal of Finance* 49:5, (1994), 1813-1840.

Cooney, J.W., B. Van Ness, and R. Van Ness R., "Do investors prefer even-eighth prices? Evidence from NYSE limit orders**,"** *Journal of Banking and Finance* 27:4, April (2003), 719-748.

Corwin, Shane A., "The Determinants of Underpricing for Seasoned Equity Offers," *Journal of Finance* 58:5, (2003), 2249-2279.

Das, Somnathi and Huai Zhang, "Rounding-up in Reported EPS, Behavioral Thresholds, and Earnings Management," *Journal of Accounting and Economics* 35, (2003), 31-50.

Gedenk, Karen and Henrik Sattler, “The Impact of Price Thresholds on Profit Contribution-Should Retailers Set 9-Ending Prices?” *Journal of Retailing* 75, (1999), 33-57.

Goodhart, Charles, and Richard Curcio, “Asset Price Discovery and Price Clustering in the Foreign Exchange Market,” *working paper*, London School of Economics, (1990).

Grossman, Sanford J. et al., “Clustering and Competition in Asset Markets”, *Journal of Law and Economics,* 40, (1997), 23-60.

Harris, Lawrence, “Price Clustering and Discreteness” *Review of Financial Studies,* 4, (1991), 389- 415.

Herrmann, Don and Wayne B. Thomas, “Rounding of Analyst Forecasts.” *The Accounting Review,* 80:3, (2005), 805-823.

Ikenberry, David and James P. Weston, “Clustering in US Stock Prices after Decimalization,” *working paper*, (2003).

Kandel, Shmuel, Oded Sarig, and Avi Wohl, “Do Investors Prefer Round Stock Prices? Evidence From Israeli IPO Auctions,” *Journal of Banking and Finance* 25, (2001), 1543-1551.

Kumar, Alok, and Charles Lee, “Retail Investor Sentiment and Return Comovements,” *The Journal of Finance*, 61:5, (2006), 2451-2486.

Malmendier, Ulrike and Devin Shanthikumar. “Are Small Investors Naïve About Incentives?” *Journal of Financial Economics*, 85, (2007), 457-489.

Niederhoffer, V., “Clustering in Stock Prices”, *Operations Research* 13, (1965), 258-265.

Niederhoffer, V., “A new look at clustering in stock prices”, *Journal of Business* 39, (1966), 309-313.

Odders-White, E. R., “On the occurrence and consequences of inaccurate trade classification,” *Journal of Financial Markets* 3, (2000), 259-286.

Op de Beeck, Hans, Johan Wagemans and Rufin Vogels, “Asymmetries in Stimulus Comparisons by Monkey and Man,” *Current Biology* 13, (2003), 1803-1808.

Osborne, M. F. M., “Periodic Structure in the Brownian Motion of Stock Prices”, *Operations Research, (*1962), 345-379.

Osler, Carol, “Currency Orders and Exchange Rate Dynamics: An Explanation for the Predictive Success of Technical Analysis,” *The Journal of Finance* 58:5, (2003), 1791-1819.

Rosch, Eleanor, “Cognitive Reference Points,” *Cognitive Psychology* 7, (1975), 532-547.

Schindler, Robert M. and Patrick N. Kirby, “Patterns of Rightmost Digits Used in Advertised Prices: Implications for Nine-Ending Effects,” *The Journal of Consumer Research,* 24:2, September, (1997), 192-201.

Schindler, Robert M. and Alan R. Wiman, “Effect of Odd Pricing on Price Recall,” *Journal of Business Research*, 19, November, (1989), 165-177.

Securities Exchange Act Release No. 42914 (June 8, 2000), 65 FR 38010 (June 19, 2000).

Shleifer, Andrei, “Inefficient Markets: An Introduction to Behavioral Finance,” (2000), Oxford University Press.

Stiving, Mark and Russell S. Winer, “An Empirical Analysis of Price Endings with Scanner Data.” *The Journal of Consumer Research* 24:1, June, (1997), 57-67.

Thomas, Jacob, “Unusual Patterns in Reported Earnings.” *The Accounting Review* 44:4, (1989), 773-787.

Tversky, Amos and Daniel Kahneman, “Availability: A Heuristic for Judging Frequency and Probability,” *Cognitive Psychology,* 5, September, (1973), 231-238.

Unger, Laura S., “Testimony Concerning The Effects of Decimalization on the Securities Markets,” U.S. Securities and Exchange Commission, May 24, 2001.

The Wall Street Journal, “Journal’s Price Tables for Stocks And Mutual Funds Are Changed,” The Wall Street Journal, April 30 2001.

Appendix. Clustering, Bid-Ask Bounce, and Buy-Sell Imbalance

Throughout the paper we focus on midpoint-based returns. In this subsection we look at raw returns, based on closing prices, and examine some of the problems with using closing-price-based raw returns. In particular, we explain and show how bid-ask bounce, in combination with previously documented price clustering, can magnify the return pattern we observe. In fact, it could even create the pattern in the absence of a true midpoint-based pattern. This section details how this occurs, and explains why we choose to focus on midpoint-based returns.

Description of the Differential Bid-Ask Bounce Effect

In general, the bid-ask bounce effect arises as follows: to compensate market participants for providing liquidity, market orders to buy typically occur at a higher price (providing the liquidity provider with more cash for their share) than do market orders to sell (allowing the liquidity provider to purchase the share with less cash). Thus the market maker's quotation of the price they are willing to pay to buy, the bid, is typically lower than the market maker's quotation of the price they are willing to accept to sell, the ask. Similarly, limit orders, which provide liquidity to the market, typically fall above or below the bid-ask midpoint depending on the direction of the order. Because of this pattern, we will typically see positive returns after a market order to sell, because this order occurred at a lower price, and we will typically see negative returns after a market order to buy, because this order occurred at a higher price. This phenomenon is referred to as bid-ask bounce.

If the last digit of the closing price is correlated with whether the closing trade is a "buy" or "sell" trade, then we may find a relationship between the last digit of closing price and returns that is driven by bid-ask bounce. It is important to note that this requires a relationship between the last digit of the closing price and whether the closing trade is a "buy" or "sell" trade; therefore we term this effect "differential bid-ask bounce." Absent market frictions or behavioral biases, there is no clear reason that the bid quote (the price at which the market maker is willing to buy shares) should fall more often on a price ending in 9 than a price ending in 1, or vice-versa for the ask quote. However given prior empirical results in the literature, explained by both market frictions and behavioral biases, there is reason to suspect that bid and ask quotes may be correlated with certain last digits.

Prior work has documented pervasive price and limit order clustering in securities markets. Before the move to decimalization, prices were most likely to fall on integer dollar amounts, second most likely to fall on half-dollar amounts and third most likely to fall on quarter-dollar amounts (Osborne, 1962, Nierderhoffer, 1965 and 1966, Harris, 1991, and Christie and Schult, 1994, Grossman et. al. 1997, Cooney, VanNess and VanNess, 2000). Ikenberry and Weston (2003) show that clustering persists after the move to decimalization, with prices clustering on nickel and dime amounts – i.e. prices ending in 0 or 5.

Given that prices cluster on round numbers, an investor who wants to place a limit order, but wants faster execution and a higher probability of execution than for a limit order at the clustered price, will tend to place their order one cent away from the clustered price. This sort of strategic order placement is likely to be higher after the move to decimalization, when an investor can now gain the benefits of faster execution for a cost of only one cent, rather than 6.25 or 12.5 cents when prices were quoted in sixteenths and eights, respectively. In order to gain this benefit, a limit order to buy would have to be placed at a higher price than the round number – i.e. at a price ending in 6 or 1. The investor placing the limit order will pay one extra cent per share, but will have their order filled before the cluster of orders at the round number, as they are offering a better execution price. In contrast, the investor has little incentive to place their limit order at a price ending in 4 or 9 – they gain only a penny, but their trade is much less likely to be executed. Given the relative dearth of orders in the range between nickel prices, the investor's order would be almost as likely to be executed if they set an even lower price (i.e. ending in 1, below the 4 but above the 0 cluster, and ending in 6, below the 9 but above the 5 cluster) thus gaining 3 cents per share for only a small cost in execution speed and probability. Thus, limit orders to buy, while most likely to fall on the clusters, are second most likely to fall just above the clusters. This implies that a market order to sell is more likely to fall just above the clusters, on a price ending in 6 or 1. Given bid-ask bounce, we would expect positive returns after a market order to sell, and hence higher returns after a price ending in a 6 or 1. Similar reasoning suggests that a limit order to sell should be placed just below the cluster, i.e. the investor accepts one penny less per share for the improvement in execution speed and probability, so that a market order to buy is more likely to fall just below the cluster, on a price ending in 4 or 9. These market orders to buy will be followed by negative returns, and hence we would observe lower returns after prices ending in 4 or 9. In this way, the bid-ask bounce, in combination with price clustering and strategic order placement, could lead to the return patterns we observe.

Evidence for the Differential Bid-Ask Bounce Effect

We now turn to the data to see if it is consistent with this story. We first examine whether the requisite price clustering occurs in our sample. Table V shows the observed level of one- and two-digit clustering in CRSP closing prices over the sample period. The table confirms Ikenberry and Weston's findings that clustering persisted after decimalization. For the full sample, 19.9% of the closing prices ended in 0, and 14.9% ended in 5. (If there were no clustering we would expect both percentages to be around 10%). Likewise there is significant clustering at the two digit level: 3.7% of the closing prices ended in 00, and 2.5% ended in 50, versus the 1% we would expect in the absence of clustering. However, while clustering remains prevalent, we find that it has declined dramatically over the years since decimalization. By 2006 there was, very roughly, only about one third as much excess clustering as in 2001. Thus we would also predict this would lead to a reduction in the differential buy-sell imbalance, and thus in the differential bid-ask bounce effect.

We cannot directly test strategic limit order placement, as we do not have access to limit order data. However, we can attempt to roughly classify closing price transactions based on the reported closing bid and ask quotes. If a closing price is below the midpoint of the bid-ask spread, it is more likely that it was a “sell,” (i.e. a trade resulting from a “sell” market order, executed at the bid price), and if it is above the midpoint, it is relatively more likely that it was a “buy.”[15] In the literature this is referred to as the “quote method” of trade classification. In untabulated analyses, we conduct this classification and find a familiar pattern: prices ending in 1, 2, or 6, i.e. just above a round number, are the most likely to be classified as sells, while those ending in 4, 8, or 9, i.e. just below a round number, are the most likely to be classified as buys. These classifications are exactly consistent with the “differential bid-ask bounce” effect described above. We also find similar imbalances based on clustering at the two-digit level.

For the full sample, a closing transaction at a price ending in 1 is classified as a sell 43.96% of the time, vs. 36.93% of the time for a closing price ending in 9, for a difference of 7.03%. Likewise, a closing price ending in 1 is classified as a buy only 52.33% of the time, vs. 59.27% of the time for a closing price ending in 9, for a difference of 6.94%. Therefore it would appear that there is about a 7% differential in the buy and sell percentages between 1 and 9.

Thus we have found evidence of substantively significant differences in the buy-sell composition of transactions based on the last digit of the closing price. This finding of differences in buy-sell imbalance based on last digit is itself very strong evidence that the differential bid-ask bounce effect will result in at least some return patterns based on raw returns, similar to the patterns we observe for midpoint-based returns. This would be the case even in the absence of any midpoint-return effect.

[15] There is a large literature on trade classification algorithms. Odders-White (2000) examines the accuracy of the “quote method” of trade classification that we use here, using the TORQ dataset of orders from a sample of NYSE firms in 1990. She finds that the method misclassifies 9.1% of the transactions and fails to classify 15.9%, due to trades occurring at the bid-ask midpoint. Our rates of misclassification may well be quite different, since this study was based on a small sample of firms pre-decimalization, but this gives us some confidence in the method.

Figure 1. Returns based upon digit characteristics of previous-day closing price
This figure shows daily midpoint-based size-adjusted returns for sub-samples grouped by specific digits of the previous day's closing stock price. The last digit, Figure 1a, indicates the rightmost digit for prices quotes in dollars and cents (decimals). The "Dime" digit, Figure 1b, indicates the second rightmost digit. "Dollar" digit, Figure 1c, indicates the third rightmost digit, the digit which is left of the decimal point. Midpoint-based Size-adjusted Excess Returns are returns based upon the midpoint of the closing bid and ask quotes, adjusted for market returns by subtracting the value-weighted return of the given firms' size quintile, using NYSE size cutoffs. In each figure, the points mark the average return for the given categories, while the bars indicate 95% confidence intervals around the point estimates.

Figure 1a: Midpoint-based excess returns by last digit of previous-day closing price

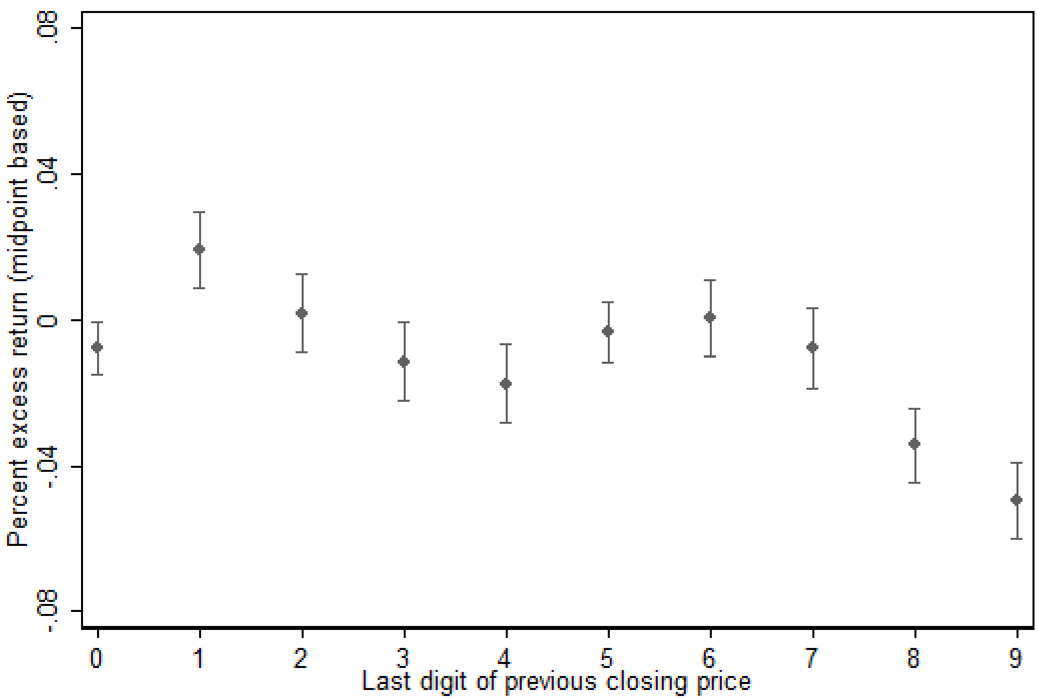

Figure 1b: Midpoint-based excess returns by "dime" digit of previous closing price

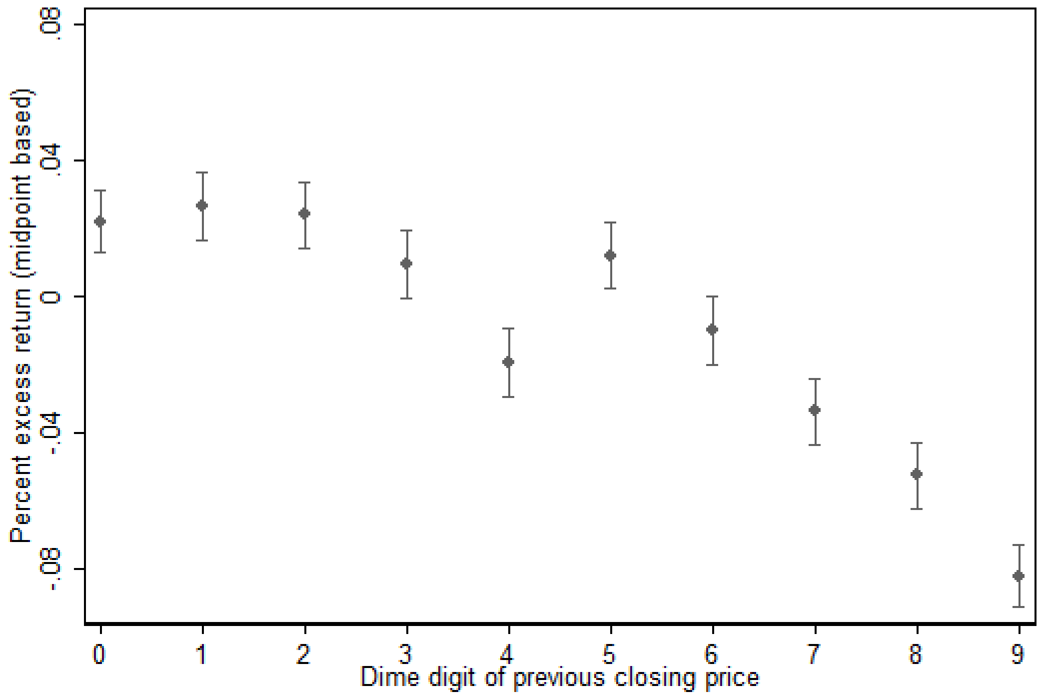

Figure 1c: Midpoint-based excess returns by "dollar" digit of previous closing price

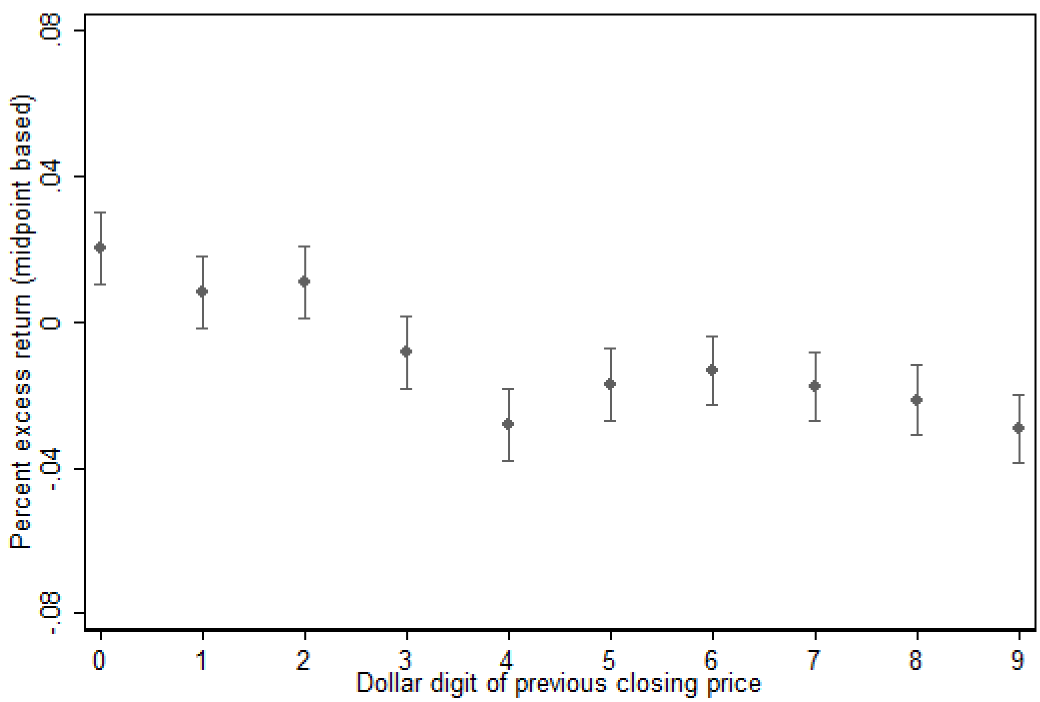

Figure 2: Midpoint-based excess returns by last two digits of previous-day closing price

This figure shows daily midpoint-based size-adjusted returns for sub-samples grouped by the last two digits of the previous day's closing stock price. Midpoint-based Size-adjusted Excess Returns are returns based upon the midpoint of the closing bid and ask quotes, adjusted for market returns by subtracting the value-weighted return of the given firms' size quintile, using NYSE size cutoffs. The figure also shows fitted regression lines, fitted separately for prices ending in 01-49 and prices ending in 51-99.

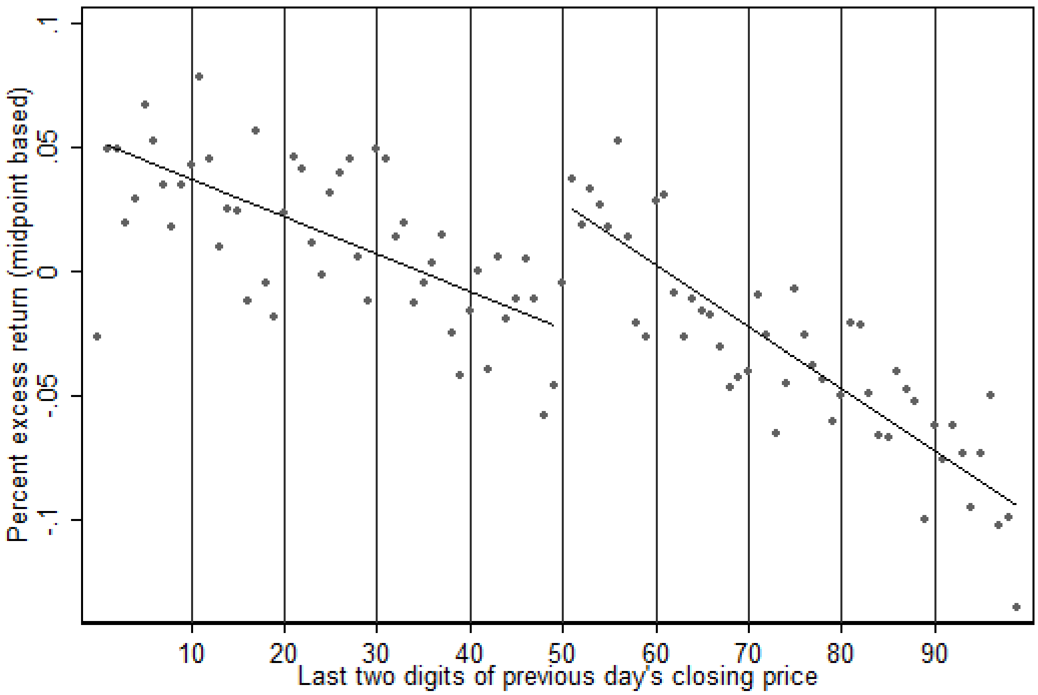

Figure 3: Midpoint-based excess returns by previous day closing price truncated to the 10 cent level

This figure shows midpoint-based excess returns for securities based upon the previous-day closing price, grouped into 10-cent intervals. For example, there is one point estimate for prices between 10.00 and 10.09 and another for prices between 10.10 and 10.19. Midpoint-based Size-adjusted Excess Returns are returns based upon the midpoint of the closing bid and ask quotes, adjusted for market returns by subtracting the value-weighted return of the given firms' size quintile, using NYSE size cutoffs. The figure also shows fitted regression lines, fitted separately for the points within each 10-dollar interval.

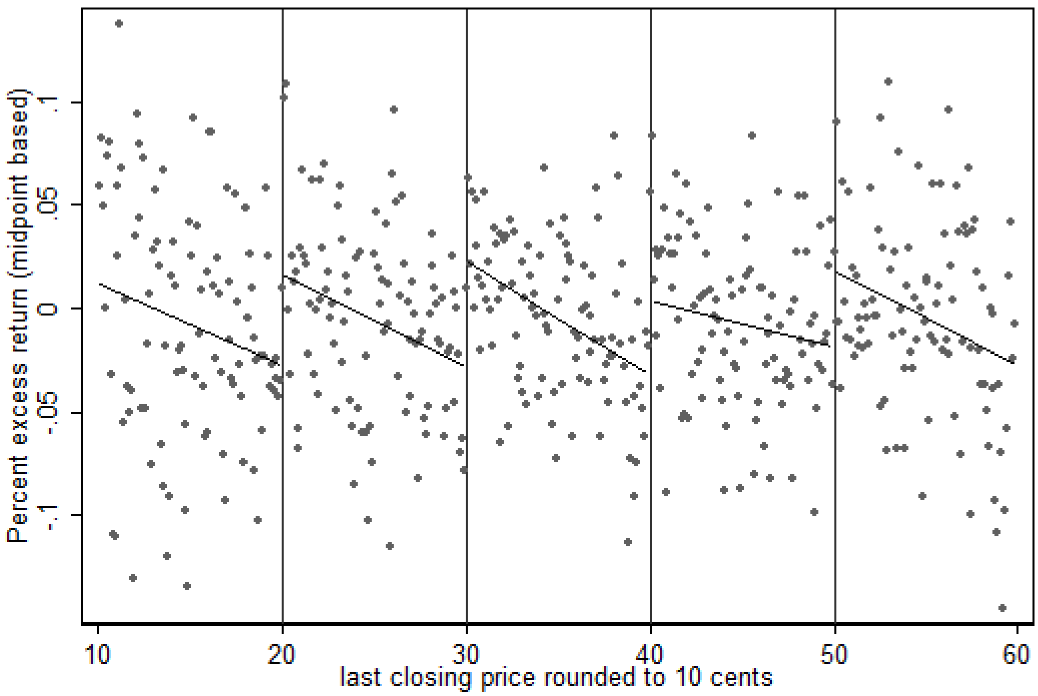

Figure 4: Fitted values from sawtooth model, for prices between $15 and $25

This figure plots fitted values from estimating a sawtooth model of the relationship between daily midpoint-based excess returns and prior day closing price. Midpoint-based size-adjusted excess returns are based upon the midpoint of the closing bid and ask quotes, adjusted for market returns by subtracting the value-weighted return of the given firms' size quintile, using NYSE size cutoffs. The model includes indicator and count variables for 5-, 10-, 25-, 50-, 100-, 500-, and 1000-cent increments.

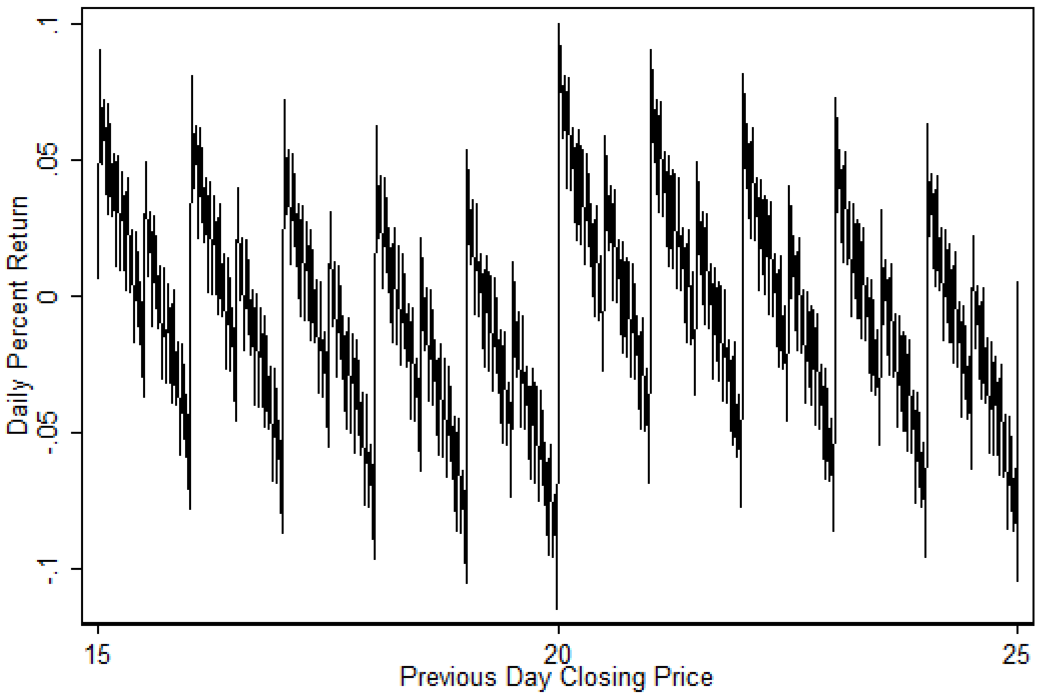

Figure 5: Average Daily Returns, by last digits of previous day closing price
This figure plots average daily midpoint-based excess returns for four sub-samples based upon the last two digits of closing price on day -1, for several days surrounding the given closing price. Midpoint-based size-adjusted excess returns are based upon the midpoint of the closing bid and ask quotes, adjusted for market returns by subtracting the value-weighted return of the given firms' size quintile, using NYSE size cutoffs.

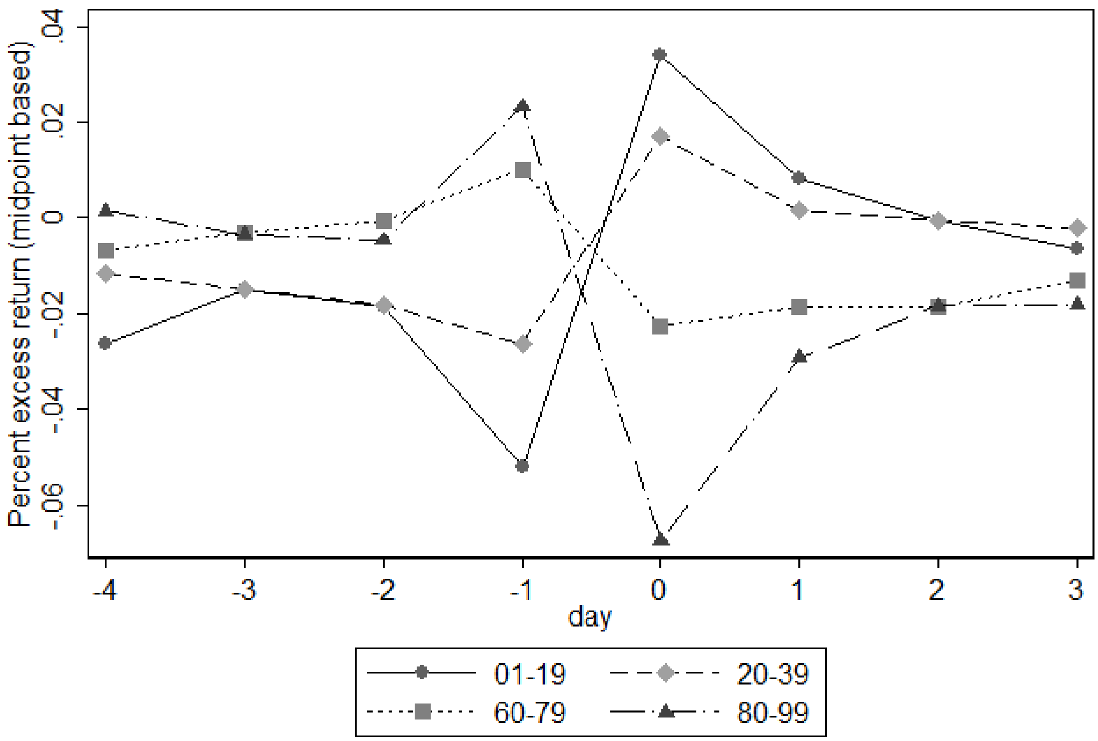

Table I. Returns by Prior Day Closing Price Digits

This table shows daily raw returns and daily midpoint-based size-adjusted returns for sub-samples grouped by specific digits of the previous day's closing stock price. The "Penny" digit indicates the rightmost digit for prices quotes in dollars and cents (decimals). For example, "Penny" digit equaling 4 indicates a price of $X.Y4. The "Dime" digit indicates the second rightmost digit. "Dollar" digit indicates the third rightmost digit, the digit which is left of the decimal point. Midpoint-based Size-adjusted Excess Returns are returns based upon the midpoint of the closing bid/ask spread, adjusted for market returns by subtracting the value-weighted return of the given firms' size quintile, using NYSE size cutoffs.

"Penny" digit	Raw Returns			Midpoint-based Size-adjusted Excess Returns	
	Sample Size	Mean	Standard Error	Mean	Standard Error
0	719,830	.043	0.0036	-.008	0.0034
1	311,778	.103	0.0052	.019	0.0049
2	297,092	.074	0.0052	.002	0.0049
3	282,203	.051	0.0053	-.012	0.0051
4	293,234	.021	0.0053	-.017	0.0051
5	545,135	.052	0.0041	-.003	0.0039
6	301,495	.083	0.0052	.000	0.0050
7	285,642	.055	0.0053	-.008	0.0051
8	308,139	.010	0.0051	-.034	0.0048
9	316,788	-.017	0.0052	-.050	0.0049
"Dime" digit	**Raw Returns**			**Midpoint-based Size-adjusted Excess Returns**	
	Sample Size	Mean	Standard Error	Mean	Standard Error
0	443,040	.083	0.0044	.022	0.0042
1	358,114	.084	0.0049	.027	0.0047
2	363,655	.083	0.0049	.024	0.0046
3	339,158	.064	0.0050	.010	0.0047
4	347,670	.034	0.0049	-.019	0.0047
5	365,505	.076	0.0049	.012	0.0046
6	336,816	.052	0.0050	-.010	0.0048
7	357,770	.024	0.0049	-.034	0.0046
8	353,796	.008	0.0048	-.053	0.0046
9	395,812	-.036	0.0046	-.082	0.0044
"Dollar" digit	**Raw Returns**			**Midpoint-based Size-adjusted Excess Returns**	
	Sample Size	Mean	Standard Error	Mean	Standard Error
0	345,415	.078	0.0049	.020	0.0047
1	331,586	.067	0.0050	.008	0.0048
2	333,216	.065	0.0049	.011	0.0046
3	339,846	.047	0.0048	-.008	0.0046
4	340,938	.020	0.0049	-.028	0.0047
5	375,381	.043	0.0049	-.017	0.0047
6	410,752	.047	0.0047	-.013	0.0045
7	414,126	.044	0.0046	-.018	0.0044
8	395,883	.039	0.0046	-.021	0.0044
9	374,193	.026	0.0048	-.029	0.0045
Full Sample	3,661,336	.047	0.0015	-0.0103	0.0015

TABLE II: Sawtooth model results

This table shows the coefficients from regressions of daily midpoint based, size-adjusted excess returns on previous day's closing price, including a log term, and a set of discrete variables based on the previous closing price, over the range of prices from 10 to 80 dollars. The coefficient on the "N" variables shows the magnitude of the "jump" in the sawtooth function at each class of round number. Also shown for comparison is the magnitude of excess clustering at each class of cluster point.

	Regression Results				Clustering Strength		
	Coefficient	t-statistic	Coefficient	t-statistic	Observed frequency	Expected frequency	Excess
"Jump" magnitudes [1]							
N5	.051	5.680	.055	5.980	0.1159	0.0800	44.9%
N10	.068	7.750	.071	7.920	0.1411	0.0800	76.4%
N25	.062	6.110	.067	6.500	0.0362	0.0200	81.1%
N50	.113	10.930	.118	11.130	0.0248	0.0100	148.0%
N100	.186	18.140	.191	18.130	0.0292	0.0080	264.5%
N500	.222	19.480	.227	19.540	0.0043	0.0010	325.1%
N1000	.242	21.010	.246	20.900	0.0040	0.0010	299.3%
Cluster point dummies [2]							
C5	-.021	-3.240	-.026	-3.890			
C10	-.024	-3.890	-.029	-4.560			
C25	-.012	-1.270	-.016	-1.660			
C50	-.051	-4.760	-.058	-5.280			
C100	-.108	-10.580	-.115	-10.960			
C500	-.097	-3.980	-.101	-4.040			
C1000	-.136	-6.120	-.140	-6.170			
Other [3]							
last closing price (LP)	-1.384	-8.690	-1.461	-8.960			
log-LP	-.007	-.660	-.003	-.260			
constant	.164	2.130	.123	1.560			
Controls(prev. day) [4]							
Market Return			1.665	11.030			
Stock Midpoint Return			.002	3.600			
Relative Volume			.064	28.490			
Relative Spread			.020	7.500			

[1] *Nk* variables (where k=5, 10, etc) contain a count of the number of "type k" cluster points at or below the prior closing price.

[2] *Ck* variables (where k=5, 10, etc) are indicator variables that are equal to one if the prior closing price is a "type k" cluster point.

[3] *LP* is the previous-day closing price, log-LP is the natural logarithm of LP.

[4] *Market Return* is the daily value weighted return for the entire market on Day *t-1*, *Stock Midpoint Return* is the return for that stock on Day *t-1* , *Relative Volume* is the log of the ratio of the trading volume at *t-1* to the median volume for the previous calendar month, and *Relative Spread* is the closing bid-ask spread at *t-1* divided by the price at *t-1* .

TABLE III. Differences in Midpoint-Based Returns By Last Digit(s)

This table shows the coefficients from regressions of daily returns on two indicator variables based on the last digit(s) of the closing price at the end of the previous day, day t-1. The results shown in the tables are the coefficients on the indicator variables and the t-statistics from the test of the null hypothesis that the coefficients are equal. Daily returns are calculated as ([bid-ask midpoint at t]-[bid-ask midpoint at t-1])/[bid-ask midpoint at t-1]), with the sample restricted to trading days with no distributions or share splits. The bid-ask midpoint at t-1 is calculated as ([bid price at t-1]+[ask price at t-1])/2.

		One-digit regression[4]			Two-digit regression[5]		
	Sample Size	lastdig=1 β1	lastdig=9 β2	t-test β1 = β2	last 2 dig 01 -09 β1	last 2 dig 91 - 99 β2	t-test β1 = β2
Full Sample	3,320,668	.028	-.041	9.82	.050	-.079	18.56
By Year							
2001	374,121	-.004	-.073	2.34	.071	-.104	6.35
2002	557,204	.013	-.058	3.26	.062	-.102	7.73
2003	519,555	.014	-.030	2.65	.065	-.080	9.04
2004	586,081	.044	-.044	5.86	.047	-.082	8.67
2005	620,736	.042	-.034	5.73	.032	-.066	7.34
2006	662,971	.037	-.025	4.89	.034	-.057	7.04
NYSE Size quintiles [1]							
Smallest	1,045,222	.045	-.055	6.67	.075	-.120	13.32
2nd	785,422	.021	-.035	3.66	.059	-.086	9.73
3rd	590,597	.024	-.048	4.84	.044	-.055	6.69
4th	479,988	.031	-.017	3.28	.009	-.036	3.08
Largest	419,439	.006	-.030	2.65	.022	-.042	4.77
Price Categories [2]							
price ≤ \$20	1,464,718	.035	-.051	6.62	.084	-.113	15.59
\$20 < price ≤ \$35	1,013,366	.022	-.036	5.23	.033	-.060	8.54
price > \$35	842,584	.025	-.028	5.63	.011	-.036	5.01
Dollar volume quintiles [3]							
Lowest	566,949	.033	-.061	4.89	.098	-.119	11.63
2nd	652,195	.034	-.048	4.98	.052	-.090	8.82
3rd	648,540	.042	-.033	4.83	.051	-.068	7.75
4th	644,703	.031	-.025	3.88	.036	-.059	6.60
Highest	618,675	-.003	-.024	1.50	.005	-.042	3.42
Exchange							
NYSE/AMEX	1,784,841	.032	-.042	9.60	.044	-.063	14.03
NASDQ	1,535,827	.026	-.037	5.15	.057	-.096	12.73

[1] *Size* is calculated as [closing price at t-1]*[shares outstanding at t-1] as reported by CRSP. Quintiles are formed monthly.

[2] *Price* categories are based on closing prices at t-1 as reported by CRSP.

[3] *Dollar volume* is calculated as [closing price at t-1] * [trading volume at t-1] as reported by CRSP.

[4] *lastdig* is the last digit of the closing price at t-1 as reported by CRSP. Prices are rounded to the nearest penny if necessary.

[5] *last 2 dig* is the last two digits of the closing price at t-1 as reported by CRSP.

TABLE IV. Differences in Intraday Returns By Last Digit(s)

This table shows the coefficients from regressions of intraday returns (calculated from 11am to close) on log of market capitalization at t-1, value weighted market return, and two dummies based on the last digit of the closing price at t-1. The results shown in the tables are the coefficients on the dummy variables and the t-statistic from the test of the null hypothesis that the coefficients are equal. Returns are calculated as ([closing price on day t]-[price at 11am on day t])/[price at 11am on day t]), with the sample restricted to trading days with no distributions or share splits.

		One-digit regression[4]			Two-digit regression[5]		
	Sample Size	lastdig=1 β1	lastdig=9 β2	t-test β1 = β2	last 2 dig 01 -09 β1	last 2 dig 91 - 99 β2	t-test β1 = β2
Full Sample	2,641,247	.003	-.015	3.05	.014	-.016	5.18
By Year							
2001	220,476	-.025	-.026	0.01	.020	-.008	0.96
2002	374,293	.027	-.037	2.76	.000	.001	0.04
2003	419,728	-.011	.000	0.58	.040	-.029	3.82
2004	495,959	.010	-.011	1.98	.015	-.020	3.33
2005	555,883	.014	-.001	1.70	.018	-.011	3.26
2006	574,908	.012	-.007	2.23	.004	-.017	2.46
NYSE Size quintiles [1]							
Smallest	567,437	.012	-.007	1.46	.022	-.037	4.58
2nd	638,020	.007	-.013	1.86	.026	-.025	4.89
3rd	536,711	.006	-.022	2.79	.018	-.010	2.94
4th	468,594	-.001	-.013	0.58	.007	.018	0.51
Largest	430,477	-.005	-.014	1.01	-.010	-.016	0.73
Price Categories [2]							
price ≤ $20	842,597	.003	-.012	1.46	.028	-.035	5.92
$20 < price ≤ $35	945,891	.005	-.019	3.24	.015	-.009	3.22
price > $35	852,759	.000	-.011	0.91	.001	-.002	0.31
Dollar volume quintiles [3]							
Lowest	276,245	.010	-.001	0.55	.054	-.054	5.61
2nd	482,596	-.003	-.008	0.34	.015	-.036	3.70
3rd	566,801	.020	-.016	2.29	.025	-.013	2.40
4th	598,550	.002	-.014	1.29	.007	-.006	0.97
Highest	636,068	.001	-.011	1.33	-.004	-.006	0.22
Exchange							
NYSE/AMEX	1,528,164	.005	-.016	2.51	.019	-.018	4.35
NASDQ	1,113,083	.002	-.009	1.20	.013	-.019	3.51

[1] *Size* is calculated as [closing price at t-1]*[shares outstanding at t-1] as reported by CRSP. Quintiles are formed monthly.

[2] *Price* categories are based on closing prices at t-1 as reported by CRSP.

[3] *Dollar volume* is calculated as [closing price at t-1] * [trading volume at t-1] as reported by CRSP.

[4] *lastdig* is the last digit in the closing price at t-1 as reported by CRSP.

[5] *last 2 dig* is the last two digits of the closing price at t-1 as reported by CRSP.

TABLE V. Controlling for Volume or Previous Returns

This table shows the coefficients from regressions of daily returns on two indicator variables based on the last digit(s) of the closing price at t-1. The results shown in the tables are the coefficients on the indicator variables and the t-statistics from the test of the null hypothesis that the coefficients are equal. Daily returns are calculated as ([bid-ask midpoint at t]-[bid-ask midpoint at t-1])/[bid-ask midpoint at t-1]), with the sample restricted to trading days with no distributions or share splits. The bid-ask midpoint at t-1 is calculated as ([bid price at t-1]+[ask price at t-1])/2. "Low" and "high" volume are defined by whether the volume exceeds the median share volume for that stock in the previous calendar month.

		One-digit regression[1]			Two-digit regression[2]		
	Sample Size	lastdig=1 β1	lastdig=9 β2	t-test β1 = β2	last 2 dig 01 -09 β1	last 2 dig 91 - 99 β2	t-test β1 = β2
Full Sample	3,320,668	.028	-.041	9.82	.050	-.079	18.56
Previous day volume							
low	1,701,176	.031	-.033	7.14	.054	-.074	14.43
high	1,614,704	.026	-.049	6.83	.045	-.084	12.04
Same day volume							
low	1,705,023	.033	-.043	11.56	.048	-.068	17.97
high	1,615,005	.025	-.037	4.90	.051	-.091	11.36
Previous day return							
negative	1,634,690	.036	-.039	7.37	.065	-.081	14.52
positive	1,685,807	.021	-.042	6.49	.035	-.077	11.70

[1] *lastdig* is the last digit in the closing price at *t-1* as reported by CRSP.

[2] *last 2 dig* is the last two digits in the closing price at *t-1* as reported by CRSP.

TABLE VI. Brokerage Account Trading

This table shows the coefficients from regressions of brokerage account trading variables on two indicator variables based on the last digit(s) of the closing price at t-1. The results shown in the tables are the coefficients on the indicator variables and the t-statistics from the test of the null hypothesis that the coefficients are equal. Daily returns are calculated as ([bid-ask midpoint at t]-[bid-ask midpoint at t-1])/[bid-ask midpoint at t-1]), with the sample restricted to trading days with no distributions or share splits. The bid-ask midpoint at t-1 is calculated as ([bid price at t-1]+[ask price at t-1])/2. The sample period is 1991-1996, during which time stock prices were quoted in fractions, primarily 1/8 increments, but 1/16 on certain lower-priced stocks. "frac" is the fractional component of the closing price reported by CRSP. The "Return control regression" column presents results for the estimation of a regression where an additional control for current-day returns is included.

		Base regression			Return control regression		
	Sample Size	frac= 1/16 or 1/8 β1	frac = 7/8 or 15/16 β2	t-test β1 = β2	frac = 1/16 or 1/8 β1	frac = 7/8 or 15/16 β2	t-test β1 = β2
Returns, Full Sample	3,402,881	.009	-.027	4.92			
Trade Imbalance[1]	716,014	-.017	-.033	3.35	-.024	-.042	3.25
Dollar Imbalance[2]	716,014	-.017	-.030	2.74	-.024	-.039	2.81
Trade Imbalance, Mean 0, Sd 1[3]	715,999	-.036	-.066	2.48	-.052	-.083	2.13
Dollar Imbalance, Mean 0, Sd 1[3]	716,008	-.036	-.061	2.07	-.051	-.078	1.87

[1] *Trade Imbalance* is defined as (# buy trades - # sell trades)/(# buy trades + # buy trades), where there is at least one trade in the database.

[2] *Dollar Imbalance* is defined as ($ bought - $ sold)/($ bought + $ sold), where there is at least one trade in the database.

[3] *Trade Imbalance and Dollar Imbalance "Mean 0, Sd 1"* variables are calculated by subtracting the firm-year mean of the given imbalance variable and normalizing by the firm-year standard deviation. Observations for which firm-year standard deviation is 0 are removed.

TABLE VII. Institutional Ownership

This table shows the coefficients from regressions of daily returns on two indicator variables based on the last digit(s) of the closing price at t-1. The results shown in the tables are the coefficients on the indicator variables and the t-statistics from the test of the null hypothesis that the coefficients are equal. Daily returns are calculated as ([bid-ask midpoint at t]-[bid-ask midpoint at t-1])/[bid-ask midpoint at t-1]), with the sample restricted to trading days with no distributions or share splits. The bid-ask midpoint at t-1 is calculated as ([bid price at t-1]+[ask price at t-1])/2. "Percentage Institutions Among Active Traders" is defined as the percentage of remaining shares owned by institutions, after removing shares held by institutions that are categorized as quasi-index or buy-and-hold.

			One-digit regression[1]			Two-digit regression[2]		
	Mean Percentage	Sample Size	lastdig=1 β1	lastdig=9 β2	t-test β1 = β2	last 2 dig 01 -09 β1	last 2 dig 91 - 99 β2	t-test β1 = β2
Full Sample		3,320,668	.028	-.041	9.82	.050	-.079	18.56
Percentage Institutional Ownership								
Quintile 1	25.6%	633,474	.014	-.040	2.76	.071	-.098	8.80
Quintile 2	50.1%	633,448	.028	-.060	5.38	.046	-.095	8.75
Quintile 3	65.3%	633,447	.033	-.033	4.25	.033	-.069	6.65
Quintile 4	77.5%	633,484	.041	-.048	5.97	.039	-.077	7.92
Quintile 5	90.0%	633,409	.017	-.029	3.32	.059	-.067	9.09
Percentage Institutions Among Active Traders								
Quintile 1	7.1%	629,649	.011	-.028	2.22	.047	-.079	7.25
Quintile 2	20.2%	629,620	.036	-.056	5.81	.048	-.086	8.58
Quintile 3	33.4%	629,567	.037	-.038	4.75	.056	-.080	8.74
Quintile 4	48.5%	629,608	.032	-.050	5.19	.034	-.081	7.43
Quintile 5	70.5%	629,608	.019	-.036	3.45	.064	-.081	9.16

[1] *lastdig* is the last digit in the closing price at *t-1* as reported by CRSP.

[2] *last 2 dig* is the last two digits in the closing price at *t-1 as* reported by CRSP.

TABLE VIII. New Highs and Lows

This table shows the coefficients from regressions of daily returns on a collection of indicator variables and interactions. The "lastdig" indicator variables are based on the last digit(s) of the closing price at t-1, while the "new" indicator variables are based upon prices over the preceding two trading days. New high (low) is an indicator variable equal to one if the closing price at t-1 is above (below) the highest (lowest) trading price during the previous two days. Daily returns are calculated as ([bid-ask midpoint at t]-[bid-ask midpoint at t-1])/[bid-ask midpoint at t-1]), with the sample restricted to trading days with no distributions or share splits. The bid-ask midpoint at t-1 is calculated as ([bid price at t-1]+[ask price at t-1])/2. Panel A shows results for estimation of regressions based upon lastdig, the last digit in the closing price on t-1 as reported by CRSP. Panel B shows results for estimation of regressions based upon last 2 dig, the last two digits in the closing price on t-1 as reported by CRSP.

Panel A. Returns relating to the last digit of closing price

		Basic Model		Intermediate Model		Full model	
	Number[1]	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
lastdig=1	312,377	0.03	5.35	0.03	5.37	0.03	3.92
lastdig=9	317,218	-0.04	-7.71	-0.04	-7.72	-0.03	-4.58
new high[3]	702,651			0.02	6.59	0.03	6.96
new low[4]	633,331			0.00	-0.68	0.00	-0.72
lastdig=1 * new high	58,819					-0.01	-0.86
lastdig=1 * new low	55,779					0.03	1.85
lastdig=9 * new high	61,771					-0.03	-2.55
lastdig=9 * new low	53,890					-0.02	-1.52

Panel B. Returns relating to the last two digits of closing price

		Basic Model		Intermediate Model		Full model	
	Number[1]	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
last 2 dig 01-09	315,420	0.05	9.5	0.05	9.54	0.04	6.88
last 2 dig 91-99	328,687	-0.08	-15.35	-0.08	-15.33	-0.08	-12.83
new high[3]	702,651			0.02	6.56	0.02	5.82
new low[4]	633,331			0.00	-0.71	-0.01	-1.47
last 2 dig 01-09 * new high	57,896					-0.01	-0.36
last 2 dig 01-09 * new low	53,773					0.04	2.49
last 2 dig 91-99 * new high	61,627					0.01	0.86
last 2 dig 91-99 * new low	54,782					0.01	0.46

Table IX. Microstructure Trends

This table shows the percentage of daily closing stock prices ending in the indicated digit(s) each year and the average closing bid-ask spread by year. Bid-ask spread is normalized by dividing by previous closing price.

		Clustering in closing prices				Closing bid-ask spread	
		One-digit clustering[1]		Two-digit clustering[2]			
Year	Sample Size	lastdig=0	lastdig=5	last 2 dig = 00	last 2 dig = 50	Mean	Standard Deviation
2001	398,823	26.5%	18.9%	5.1%	3.3%	0.818	1.634
2002	582,193	23.9%	17.3%	4.5%	2.9%	0.746	0.929
2003	586,352	20.5%	15.5%	3.6%	2.5%	0.368	0.600
2004	686,476	18.7%	14.3%	3.3%	2.2%	0.235	12.262
2005	705,520	16.9%	13.3%	2.9%	2.0%	0.197	0.320
2006	726,608	15.4%	12.4%	2.5%	1.8%	0.165	0.257
Total	3685972	19.7%	14.9%	3.5%	2.4%	0.379	5.346

[1] *lastdig* is the last digit in the closing price reported by CRSP

[2] *last 2 dig* is the last two digits in the closing price reported by CRSP