# 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

**Abstract**

We document consistent differences in daily stock returns related to round numbers. In particular, we examine returns following previous day closing prices that are just above or just below round number benchmarks. We find, for both 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 a round number benchmark. For example, at the one-digit level we find that returns following "9-ending" prices, which are just below a round number, such as $25.49, are significantly lower than returns following "1-ending" prices such as $25.51, just above the round number. Our results are based on midpoint-based returns, controlling for bid/ask bounce, and are robust for a wide collection of subsamples based on year, firm size, trading volume and exchange. While the magnitude of the return difference varies depending on the type of round number examined, or the particular subsample used, 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 the Third Annual Early Career Women in Finance Mini-Conference for helpful discussions and feedback.

## I. Introduction

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.

This phenomenon appears to be related to the well known tendency of prices and orders to cluster at round numbers, though the return pattern is not driven by clustering itself. Since decimalization, these cluster points are prices that end in one or more zeros, or, to a lesser extent, those that end in 5 or 25. The clustering is stronger for the "rounder" prices (such as $25.00), than for prices that are "less round" (such as $26.25). Likewise, we find that the return differentials 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. While the return patterns we document appear to be related to clustering, they are do not appear to be a simple consequence of clustering---most models of high-frequency price dynamics that include clustering will not generate the patterns we document.

Our investigation begins with the observation that U.S. daily stock returns differ based on the last digit of the previous closing price. Figure 1 demonstrates the average returns since decimalization occurred in 2001. Strikingly, we find that similar (and even stronger) patterns are observed at other scale levels.

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, in particular whether those digits place price just above or just below a round number. For example, do returns differ for a stock with a closing price just above or below $24.10, at $24.11 or $24.09, or do returns differ for a stock with a closing price just above or below $25.00, at $24.99 or $25.01? Our results show that returns are significantly related to closing prices. Returns are significantly higher following a closing price just above a round number, such as $24.11, than for a closing price just below a round number, such as $24.09.

To give an idea of the magnitude of the returns differences we find, let us focus on returns following prices just above or below a round dollar barrier. 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 would work out to an annualized rate of over 38% per year. While it is most likely impossible to exclusively trade on this pattern and make a profit, given the almost-daily rebalancing required, these patterns may have a significantly positive impact on returns if exploited for market timing purposes.

These differences are significantly weaker for the largest stocks, but are strong for medium size and small stocks. The phenomenon has weakened only moderately since 2001, despite significant changes in market microstructure. In particular, there have been 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. Our cross-sectional analyses indicate that the pattern is not driven by firm size or liquidity, and is robust in subsets based on institutional ownership. We hope to explore alternative potential explanations for the return pattern in future work, however our work to date has shown that many microstructure-based explanations are insufficient to explain the return pattern.

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 find them unable to describe the patterns we describe. Section 6 concludes.

## II. Background and Prior Literature

Prior work on round numbers has focused on "cognitive reference points" (Rosch (1975)). 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[1] (Schindler and Wiman, 1989). These results indicate that 5- and 0-ending numbers may be more cognitively accessible than other numbers and could represent reference points. The effects of such 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[2] and several explanations based on consumers' cognitive biases have been proposed to explain this finding. Stiving and Winer (1997) and Bizer and Schindler (2005) find evidence of a 'truncation effect' in which consumers tend to put more emphasis on the left-hand digits of numbers when comparing two prices. Schindler and Kirby (1997) argue that consumers may be subject to a 'perceived gain effect' in which 9-ending (1-ending) prices are perceived as a discount (premium) relative to zero-ending prices because zero-ending numbers tend to be cognitive reference points. Finally, there is some evidence that consumers are subject to an 'image effect,' having been conditioned to view nine-ending prices as a signal of a sale or bargain and zero-ending prices as a signal of quality. (See Bizer and Schindler (2005), Stiving and Winer (1997), and Schindler and Kirby

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

[2] 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.

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

In the accounting literature, Herrmann and Thomas (2005) 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[3]. 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.

Finally, in the finance literature, researchers have documented price clustering on round numbers in many settings. Price clustering occurs in US equity markets (Osborne (1962), Nierderhoffer (1965 and 1966), Harris (1991), Christie and Schult (1994), Grossman et. al. (1997), Cooney, VanNess and VanNess (2000)), foreign equity markets (Grossman et. al. (1997), Aitken et. al.), in Israeli IPO auctions (Kanhel, Sarig and Wohl (2001)), US SEOs (Corwin (2003)), currency markets (Goodhart and Curcio (1991), Osler (2003)) and a variety of other settings.

Yet prior literature has not examined whether round numbers or ending digits in stock prices have any effect on consumer (investor) behavior. This may partially be due to the lack of a mapping between numbers that individuals see in most settings, typically in decimal form, and the way prices had been quoted on U.S. exchanges historically. Before 2001, most exchanges quoted prices in eighths or sixteenths, i.e. 14 3/8 instead of $14.375. One sixteenth is 6.25 cents, and does not map to prior work on individual responses to number characteristics. We exploit the move to decimalization to test whether the market responds differently to a stock depending on the characteristics of its closing price.

---

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

Our paper also relates to a large literature in behavioral finance on the affects of non-standard preferences, psychological biases or 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) find that overall market sentiment predicts returns based on a large set of firm characteristics for a broad set of stocks; however sentiment has almost no effect on the returns of the largest stocks. In a related paper, Kumar and Lee (2006) show that retail investor sentiment has a significant impact on the returns of firms, but primarily those with higher arbitrage costs and/or higher concentrations of retail investors. Both of these papers show the significant impact of investor psychology on returns (and suggest that better understanding this psychology is important for understanding the cross-section of returns), but also show the most significant impacts for firms with higher arbitrage costs and less sophisticated investor bases. However, because the use of cognitive reference points is ingrained[4], it is quite likely that even the most sophisticated investors will unconsciously rely on these reference points whenever making subjective judgments. The effect of round numbers is likely to persist unless investors make an explicit effort to overcome their use of round-number reference points. 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.

Finally, while returns patterns based on number perceptions have not been explored, there is a large literature in finance on the phenomenon of price clustering in financial markets. In U.S. equity markets, previous to decimalization, prices were most 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. In particular, transaction prices

[4] 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.

tend to cluster at “nickels” and even more so at “dimes”. There is even stronger clustering at the two-digit level, at “dollar” or “50-cent” numbers. While price clustering may arise due to market microstructure reasons, such as reducing price negotiation costs, we control for the microstructure effects of clustering, rather than focusing on clustering itself. The Appendix discusses in more detail the prior price clustering literature, and how price clustering may affect raw security returns.

## III. Data

Our primary data sample is limited to common stock for US firms, trading on the New York Stock Exchange, American Stock Exchange 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. 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 removed stocks that had a median trading volume of less than 20,000 shares over the previous calendar month, or those with 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 bid and ask quotes at the closing, 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 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.

## 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 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, in percentage form, is shown 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.

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 that appears to be 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 round numbers and clustering. 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 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 any 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 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 right panel of table II also 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 this 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 statitstical 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

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

where $\text{Return}_{i,t}$ is the midpoint-based daily percent return for stock *i*, $\text{LD1}_{i,t}$ is an indicator variable that equals one if the last digit of the closing price at time *t-1* is 1, and $\text{LD9}_{i,t}$ 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 $\beta_1$ and $\beta_2$ 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 $\beta_1$ and $\beta_2$ 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:

$$\text{Return}_{i,t} = \alpha + \beta_1 \text{ LD2LOW}_{i,t-1} + \beta_2 \text{ LD2HI}_{i,t-1} + \varepsilon_{it}, \quad (2)$$

where $LD2LOW_{i,t,}$ and $LD2HI_{i,t}$ 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 (1) and (2). The columns in the table under the heading "one-digit regression" show the estimated values of $\beta_1$ and $\beta_2$, 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 other rows in the table show the results of estimating Equations (1) and (2) 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; every estimate but one for $\beta_1$ is positive, every estimate of $\beta_2$ is negative, and the t-statistics for the difference between the two are statistically significant in every 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 equal sized.) 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 trading volume quintiles, the lowest trading volume group shows the largest return difference for both regressions. The stocks with the highest trading volume show the smallest effect, but the differences between $\beta_1$ and $\beta_2$ are still significant in both regressions.

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

**Robustness.**

The general patterns described above appear quite robust. The pattern appears to hold in many subsamples, as shown in Table III. We also find that the results are robust to including additional or alternative control variables, including previous market return, previous stock-specific return, spread, price, etc. None of these control variables changes the basic results qualitatively. This is not surprising, since last digit is only very weakly correlated with other variables. The results also appear robust to using various other return variables, including those based on transaction prices as well as midpoints, and including various ways of controlling for size and market effects. All of these measures produced similar qualitative patterns of returns.

**Patterns Over Time**

In this section we present some results on the pattern of returns over time. One question we wish to address is whether the pattern of return differences is due to price changes from the close to the following day's open (or soon thereafter), or whether the pattern persists for a longer time.

We begin examining returns between 11AM and closing on the current day, conditional on the last digit(s) of the previous reported close.[5] 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 open. 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.

The returns in this table are not 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 last digits. 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 return.

Once again we find that the familiar pattern of returns holds for these intra-day returns, although with smaller magnitudes. Table IV shows the regression results for the

[5] The choice of 11AM is somewhat arbitrary, but we found similar results using 10AM or 12PM instead.

11AM-to-close return, for the entire sample and for sub-samples, analogous to Table III. The first line shows the estimates for the whole sample. Once again, the estimate of $\beta_1$ is positive (although small), and the estimate for $\beta_2$ is negative. We again reject the hypothesis that the coefficients are equal with a high degree of confidence (t=3.52). For the two digit regression, the estimated differences are still larger and more statistically significant than for 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 ($\beta_1$ - $\beta_2$) has the expected positive sign in all sub-samples with the exception of 2001 and 2003 in the single-digit regressions, and the 4$^{th}$ size quintile and 5$^{th}$ volume quartile 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.

Next, we looked at the pattern of returns over several days before and after the day the conditioning "last digits" were recorded. Figure 5 shows the pattern of returns over a range of 8 days, numbered from 0 to 7, conditional on the closing price on "day 3." The returns are calculated for four groups of stocks, based on the last two digits of the closing price on day 3. The groups are those with prices ending in 01-19, 20-39, 60-79, and 80-99.

The large up and down "spikes" on day 4 represent the differential returns already extensively documented above. As expected, they are unusually high for the 01-19 group, somewhat high for the 20-39 group, somewhat low for the 60-79 group, and much lower for the 80-99 group.

What the figure also shows is 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 "3," the day that the conditioning price is recorded. It would seem that the large differences on day "4" are at least in part a reversal of the differences opposite differences on day "3," and to a lesser extent on days before day 3. We do not have a good explanation for what causes these differences.

The figure also shows that the return differences continue for several days afterwards, fading out only slowly over time. Statistical tests show that these differences remain significant 3 days after the conditioning closing price. This reinforces the results on intra-day returns, showing that the return differences persist over time, rather than occurring in the trades immediately after the previous close.

## V. Considering Possible Explanations

Unfortunately, many of the possible explanations for the return pattern we document are poorly specified. For example, we described in Section II a large set of prior literature in psychology and marketing showing that individuals use round numbers as cognitive reference points, which play a vital role in number perceptions. However, how the use of cognitive reference points will translate into valuation or behavior can vary from situation to situation. For example, consumers may see $1.50 as a discount from the reference point of $2.00 at a store, or may see $1.50 as a failure to bring a price down to $1.00 in a bargaining situation. The particular reference point an individual refers to may vary depending on details of the situation. Similarly, it is not completely clear how behavioral tendencies would translate into returns patterns. However, many of the alternate explanations for the pattern we document do have clearly testable predictions, and so our approach is to examine alternate explanations.

### 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. First, the return patterns we document are based on "midpoint returns" thereby controlling for bid-ask bounce effects. We further examine intra-day returns to remove any affect of overnight or previous day orders which are handled during the opening auction. Second, even more strongly than excluding just the opening two hours of trade, the differences in returns continue into the following few days. Third, the patterns are robust in subsamples of more liquid firms, such as large firms and firms with high trading volume (see Table III). Finally, 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 and Nasdaq, 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 prevelant, 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. Table V displays statistics for the level of quote clustering and closing bid-ask spreads over our sample period, for the years 2001 through 2006, and each decrease dramatically. Over this same period of time, the return difference we document has been remarkably stable. While the difference in returns for the last two digits of closing price has decreased somewhat over the years, it remained at roughly 25% annualized for 2006, while the returns differences based on the last one digit of closing price have not diminished over the years.

**Clustering**

One possible explanation is that our patterns are a result of clustering in limit orders. However, none of the prior explanations for clustering have had explicit returns predictions, and to our knowledge no analytical models of clustering have returns predictions. As such, we take two approaches to test whether clustering may explain our returns patterns. First, we run a large set of simulations with clustered asset prices, to determine if they produce similar returns patterns. Second, we more qualitatively compare the decrease in clustering over time with the return patterns we document.

The most obvious mechanism for both clustering and a returns pattern, intuitively, is that investors round their orders – i.e. they tend to place orders that are round numbers. For example, if investors round their orders, then prices just above a cluster point could be "rounded down" prices while prices just below a cluster point could be "rounded up" prices. We created two main simulations to see if 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 differential returns for the next day, related to the closing prices around round numbers, in any variations of this test.

Our second model is based on a 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 cluster points until the stale orders are filled. A stock price ending in 1 could move up four cents before getting "stuck" at a 5-ending round number or move down one cent before encountering resistance at a zero-ending round number. This phenomenon might result in positive average returns since the stock price can move up more freely than it can move down. A stock price that ends in 9 would experience the opposite effect – it could move down to the next 5-ending price more freely than it could move up past the adjacent 0-ending price, 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 either.

More qualitatively, Table V displays the decrease in clustering over time. In 2006, the frequency of each of the cluster points listed had decreased to roughly half of it's 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.

**Technical Analysis**

Another possible explanation for this pattern is that it is essentially a self-fulfilling prophecy driven by technical analysis. Technical analysis is the practice of trading stocks based on price patterns (rather than company fundamentals or valuation alone). Technical analysts often speak of “resistance” and “support” levels, which are most often round numbers. The idea is that a stock’s price has natural resistance and support levels which it will stay within. If price increases to just under a resistance level, it is likely to go down. If, however, price breaks upwards through a resistance level, then technical analysts expect that it will continue to go upwards, as the stock now has a new resistance level. It is difficult, if not impossible, to create explicit tests for this within the universe of common stocks, since 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. We have attempted to construct tests based on prior price movements – for example, are the higher returns after a $25.01 price due to stocks that have just come up through a resistance level? We have failed to find any strong patterns supporting the technical analysis story, however that is clearly not conclusive.

We take a different approach to testing the possibility that technical analysts drive the round-number return pattern. Technical analysis does not apply for mutual funds, since the value of mutual fund assets is based on the securities held by the fund, and not the fund’s share price. In addition, the constraints and costs that many mutual funds place on traders makes them an expensive security to trade for technical analysts. However, mutual funds do report a “net asset value” (NAV) at the end of each day. An investor looking to invest in a mutual fund will see NAV listed along with fund name and other characteristics, whether looking in the newspaper, online, or at a financial adviser’s office. If the pattern we document is due to investor reactions other than technical analysts, then the flow of money into the mutual fund (fund flow), capturing how many investors are deciding to put new money into the fund on a given day, may depend on closing net asset values for the fund. While this analysis is still in preliminary stages, we find using monthly fund flow data that there is a strong and significant fund flow pattern based on the last three digits of NAV at the end of the prior month, analogous to the returns pattern we document. We are currently working to obtain daily fund flow data to further strengthen this test, and plan to include details in future versions of this working paper.

**Psychological/Behavioral Explanations**

While there are a whole host of psychological and behavioral explanations for the return pattern we document, based on the use of cognitive reference points in valuing a stock or responding to a stock price, we attempt to test one particular behavioral explanation here. One idea put forth in the marketing literature (Bizer and Schindler (2005)) is that investors simply "truncate" numbers, ignoring the least significant digits. If this were the mechanism driving our results, we would expect to see similar patterns following prices *at* the round numbers as *just above* the round numbers. For example, the return following a price of 30.00 would be similar to that following a price of 30.01 since they would both be truncated to 30.00, but the return for 30.00 would differ significantly from the return following a price of 29.99, since they would be truncated to prices a full dollar apart. When we examine the returns at the round numbers specifically, we find that in fact the returns at the round numbers appear to be around average between the returns for the prices just above the given benchmark and the prices just below. This can be seen in the results of the sawtooth regression in Table II, and it has been verified in other untabulated regressions. This suggests that while the use of cognitive reference points may still be driving our results, *truncation*, per se, is not.

## VI. Conclusion

Individuals tend to use round numbers as cognitive reference points, as documented in psychology research. This affects their perceptions and reactions in a variety of settings, most notably in consumption decisions, when faced with prices. Although these number perceptions, differing from simply the true value of the number, affect individuals in a variety of settings, we might reasonably doubt that these same perceptions would have an impact on equity markets, particularly in aggregate. While prior literature has documented various ways in which individual investors trade sub-optimally, it would be 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, prior work in finance has shown price clustering in a variety of market settings. Clustering may or may not be due to cognitive biases in market participants' perception of, and preference for, certain numbers. Ikenberry and Weston (2003) provide evidence that existing explanations for clustering with fully rational agents, such as price clustering as a mechanism for reducing negotiation costs, fail to explain the magnitude of clustering after decimalization. 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.

Our results suggest, even more strongly than prior work on price clustering, that these number perceptions affect pricing in equity markets. We exploit the move to decimalization to conduct a straightforward test of the impact of number perceptions on returns and find a strong consistent return pattern. Returns following prices just above a round-number price are significantly higher than for other prices, while returns just below a round-number price are significantly lower. This pattern persists using midpoint-based returns to control for bid-ask bounce and for a wide collection of subsamples, for the largest stocks, stocks with the highest trade volumes and stocks with the lowest spreads.

While we cannot conclusively attribute these returns to any particular explanation, our results cast doubt on several explanations, such as market microstructure drivers, clustering, and technical analysis feedback trading. Market participants may prefer round number prices for a variety of reasons, but even beyond their preference for round number prices, we find a significant return response to prices just above and below the round numbers.

**References**

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

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.

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.

Fama, Eugene and Kenneth R. French, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics* 33, (1993), 3-56.

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.

Lease, Ronald C., Ronald W. Masulis, and John R. Page, “An Investigation of Market Microstructure Impacts on Event Study Returns,” *The Journal of Finance*, 46:4, (1991), 1523-1536.

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 occurence 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 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.”[6] 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.

---

[6] 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 1a: Midpoint-based excess returns by last digit of previous-day closing price**.

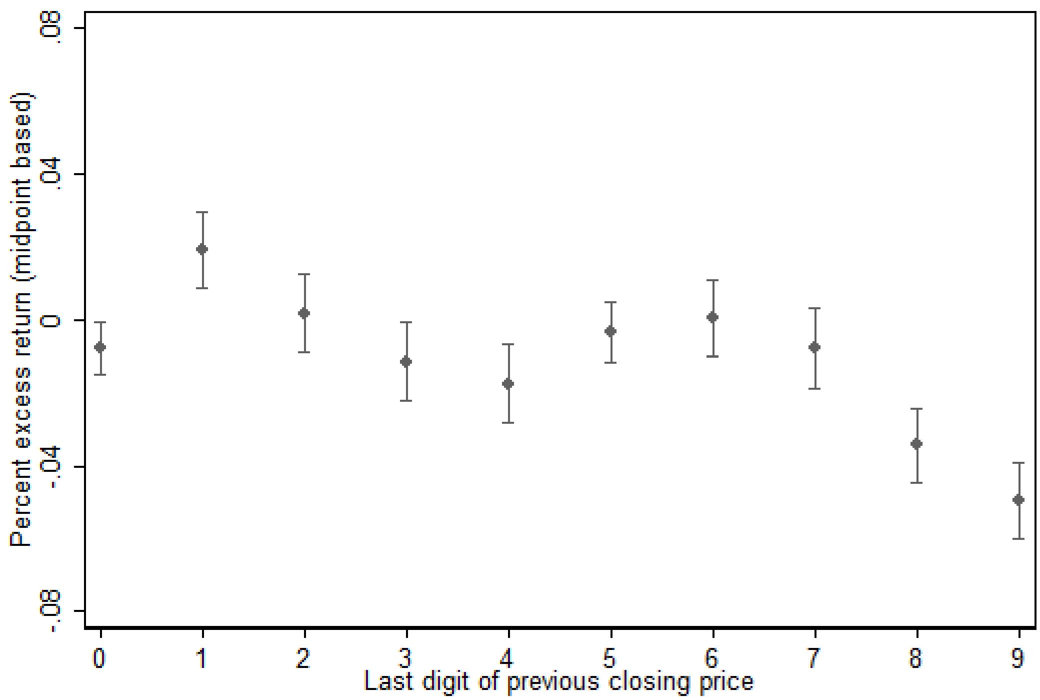



**Figure 1b: Midpoint-based excess returns by "dime" digit of previous closing price**.

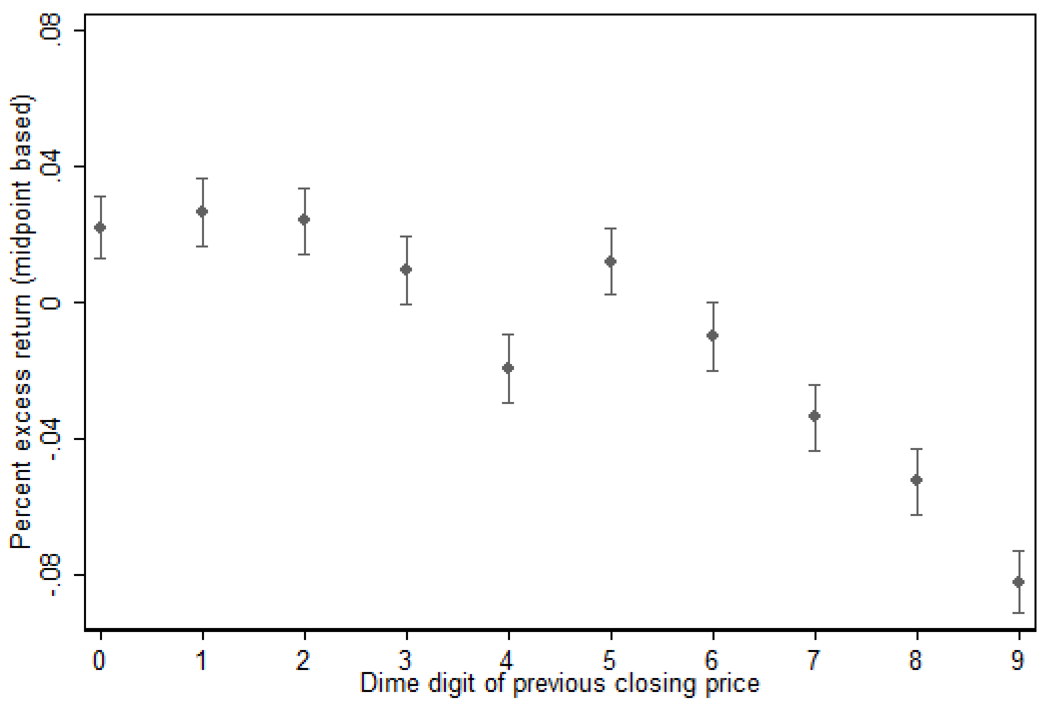


**Figure 1c: Midpoint-based excess returns by "dollar" digit of previous closing price**.

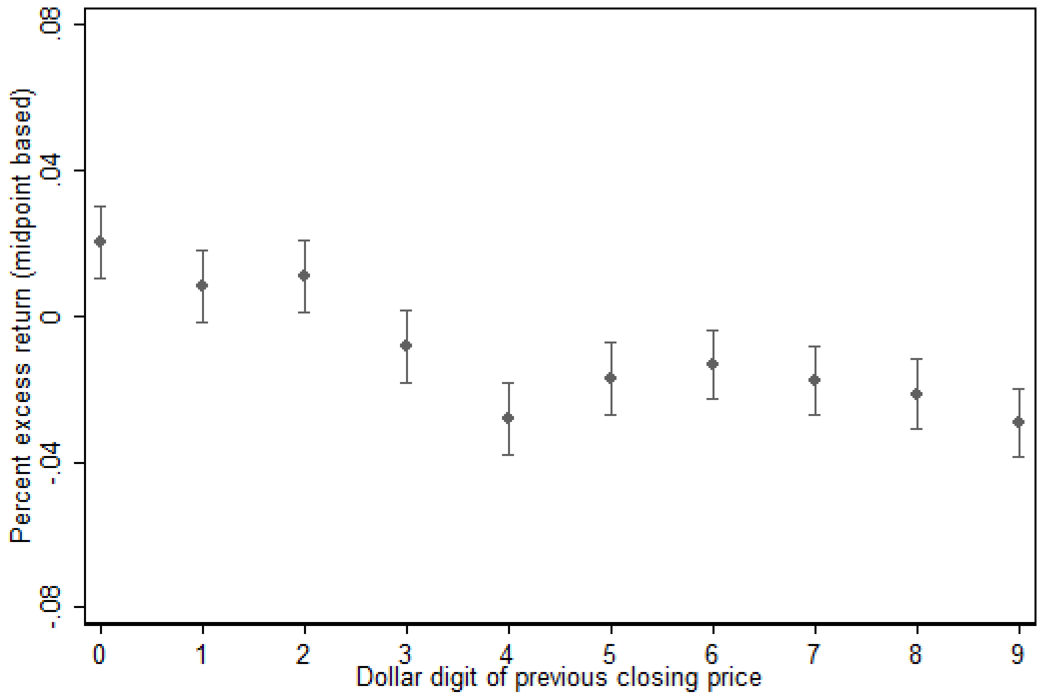



**Figure 2: Midpoint-based excess returns by last two digits of previous-day closing price**.

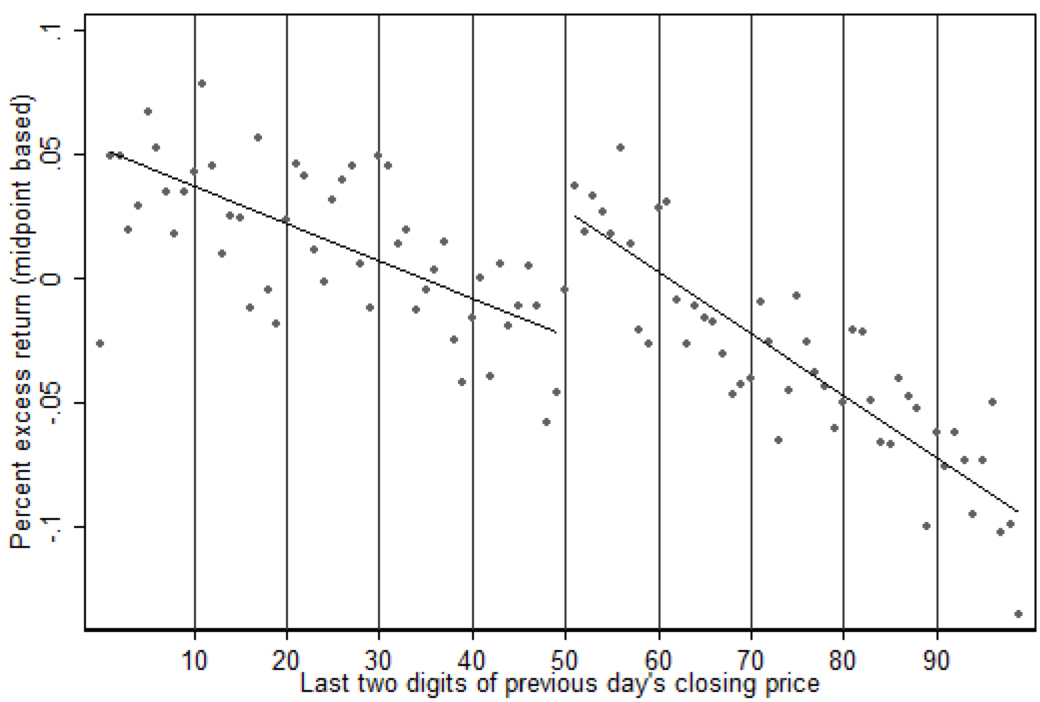


**Figure 3: Midpoint-based excess returns by previous day closing price rounded to the nearest 10 cents.**

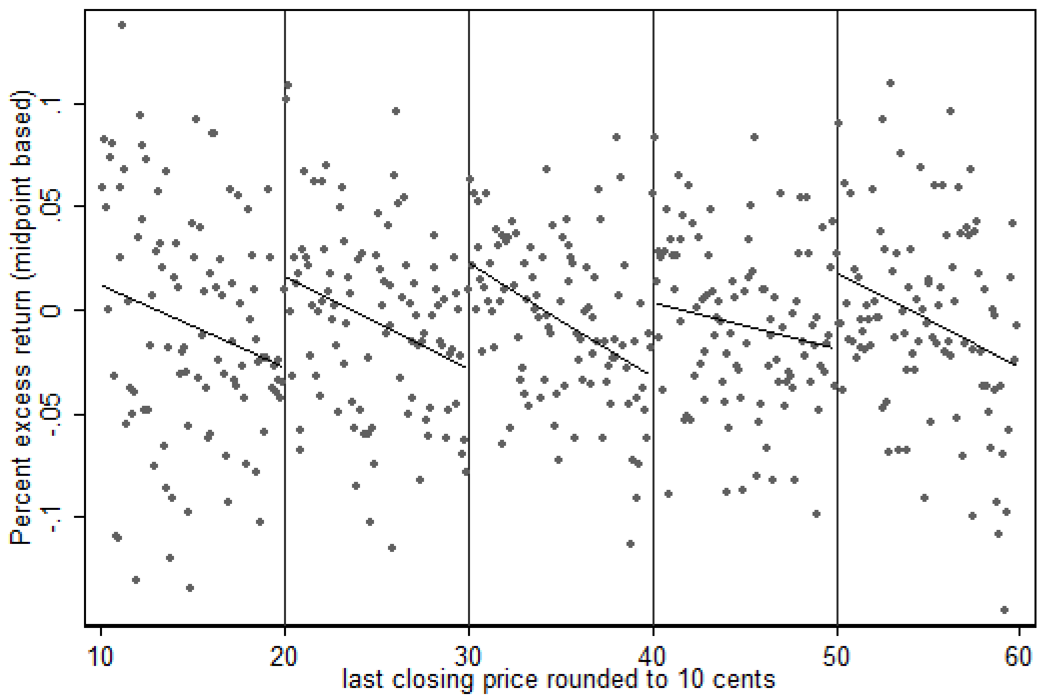



**Figure 4: Fitted values from sawtooth model, for prices between $15 and $25**

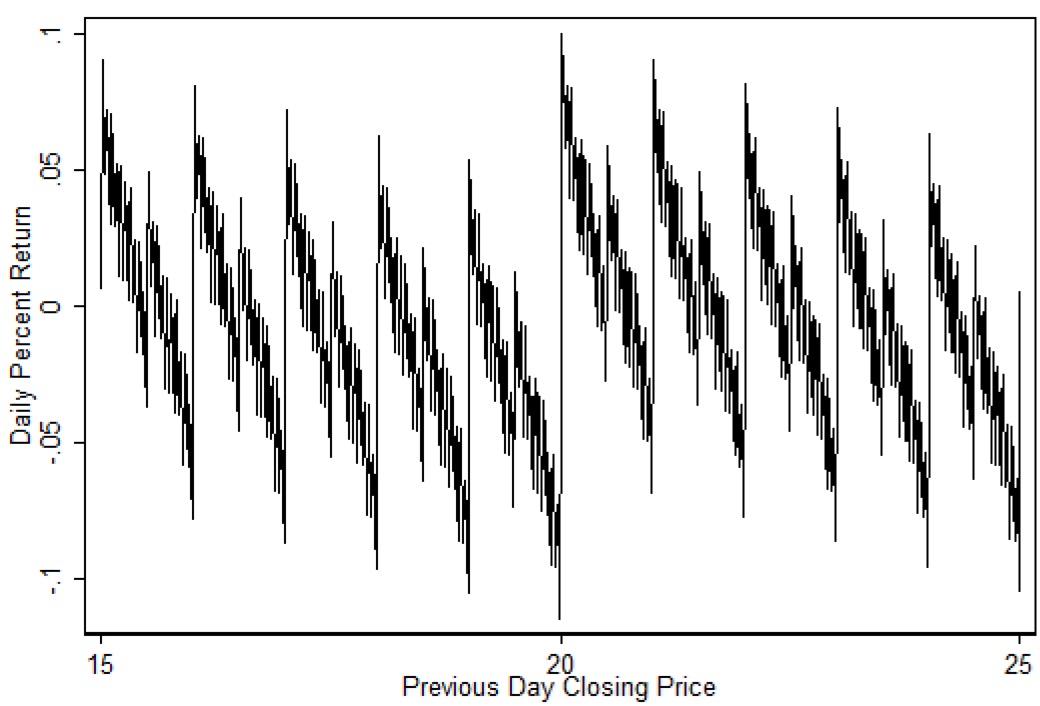
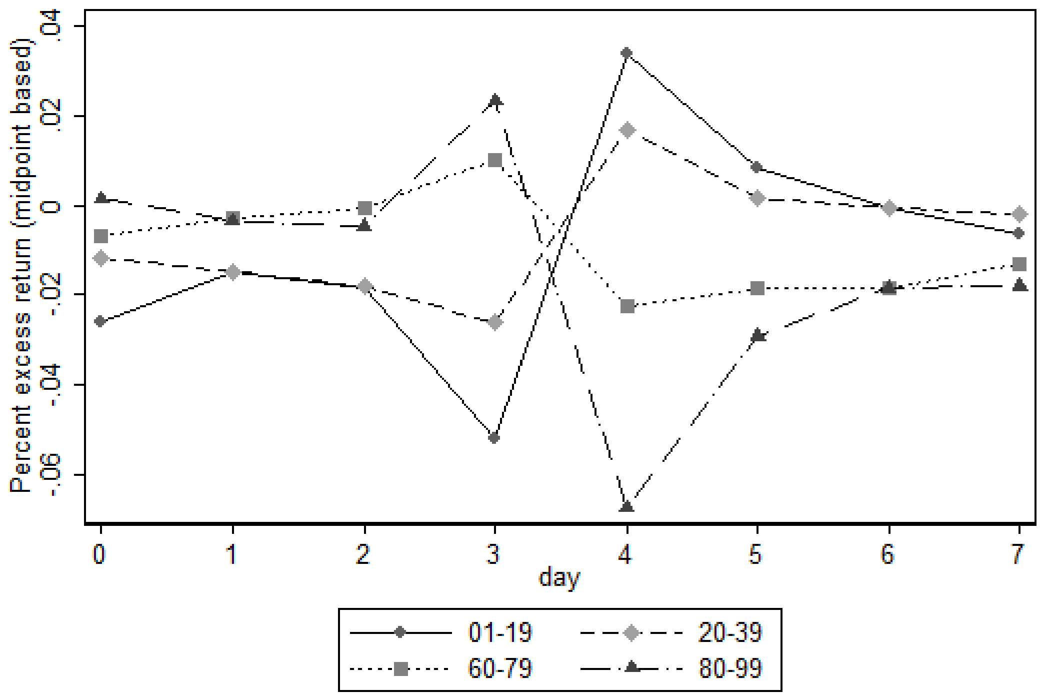


**Figure 5: Returns over time, by last digits of closing price on day 3**



## Table I. Returns by Last Digits

This table shows daily raw returns and daily midpoint-based size-adjusted returns grouped by the last digit of the previous day's closing stock price. SE is the standard error of the mean.

| "Penny" digit | Raw Returns | | | Midpoint-based Size-adjusted Excess Returns | |
|---|---|---|---|---|---|
| | N | Mean | SE | Mean | SE |
| 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 | |
| | N | Mean | SE | Mean | SE |
| 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 | |
| | N | Mean | SE | Mean | SE |
| 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 |

## <u>TABLE II: Sawtooth model results</u>

This table shows the coefficients from regressions of daily midpoint based, size-adjusted excess returns on last closing price, including a log term, and a set of discrete variables based on the previous closing price, as described in the text, over the range of prices from 10 to 80. 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 | Observed frequency | Expected frequency | Excess |
| *"Jump" magnitudes* | | | | | |
| N5 | .051 | 5.680 | 0.1159 | 0.0800 | 44.9% |
| N10 | .068 | 7.750 | 0.1411 | 0.0800 | 76.4% |
| N25 | .062 | 6.110 | 0.0362 | 0.0200 | 81.1% |
| N50 | .113 | 10.930 | 0.0248 | 0.0100 | 148.0% |
| N100 | .186 | 18.140 | 0.0292 | 0.0080 | 264.5% |
| N500 | .222 | 19.480 | 0.0043 | 0.0010 | 325.1% |
| N1000 | .242 | 21.010 | 0.0040 | 0.0010 | 299.3% |
| *Cluster point dummies* | | | | | |
| C5 | -.021 | -3.240 | | | |
| C10 | -.024 | -3.890 | | | |
| C25 | -.012 | -1.270 | | | |
| C50 | -.051 | -4.760 | | | |
| C100 | -.108 | -10.580 | | | |
| C500 | -.097 | -3.980 | | | |
| C1000 | -.136 | -6.120 | | | |
| *Other* | | | | | |
| last closing price (LP) | -1.384 | -8.690 | | | |
| log-LP | -.007 | -.660 | | | |
| constant | .164 | 2.130 | | | |

## 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 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]). 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 $\beta1$ | lastdig=9 $\beta2$ | t-test $\beta1 = \beta2$ | last 2 dig 01 -09 $\beta1$ | last 2 dig 91 - 99 $\beta2$ | t-test $\beta1 = \beta2$ |
| 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 | 695,185 | .039 | -.057 | 5.55 | .090 | -.110 | 11.94 |
| 2nd | 675,902 | .034 | -.046 | 4.92 | .044 | -.105 | 9.37 |
| 3rd | 663,843 | .025 | -.032 | 3.64 | .059 | -.072 | 8.55 |
| 4th | 655,981 | .035 | -.034 | 4.72 | .045 | -.056 | 6.89 |
| Highest | 629,757 | .005 | -.033 | 2.69 | .004 | -.045 | 3.48 |
| *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 |

Note: *t* denotes one trading day

[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 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]).

| | Sample Size | One-digit regression[4] | | | Two-digit regression[5] | | |
|---|---|---|---|---|---|---|---|
| | | 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 | 2,774,001 | .003 | -.016 | 3.52 | .013 | -.023 | 6.42 |
| *By Year* | | | | | | | |
| 2001 | 238,389 | -.034 | -.027 | 0.22 | .034 | -.028 | 2.33 |
| 2002 | 399,563 | .027 | -.031 | 2.78 | -.012 | -.019 | 0.35 |
| 2003 | 452,447 | -.005 | -.002 | 0.20 | .033 | -.039 | 4.55 |
| 2004 | 518,167 | .005 | -.017 | 2.13 | .012 | -.020 | 3.02 |
| 2005 | 575,348 | .014 | -.004 | 2.08 | .018 | -.009 | 3.04 |
| 2006 | 590,087 | .015 | -.007 | 2.68 | .005 | -.019 | 2.97 |
| *NYSE Size quintiles*[1] | | | | | | | |
| Smallest | 681,241 | .013 | -.017 | 2.48 | .019 | -.043 | 5.20 |
| 2nd | 660,027 | .010 | -.011 | 1.98 | .028 | -.028 | 5.32 |
| 3rd | 539,034 | .006 | -.024 | 2.99 | .016 | -.010 | 2.68 |
| 4th | 465,765 | -.008 | -.015 | 0.37 | -.006 | .001 | 0.37 |
| Largest | 424,852 | -.005 | -.007 | 0.23 | -.004 | -.017 | 1.51 |
| *Price Categories*[2] | | | | | | | |
| price ≤ $20 | 928,574 | .006 | -.016 | 2.03 | .025 | -.046 | 6.76 |
| $20 < price ≤ $35 | 985,335 | .005 | -.022 | 3.68 | .015 | -.008 | 3.18 |
| price > $35 | 860,092 | -.002 | -.008 | 0.61 | -.002 | -.012 | 0.87 |
| *Dollar volume quartiles*[3] | | | | | | | |
| Lowest | 554,800 | .008 | -.016 | 1.70 | .038 | -.053 | 6.64 |
| 2nd | 554,800 | .007 | -.017 | 1.78 | .009 | -.035 | 3.31 |
| 3rd | 554,800 | .001 | -.010 | 0.78 | .019 | -.022 | 2.82 |
| 4th | 554,800 | .010 | -.020 | 2.63 | .015 | -.003 | 1.56 |
| Highest | 554,800 | -.002 | -.003 | 0.14 | -.010 | -.006 | 0.41 |
| *Exchange* | | | | | | | |
| NYSE/AMEX | 1,554,322 | .004 | -.015 | 2.45 | .015 | -.024 | 5.21 |
| NASDQ | 1,219,679 | .003 | -.014 | 1.92 | .014 | -.023 | 4.23 |

Note: *t* denotes one trading day

[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. 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.

| | | Clustering in closing prices | | | | Closing bid-ask spread | |
|---|---|---|---|---|---|---|---|
| Year | Number of Observations | One-digit clustering[1] | | Two-digit clustering[2] | | Mean | Standard Deviation |
| | | lastdig=0 | lastdig=5 | last 2 dig=00 | last 2 dig=50 | | |
| 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