125.899 Research Thesis

Investor psychological bias towards number
preferences in stock price endings:
Rationality Vs Irrationality

**A research thesis submitted in fulfillment of the requirement
of the Degree of Masters in Business Studies (Fin) at Massey
University**

Name: Amanda Ling Qian Wang

ID number: 05128978

Due Date: 15/04/2011

Abstract

Consumer reactions towards products that end with $X.99 have been heavily researched in the marketing literature. My study has found psychological bias towards numbers in finance prices, where there is a positive return for prices ending in $X.01 and a negative return for prices ending in $X.99 for the American and Chinese Markets, with the return difference annualised to 22.65% and 54.43% per year, respectively. I find there are more buyer initiated stocks for stocks ending with the digit 9. This is not the same as consumer psychological bias, where consumers consider prices that end with 9 to be much cheaper than those ending in other digits. Rather, the case here is a rationality response to psychological bias. This would also explain why I did not find excessive buys in the Chinese market in relation to *lucky* numbers, as is found in the marketing literature.

Acknowledgements

First, I would like to thank my two supervisors, Professor Philip Stork and Associate-Professor John Pinfold, for their enlightening comments and support throughout the process of my thesis. I would also like to thank Min Bai for his help in data downloading and processing. Last, but not least, I would like to thank the staff of the Finance and Economics Department of Massey University for their supportive advice during the completion of my thesis.

Table of Contents

1. INTRODUCTION

Investor rationality has always been a topic explored in behavioural finance, as there are anomalies in the stock market caused by investor irrationality. In traditional finance theory, the efficient market hypothesis asserts that investors cannot make any abnormal returns given the same amount of risk. This is due to the assumption that investors will be rational in choosing those stocks that reflect the fundamental value of the firm. When a stock deviates from this fundamental value, there should instantly be many rational arbitragers that correct this phenomenon. Thus, it would be unrealistic under the efficient market hypothesis to generate excess returns given the same amount of risk.

As can be recalled from advertisements, around 99.99% of the time advertised product pricing is $X.99. This pricing method is used to stimulate consumer sales, as humans process the stock price from left to right, however, prices ending with $X.99 can effectively change the left-most digit of the price. In products that are highly priced, this method may even be used to change the digits before the decimal point; for example, $9,999.99 rather than $10,000.00. Potential customers will perceive the former price as being much cheaper than the latter one.

In my research, I investigate whether investors show the same irrationality towards stock price endings as is documented in the marketing literature, and whether this influences stock returns. My research is, to my knowledge, the first to document

intraday returns with respect to different cent digit endings and attempt to determine whether there is a connection between investor buying behaviour and return effects. I find this research topic very fascinating, because there may be a psychological bias regarding cent digit variations, which may affect stock prices. Differences in cent digits may generate abnormal returns, which may be due to either investor irrationality, or rationality. While much of the literature in this area documents investor irrationality due to personal beliefs and attitudes, there are still many research gaps in the area of behavioral bias towards numbers. The contribution of my research to this topic is the finding that investors are rational towards stock price purchasing, and that there are differences in return with respect to cent digit endings. It is also found that this effect persists for longer than 24 hours.

2. LITERATURE REVIEW

In this research I examine whether or not financial markets are prone to known psychological effects associated with numbers; namely the left digit effect, as is found in consumer behaviour. This minor difference in cent digits can create a difference in the return on share price, which will be explained later in this research report. Furthermore, I examine the lucky numbers effect in the Chinese market. The topic of how numbering has different effects on the human brain was first examined in the field of psychology in terms of brain function. Insights into the memory centre of the human brain help to explain why some anomalies in different pricing strategies have different effects on consumer and investor behaviour.

2.1 Left digit effect

The left digit effect is a pricing effect associated with the brain's function of left to right processing and memory capacity. The left digit effect is mostly concerned with numbers that end with a 9, because when one more number is added to such a figure there will be an increase in the leftmost digit. This can lead the brain to believe numbers ending with a 9 are substantially cheaper than whole numbers that are, in fact, only one price unit higher. In other words, in retail pricing of consumer products, the consumer may observe $99.99 as being substantially cheaper than $100.00. This is because the former amount is considered to be under a hundred dollars and the

latter is seen as being over a hundred dollars. This difference of 1 cent can cause the leftmost digit to change, resulting in more consumer sales of the former product.

2.1.1 Psychology-based findings in support of the left digit effect

Hayes (1952) first experimented on the span of immediate memory, using words, sequences of binary digits, numbers, and numbers with decimal places, alphabetical letters and also a combination of these, to test the memory capacity of the human brain. From his experimental findings it is indicated that, as these numbers and letters are combined, there will be a sharp decrease in the number of items recalled in immediate human memory. The area of his research that is related to my study indicated that 42 items are recalled when remembering binary digits, however, when a decimal place is added, the number remembered will decrease to around ten or eleven items. When applied to the left digit effect, this finding suggests that people are more likely to remember the digits in front of the decimal place, than those following it.

Hayes' (1952) findings on the memory capacity of the human brain has encouraged many others in the field of psychology to examine human memory and stimulus in terms of colour, sound, taste, and literature (Bousfield & Cohen, 1955; Hieder, 1972; Pollack, 1953; Rosch, 1975). The most interesting of these preliminary findings is that of Rosch (1975), where she finds that people remember items using cognitive reference points. In the case of numbers, multiples of 10 are the cognitive reference points used. She also finds that the human brain does not automatically round numbers, instead grouping the numbers using the first digit of the sequence.

2.1.2 Marketing literature on numbers that end with nine

The psychology findings regarding human memory capacity and left to right processing have are used as a foundation for business studies in consumer behaviour. This has mainly occurred in the area of marketing research, with prices ending in a nine being heavily reseached.

Brenner and Brenner (1982) find that the biological constraints of the human memory have implications for consumer behaviour, especially in the area of advertising. They explain the reasoning for ending prices with the digit nine in advertising as being due to the limited memory capacity of the human brain. They find that, when people are exposed to a wide range of information such as many price numbers, the brain will prioritise the value of these messages. For example, if a product is priced at $287.45 the first leftmost digit (2) is more informationally significant then the second digit (8), and so on. It is more likely that consumers will recall the price as being over $200 than it is for them to recall the price as being $280. It is rare that, after being exposed to many prices, people can remember numbers after the decimal place (Coulter, 2001). The left digit effect assumes that consumers remember prices with the leftmost digit having the highest priority, followed by the second leftmost digit, with the memory process always being from left to right (Coulter, 2001). From this theory, the left digit effect assumes that consumers will consider $199.99 to be much cheaper than $200.00, as these two prices differ in their leftmost digit. As a result, the former will be considered to be substantially cheaper than the latter, regardless of there only being a one cent difference. Schindler and Wiman (1989) refer to this type of pricing strategy that uses prices just below a round number as *odd pricing*.

Since the development of the odd pricing strategy, many marketing academics have tested this theory on product sales, all finding evidence of excessive sales volumes for products with prices ending in 0.99 or 9 (Gedenk and Sattler, 1999; Manning & Sprott, 2009; Schindler, 2006; Schindler and Kirby, 1997; Thomas and Morwitz, 2005; Stiving and Winer, 1997). Prices ending with 99 display dicounted price signals. A large body of literature examines and demonstrates this underestimation effect. Here, I outline some of the marketing articles that find significant effects and are related to the current research topic.

Schindler and Wiman (1989) study odd pricing, with a number of test subjects being exposed to a number of different pricing levels of product pictures, some being round numbers and some being odd pricing. They retest the subjects after two days, finding that the test subjects can remember round number pricing better than odd pricing, and that odd pricing is underestimated even though there is only a 1 cent difference in the actual pricing. These results are consistent with the findings of Lambert (1975) and Schindler and Kibarian (1997). Furthermore, they find that the leftmost digits of prices are more accurately recalled. This is consistent with the findings of Coulter (2001), in his examination of the left digit effect in advertising. He finds that the test subjects perceive very different pricing when the leftmost digit changes, also perceiving this difference to be a discount. These theories have been applied to sales volume by Stiving and Winer (1997), who prove that decimal digits of a product price can predict sales volume more accurately that the actual price prior to the decimal point and that there are also more sales when the decimal digit is 99.

Nonetheless, Monroe (1973) also mentions that, although the number nine is over-represented and that there are discounts perceived in the pricing, there is a positive relation with quality price perception. This means that, with certain products, more expensive pricing may signal better quality to consumers. The effect of price and quality perception is mainly seen in the case of high priced luxury goods.

2.1.3 Finance research into stock prices that end with nine

In finance, articles based on investor behaviour in relation to different numbers are quite limited. Differently to marketing, finance is concerned with risk and return and whether investors react to stock prices in the same way as consumers are found to in the marketing literature. The question remains as to whether deviations in cent digits will affect stock sales and return, given the same level of risk.

Bagnoli et al. (2006) examine net sales in those shares that end with nine in the overnight period from close to open. In their hypothesis development they refer to the marketing literature regarding nine-ending product pricing as a discount signal and use overnight returns as an indication as to whether the stock is a buy, or a sell. The data they select are for companies listed on both the New York Stock Exchange (NYSE) and the NASDAQ in 2002. They hypothesize that a positive return will indicate a net buy and that a negative return will indicate a net sell, as they classify a negative return as an act of many investors selling the shares. They use all stocks that have the end digit of nine and examine the next day opening prices for raw returns.

After the analysing the end of day prices using return regressions, they conclude that stock prices ending in nine have a significant negative return in the overnight period and, thus, is a net sell stock (Bagnoli et al., 2006).

In an extention of this research Bagnoli et al. (2006) and Johnson et al. (2008) find return effects towards individual cent digits for the overnight period of close to open. By using individual cent digits they disregard the prices that are to the left of the decimal place and look for patterns in returns for the cent digits. They find that, when a stock price is just above a round number; i.e. $9.01 or $9.02; the stock has a positive overnight return. On the other hand, when the price is just below a round number; i.e. 8.99 or 8.98; the stock will yield an overall negative return. This effect is referred to as the *round numbers effect* in my current research.

Johnsons et al. (2008) use similar data and methodologies to those used by Bagnoli et al. (2006) to test different return behaviour in relation to individual cent digit endings of stock prices, focusing on those stocks that are priced just above, or below, a round number. Furthermore, they use combined cent digit endings to test only the last cent digit. Through combining cent digits, there are then only 10 cent digits to be tested, being the rightmost digit of the security price, which is the second decimal place of every price. The combined cent digit 0 includes all of the cent digits from X.00, X.10….to X.91, with the same system used to obtain the other combined cent digits from 0 to 9. They choose the years 2001 to 2006 and stocks from the NYSE for their data timeframe, as 2001 saw the change in American stock pricing from one-eighth to decimal pricing. After computing the overnight return, Johnson et al. (2008) findings

are very similar to those of Bagnoli et al. (2006), wherein both researches find that prices ending with 9 or 99 have a significant negative return.

Furthermore, Johnson et al. (2008) examine buy and sell volumes in an attempt to explain the return, finding a "…7% differential in the buy and sell percentages between 1 and 9" (p.34). They use the midpoint method proposed by Huang and Stoll (1997) to identify whether the stock is buyer initiated, or seller initiated. The classification of a buyer initiated stock is one where the investor that closes the trade is placing a buy order, and vice versa for a seller initiated stock. A clear example of a seller initiated stock would be where an investor puts in a limit order to buy at 4:30pm and the order is then matched with either a market order, or another limit order, to sell at 4:35pm. This will be classified as a seller initiated stock. When using the midpoint method to define trade direction, one uses the stock price and compares it with the midpoint of the bid and ask prices. If the stock price is above the midpoint of the bid and ask prices it will be classified as buyer initiated, and if it is below the midpoint of the bid and ask prices it will be classified as seller initiated. Johnson et al. (2008) conclude that it is unlikely that the difference in the buy and sell volume is a driver for the return result. Furthermore, Johnson et al. (2008) recommend that, due to market makers bid prices generally being lower than the ask price, limit orders will fall below the midpoint for the buy price and above the midpoint for the sell price. This may result in a positive gain (loss) in the market sell (buy) orders.

Bhattacharya et al. (2010) finds the existence of the left digit effect of security trading in intraday data. The left digit effect of security trading is similar to the psychological

effect that is present in marketing, whereby buyers remember and process the first digits of a price left to right and perceive prices that end with .99 to be much cheaper than whole dollar prices. Such unrounded prices will achieve greater sales volumes (Manning & Sprott, 2009). This is an act of irrationality by investors. There were a total of 100 million data points for 100 companies over a 6 year period starting from 2001. With respect to defining trade direction, Bhattacharya et al. (2010) use the same methodology as Johnson et al. (2008), with reference to Huang and Stoll (1997), using the midpoint method to infer intraday trading direction. Again, the prices that are higher (lower) than the midpoint are buyer (seller) initiated stocks. The only difference in the methodology used by Bhattacharya et al. (2010) from the previous two researches is that they use the bid and ask prices one second before the open price as a reference point. This adaptation is in reference to the finding by Henker and Wang (2006) of a one second lag time in the quotes.

Bhattacharya et al. (2010) find strong supporting evidence of the left digit effect using the number of buyer initiated and seller initiated stocks. They use the buy to sell ratio to describe the differences in investor behaviour. The buy to sell ratio is the number of classified buyer initiated stocks divided by the number of seller initiated stocks, with a big difference found in the buy/sell ratio, which varies from 0.85 for prices that end with X.01 to 1.6 in the buy/sell ratio for prices that end with X.99. This difference in the ratio of 0.8 means that there are around 80% more buys than sells in cent digits that end with .99 (Bhattacharya et al., 2010). Their unconditional tests were all significant at the 1% level. They conclude that when investors see a stock price of X.99 it is considered to be cheaper than a round priced stock, such that they will put in an order to purchase the stock. Moreover, they cite Johnson et al.'s (2008), stating

that "According to the left-digit effect hypothesis, traders susceptible to the left digit-effect buy below an integer and sell at or above the integer. This would mean buying below an integer would cause a loss, and buying above the integer – which is what traders who exploit the left-digit effect do – would cause a profit. This is exactly what Johnson, Johnson and Shanthikumar (2007) find" (Bhattacharya et al., 2008).

2.1.4 Difference data and methodology used in determining trade direction may drive the different conclusions in the financial literature

Summarising the three researches that examine security prices that end with 9 reveal that confounding conclusions and effects arise for numbers that end with 9. Different conclusions are found regarding net sales of stocks that end with 9 in Bagnoli et al. (2006) and net buys of stocks in Bhattacharya et al. (2010) for price digits that end with X.99 and 9. Johnson et al. (2008) refers to Bagnoli et al. (2006) as having a similar finding of significant net returns for stocks that end with 9, but did not comment on investor behavior in terms of the trading direction. Bhattacharya et al. (2010) refers to Johnson et al. (2007) as providing supportive evidence of the left digit effect.

There are some distinct differences in the data and methodology used in these three articles, which will be explained below.

The first of these differences is the type of data; either intraday, or end of day, stock prices. Wang (2009) finds no significant evidence of the left digit effect for all three

indices in my ratio analysis, with the buy/sell ratio difference between X.99 and X.01 varying between 0.1 and 0.2. I test for the left digit effect and the round numbers effect for end of day stock prices in Europe, Australia, and the US using the close and open price of the stocks in the stock index. This is consistent with Johnson et al. (2008), who find an approximately 7% difference in buy volumes between stocks that end with 9 and those that end with 1. My findings for the round numbers effect are consistent with the findings of Johnson et al. (2008), who find significant positive returns for stock prices that just exceed a round number and negative returns for stock prices that are just below a round number. This difference in the left digit effect, as mentioned in Wang (2009), may be due to the call auction at the end of day markets and attempts to make the market more efficient, as well as large amounts of information flowing in overnight, which effect may not be as obvious.

Second, different methodologies are used to classify the trading direction. I find some confounding methodology issues with the classification of buys and sells in the above mentioned literatures. Bagnoli et al. (2006) use returns as a classification of buy and sell, and use negative returns as an indication of net sales. Bhattacharya et al. (2010) and Johnson et al. (2008) use the midpoint of the bid and ask prices as an interpretation of buys and sells. In order for Bhattacharya et al.'s (2010) left digit effect theory and statistical tests to be applicable, the classification of whether a stock is buyer initiated or seller initiated needs to be clear. Bhattacharya et al. (2008) refer to Huang and Stoll (1997) using the midpoint of the bid and ask prices and looking at the price where "…trades above the midpoint are classified as buyer-initiated, trades below the midpoint is seller initiated, and trades equal to the midpoint are disregarded" (Bhattacharya et al., 2010, p.1).

17

Here the term *buyer initiated* refers to the situation where the buyer closes the deal on the order. This is the method used for looking at bid ask spreads and does not indicate the trade direction if security prices stay the same for a number of entries. With intraday data there are, however, different time periods of one second, five seconds, and ten seconds, with these time periods able to be used to draw different conclusions. A classic example using one second tick data occurs when the price and bid ask quotes do not change for a few seconds. This is especially prevalent for illiquid stocks. Using the Huang and Stoll (1997) methodology, which records a buy or sell every second irrespective of whether a trade has occurred, this may lead the results to have multiple buy or sell recordings. For example, if there is an open price of $8.99 and the midpoint of the bid and ask quotes is $8.985 and we assume this price remains constant for five seconds, there will be five recordings of buy, as the opening price is higher than the midpoint of the bid and ask spread. More details of this methodology will be explained in the methodology section below.

There are two other methods for inferring the trade direction from the tick data. The methodology proposed by Lee and Ready (1991) is the only methodology that acknowledges and classifies the trades when prices stay the same. Instead of using the midpoint of the bid and ask prices to infer the trade direction, as in Huang and Stoll (1997), Lee and Ready (1991) use the next price to assign buyer and seller initiated trades. If the next price is higher than the current price it is classified as a buy, and vice versa. There are two more categories of buy and ask assigned to trade direction when the price does not change. When prices stay the same over a few seconds, Lee

and Ready (1991) examine the last trading direction and assign that to the stock price. A zero-buy, which is also a buyer initiated stock, is assigned to the stock direction if the last trading direction was a buy, and a zero-sell is assigned to stocks if the previous change in price was seller initiated. They state that, when classifying trade direction using midpoints of the bid/ask quotes (reverse tick test) as described in Hasbrouck (1988), it is possible to obtain inaccurate trade directions when the next price is within the bid/ask spread.

The second methodology used to infer trade direction is similar to that of Huang and Stoll (1997), where the midpoint of the bid and ask quotes is used to infer trade direction. The only difference between these methodologies is in the lag time used in recording the bid and ask quotes. Henker and Wang (2006) state that there is a delay in the prices associated with the bid and ask spread. In their findings they indicate that there is a one second delay in the pricing of securities in intraday prices. Thus, the bid and ask quotes that are associated with each price come one second before the actual pricing. Bhattacharya et al. (2010) use this methodology where they use the bid and ask prices one second prior to the security prices to infer trade direction. Bessembinder (2003), however, finds a zero second delay in security pricing, which means that the bid/ask and open prices should appear simultaneously and that some bid ask prices can even appear immediately after the open price. This methodology is adopted by Johnson et al. (2008), where they use the simultaneous bid and ask prices and use the next open price of the security price as an indication of trade direction.

Summarising the research on the left digit effect it can be seen that there are some gaps in the financial literature. First, a standardised dataset and methodology should be used to test for the round numbers effect and the left digit effect, so that the results of different researches are comparable. Furthermore, the definition of a buy and sell needs to be clarified and it is also necessary to determine whether there are any relations between the round numbers effect and the left digit effect for end of day data. Moreover, are the differences in the buy and sell volumes and returns caused by investors' irrational behavior, as found in the marketing literature?

2.2 Lucky numbers effect

The lucky numbers effect is inferred directly in the Asian market. Many Asians, especially Chinese, seem to be superstitious in terms of the numbering of prices, which days are lucky or unlucky, and are especially sensitive to the digits four and eight. In Chinese culture, four means death and eight means prosperity. The numbering effects of lucky and unlucky numbers on consumer retail is researched, however, there are no relevant articles in the area of finance that are directly related to this phenomenon.

There is specific marketing research on lucky number pricing in relation to the Chinese market. Simmons and Schindler (2003) find that cultural superstitions have an explicit effect on Chinese market product pricing, with the number eight over-represented in the Chinese market, while the number four is underrepresented. This is

because there are distinct differences among Chinese and Western cultures, with Chinese people believing that supernatural factors cannot be separated from human activity. When superstition is applied to number pricing, the Chinese are most sensitive to the numbers four and eight. The number four has the same pronunciation as the word *death*, while eight is seen to mean prosperity and wealth (Lip, 1992). Although the number nine indicates longevity in Chinese, it does not have that much relevance in commerce, especially in the case of high priced products (Simmons & Schindler, 2003). Simmons and Schindler (2003) examine end digits and find that varying them slightly does not cause much difference in the price level. For example, when comparing ￥25,888.88 and ￥25,500.00 many Chinese consumers and managers would choose ￥25,888.88 as the better alternative, even though it is more expensive. This phenomenon is explained as being due to differences in culture and the signal theory of product pricing (Erevelles, Roy & Yip, 2001).

2.3 Behavioural finance findings of other number effects

Behavioural finance focuses on how an investor's psychology and beliefs have an impact on their actions and also aims to explain some anomalies that are present in the market, but cannot be explained by other traditional theories (Baberis & Huang, 2001; Barberis & Thaler, 2002; Fuller, 1998; Shiller, 2003).

2.3.1 Economic theory linkage to behavioural finance

In order to understand behavioural finance it is necessary to examine the prospect theory of economics, first developed by Kahneman and Tversky (1979). According to

the economic theory of utility, the individual, given probabilities and outcomes, will choose the alternative that offers them the highest level of satisfaction and happiness (Markowitz, 1952). Prospect theory, on the other hand, states that given psychological biases and preferences, losses have more emotional impact than gains, and that individuals are not risk adverse when they think their losses are limited (Kahneman & Tversky, 1979).

From a behavioural finance point of view, the over and underestimation of risk along with cognitive biases directly impacts on investor behaviour. The stock market has many agents who do not react in a fully rational manner to new information. As is obvious, as humans it is impossible to have fully rational expectations regardless if whether the individual is a Chief Financial Officer, or someone without financial knowledge (Barberis & Thaler, 2002). These irrational behaviours can deviate stocks prices away from the fundamental value and cause constant excess returns. For example, De Bodt and Thaler (1985) find that individuals overreact to new information and that having a portfolio full of past extreme negative return stocks will generate a positive return of around 25% higher than a portfolio of extreme past winners.

The main focus of behavioural finance in my thesis is in the area of psychological bias, where numbering significance creates a bias in investor decisions of which shares to purchase, sell or hold, without focusing on the fundamental value of the stocks. There are many findings in behavioural finance which relate to human psychology bias in investment choices. Personal beliefs have an impact on how

investors behave and there are some effects which have an unintentional impact on investor decisions. Barber and Odean (2001) find that male investors are overconfident and trade more frequently than female investors, while generating a loss in return compared to female investors. National happiness due to such factors as success in sport has proven to have a positive effect on the national index on the following day (Edmans, Garcia & Norli, 2007).

2.3.2 Other anomalies associated with numbers in behavioural finance

Clustering

Clustering is a numbers anomaly associated with irrational investor behaviour. Frequency clustering in stock prices is observed especially with round dollars and half dollars. As defined by Harris (1991, p.390), "price clustering occurs because traders use a discrete set of prices to specify the terms of their trades". There are many articles on clustering of trading in whole number end digits (Harris, 1991; Ikenberry & Weston, 2007; Nierderhoffer, 1965), where traders set the round and half numbers in order to simplify their trading process of bids and asks, which in turn decreases their trading costs. The above researches observe that prices fall on these integers more often and, as mentioned in the psychology journals, humans remember whole numbers and are more likely to place orders on whole numbers. This is in line with the behavioural finance school of thought.

Psychological barriers

Psychological barriers in finance relate to the tendency of investors to allocate special significance to certain numbers, which creates resistence or support levels that a price or index must break through. An example of this can be seen in stock indices where, normally, the psychological barriers would be at the 1000's. When the index is below the barrier there will be a resistence level to break through, causing the index to have a tenderncy to fall rather than rise. Once the barrier has been breached, there will be a support level and the index is likely to rise rather than break through and fall below the barrier.

Numerous articles find psychological barriers toward numbers in financial markets in relation to stock price, foreign exchange, commodity pricing, and stock indices (Aggarwal & Lucey, 2007; De Ceuster, Dhaene & Schatteman, 1998; De Grauwe & Decupere, 1992 Donaldson & Kim, 1993; Koedijk & Stork, 1994; Mitchell, 2001; Ley & Varian, 1994). Psychological barriers are formed due to investor expectations towards different numbers. When a price reaches above (below) the price barrier, investor expectations can push the prices, or indices, to a greater tendency to rise (fall). These researches do not, however, examine buyer behaviour towards these psychological numbers and do not find predictability in return patterns to these barriers, which means that there is no violation of the efficient market hypothesis.

Both of the effects described above are signals of irrationality, which are cited in financial researches as signals of inefficiency in the stock market (Harris, 1991; Mitchell, 2001).

2.4 Counter arguments for number anomalies in the stock market and the left digit effect

On the other side of the debate, the observations in the marketing literature regarding behavior towards prices that end with nine should not be observable in relation to stock investors, due to the different expectations that are present in regards to products and stocks.

2.4.1 Investor rationality explanation in traditional finance theory

Consumer reactions to these product pricing theories appear to be rather irrational, however, many traditional finance theories are based on the assumption of rational investors (Fama, 1970; Fuller, 1998; Muhammad, 2009; Shiller, 2003). The efficient market hypothesis developed by Fama, (1970) has dominated the finance sector in terms of security valuation for the past 40 years and has become the bedrock for many security valuation models. For example, many models, including the Capital Asset Pricing Model that uses beta as a risk measure (Merton, 1973) and the consumption betas of individual stocks to determine individual stock risks in order to determine the

appropriate stock price (Breeden & Litzenberger, 1978), are based on the assumption of investor rationality (Fama, 1970).

Investor rationality refers to the assumption that investors will only choose to purchase a share if it is at, or below, the fundamental price, or if the return is larger than the risk taken. Under this theory it is assumed that psychological biases do not influence the investor's decision to purchase, or sell, shares. The efficient market hypothesis asserts that rational traders will process information in the *right* way, meaning that they should have homogeneous expectations. Furthermore, Fama (1970) asserts that even if there are irrational investors in the market, their irrational actions will be cancelled out by the majority, who are rational traders. A typical example of this is where a security is priced above the fundamental value and some irrational investors purchase these shares. It is assumed that the majority of investors will be rational and will be selling their shares, which should cancel out the effect of the irrational trades.

3. HYPOTHESIS DEVELOPMENT

In this thesis, I intend to find whether the investors react irrational to numbers of stock price endings. This is because although behavioral finance has documented many anomalies towards investor behaviour to many events, the endings of stock price should not have such an irrational effect that is shown in marketing literatures. The reason for this is shares has an investment value that is different to consumer products which are brought for consumption and most investors sell their shares at a later date thus should not be influenced by this pattern observed in consumer behaviour. If my assumption is right, the returns effect observed by Johnsons et al. (2008) should not be caused by the left digit effect. I intend to first test if this returns effect exists for intraday stock prices in one second intervals using the same dataset as testing the left digit effect. If there is this returns effect, I also would look at the timeframe that this effect would last until and at the same time determine if there was a left digit effect present at the same time. Also, there is currently no research in finance on investor behaviour towards lucky numbers. Assuming that investors do react the same as marketing literatures where they are irrational towards stock price pricing, the lucky numbers effect should be observed in Chinese market.

From the literature review it becomes evident that the two countries to be examined in this study will be the US and China. I have chosen the US as my first market because the previous finance researches on number effects examine the US market. All three studies outlined above use stock and data from the New York Stock Exchange (NYSE), so I have chosen to use data from the NYSE in my thesis.

I include the Chinese market in my hypothesis development and examination due to the consumer characteristics of the lucky numbers effect as a result of Chinese cultural superstition. Furthermore, the Chinese stock market is the fastest growing equity market in the developing countries. The market mechanism of the Shanghai Stock Exchange (SSE) is similar to that of the NYSE, with a continuous auction market mechanism. Thus, the two markets can be used as robustness checks for each other.

Hypothesis 1

Johnson et al. (2008) find statistical significance in negative returns for numbers below an integer and also for decimal endings of X9, and a positive return for numbers above an integer and also for decimal endings of X1. This should also be found in the intraday data. Due to time span differences and the presence of call auctions and information flows during the overnight period, the intraday return is, however, expected to be smaller than that of the end of day returns when the timeframe is within one minute. Due to the psychological barriers present in stock indices, foreign exchange, and stock commodity pricing, this returns effect may be present for stock prices in the dollar level. This may cause the return effect to persist for a length of time, regardless of whether or not the stock price has changed.

H1: There will be a returns effect for different cent digit endings in the US market and the Chinese market for intraday stock prices and the timeframe of this effect should last for one minute or longer.

Hypothesis 2

The left digit effect is found to exist for intraday stock prices in the US stock market (Bhattacharya et al., 2010). There are some confounding results regarding investor behaviour towards digits ending in nine in stock prices. Assuming the marketing literature is universal, the Chinese market should exhibit the same effect.

H2: There will be a left digit effect in both the US market and the Chinese market for both the individual cent digits and the combined cent digits

Hypothesis 3

There are confounding methodologies in regards to how to classify buyer initiated and seller initiated trades (Huang & Stoll, 1997; Henker & Wang 2006; Lee & Ready, 1991). There may be potential bias and difference in the results produced for the left digit effect when different methodologies are used to define trade direction.

H3: There will be different proportions in the buy/sell ratio when using different methodologies with respect to each cent digit.

Hypothesis 4

The phenomenon of clustering is observed with index and stock prices at the round number. This should also be observed in cent digit endings.

H4: There will be clustering of cent digits observed in digits that end with zero and five.

Hypothesis 5

In terms of Hypothesis 5, there is marketing literature (Erevelles, Roy & Yip, 2001; Simmons & Schindler, 2003) that finds different consumer behaviours, as well as significant findings regarding the effect of number superstitions in the Chinese market. Assuming investors react irrationally towards stock prices as observed in marketing literatures, this effect may be present for the Chinese stock market.

H5: There will be an additional lucky numbers effect and round numbers effect in the Chinese market with excessive buys at ¥8.88 and excessive sells at ¥8.44.

4. DATA AND METHODOLOGY

4.1 Data and extraction

I have selected intraday data of stocks prices and buy and sell quotes that are traded in intervals of one second, using a similar method to Bhattacharya et al. (2010), with a random selection of 100 companies listed on the New York Stock Exchange (NYSE) and 87 companies with A-shares listed on the Shanghai Stock Exchange (SSE), all data being for 2005. I have chosen A-shares only in the Chinese market as, according to Chinese stock market regulations, only Chinese citizens are allowed to purchase A-shares. Choosing A-shares will best reflect the behaviour of Chinese investors. The choice of data from 2005 follows Bhattacharya et al.'s (2010) random selection of data from 2001 to 2006. In the robustness test I have randomly selected another 100 stocks from the NYSE from 2005 for examination of the left digit effect.

The data is drawn from the Sirca database and Java SE2 is applied to the raw data to obtain the first initial output, using SAS and Excel to conduct the statistical tests. I remove all dates and times from the data, as events and information flow should not have an effect on investor decisions to buy or sell on certain cent digits.

4.1.1 Data extraction for the return effect

To test the return effect on different cent digits, I have extracted the next price at one second, one minute, thirty minutes, one hour and twenty-four hours with respect to ending cent digits at time t, to test whether these return effects can persist over a long period of time. The data set after the data extraction consists of around nine million data points from the SSE and thirty-one million data points from the NYSE.

4.2 Methodologies to infer trade direction

There are two types of trade direction. One being buyer initiated and the other being seller initiated. From the literature review it is recalled that a buyer (seller) initiated stock occurs when the buyer (seller) of the stock is the person who closes the trade. As there are different types of methodologies to infer trade direction, I extract data for all three methods to test for the left digit effect. This use of multiple methods is used to check whether the left digit effect is prone to bias when using different methodologies, and as a robustness check for inferring trade direction.

4.2.1 Lee and Ready (1991) methodology

According to Lee and Ready (1991), if the next price is higher (lower) than the current price, it is considered to be buyer initiated (seller initiated). If the price stays the same for more than one second I look at the previous change in price and, if it was buyer initiated (seller initiated), it would be considered to be a zero buy (zero sell).

Zero buys are classified as buyer initiated and zero sells are classified as seller initiated. There are nine million data points in the SSE data set and thirty-one million data points in the NYSE data set. There are also an additional thirty-six million data points used in my robustness test.

Lee and Ready (1991) state that using the next price to determine the trade direction is appropriate when the change in price is within the current bid/ask range. Furthermore, theirs is the only research that classifies trade direction when there are no trades present, or where prices remain the same. Nonetheless, there is a 25% error rate with this methodology and they state that, when compared to the midpoint of the bid/ask spread as a point to infer trade direction, 92% of the results are found to be the same in when classifying the data as buy and sell prices, however, different results are found for zero-buys and zero-sells (Lee & Ready, 1991). This methodology for inferring trade direction will be referred to as Methodology 1 in this thesis.

4.2.2 Huang and Stoll (1997) midpoint method with one second lag

The second methodology used here follows Huang and Stoll (1997), as well as being suggested by Henker and Wang (2006) and also used by Bhattacharya et al. (2010). It uses the midpoint of the bid and ask price one second prior to the price in order to infer whether it is a buy or sell price. When the price is above (below) the midpoint of the bid and ask price in the previous second, it will be considered to be a buy (sell). With respect to prices that remain the same over a period of time, and where the price is still above (below) the midpoint of the bid and ask prices, it will still be considered

to be a buy (sell). All prices that equal the midpoint of the bid and ask price are disregarded and removed from the data set. Huang and Stoll (1997) state that this may eliminate the 25% error rate from using the Lee and Ready (1991) methodology. After extraction I have around eight million data points for stocks in the SSE and twenty-six million data points left from the NYSE stocks. This methodology will be referred to as Methodology 2 for the remainder of my thesis.

4.2.3. Zero lag time for the midpoint method

The third methodology refers to the zero lag time in bid and asks prices, as mentioned in Bessembinder (2003). Using this methodology, I have extracted the bid and ask prices that are in the current second of the price, which has to be before the next price, and I use the midpoint of the bid and ask prices at that time as the midpoint reference. As with with Methodology 2, the next price that equals the midpoint of the bid and ask is disregarded. The assignment of whether a stock is a buy or a sell is also the same as with the previous methodology, with a stock price above (below) the midpoint classified as a buy (sell). After extraction I have around eight million data points for stocks in the SSE and twenty-seven million data points for the NYSE.

4.3 Formulas and statistical tests

In the remainder of thesis I refer to two terms. One is *individual cent digits* and the other is *combined cent digit*. The term individual cent digit refers to 100 cent digit endings such as $X.00, $X.01…$X.99. Combined cent digit refers to the case where I

only use the last digit of the cent decimal to divide the individual cent digits into ten groups. For example, the combined cent digit X1 includes all of the cent digits from X.01, including X.11 through to X.91.

In order to test Hypothesis 1 I use Equation (1) to calculate the return percentage for each of the four time periods.

Return percentage=((next price(t)-price)/price)*100 $\qquad$ (1)

where t is the next price at one second, one minute, thirty minutes, one hour, and one day.

To test for statistical significance of the returns I use the simple straight line regression in Equation (2), assigning individual dummies from 0 to 99 for the cent digits. Furthermore, I assign the combined dummies for all the combined cent digits; e.g. dummy X1…dummy X9 from 0 to 9, as shown in Equation (3).

For the Chinese data, two separate dummy variables are used to test the lucky numbers effect for prices ending in 4.44 (four) and 8.88 (eight) in order to test for the existence of a lucky number effect with regards to differences in return in Equation 4. There are many lucky number combinations available in Chinese culture, however, 4.44 and 8.88 are, respectively, the most unlucky and lucky. If there is no effect present for these two stock prices, then we can conclude that there is no lucky numbers effect in the Chinese market.

Return (t) = C+βD01+βd02+…….. βd99+ ε   (2)

Return (t) = C+βcombined0+βcombined1+…….. βcombined9+ ε   (3)

Return (t) = C+βD01+βd02+…….. βd99+ β4.44+β8.88+ ε   (4)

where t is the next price at one second, one minute, thirty minutes, one hour, and one day. In the combined digits I have omitted all endings with 7, as these all have insignificant results in the previous studies reviewed above. Furthermore, it will be impossible to run a regression with all of the variables in the dataset.

In order to test Hypothesis 4 for clustering and frequency, I extract the number of times a price lands on the combined cent digits and construct a graph to document the significance of clustering occurring at X0 and X5.

In order to test Hypotheses 2, 3, and 5 I use the buy to sell ratio to determine whether they are statistically different from each other. This is because clustering occurs with cent digits endings in zero and five, meaning that the amount of buys for each cent digit may differ significantly. The buy to sell ratio is the best way to determine whether the proportions of buys relative to sells are the same for each digit price. I use Equation (5) to determine the buy/sell ratio and test for statistical significance. I use Equation (6) in a straight line regression on the four individual cent digits buy/sell ratios. The four cent digits that I examine are X.01, X.49. X.51, and X.99. I use Equation (7) to test the ratio regression for significance in combined cent digits in order to determine whether there is a pattern in the buy to sell ratios for the individual

cent digits. I also use the statistical ANOVA test, as it can test for differences in variance in the combined groups and does not regard the mean. This is because it reduces the chance of rejecting the null hypothesis when it is true.

Buy/sell ratio = Number of buy/ number of sell $\qquad$ (5)

Buy/sell ratio = $C+\beta D01+\beta d49+\beta d51+\beta d99+ \varepsilon$ $\qquad$ (6)

Buy sell ratio = $C+\beta combined1+\beta combined2\ldots\ldots. \beta combined9+ \varepsilon$ $\qquad$ (7)

5. RESULTS

5.1 Return results

H1: There will be a round numbers effect in the US market and the Chinese market for intraday stock prices and the timeframe of this effect should last for one minute or longer.

My finding is that returns differ significantly based on cent digit endings.

Time frame one second

Table 1 (full tables are available in the Appendix as Tables 1 and 2), show the accumulated return for the combined digit endings, with respect to returns in percentages within the next second for the NYSE and the Shanghai Stock Exchange. As returns within the next second are quite small and apply to many prices, the price stays the same. Thus, stocks are less likely to be tradable due to the degree of price fluctuations. I have only shown the combined digit decimal returns.

Table 1: Combined returns of NYSE and SSE for one second intervals

	NYSE RETURN RESULTS			SSE RETURN RESULTS		
variable	Coefficient	t-value	p-value	Coefficient	t-value	p-value
Intercept	0.0001559	3.69	0.0002	-0.00184	-6.69	<.0001
combined1	0.00139	23.49	<.0001***	0.01073	27.66	<.0001***
combined4	-0.00104	-17.41	<.0001***	-0.00417	-10.7	<.0001***
combined6	0.0007316	12.33	<.0001***	0.00697	18.3	<.0001***
combined9	-0.0015	-25.47	<.0001***	-0.00329	-8.58	<.0001***

Note: *, **, and *** represent statistical significance at the 10%, 5%, and 1% level, respectively.

The results from both countries' markets indicate that there is a round numbers effect in the one second time frame. There are significant positive returns for all prices that end with X1 on decimal digits over the two stock exchanges (0.139 basis points for the NYSE and 1.07 basis points for the SSE) and significant negative returns for the combined nine decimal digits (-0.1 basis points for the NYSE and -0.3 basis points for the SSE). This is a difference of 0.239 basis points for the NYSE and of 1.37 basis points for the Shanghai Stock Exchange in the one second period alone. Furthermore, a second set of prices display the same effect, being prices that are just below, or above, the $X.50 margin. As seen from Table 1, for are significant negative returns for prices that end with X4 decimal digits for both exchanges (-0.104 basis points for the NYSE and 0.417 basis points for the SSE), while prices that end with X6 show significant positive returns (0.07 basis points for the NYSE and 0.6 basis points for the SSE). This is consistent with the findings of Johnson et al. (2008).

Aggregated timeframe results

Table 2 is a modified table showing how long this round numbers effect for decimal points that ends with .01, .49, .51, and .99 persists for four different time periods of the next price, at one minute, thirty minutes, sixty minutes, and twenty-four hours for the NYSE.

Table 2: Summary of aggregated return over 24 hours for the NYSE

Variable		Coefficient	P-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
NYSE Number of Observations Read 31,974,029									
		1 minute		30 minute		60 minute		24 hours	
Intercept		0.00111	<.0001	0.00901	<.0001	0.01837	<.0001	0.03887	<.0001
d01		0.00367	<.0001***	-0.00203	0.0339**	-0.00788	<.0001***	0.03236	<.0001***
d99		-0.00404	<.0001***	-0.01129	<.0001***	-0.01465	<.0001***	-0.04934	<.0001***
DIFFERENCE	0.00771			0.00926		0.00677		0.0817	
d51		0.00365	<.0001***	0.0101	<.0001***	0.00556	<.0001***	-0.01339	0.0001 ***
d49		-0.00302	<.0001***	-0.0046	<.0001***	-0.01237	<.0001***	-0.03417	<.0001***
DIFFERENCE	0.00667			0.0147		0.01793		0.02078	

Note: *, **, and *** represent significance at the 10%, 5%, and 1% levels, respectively.

As shown in Table 2, this round number return effect for stock prices persists for 24 hours, and even longer. The degree of difference in returns alters over time, however, in both cases, as time is lengthened the difference in the returns increase. The difference in return between buying stocks that end with .01 and .99 is 0.77, 0.92, 0.67, and 8.17 basis points for the time frame of 1 minute, 30 minutes, 60 minutes, and 24 hours, respectively. The 24 hour difference when aggregated results in an annualised return difference is 22.65% assuming 250 trading days annually. As the NYSE did not have an index at the time the data is drawn from, I use the S&P 500 index as a reference point. At the beginning of 2005 the S&P 500 index stands at 1202.08 and on December 30th the index closes at 1248.29, with the year's return being 3.84%. Although this is not the NYSE index, it may give an indication of the expected return over the year and any abnormal return caused by the round numbers effect.

Differences in returns between buying stocks that end with .49 and .51 are 0.67, 1.47, 1.79, and 2.07 basis points for the time frame of 1 minute, 30 minutes, 60 minutes, and 24 hours, respectively. This is an annualised return of 6.98% over 250 trading days. All of the variables are significantly different at 1%, with the exception of D01 at the 30 minute interval (statistically significant at the 5% level).

Appendix Table 1 shows the results of the aggregated returns for the NYSE for cent digits from 01 to 99. As seen in the table, stocks that end with 1 have mostly positive returns, while stocks that end with 9 have mostly negative returns.

Table 3 shows a modified summary of the round numbers effect in the SSE for decimal points endings of .01, .49, .51, and .99, which persist for the four different time periods of next price at one minute, thirty minutes, sixty minutes, and twenty-four hours, respectively.

Table 3: Summary of aggregated returns over 24 hours for the SSE

SSE Number of Observations Read 8836979								
Variable		Coefficient	P-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
		1 minute		30 minute		60 minute		24 hours	
Intercept		0.00141	<.0001	-0.01885	<.0001	-0.03382	<.0001	-0.15293	<.0001
d01		0.01429	<.0001***	0.04413	<.0001***	0.05999	<.0001***	0.14346	<.0001***
d99		-0.00647	<.0001***	-0.01101	0.0262**	-0.00394	0.5297	-0.0307	0.009***
DIFFERENCE		0.02076		0.05514		0.06393		0.17416	
d51		0.00911	<.0001***	0.00791	0.1063	0.01675	0.0069***	-0.11068	<.0001***
d49		-0.0058	<.0001***	-0.00806	0.0933*	-0.00417	0.4931	-0.07428	<.0001***
DIFFERENCE		0.01491		0.01597		0.02092		-0.0364	

Note: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

As seen from this table, the aggregated difference in returns for the SSE between prices that end with .01 and .99 also increases over time with differences of 2.08, 5.51, 6.4, and 17.4 basis points for the time frames of 1 minute, 30 minutes, 60 minutes, and 24 hours, respectively. This is an annualised return of 54.43% for 250 trading days in the SSE. In the same year, the Shanghai Stock Index has made a return of 9.08%, which may be an indication of the abnormal returns generated by the returns effect. All values are statistically significant at the 5% level, with the exception of those stocks ending in .99 at the 60 minute interval. Furthermore, the difference in returns between buying stocks that end with .49 and .51 is 1.5, 1.6, 2.01, and -3.6 basis points for the time frames of 1 minute, 30 minutes, 60 minutes, and 24 hours, respectively. Here we see that the round numbers effect for differences in returns is not present for digits of .51 and .49 and that the values of difference at 30 minutes and 60 minutes are not statistically significant, although the right hypothesized signs of the returns are indicated.

Appendix Table 2 presents the results of the aggregated returns for the SSE for cent digits from .01 to 0.99. As seen from the table, many cent digits have returns that are statistically significant at the 1% level, however, it is worth mentioning that, for China, only a slight round numbers effect is present after the first hour, making the aggregated return for 24 hours insignificant. Iregard those returns that are significant, but that have different positive or negative returns as having no effect. For the cent digits from .01 to .10 and .90 to .99, I find another effect. Cent digit stocks that start with .1X display positive returns and also have higher returns in relation to cent digits that start with .9X.

5.2 Left digit effect results

Hypothesis 2: There will be a left digit effect in the NYSE and the SSE for individual cent digits and combined cent digits.

I found no left digit effect for the individual cent digits for the NYSE and SSE, however, I have found significant difference in proportion of buy to sell in the combined cent digits.

Hypothesis 3: There will be different proportions in the buy/sell ratio when using different methodologies with respect to each cent digit.

I have found that different methodologies significantly affect the proportions of buy/sell ratio and its statistical significance.

I use three different methodologies to extract data, as mentioned in the above. The first of these follows Lee and Ready (1991) (Methodology 1), where I use the next price to infer trade direction. If the next price is higher (lower) than the previous price, it would be considered a buy (sell). When prices stay the same over certain cent digits, I use the previous trade direction and assign it to the current trade direction to determine a buy, or a sell. Methodology 2 follows Bhattachayra et al. (2010), such that, if the next price is above (below) the midpoint of the previous bid and ask, it would be considered a buy (sell) regardless of whether prices stay the same. Methodology 3 holds that, if the next price is above (below) the midpoint of the

43

current bid and asks the sale will be considered a buy (sell) regardless of whether or not prices remain the same.

New York Stock Exchange

Figures 1 to6 (see Appendix) shows the buy/sell ratio for each cent digit in the NYSE and also the buy/sell ratio for the combined cent digits. The four digits that I will refer to specifically in my study, as displayed in Figures 1 to 6, are X.01, X.49, X.51, and X.99, as also focused on by Bhattacharya et al. (2008) in their unconditional test. I use the averages of the combined cent digits to construct a line graph to determine whether this effect is evened out. My results do not replicate those of Bhattacharya et al. (2008), with differences in the buy/sell ratio of 0.8 using all three methodologies, however, my results were consistent with those of Johnson et al. (2008), with a difference of around 5% with respect to the buy/sell ratio.

There are differences to the buy/sell ratio with respect to differences in methodology. Figure 1 is the primary bar graph results from Methodology 1 in the buy/sell ratio for each cent digit. Using Methodology 1, the buy/sell ratios are 1.136, 1.185, 1.136, and 1.194 for the four cent digits of X.01, X.49, X.51, and X.99, respectively, where the difference in the buy/sell ratio between digits X.01 and X.99 and digits X.49 and X.51 are 0.058 and 0.049, respectively. Figure 2 is the line graph of the average combined ratios. As seen from the graph, the buy to sell ratios are very similar across the cent endings, with the exceptions of a small high spike for X4 and X9 and the difference between X1 and X9 of 0.028.

Figure 3 shows the buy/sell ratios for Methodology 2 applied to the NYSE data. The ratios are 1.138, 1.213, 1.130, and 1.172 for the four cent digits of X.01, X.49, X.51, and X.99, respectively. The difference in the buy/sell ratio between digits X.01 and X.99 and digits X.49 and X.51 are 0.034 and 0.083, respectively. The combined cent digits line graph (Figure 4) shows a more obvious pattern towards the different cent digits with higher ratios for X4 and X9, and lower ratios for X1 and X6. The difference between the highest ratio and the lowest ratio (X9 and X1, respectively) is 0.0641.

The buy/sell ratio results for Methodology 3 are shown in Figures 5 and 6. The ratios are 1.128, 1.196, 1.126, and 1.216 for the four cent digits of X.01, X.49, X.51, and X.99, respectively. The differences in the buy/sell ratio between digits X.01 and X.99 and X.49 and X.51 are 0.088 and 0.07, respectively. The pattern generated in Graph 6 for the combined digits is similar to that for Methodology 2, however, this is to a lesser degree.

With all three methodologies, however, the difference in the buy to sell ratios remains less than 0.1, which is much less than found in Bhattacharya et al. (2008), of the left digit effect in differences in buy/sell ratio for stock ending with X.01 (0.83) and X.99 (1.63), and X.49 (1.57) and X.51 (0.84) of 0.8 and 0.73, respectively. Furthermore, their combined line graph has a 0.5 difference in the buy/sell ratio.

Table 4 presents the results of the regression analysis of the buy/sell ratio using the three methodologies over the cent digits of X.01, X.49, X.51, and X.99. As seen from Table 4, although the four cent digits have the expected signs, most of the values

remain insignificant when regressed individually. Methodology 1 has significant values at X.49 (5% level) and X.99 (1% level), while Methodology 2 has only one significant value (X.49 at the 1% level). Methodology 3 finds that X.49 (5% level) and X.99 (5% level) are significant, however, none of the methodologies show significance for cent digits that end with X.01 and X.51, which is in violation of the left digit effect proposed by Bhattacharya et al. (2008).

Table 4: Ratio regression results for the NYSE in the three methodologies

Variable	Methodology 1		Methodology 2		Methodology 3	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	1.14577	<.0001	1.14527	<.0001	1.15216	<.0001
d01	-0.01026	0.4041	-0.00693	0.7747	-0.02381	0.1915
d49	0.03877	0.0021**	0.06804	0.0059***	0.0441	0.0167**
d51	-0.00994	0.4192	-0.01558	0.5203	-0.02661	0.1449
d99	0.04866	0.0001***	0.02671	0.2714	0.06446	0.0006*

Note: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5 shows the regression results of the buy/sell ratio for the NYSE for the three methodologies. In contrast to the single cent digits, the combined digits present a pattern using all three methodologies. Although the variation in the ratio is quite limited, there are higher buy/sell ratios at all digits that end with X9 and X4, with lower ratios for those ending in X1 and X6.

Table 5: Combined ratio regression of the three methods for the NYSE

Variable	Methodology 1		Methodology 2		Methodology 3	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	1.14469	<.0001	1.16557	<.0001***	1.15985	<.0001***
combined1	-0.0005575	0.8779	-0.04701	<.0001***	-0.02388	0.0001***
combined2	-0.00501	0.1698	-0.03135	<.0001***	-0.02905	<.0001***
combined3	-0.00184	0.6127	-0.02167	<.0001***	-0.00804	0.1752
combined4	0.01683	<.0001***	0.01459	0.0044***	0.00758	0.201
combined5	-0.01075	0.0038	-0.01891	0.0003***	-0.00604	0.3075
combined6	-0.00908	0.0139**	-0.05047	<.0001***	-0.01988	0.0011***
combined7	-0.00187	0.6071	-0.02911	<.0001***	-0.01449	0.0157**
combined8	0.00266	0.4638	-0.02889	<.0001***	-0.00146	0.8039
combined9	0.02708	<.0001***	0.01707	0.0009***	0.02416	<.0001***

Note: *, **, and *** represent statistical significance at the 10%, 5% and 1% levels, respectively.

As seen from this analysis, Methodology 2 shows significance for all of the variables at the 1% level. When using Methodology 1, only X4, X6, and X9 are found to be significant at the 1% level. When adopting Methodology 3, X1, X2, X6, and X9 are all found to be significant at the 1% level. The implication of these findings is that using different methods to determine whether the stock is a buy or a sell will differentiate the results to a significant degree. This will be discussed further in the discussion section of this thesis.

Table 6 shows the results of the ANOVA testing of the differences in the combined digit results. The groupings I test to analyse the difference in the means of the ratios are drawn from X0 to X9. All three methods show that each group's mean is significantly different from the others, with differences in the mean of around 0.1. As previously mentioned, however, I am not able to replicate the results of Bhattacharya et al. (2008). This will be discussed further in the discussion section of the thesis.

Table 6: ANOVA results for the NYSE from testing for differences in variance among the combined cent digits.

	ANOVA				
	Source of Variation	*SS*	*MS*	*F*	*P-value*
	Between Groups	0.045	0.00562	59.3066	1.10E-30
	Within Groups	0.0077	9.476E-05		
Methodology 1	Total	0.0526			
	Between Groups	0.0122	0.0015254	24.8588	7.35E-19
	Within Groups	0.005	6.136E-05		
Methodology 2	Total	0.0172			
	Between Groups	0.022	0.0027526	15.0477	2.74E-13
	Within Groups	0.0148	0.0001829		
Methodology 3	Total	0.0368			

Shanghai Stock Exchange

Figures 7 to 12 present the graphical analysis of the buy/sell ratios for the stocks in the SSE.

Figure 7 presents the bar graph results of Methodology 1 in terms of the buy/sell ratio for every cent digit, where the buy/sell ratios are 0.990, 0.992, 0.984, and 1.000 for the four cent digits of X.01, X.49, X.51, and X.99, respectively; where the difference in the buy/sell ratio between digits X.01 and X.99, and X.49 and X.51 are 0.001 and 0.008, respectively. Figure 8 is the line graph of the average combined ratio. There is no obvious pattern to this graph, with spikes of high buy/sell ratios at X2, X4, and X9 and the lowest point at X8. These results are in violation of the left digit effect, with the difference between X1 and X9 being 0.018.

Figure 9 shows the buy/sell ratios for Methodology 2 in the SSE with buy/sell ratios at 0.873, 1.048, 0.864, and 1.039 for the four cent digits of X.01, X.49, X.51, and X.99, respectively. The difference in the buy/sell ratio between digits X.01 and X.99, and X.49 and X.51 are 0.165 and 0.184, respectively. The pattern seen in this graph favours the left digit effect more, with obvious spikes at digits that end with 9. In the combined cent digit line graph in Figure 10 there is a more obvious pattern towards the different cent digits, with higher ratios at X4 and X9 and lower ratios at X1 and X6. The differences between the buy/sell ratio of X1 and X9 are at 0.162, while X4 has the highest buy/sell ratio of 1.021.

Figures 11 and 12 present the results for Methodology 3 in the SSE. In Graph 11 the individual buy/sell ratios are 0.900, 1.031, 0.885, and 1.056 for the four cent digits of X.01, X.49, X.51, and X.99, respectively. The differences in the buy/sell ratios between digits X.01 and X.99, and X.49 and X.51 are 0.156 and 0.146, respectively. As indicated by the combined line graph, the pattern is still similar to that uncovered by Methodology 2, with spikes in the buy/sell ratio at X4 and X9. The difference in the buy/sell ratios of X1 and X9 is 0.138.

The SSE stock data presents similar findings to those of the NYSE stock data, where there are still higher buy/sell ratios in stocks that end with nine. I still cannot replicate the findings of Bhattacharya et al. (2008).

Table 7 shows the individual buy/sell ratio regression analysis of the SSE for the three methodologies. As seen from the graph, there are no significant results from

Methodology 1, only one significant value from Methodology 2 (d51), with Methodology 3 finding only two significant results (d51 and d99 at the 5% level). For all three methods, however, neither d01 nor d49 are significant. Thus, the results are in violation of the left digit effect.

Table 7: Individual cent digit regression analysis for the SSE

Variable	Methodology 1		Methodology 2		Methodology 3	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	0.98339	<.0001	0.96604	<.0001	0.96836	<.0001
d01	0.00626	0.677	-0.09342	0.1133	-0.06816	0.1043
d49	0.00861	0.5666	0.0823	0.1624	0.06232	0.137
d51	0.00011088	0.9941	-0.10246	0.0829*	-0.08344	0.0475**
d99	0.01616	0.2834	0.07333	0.2128	0.08723	0.0385**

Note: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 8 shows the combined ratio analysis for the SSE under the three methodologies. As seen from the table, similarly to the results for the NYSE, the three different methodologies yield results at different significance levels.

Table 8: Combined ratio analysis of the SSE for three methodologies

Variable	Methodology 1		Methodology 2		Methodology 3	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
Intercept	0.98282	<.0001	1.0375	<.0001	1.00281	<.0001
combined1	-0.00465	0.3947	-0.17973	<.0001***	-0.11616	<.0001***
combined2	0.01395	0.0119**	-0.09444	<.0001***	-0.06321	<.0001***
combined3	-0.00866	0.1143	-0.12553	<.0001***	-0.06795	<.0001***
combined4	0.00634	0.2462	-0.01652	0.0623*	-0.00587	0.4382
combined5	0.00327	0.5492	-0.03766	<.0001***	-0.01351	0.0763*
combined6	0.00133	0.8066	-0.13137	<.0001***	-0.06819	<.0001***
combined7	-0.00258	0.6356	-0.05701	<.0001***	-0.02143	0.0055***
combined8	-0.01358	0.0143**	-0.07262	<.0001***	-0.01659	0.0302**
combined9	0.01349	0.0148**	-0.01803	0.0422**	0.02139	0.0056***

Note: *, **, and *** represent statistical significance at the 10%, 5% and 1% levels, respectively

As shown by this table, Methodology 1 finds only three significant values compared to nine significant values found under Methodology 2, and no significant values found for the combined4 grouping under Methodology 3. Most of the coefficients found are negative. This phenomenon will be discussed in the discussion section.

Table 9 shows the ANOVA results for the three methodologies. The results show that the means of the buy/sell ratios of the combined cent digits are significantly different from each other.

Table 9: ANOVA results for the SSE when testing for differences in variance among the combined cent digits

	ANOVA				
	Source of Variation	SS	MS	F	P-value
Methodology 1	Between Groups	0.2509	0.0313651	72.7722	9.1E-34
	Within Groups	0.0349	0.000431		
	Total	0.2858			
Methodology 2	Between Groups	0.0071	0.0008851	5.7074	9.35E-06
	Within Groups	0.0126	0.0001551		
	Total	0.0196			
Methodology 3	Between Groups	0.1444	0.0180446	76.5949	1.49E-34
	Within Groups	0.0191	0.0002356		
	Total	0.1634			

Overall, from the results presented in the above graphs and tables, I cannot conclude that there is a left digit effect for the single cent digits, as proposed by Bhattacharya et al. (2008). The pattern variation in the combined cent digits may not be due to investor behaviour patterns in perceiving these stocks as being cheaper, but may rather be due to market mechanisms. Although there is a slight variation in the buy/sell ratio and a constant pattern for digits that end in .99 and .01, the degree of variation for both samples (The NYSE and the SSE) is very small compared with that found by Bhattacharya et al. (2008). Furthermore, when applying different methods to the intraday data to determine whether a stock is a buy or a sell different results are generated.

5.3 Clustering

Hypothesis 4: Clustering will be observed in digits that end with zero and five.

I have found clustering in cent digit endings that ends with zero and five.

I have observed clustering in both countries for stocks that end with zero, especially in round dollars, and also on stock prices that ends with X.50. Figures 13 and 14 show the frequency with which stock prices fall on the particular price. As seen from the two figures, in both countries, prices stop most frequently on prices that end with X.00 and second most frequently on prices that end with X.50. As observed by both trends, it can be clearly seen that there are higher frequencies with all prices that have cent digits that end with zero, and next with cent digits that end with five. After the higher frequencies observed in cent digits that end with zero and five, the next highest frequency occurs on both sides of the spikes, forming a wave-shaped trend. This means that the price that is next likely to be stopped on after cent digits that end with zero and five are the cent digits adjacent to them. For cent digits that end with zero, this would be cent digits that end with nine or one, and for cent digits that end in five, the next frequency down would be stock prices ending with four or six.

Figure 13: NYSE cent digit frequency
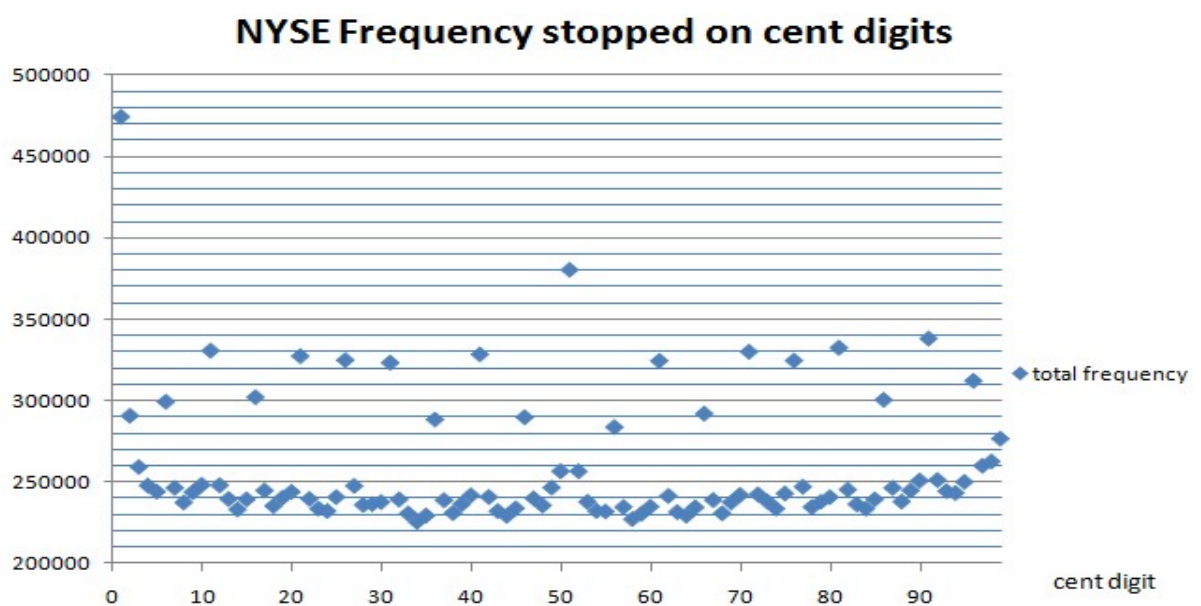
Figure 14: SSE cent digit frequency

SSE frequency stopped on cent digits
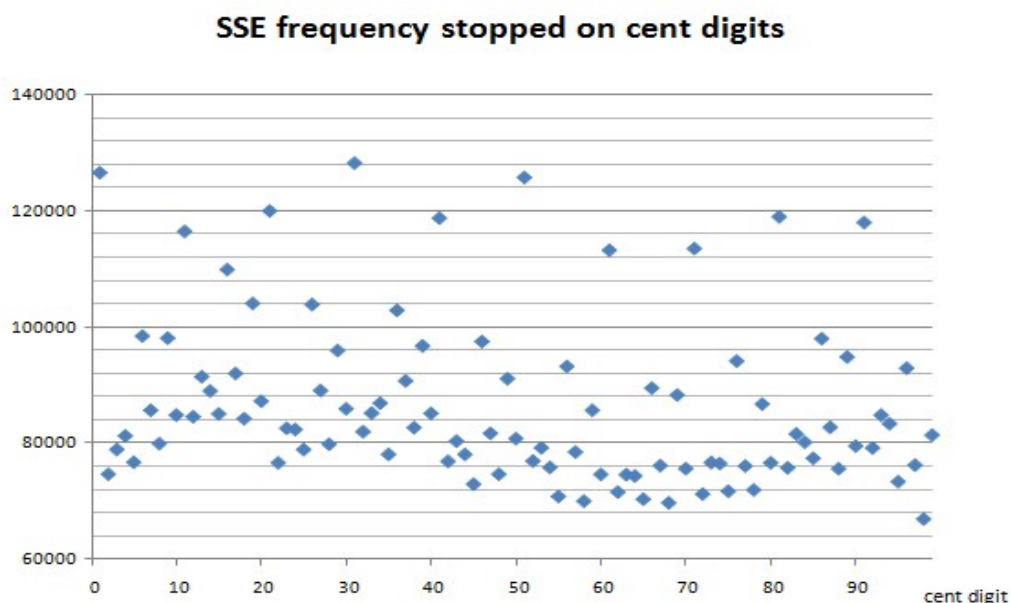
Table 10 presents the results for the frequency of occurrence for the cent digit endings in the NYSE and the SSE. As seen from the table, the coefficients of combined0 and the combined5 are significantly higher than the coefficient frequency observed in other cent digits. This is statistically significant at the 1% level for both markets. Furthermore, all individual cent digits in the NYSE are statistically significant at the 1% level and, as seen in the coefficients, cent digits that end with 00 and 50 are much higher than are the other cent digits. Similar results are found in the SSE, with high positive coefficients for stocks that end with 00 and 50, both statistically significant at the 5% level, however, the other cent digits remain insignificant.

Table 10: Frequency regression results for the NYSE and the SSE

Variable	NYSE frequency		SSE frequency	
	Coefficient	p-value	coefficient	p-value
combined0	92432	<.0001***	55150	<.0001***
combined1	6684	0.1133	12098	0.007***
combined2	1511	0.7023	16693	0.0003***
combined3	-2818	0.4763	15942	0.0005***
combined4	1776	0.6532	10693	0.0166**
combined5	64855	<.0001***	33219	<.0001***
combined6	7616	0.0565*	18045	<.0001***
combined8	6249	0.1163	27469	<.0001***
combined9	5678	0.1777	15716	0.0005***
d00	144840	<.0001***	42208	0.0152**
d99	65421	<.0001***	-9183	0.5917
d01	47157	<.0001***	-9753	0.569
d50	50951	<.0001***	41386	0.0172**

Note: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

5.4 Lucky numbers results

Alternative Hypothesis 5: There will be an additional lucky numbers effect and round numbers effect in the Chinese market with excessive buys at ¥8.88 and excessive sells at ¥8.44.

Do not reject the null hypothesis. There is no lucky numbers effect present for the Chinese market.

Similar to the findings observed for the left digit effect, I was not able to find any significant difference in the buy/sell ratios with respect to lucky numbers in the Chinese market. In Figures 7, 9, and 11 two additional bars are included, labelled *four*

and *eight,* being the dummy variable for stock prices of ¥4.44 and ¥8.88. From these three figures it can be seen that there is little difference among the buy/sell ratios for stocks that are considered lucky, or unlucky, according to their pricing. Table 11 shows the regressed results of the buy/sell ratio of lucky numbers that end with 4 or 8, and the two stock prices of ¥4.44 and ¥8.88.

Table 11: Regression results for endings of four, eight, combined four, and combined eight

Variable	Methodology 1		Methodology 2		Methodology 3	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
four	0.00475	0.7458	0.06989	0.2449	0.01181	0.7826
eight	-0.03061	0.0389**	0.07089	0.2383	0.10234	0.0184**
combined4	0.00634	0.2462	-0.01652	0.0623	-0.00587	0.4382
combined8	-0.01358	0.0143**	-0.07262	<.0001	-0.01659	0.0302**

Note: *, **, and *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

From these results, it can be concluded that there is no lucky numbers effect with the combined4 and combined8, and there is even a reversed effect on these prices' buy/sell ratios.

6. ROBUSTNESS TEST

The return results are consistent with the previous findings in the literature, with the NYSE generating similar results to the SSE. This confirms the results of Johnson et al. (2008), thus the return results are robust. Nonetheless, I was unable to replicate the findings of Bhattacharya et al. (2010) in regards to the left digit effect in either of these countries. In order to test the robustness of my results for the left digit effect, I expand my data set and randomly select another 100 companies from the NYSE, downloading the intraday data for these companies for 2005. After data extraction I have around 36 million data points in this new data set.

Figures 15 to 20 (see Appendix) present bar and line graphs of the buy/sell ratios for the individual cent digits and the combined cent digits for the three methodologies used to determine trade direction. As similar results were generated to those already obtained from the first NYSE data, I will describe these results only briefly.

The same pattern and variation is seen, which is consistent with my previous findings, where the difference in the buy to sell ratio in individual cent digits between X.01 and X.99 remains within 0.1. The same result is generated for the combined cent digits, where the buy to sell ratio difference between X1 and X9 is no larger than 0.1. From the results of the robustness data testing, I conclude that my results for the left digit effect are robust.

7. DISCUSSION

Some points from the results need to be discussed. I find evidence of the returns effect for different cent digit endings. I do not find evidence of the left digit effect for individual cent digits and have not been able to replicate the findings of Bhattacharya et al. (2010). I find clustering in the occurrence of cent digit endings X0 and X5 and no lucky numbers effect for the SSE. I cannot conclude the existence of any relation among the return effects.

7.1 Explanations of the returns effect

As seen from the returns results, for both countries there is a pattern of higher returns for stock prices that are just above a round number, and negative returns for stock prices that are just below a round number for both countries. This effect is present for the time frame of one second, one minute, thirty minutes, one hour, and even after twenty-four hours. To my knowledge, this is the first study to examine these intraday returns in respect to cent digit endings.

These differences in returns for different cent digits are difficult to explain in terms of traditional finance theory. This phenomenon is caused by irrationality in the market, however, under an efficient market there should not be any abnormal returns. The psychological bias in the round numbers effect can be contributed to the difference in return on my results. The NYSE has an aggregated return difference of 8.17 basis points per day between stocks that end with X.99 and X.01. This annualises to a difference of 22.65%, assuming a 250 trading days per annum. In the case of China, it

annualises up to a 54.43% difference in returns per annum, assuming 250 trading days. As seen in the results table, this return effect can persist for more than 24 hours and does not reduce over the time frame of this examination, but instead yield a bigger difference in returns. Although I have not included the effects of transaction costs, the annualised return for the return effects are above the S&P 500 stock index return (3.84%) and the Shanghai Stock Index (9.08%) and can, therefore, be considered as abnormal returns. It is difficult to determine whether this is in violation of the efficient market hypothesis, as I do not include transaction costs in the analysis. The larger return on the SSE may, however, signal stock market inefficiency when compared to the NYSE.

Behavioural finance explanations

There may be some explanations in behavioural finance able to justify this difference in returns. The first of these possible explanations is the psychological barrier of numbers. As mentioned in the literature review, this issue is observed in many researches into stock prices, stock indices, foreign exchange markets, and commodity pricing (Aggarwal & Lucey, 2007; De Ceuster, Dhaene & Schatteman, 1998; De Grauwe & Decupere, 1992; Donaldson & Kim, 1993; Koedijk & Stork, 1994; Mitchell, 2001). The resistance levels found in the previous studies of psychological barriers are usually multiples of tens, hundreds, and thousands, however, no predictability of this issue has been found. It has, therefore, been concluded that this is not in automatic violation of the efficient market hypothesis. My results have shown a systematic pattern, which is consistent with the findings of Johnson et al. (2008) and Wang (2009) of end of day trading patterns towards stock prices, with my findings

showing that stocks that are just above (below) any whole number, or X.50, will have significant positive (negative) returns. This means that the price of these stocks once reaching the X.00 pricing level is more likely to rise, thus generating a positive return. Meanwhile, for those stocks that drop below this point, the stock price is has a greater likelihood of falling, which is consistent with my findings in this report.

The second possible explanation for this return effect is the left digit effect, as proposed by Bhattacharya et al. (2010) and in reference to Johnson et al. (2008) as providing proof of the left digit effect. Due to the insignificance of my findings in the individual ratio analysis of the left digit effect, I cannot conclude that there is a significant relation among the two effects for the individual digits. In the combined regression results for the ratio analysis, however, I find a relation between the left digit effect and the returns effect, where the buyer initiated transactions (higher than the buy/sell ratio) are likely to have a negative return, and the seller initiated transactions have a positive return. As these effects persist for 24 hours it is, however, less likely that the returns effect is caused by the left digit effect.

The third possible explanation is generated by Bagloni et al. (2006), who documented that the return of stocks ending with nine is negative due to the action of net selling during the overnight period, as investors regard stocks that end with nine to be a signal of poor quality. My results with respect to the proportion of buyer initiated and seller initiated prices reject this explanation, as there are more buyer initiated stocks with numbers that end in nine than other digit endings. The overnight period for the NYSE is, however, followed by an opening auction and other studies confirm that investors tend to cluster trades on X.00 and X.50, as these amounts are easy memory

reference points. These levels also reflect one of the psychological biases of many investors. The overnight trading is generally in limit orders and is executed at the opening auction. Due to the opening call auction mechanism, there may be a number of rational investors that are aware of this clustering and are willing to sell the stock at X.99 due to clustering of orders during the overnight period. This will also result in the shares being classified as buyer initiated. This may confirm that there is net selling during the overnight period, as investors are more willing than usual to sell at X.99 in order to be at the front of the queue. This behaviour will be discussed further in the following section.

7.2. Explanations of differences in buyer initiated shares relative to seller initiated shares for stocks that end with nine and one

I find that there are higher (lower) proportions of buyer initiated shares relative to seller initiated shares for stocks that end with nine (one). The left digit effect is not statistically significant for individual cent digit results. The degree of variation in the combined cent digits is quite small when compared to the finding of Bhattacharya et al. (2010) of a buy to sell ratio difference of 0.1, which is consistent across all three methodologies. The presence of this small difference in the buy to sell ratio should not be due to the marketing left digit effect. I cannot conclude from my results that the difference in the buy and sell volume is due to investor preferences for stock that appears cheaper. The pattern in the buy to sell ratio may have other causes.

7.2.1 Methodology differences generate similar results for the number of buys and sells

When I use the three different methodologies to determine tick data trade direction, there are small variations within the generated results. There is no determination as to which methodology is the most appropriate, however, with all three methodologies generating similar results, my results for the amount of buys and sells should be robust.

When using the Lee and Ready (1991) methodology, using net price as the trade reference, there is a 25% error rate, which they especially regard when the price does not change. When using the midpoint of the bid and ask as an indication of the trade direction, there is a deletion of data of around 20% of those prices that are equal to the midpoint of the bid and ask (Bhattacharya et al., 2008; Henker & Wang, 2006; Huang and Stoll, 1997). With respect to time lags, there is a one second lag and a zero lag proposed for the bid and ask prices, however, under all three methodologies I generate similar results and have not been able to replicate Bhattacharya et al.'s (2008) significant result of the left digit effect in individual digits, nor the degree of variation in the buy/sell ratio among the cent digits. Nonetheless, under the combined cent digits, there is a higher (lower) number of buys relative to sells for stocks that end with nine (one). This is statistically significant using both the regression and the ANOVA tests.

Although similar results were generated, a difference is observed in the level of statistical significance in Tables 5 and 8 for the three methodologies when applied to the two stock exchanges. Methodology 2, as the midpoint method with a one second time lag in bid and ask prices, yields the most significant results of the three methodologies. Nonetheless, there is no difference in the ANOVA tests of difference in variance, with all cent digit groups being significantly different from each other. More research is needed in this area of defining the trade direction of intraday data.

7.2.2 Investors do not react in the same way as consumers

As I am not able to find a statistically significant increase in the amount of buys occurring for cent digits ending in X.99, I can conclude that this psychological bias in the proportion of buys and sells of a stock is not the same as consumer behaviour towards products that end with X.99.

The explanation given by Bhattacharya et al. (2010) for differences in the buy/sell ratio is that there are excessive investors buying stocks that end in X.99, as such prices appear much cheaper than stocks that end with X.00 or X.01. In my study of the left digit effect, however, the variation to the buy/sell ratio is found to be quite limited, with a difference of around 10% for all three methodologies. Due to this I cannot conclude that this is a result of investor behaviour, as related to marketing literature.

This conclusion arises due to two main reasons. First, investors who purchase shares should not behave in the same way as consumers do when purchasing products. When a consumer is purchasing a product, they do so to satisfy a want or need. They will use and consume the product and will not consider the residual, or resale, value of the product. From this point of view, consumers faced with similar products that have the same function are likely to choose the one that is of a lower price. As mentioned in the marketing literature, consumer behaviour of excess buying of products that end with nine occurs because these prices appear much cheaper from left to right processing. When the product is a share, however, consumers will, if acting rationally, have different preferences in stocks. When purchasing a stock, the value of the product is used as a store of wealth for future appreciation and dividends, and creates an increase in personal wealth. From this perspective, the incentive of buying a stock that appears much cheaper does not enter into investor stock purchasing behavior. Furthermore, given the time an investor has to choose a stock, and considering the volatility of stock prices, cent digit endings should not have a major impact on investor decisions.

The lucky numbers hypothesis proposed in this research confirms this, as Chinese consumers have a special tendency towards prices that end with eight, as such prices are considered to be lucky. This effect is not observed in the Chinese stock market, with stock prices consisting of all fours yielding similar buy/sell ratios to stock prices that consist of all eights. My results for the lucky numbers effect are highly dependent on the methodology used to derive the numbers of buy and sell. This finding is consistent with my explanation of investor rationality towards stock prices and this

finding confirms that the increase in the number of buys and sells may not be due to investor preference, or consumer choice, and may instead be due to some other cause.

The second reason why investors do not react to stock prices in the same way as do consumers of products; as found in the marketing literature; is due to the market mechanisms of continuous trading and electronic trading. Consumer products have a price, which do not fluctuate frequently, and do not have bid and ask prices. For the major stocks, especially ones that are more liquid, stock prices change almost every second. When a market day trader is examining these prices, and they see an open price of X.99, even if they submit an immediate market order, when the order is executed it may be at a different price to that observed when the decision to bid was made. Furthermore, when stocks have the open price of X.99, when buying the share, the ask price may be a cent, or half a cent, above the price previously displayed, depending on the bid ask spread. Thus, investors will not be looking at the open price, but rather at the bid and ask prices that are offered.

The SSE and the NYSE experienced opposite markets in 2005, with the SSE experiencing a bear market and the NYSE experiencing a bull market during that year. My results for the two markets generate similar differences in the buy/sell ratios between stocks that end with one and nine. Thus, my results in regards to the left digit effect should not be affected by the current market state, or macro-economic events.

7.2.3 Explanation of difference in buy and sell volumes in terms of clustering and undercutting

Undercutting is a phenomenon where investors that are trading are willing to give up one or two cents as an opportunity cost for a faster execution of their limit orders. Bhattacharya et al. (2010) finds that the effect of undercutting is not enough to account for the buy to sell ratio imbalance of 0.8 between the combined cent digits that end with nine and one. My results from the current research are, however, different to those findings and the left digit effect may be a result of undercutting.

The difference in ratios between stocks that end with one and nine, and also with four and six, may be explained by other theories that are more logical. Clustering in the cent digits is especially observed for X.00 and X.50. There is almost twice the possibility for stocks to fall on these prices, which is consistent with the previous study findings (Harris, 1991; Ikenberry & Weston, 2007; Mitchell, 2001; Nierderhoffer, 1965).

As mentioned by Mitchell (2001), clustering can be caused by individual number preferences. The discipline of psychology has found that this is due to the human memory remembering and using whole numbers more easily and readily than mixed numbers. This act of clustering can be due to investor expectations. Nonetheless, the previous articles do not find any differences in return due to clustering. My results show that there is a difference in the buy/sell ratio observed in both countries and, furthermore, clustering is observed at every interval of cent digits that end in zero and five. The clustering frequency is highest at the cent digit of X.00 and X.50. Under these findings there may be a relation between clustering, stock returns, and differences in the number of buy and sell results in cent digit for the three different hypothesizes. Clarification of whether the stock is buyer initiated, or seller initiated, is

66

necessary. A stock is classified as buyer (seller) initiated if it is the buyer (seller) who closes the trade. Normally, a stock is buyer initiated if it is triggered by a market order to buy.

Limit orders can be a cause of the observed return difference and the buy sell imbalance. With the advancement of electronic trading and market mechanisms, many traders are familiar with the term of limit orders. When submitting a limit order to buy, the electronic trading system will execute the order when the price matched is either at, or lower than, the limit price and is triggered by a market order to sell (Domowitz, 1990). Similarly, with a limit order to sell, the trading system will allow orders to be executed at, or higher than, the sell price and this can be offset by market orders to buy. The previous clustering literature (Harris, 1991; Ikenberry & Weston, 2007; Mitchell, 2001; Nierderhoffer, 1965) documents the likelihood of investors setting their limit orders on whole numbers such as X.00 or X.50, as these prices are easier to remember than are X.01 or X.51. This is also due to the low transaction costs associated with executing the order at the cluster, with a large number of orders clustering at these cent digits.

There is a distinct pattern in the frequency of appearances, wherein the stocks that end with zero (five) is most likely to occur, followed by one and nine (four and six), which are directly beside the first two cent digits. This may be a possible explanation for the differences in the proportion of buy and sells, and also their returns. The continuous auction system of the NYSE and the SSE all have orders of buy and sell clusters at the round numbers X.00 and X.50. When orders cluster at the round number, those investors who want to put in a limit order at X.00 may need to wait in

the queue, and their order may not be executed due to the large number of orders preceding them in the queue.

Assuming rationality in investors, clustering of orders on round numbers (Harris, 1991; Ikenberry & Weston, 2007; Nierderhoffer, 1965) can lead to those investors with limit orders with buyers (sellers) who place limit orders to buy (sell) will be likely to set the price one cent above (below) the round number to obtain faster execution of their orders. This theory of investors' rational behaviour in limit orders is justified in Johnson et al. (2008, pp. 33), and it is determined that "…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 or1", and vice versa. This presents a probable explanation for there being more buys placed by investors via limit orders on prices just above a round number, and more sells placed through orders that are just below the cluster or for stocks that end with nine or four,. This is done so that faster execution of the orders can be achieved.

Looking at the stock prices of both countries, the one cent difference in the stock price is not of much importance and does not affect the overall return of the stock. Thus, this will be why many investors have an incentive to set limit buy (sell) orders one cent, or two cents, above (below) the clustering price for a guaranteed execution at a controlled price.

In order to prove this explanation of investor behaviour I randomly select five companies in my sample from the NYSE over two weeks and download the intraday trading data in order to determine how many share orders came in at each of the bid and ask prices during the period. Defined as above, a bid price is the price that the investor gets if they place a market sell order, and the ask price is the price that investors trade on if they place a market buy order. Thus, orders that come in on bid prices are orders to sell, and orders that come in on ask prices are orders to buy. Although, I cannot distinguish which are the limit orders, or whether those orders were executed, from the data extracted, this method does give an overview of how many investors want to buy and how many want to sell at prices ending with digits from zero to nine.

Figure 21 (see Appendix) shows the accumulated orders in terms of the number of shares for a stock with combined cent digit endings from zero to nine over these fourteen days. Figure 22 shows the accumulated orders for each individual cent digit for the same five companies over the fourteen days. From the two figures it can be seen that the clustering of orders at every cent that ends with zero and five is observed, with the digits adjacent to these having clusters of orders. In respect to the buying and selling of the stocks, it is obvious that when the price ends with one, there are significantly more buy orders than sell orders at the ask price, and vice versa, for stocks that end with nine. Similar patterns are observed in stocks that end with four and six. This conforms with the theory of rational investors and that they use limit orders to buy stocks that are one cent away from the cluster, so that their orders can be executed more efficiently. While investors are placing limit orders to buy (sell) at

X.01 (X.99), these orders should be triggered by market orders to sell (buy), thus, resulting in the stock being defined as buyer initiated (seller initiated).

8. LIMITATIONS

There are some limitations in this thesis regarding data extraction and selection, which are outlined below.

As I remove all times and dates for macro-economic events from the data, the impact of significant events, or dividend payouts, may cause prices to decline, or increase, over a day or two, which may have an impact on the return results and the number of buy or sells of a share. The macro-economic events, however, should not cause great concern as, when information comes in, a buyer still has the choice of deciding which cent digit to sell the stock on.

The intraday data I use is drawn from only one year of intraday data. An extension of the time period used for the data set may generate different results, as decimalisation of stock prices in the US occurred in 2001, and the left digit effect, as observed Bhattacharya et al. (2010), may be occur at the beginning of decimalisation, or may be due to different market trends. Nonetheless, the SSE and the NYSE experience opposite trends in stock indices, with differences in buy to sell ratios still remaining around 0.1. The total number data points for my data set amount to over 70 million. When compared to the 100 million data points in the data set of Bhattacharya et al. (2010) the results should not differ significantly.

I use raw returns rather than the midpoint of the bid and ask to calculate the returns effect. This may significantly reduce the difference in the returns within one minute.

When the return period is extended to twenty-four hours however, no obvious difference in the return results is generated by the bid ask midpoint and the raw return method.

9. CONCLUSION

In conclusion, this thesis investigates the investor psychological bias towards numbers in the financial market. I find a returns effect towards different cent digits where, for both countries' markets (the NYSE for the US and the SSE for China), prices that end with X.01 yield a much higher return than those that end with X.99. This effect persists at least 24 hours, with an annualised return difference of 22.65% for the NYSE and 54.43% for the SSE. My findings are consistent with those of Johnson et al. (2008) and the findings regarding end of day stock prices (Wang, 2009). This phenomenon can generate abnormal returns when not accounting for transaction costs, which is in opposition to the efficient market hypothesis, as returns are predictable on observation of cent digit endings. More research is needed in this area to account for transaction costs and to determine whether this is in automatic violation of the efficient market hypothesis.

There are more buyer initiated stocks for stocks ending with nine and less for stocks that end with one. Nevertheless, I have not found a significant left digit effect in individual cent digits for either country, where the degree of difference in the buy/sell ratios remains within 0.1 and is consistent with Johnson et al.'s (2008) finding of a difference of around 7% and Wang's (2009) finding, but is in contradiction to the finding of Bhattacharya et al (2010). Similarly, for the combined cent digits, although the results are significant, the difference in the buy to sell ratio remains within 0.1 basis points. There are three different methodologies used in this research to define trade direction. All three methods generate similar results to the number of buy and sells, with the different methods not affecting the significance of the left digit effect. It

would be difficult to determine which method is the most appropriate, as all three methods are used in research papers published in top journals. From the small difference in the results I can conclude, however, that my results are not biased by methodological differences. This small difference in the buy pattern may not be caused by irrational retail consumer behavior, as claimed in the marketing literature and, as concluded by Bhattacharya et al. (2010), it may not be caused by the left digit effect. I did not find a lucky numbers effect in stock prices in China. This seems logical given that my research concludes that the differences in buy and sell volumes in cent digit endings is not due to marketing consumer behavior, but instead due to rational incentives.

My findings contradict Bagnoli et al. (2006), in their conclusion of a net sell during the overnight period for stocks ending with nine, because the combined cent digits from my research have generated a significant result for more buyer initiated stocks ending with nine than for those ending with other digits. When explained in terms of undercutting there may, however, be more investors putting in sell orders for stocks that end with nine, so that such orders will be executed faster in the opening call auction, with the net sell making sense when considered in terms of this. Clustering is also observed in both countries, with almost twice the frequency of stock prices occurring with X.00 and X.50 than with other digit endings. From my discussion it is seen that investors like to put orders in on prices that end with X.00 and X.50, and that it seems rational for investors to place limit orders one cent away from the cluster to obtain faster execution of their orders. From the return results it is also suggested that this action may be more rational than is found in the marketing literatures in relation to consumer retail behaviour. This can be a cause of the different buy to sell

volume for cent digits ending with nine and one. More research is needed in this area to investigate the relation between clustering and the number of buy and sells towards the different cent digits.

10. BIBLIOGRAPHY

Aggarwal, R. & Lucey, B. M. (2007). Psychological barriers in gold prices? *Review of Financial Economics, 16*(2), 217-230.

Baberis, N. & Huang, M. (2001). Mental Accounting, Loss Aversion,and Individual Stock Returns. *The Journal of Finance, 56(4)*, 1247-1291.

Bagnoli, M., Park, J. & Watts, S. G. (2006). Nines in the Endings of Stock Prices. Working paper. Retrieved 01/07/2009 from SSRN working paper series: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=728544

Barber, B. M. & Odean, T. (2001). Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment. *The Quarterly Journal of Economics, 116*(1), 261-292.

Barberis, N. & Thaler, R. (2002). *A Survey of Behavioural Finance.* Retrieved 07 22, 2009, from NBER working paper series: http://www.nber.org/papers/w9222

Bessembinder, H. (2003). Issues in assessing trade execution costs. *Journal of Financial Markets*, *6(3),* 233–257.

Bhattacharya, U., Holden, C. W. & Jacobson, S. (2008). Penny Wise, Dollar Foolish: The Left-Digit Effect in Security Trading . Working Paper. Retrieved 20/07/2009 from SSRN working paper series: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1303700

Bhattacharya, U., Holden, C. W. & Jacobson, S. (2010). Penny Wise, Dollar Foolish: The Left-Digit Effect in Security Trading. Working paper. Retrieved 03/06/2010 from SSRN working paper series: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1303700

Bousfield, W. A. & Cohen, B. H. (1955). The occurrence of clustering in the recall of randomly arranged words of different frequencies-of-usage. *Journal of General Psychology, 52*, 83-95.

Breeden, D. T. & Litzenberger, R. H. (1978). Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business, 51*(4), 621-651.

Brenner, G. A. & Brenner, R. (1982). Memory and markets, or why are you paying $2.99 for a widget? *Journal of Business, 55*(1), 147-158.

Coulter, K. S. (2001). Odd-ending price underestimation: An experimental examination of left-to-right processing effects. *Journal of Product and Brand Management,10(5)*, 276-292.

De Bondt, W. F. & Thaler, R. (1985). Does the stock market overreact? *Journal of finance, 40*(3), 793-805.

De Ceuster, M. J., Dhaene, G. & Schatteman, T. (1998). On the hypothesis of psychological barriers in stock markets and Benford's Law. *Journal of Empirical Finance, 5(3)*, 263-279.

De Grauwe, P. & Decupere, D. (1992). Psychological barriers in the foreign exchange markets. *Journal of International and Comparative Economics, 9*, 87-101.

Domowitz, I. (1990). The mechanics of automated trade execution systems. *Journal of Financial Intermediation, 1*(2), 167-194.

Donaldson, R. G. & Kim, H. Y. (1993). Price barriers in the Dow–Jones industrial average. *Journal of Financial and Quantitative Analysis, 28*(3), 313-330.

Edmans, A., Garcia, D. & Norli, Ø. (2007). Sports sentiment and stock returns. *Journal of Finance, 62*(4), 1967-1998.

Erevelles, S., Roy, A. & Yip, L. S. (2001). The universality of the signal theory for products and services. *Journal of Business Research, 52*(2), 175-187.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance, 25*(2), 383-417.

Fuller, R. J. (1998). Behavioral Finance and the Sources of Alpha. *Journal of Pension Plan Investing, 2*(3), 15-22.

Gedenk, K. & Sattler, H. (1999). The Impact of Price Thresholds on Profit Contribution-Should retailers set 9-ending prices. *Journal of Retailing, 75(1)*, 33-57.

Harris, L. (1989). A Day-End Transaction Price Anomaly. *The Journal of Financial and Quantitative Analysis, 24(1)*, 29-45.

Harris, L. (1991). Stock price clustering and discreteness. *Review of financial studies, 4*(3), 389-415.

Hayes, R. J. (1952). *Memory span for several vocabularies as a function of vocabulary size.* Quarterly Progress Report, Massachusetts Institute of Technology, Acoustics Laboratory, Cambridge.

Heider, E. H. (1972). Universals in color naming and memory. *Journal of Experimental Psychology, 93*, 10-20.

Henker, T. & Wang, J. X. (2006). On the importance of timing specifications in market microstructure research. *Journal of Financial Markets, 9*(2), 162-179.

Huang, R. D. & Stoll, H. R. (1997). The components of the bid-ask spread: a general approach. *The Review of Financial Studies, 10*(4), 995-1034.

Ikenberry, D. & Weston, J. (2007). Clustering in U.S. stock prices after decimalization. *European Financial Management, 14*(1), 30-54.

Johnson, E., Johnson, N. B. & Shanthikumar, D. (2008). Round Numbers and Security Returns. Retreived 22/04/2009 from : http://www.people.hbs.edu/dshanthikumar/RoundNumbers.pdf

Kahneman, D. & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Journal of Econometrica, 47*(2), 263-292.

Koedijk, K. G. & Stork, P. A. (1994). Should We Care? Psychological barriers in stock markets. *Economic Letters, 44*, 427-432.

Lambert, Z. V. (1975). Perceived prices as to odd and even price endings. *Journal of Retailing, 51*, 13-22.

Lee, C. M. & Ready, M. J. (1991). Inferring Trade Direction from Intraday Data. *The Journal of Finance, 46(2)*, 733-746.

Ley, E. & Varian, H. (1994). Are there psychological barriers in the Dow-Jones index? *Applied Financial Economics, 4*, 217-224.

Lip, E. (1992). *Chinese Numbers: Sigificance, symbolism and traditions.* Singapore: Times Books International.

Madhavan, A. (1992). Trading Mechanisms in Securities Markets. *The Journal of Finance, 47*(2), 607-641.

Manning, K. C. & Sprott, D. E. (2009). Price endings, Left Digit effects, and Choice. *Journal of Consumer Research, 36(2)*, 328-335.

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy, 60*, 151-158.

Merton, R. C. (1973). An intertemperal capital asset pricing model. *Journal of Econometrica, 41*(5), 867-887.

Mitchell, J. (2001). Clustering and Psychological Barriers: The Importance of Numbers. *Journal of Futures Markets, 21*(5), 395-428.

Monroe, K. B. (1973, 2). Buyers' Subjective Perceptions of Price. *Journal of Marketing Research*, 70-80.

Monroe, K. B. & Lee, A. Y. (1999). Remembering Versus Knowing: Issues in Buyers' Processing of Price Information. *Journal of The Academy of Marketing Science, 27*(2), 207-225.

Muhammad, N. M. (2009). Behavioural finance vs traditional finance. *Advanced Management Journal*, 1-10.

Niederhoffer, V. (1965). Clustering of stock prices. *Operations research, 13*, 258-265.

Pollack, I. (1953). The information of elementary auditory displays. *Journal of the Acoustical Society of America, 25*, 765-769.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology, 7*, 532-547.

Schindler, R. M. & Kibarian, T. M. (1997). Testing for Perceptual Underestimation of 9-Ending Prices. *Advances in consumer research, 20*, 580-585.

Schindler, R. M. & Wiman, A. R. (1989). Effect of Odd pricing on price recall. *Journal of Business Research, 19*, 165-177.

Shiller, R. J. (2003). From efficient markets theory to behavioural finance. *Journal of Economic Prospectives, 17*(1), 83-104.

Shindler, R. M. (2006). The 99 price ending as a signal of a low-price appeal. *Journal of Retailing, 82(1)*, 71-77.

Shindler, R. M. & Kirby, P. N. (1997). Patterns of Rightmost Digits Used in Advertising Prices: Implications for Nine-Ending effects. *Journal of Consumer Research, 24*, September, 192-201.

Shindler, R. M. & Wiman, A. R. (1989). Effect of Odd Pricing on Price Recall. *Journal of Business Research, 19*, 165-177.

Simmons, L. C. & Schindler, R. M. (2003). Cultural Superstitions and the price endings used in chinese advertising. *Journal of International Marketing, 2*, 101-111.

Stiving, M. & Winer, R. S. (1997, June). An Emperical Analysis of Price Endings with Scanner Data. *Journal of Consumer Research, 24*, 57-67.

Thomas, M. & Morwitz, V. (2005). Penny Wise and Pound Foolish: The Left-Digit Effect in Price Cognition. *Journal of Consumer Research, 32,*June, 54-64.

Wang, A. L. Q. (2009). *Behavior of different cent digits in security trading for America, Australia and Europe* (Unpublished thesis). Massey University, Auckland, New Zealand.

11. APPENDIX

Table 1: Full table of return percentages for individual cent digits over 24 hours in the NYSE

	Coefficient	P-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
	1 minute		30 minute		60 minute		24 hours	
d01	0.00341	<.0001	0.000774	0.5212	-0.00395	0.0141	0.03085	<.0001
d02	0.00241	<.0001	0.00427	0.0007	0.000572	0.7329	0.02342	<.0001
d03	-0.00012	0.6455	-0.00174	0.1728	-0.00151	0.375	0.02061	<.0001
d04	-0.00173	<.0001	-0.00032	0.8018	0.00106	0.5352	0.01264	0.0037
d05	-0.00012	0.624	0.00253	0.0348	0.00111	0.4873	0.02943	<.0001
d06	0.000971	0.0002	0.00686	<.0001	0.00793	<.0001	0.04012	<.0001
d07	0.000108	0.6826	0.00787	<.0001	0.01422	<.0001	0.0435	<.0001
d08	-0.0016	<.0001	0.0063	<.0001	0.01194	<.0001	0.04321	<.0001
d09	-0.00243	<.0001	0.00268	0.0354	0.00756	<.0001	0.04394	<.0001
d10	-0.00055	0.0216	0.00331	0.0045	0.00394	0.0111	0.05395	<.0001
d11	0.00223	<.0001	0.01104	<.0001	0.01053	<.0001	0.06345	<.0001
d12	0.00136	<.0001	0.00994	<.0001	0.01194	<.0001	0.06634	<.0001
d13	-0.00048	0.0731	0.00831	<.0001	0.01244	<.0001	0.05529	<.0001
d14	-0.00219	<.0001	0.00664	<.0001	0.01085	<.0001	0.0628	<.0001
d15	-0.00087	0.0004	0.00477	<.0001	0.00967	<.0001	0.06554	<.0001
d16	0.00097	0.0002	0.01042	<.0001	0.01306	<.0001	0.0597	<.0001
d17	-0.0003	0.2638	0.00598	<.0001	0.01025	<.0001	0.06731	<.0001
d18	-0.00174	<.0001	0.00106	0.4108	0.00413	0.0164	0.05423	<.0001
d19	-0.00225	<.0001	-0.00207	0.1058	0.00313	0.0666	0.04977	<.0001
d20	-0.00096	<.0001	0.000348	0.7655	0.00265	0.0884	0.03475	<.0001
d21	0.00223	<.0001	0.0035	0.0066	0.00185	0.2823	0.02786	<.0001
d22	0.000383	0.15	0.00341	0.0086	0.00572	0.001	0.00703	0.1115
d23	-0.00121	<.0001	0.00535	<.0001	0.01175	<.0001	0.00563	0.2032
d24	-0.00254	<.0001	-0.00253	0.0485	0.000452	0.7919	-0.01167	0.0075
d25	-0.00038	0.1101	-0.00143	0.2221	0.00159	0.3072	0.00367	0.3554
d26	0.00249	<.0001	0.00892	<.0001	0.0131	<.0001	0.00368	0.3952
d27	0.000713	0.0072	0.00517	<.0001	0.01125	<.0001	0.01072	0.015
d28	-0.00063	0.017	-0.00484	0.0002	-0.00023	0.8958	-0.01609	0.0003
d29	-0.00228	<.0001	-0.00408	0.0016	0.00103	0.5502	-0.01521	0.0005
d30	0.000763	0.0015	-0.0011	0.3472	0.00217	0.1648	0.000809	0.8393
d31	0.00268	<.0001	0.00541	<.0001	0.00769	<.0001	0.01359	0.0019
d32	-0.00018	0.5133	0.00561	<.0001	0.00864	<.0001	0.01318	0.003
d33	-0.00082	0.0024	8.27E-05	0.95	0.00429	0.0146	0.000789	0.8601
d34	-0.00211	<.0001	0.000707	0.5886	0.00273	0.1178	-0.00724	0.1036
d35	0.000149	0.55	-0.00059	0.6267	0.00163	0.3155	-0.00387	0.3479

d36	0.00176	<.0001	0.00616	<.0001	0.00823	<.0001	0.01859	<.0001
d37	5.8E-05	0.8285	-0.0021	0.1086	0.00179	0.3039	-0.00143	0.7472
d38	-0.00144	<.0001	-0.00377	0.0037	0.00404	0.0194	0.00667	0.1301
d39	-0.00284	<.0001	-0.00394	0.0021	0.0043	0.012	-0.0053	0.2246
d40	-0.00033	0.1676	0.00767	<.0001	0.01158	<.0001	-0.00957	0.0159
d41	0.00212	<.0001	0.01604	<.0001	0.02055	<.0001	-0.01339	0.0022
d42	0.000938	0.0005	0.0089	<.0001	0.01507	<.0001	-0.00589	0.1846
d43	-0.00121	<.0001	0.00244	0.0626	0.00759	<.0001	0.00569	0.2018
d44	-0.00165	<.0001	0.0031	0.0172	0.00523	0.0026	0.00786	0.0754
d45	-0.00109	<.0001	0.00167	0.1688	0.00621	0.0001	0.00327	0.4277
d46	0.000678	0.0103	0.000478	0.7106	0.00447	0.0093	-5.5E-05	0.99
d47	-0.00105	<.0001	0.00266	0.0399	0.0037	0.0327	-0.01548	0.0004
d48	-0.0027	<.0001	0.00146	0.2527	0.000957	0.5744	-0.0174	<.0001
d49	-0.00327	<.0001	-0.00179	0.1538	-0.00844	<.0001	-0.03569	<.0001
d50	-7.3E-05	0.749	0.00242	0.0305	-0.00289	0.0524	-0.03005	<.0001
d51	0.0034	<.0001	0.0129	<.0001	0.00948	<.0001	-0.0149	0.0005
d52	0.00316	<.0001	0.01199	<.0001	0.01518	<.0001	-0.00087	0.8432
d53	0.000965	0.0003	0.01104	<.0001	0.01205	<.0001	-0.01374	0.002
d54	-0.00165	<.0001	0.00592	<.0001	0.0058	0.0009	-0.01122	0.0114
d55	-8E-05	0.7505	0.00592	<.0001	0.00743	<.0001	0.0102	0.0141
d56	0.00152	<.0001	0.01028	<.0001	0.0074	<.0001	-0.00758	0.0861
d57	0.000242	0.3681	0.01043	<.0001	0.00838	<.0001	-0.01442	0.0012
d58	-0.00116	<.0001	0.00414	0.0016	0.0037	0.0343	-0.02151	<.0001
d59	-0.00221	<.0001	0.000482	0.7103	0.00184	0.2869	-0.01509	0.0006
d60	0.000165	0.4928	0.00567	<.0001	0.00458	0.0034	-0.01438	0.0003
d61	0.00211	<.0001	0.01295	<.0001	0.01326	<.0001	0.00406	0.3533
d62	0.000627	0.0194	0.00952	<.0001	0.01071	<.0001	-0.00599	0.178
d63	-0.00037	0.1661	0.00753	<.0001	0.00876	<.0001	-0.01369	0.0022
d64	-0.00231	<.0001	0.00572	<.0001	0.0082	<.0001	-0.00776	0.0794
d65	-0.00056	0.0245	0.00335	0.0057	0.00209	0.1975	-0.01253	0.0024
d66	0.0018	<.0001	0.00567	<.0001	0.00581	0.0008	-0.00567	0.1974
d67	0.000281	0.2941	0.00953	<.0001	0.01091	<.0001	0.00389	0.3814
d68	-0.00156	<.0001	0.00524	<.0001	0.00588	0.0007	0.00104	0.8136
d69	-0.00252	<.0001	0.00361	0.0049	0.00522	0.0023	0.013	0.0029
d70	-0.00051	0.032	0.00141	0.2263	0.00241	0.1215	-0.01258	0.0015
d71	0.00185	<.0001	0.00555	<.0001	0.01	<0.0001	0.0852	0.0514
d72	0.000123	0.6417	0.00604	<.0001	0.00777	<.0001	0.00362	0.4107
d73	-0.00147	<.0001	-0.00532	<.0001	-0.00753	<.0001	-0.00317	0.4748
d74	-0.00295	<.0001	-0.00773	<.0001	-0.01415	<.0001	-0.00853	0.0508
d75	-0.00029	0.226	-0.00107	0.3634	-0.01181	<.0001	-0.02598	<.0001
d76	0.00201	<.0001	-0.00276	0.0309	-0.00805	<.0001	-0.01817	<.0001
d77	1.53E-05	0.9542	-0.00044	0.7349	0.00192	0.2679	-0.01039	0.0187
d78	-0.00043	0.105	-0.00471	0.0003	-0.0054	0.0018	-0.03636	<.0001
d79	-0.00244	<.0001	-0.00106	0.4119	-0.00341	0.0474	-0.03363	<.0001
d80	0.000669	0.005	0.0036	0.002	0.00391	0.0117	-0.0311	<.0001

d81	0.00298	<.0001	0.0112	<.0001	0.01341	<.0001	-0.04327	<.0001
d82	0.00135	<.0001	0.00358	0.0057	0.00219	0.2047	-0.05244	<.0001
d83	-0.0007	0.0085	0.00311	0.0169	0.00337	0.0524	-0.05342	<.0001
d84	-0.00192	<.0001	-0.00277	0.0321	-0.0038	0.0273	-0.05355	<.0001
d85	0.000119	0.6267	-0.00063	0.6004	-0.00471	0.0032	-0.03757	<.0001
d86	0.000932	0.0004	0.00272	0.0333	-0.00061	0.7218	-0.02348	<.0001
d87	4.11E-07	0.9988	0.00178	0.17	-0.00107	0.5366	-0.02384	<.0001
d88	-0.00165	<.0001	-0.00049	0.6995	-0.00178	0.2981	-0.0309	<.0001
d89	-0.00195	<.0001	-0.00136	0.2852	-0.00306	0.0709	-0.0353	<.0001
d90	-0.00024	0.3024	0.00328	0.0046	0.00711	<.0001	-0.01906	<.0001
d91	0.00282	<.0001	0.00678	<.0001	0.0045	0.0078	-0.00954	0.0269
d92	-0.00104	<.0001	2.72E-05	0.9831	-0.00093	0.5871	0.02485	<.0001
d93	-0.00057	0.0308	-0.00521	<.0001	-0.00308	0.0715	-0.03007	<.0001
d94	-0.00196	<.0001	-0.00621	<.0001	-0.00559	0.001	-0.05877	<.0001
d95	-0.00074	0.0024	-0.00369	0.0018	-0.00468	0.0031	-0.05407	<.0001
d96	-0.00017	0.5139	0.000668	0.5936	-0.00067	0.6869	-0.05141	<.0001
d97	-0.00121	<.0001	-0.00279	0.0254	-0.00821	<.0001	-0.05113	<.0001
d98	-0.0032	<.0001	-0.00629	<.0001	-0.00773	<.0001	-0.05155	<.0001
d99	-0.00429	<.0001	-0.00848	<.0001	-0.01072	<.0001	-0.05085	<.0001

Table 2: Full table of return percentageS for individual cent digits over 24 hours in the SSE

	Coefficient	P-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
	1 minute		30 minute		60 minute		24 hours	
d01	0.00915	<.0001	0.02476	<.0001	0.03979	<.0001	0.10292	<.0001
d02	0.00509	0.0014	0.02758	<.0001	0.04897	<.0001	0.10576	<.0001
d03	-0.00248	0.1149	0.00442	0.4651	0.01259	0.1005	0.04546	0.0016
d04	-0.00846	<.0001	-0.00378	0.5406	-0.00945	0.2274	0.08224	<.0001
d05	-0.00418	0.0051	-0.01216	0.034	-0.02038	0.005	0.0451	0.0009
d06	0.00169	0.2765	-0.01592	0.0077	-0.02166	0.0042	0.04649	0.0011
d07	-0.00567	0.0004	-0.0035	0.5664	0.00154	0.8422	0.05209	0.0003
d08	-0.00698	<.0001	-0.01865	0.0012	-0.03152	<.0001	-0.00934	0.4941
d09	-0.01041	<.0001	-0.02631	<.0001	-0.04239	<.0001	-0.00662	0.6418
d10	-0.00067	0.6404	-0.01438	0.0087	-0.02044	0.0032	-0.00179	0.8905
d11	-0.00373	0.0167	-0.04467	<.0001	-0.02389	0.0016	0.10458	<.0001
d12	0.000192	0.8999	-0.03405	<.0001	-0.05228	<.0001	0.13864	<.0001
d13	-0.00019	0.8995	-0.02774	<.0001	-0.05384	<.0001	0.09635	<.0001
d14	-0.0131	<.0001	-0.02219	0.0002	-0.04571	<.0001	0.12882	<.0001
d15	-0.01002	<.0001	-0.03051	<.0001	-0.04342	<.0001	0.09912	<.0001
d16	-0.00461	0.0024	-0.0895	<.0001	-0.16445	<.0001	-0.03538	0.0109
d17	-0.00559	0.0003	-0.07012	<.0001	-0.10883	<.0001	-0.07336	<.0001
d18	-0.00712	<.0001	-0.02731	<.0001	-0.02748	0.0001	0.03579	0.0078
d19	-0.01066	<.0001	0.00273	0.6464	0.02039	0.0069	0.11462	<.0001
d20	-0.00277	0.0511	0.01672	0.0022	0.0328	<.0001	0.08055	<.0001
d21	0.00499	0.0019	0.02631	<.0001	0.03761	<.0001	0.06506	<.0001
d22	0.00126	0.4221	-0.00998	0.0988	-0.00317	0.6788	0.066	<.0001
d23	-0.00055	0.7257	-0.00743	0.2204	0.0062	0.4189	0.05754	<.0001
d24	-0.00962	<.0001	-0.01331	0.0302	0.00564	0.4683	-0.00045	0.9755
d25	-0.00487	0.001	0.00905	0.1105	0.0264	0.0002	0.08051	<.0001
d26	-0.00034	0.8272	0.01357	0.0217	0.02123	0.0046	0.019	0.176
d27	-0.00417	0.0087	0.00925	0.1303	0.01636	0.0346	0.02316	0.111
d28	-0.00479	0.0015	0.02261	<.0001	0.0179	0.0147	0.00402	0.7703
d29	-0.00886	<.0001	0.00396	0.5078	0.01364	0.0717	-0.0182	0.2003
d30	-0.0035	0.012	-0.00455	0.396	-0.00705	0.2996	-0.04445	0.0005
d31	0.00319	0.0426	-0.00306	0.6139	0.0052	0.4979	-0.02614	0.0694
d32	-0.00336	0.031	-0.01559	0.0093	-0.02376	0.0018	-0.07904	<.0001
d33	-0.00678	<.0001	-0.04943	<.0001	-0.07082	<.0001	-0.1757	<.0001
d34	-0.01143	<.0001	-0.03288	<.0001	-0.05	<.0001	-0.11321	<.0001
d35	-0.00501	0.0007	-0.03101	<.0001	-0.05923	<.0001	-0.14748	<.0001
d36	0.000243	0.8742	-0.0439	<.0001	-0.05597	<.0001	-0.16463	<.0001
d37	-0.00967	<.0001	-0.07295	<.0001	-0.07855	<.0001	-0.14378	<.0001
d38	-0.01211	<.0001	-0.05432	<.0001	-0.05409	<.0001	-0.12353	<.0001
d39	-0.01027	<.0001	-0.02314	0.0001	-0.00657	0.387	-0.07822	<.0001

d40	-0.003	0.0353	-0.0289	<.0001	-0.01423	0.0404	-0.02855	0.0284
d41	-0.00087	0.589	-0.0123	0.0473	0.00555	0.4797	0.04301	0.0035
d42	-0.00315	0.0477	-0.00494	0.419	0.02041	0.0084	-0.08357	<.0001
d43	-0.00616	0.0001	-0.0105	0.0889	0.00922	0.238	-0.1507	<.0001
d44	-0.01654	<.0001	-0.01836	0.0078	-0.00662	0.4482	-0.14597	<.0001
d45	-0.00979	<.0001	-0.02902	<.0001	-0.01194	0.1026	-0.13062	<.0001
d46	0.0044	0.0053	-0.01232	0.0424	-0.01333	0.0828	-0.13569	<.0001
d47	-0.00583	0.0003	-0.02641	<.0001	-0.03295	<.0001	-0.14741	<.0001
d48	-0.00678	<.0001	-0.02891	<.0001	-0.04474	<.0001	-0.13155	<.0001
d49	-0.01093	<.0001	-0.02743	<.0001	-0.02437	0.0016	-0.11483	<.0001
d50	-0.00189	0.1778	-0.01588	0.0033	-0.01446	0.0346	-0.12473	<.0001
d51	0.00398	0.0133	-0.01147	0.0635	-0.00345	0.6593	-0.15123	<.0001
d52	-0.00062	0.6981	-0.02754	<.0001	-0.01859	0.0169	-0.04261	0.0035
d53	-0.0053	0.001	-0.03241	<.0001	-0.00684	0.3845	-0.00286	0.8463
d54	-0.01178	<.0001	-0.02552	<.0001	-0.02146	0.0076	0.02337	0.1213
d55	-0.01052	<.0001	-0.02071	0.0004	-0.00928	0.2099	-0.02993	0.0311
d56	0.0021	0.1901	-0.00749	0.2235	0.0112	0.1505	0.03231	0.0271
d57	-0.00992	<.0001	-0.01176	0.0651	0.00683	0.3976	0.04791	0.0016
d58	-0.01066	<.0001	-0.0183	0.0022	-0.00916	0.2274	0.03638	0.0106
d59	-0.01067	<.0001	-0.01267	0.0425	0.01164	0.1413	0.02839	0.0557
d60	-0.00249	0.0853	-0.00906	0.1029	-0.0087	0.2166	0.02058	0.119
d61	0.00627	0.0001	0.00119	0.8512	-0.0096	0.2327	0.03497	0.0205
d62	-0.00309	0.0568	-0.02255	0.0003	-0.03562	<.0001	-0.00375	0.8007
d63	-0.00734	<.0001	-0.06465	<.0001	-0.0887	<.0001	-0.24974	<.0001
d64	-0.01488	<.0001	-0.0233	0.0003	-0.01532	0.0578	-0.03935	0.0094
d65	-0.01324	<.0001	-0.00765	0.1961	-0.0083	0.2681	-0.02023	0.1502
d66	0.00576	0.0004	0.01625	0.0089	0.01693	0.0313	-0.00552	0.7081
d67	-0.0038	0.0224	0.0082	0.2	-0.00654	0.4197	-0.03502	0.0212
d68	-0.01039	<.0001	-0.02564	<.0001	-0.03756	<.0001	-0.1027	<.0001
d69	-0.01077	<.0001	-0.03726	<.0001	-0.04687	<.0001	-0.1228	<.0001
d70	-0.0041	0.0045	-0.06086	<.0001	-0.07418	<.0001	-0.15446	<.0001
d71	0.00429	0.0094	-0.06701	<.0001	-0.07139	<.0001	-0.09485	<.0001
d72	-0.00053	0.745	-0.02859	<.0001	-0.03176	<.0001	-0.08604	<.0001
d73	-0.00109	0.4969	-0.02353	0.0001	-0.03536	<.0001	-0.09095	<.0001
d74	-0.0142	<.0001	-0.00836	0.1859	-0.03467	<.0001	-0.10666	<.0001
d75	-0.00722	<.0001	-0.00174	0.7653	-0.00344	0.6407	-0.10725	<.0001
d76	0.00252	0.1177	-0.00363	0.5572	-0.00895	0.2537	-0.10449	<.0001
d77	-0.00492	0.0027	0.01747	0.0057	0.00877	0.2726	-0.06236	<.0001
d78	-0.00871	<.0001	0.01159	0.0517	0.00379	0.6157	-0.1051	<.0001
d79	-0.01292	<.0001	-0.00947	0.1248	-0.02675	0.0006	-0.1123	<.0001
d80	-0.00179	0.208	-0.00843	0.1229	-0.03021	<.0001	-0.09251	<.0001
d81	0.00415	0.0101	0.00195	0.7535	0.000144	0.9854	-0.00597	0.6853
d82	-0.00178	0.259	-0.06799	<.0001	-0.07728	<.0001	-0.22798	<.0001
d83	-0.00018	0.9112	-0.00902	0.1394	-0.00569	0.4619	-0.01329	0.3594
d84	-0.01158	<.0001	-0.00954	0.1215	0.01881	0.016	0.07713	<.0001

d85	-0.00805	<.0001	-0.01026	0.0751	-0.00151	0.836	-0.03208	0.0191
d86	0.0013	0.4075	0.01209	0.0456	0.02296	0.0027	0.000482	0.9732
d87	-0.0053	0.001	0.01592	0.0104	0.04265	<.0001	0.00626	0.6714
d88	-0.01283	<.0001	-0.02326	<.0001	-0.01652	0.0268	-0.0647	<.0001
d89	-0.00922	<.0001	-0.03139	<.0001	-0.01734	0.025	-0.04145	0.0043
d90	-0.00578	<.0001	-0.02495	<.0001	-0.00524	0.45	-0.12136	<.0001
d91	0.000765	0.6302	-0.02447	<.0001	-0.0152	0.0497	-0.062	<.0001
d92	0.000974	0.5318	0.04002	<.0001	0.03213	<.0001	0.11157	<.0001
d93	-0.00908	<.0001	-0.03469	<.0001	-0.03404	<.0001	-0.23643	<.0001
d94	-0.01707	<.0001	-0.0441	<.0001	-0.04102	<.0001	-0.21106	<.0001
d95	-0.01319	<.0001	-0.05888	<.0001	-0.07	<.0001	-0.1656	<.0001
d96	-0.00589	0.0002	-0.05901	<.0001	-0.07059	<.0001	-0.12433	<.0001
d97	-0.00837	<.0001	-0.06387	<.0001	-0.08695	<.0001	-0.13756	<.0001
d98	-0.01223	<.0001	-0.04784	<.0001	-0.06305	<.0001	-0.13337	<.0001
d99	-0.01161	<.0001	-0.03038	<.0001	-0.02414	0.0022	-0.07125	<.0001
four	0.0054	0.0716	0.01087	0.3462	0.02161	0.1391	-0.14872	<.0001
eight	0.01052	0.0586	0.10229	<.0001	0.10535	0.0001	-0.08957	0.0781

Figure 1: Bar graph of buy/sell ratio for individual cent digits of the NYSE -
Methodology 1
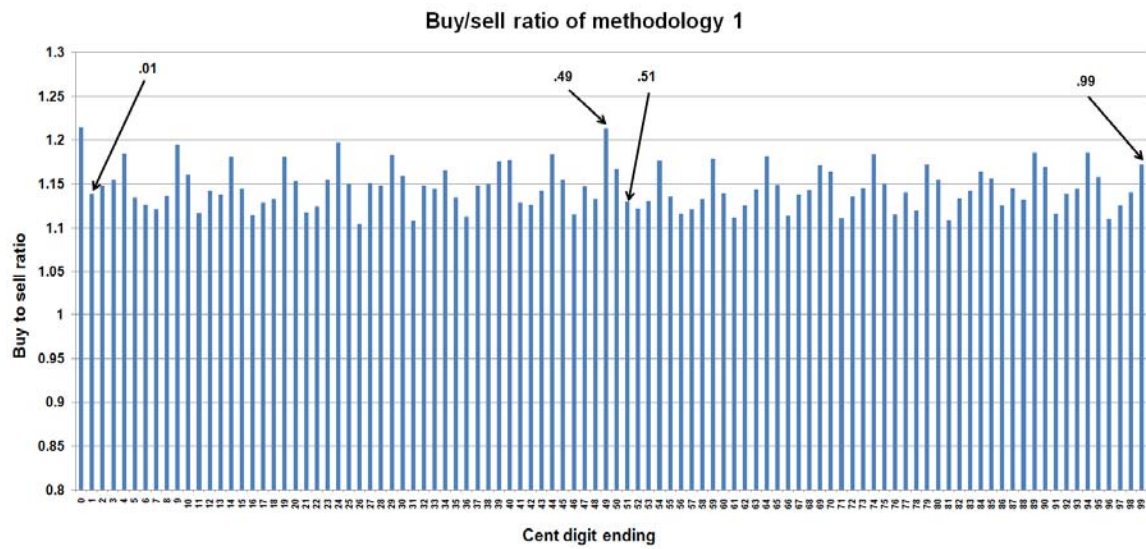
Figure 2: Line graph of buy/sell ratio for average combined cent digits of the
NYSE - Methodology 1
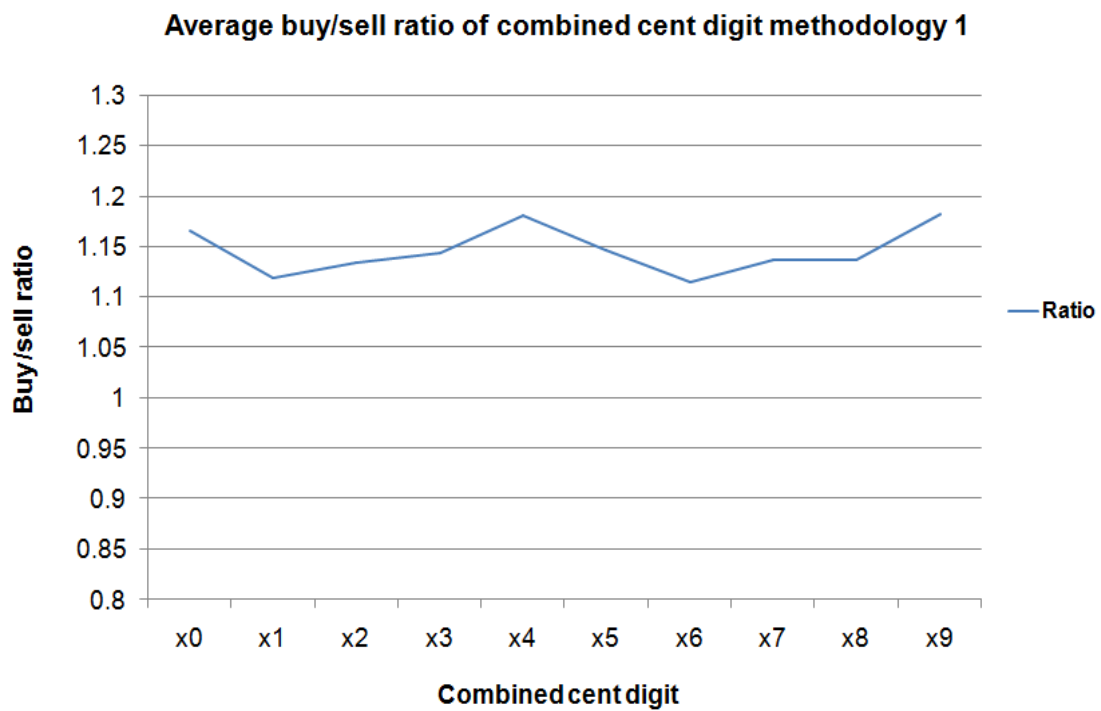
Figure 3: Bar graph of buy/sell ratio for individual cent digits of the NYSE - Methodology 2
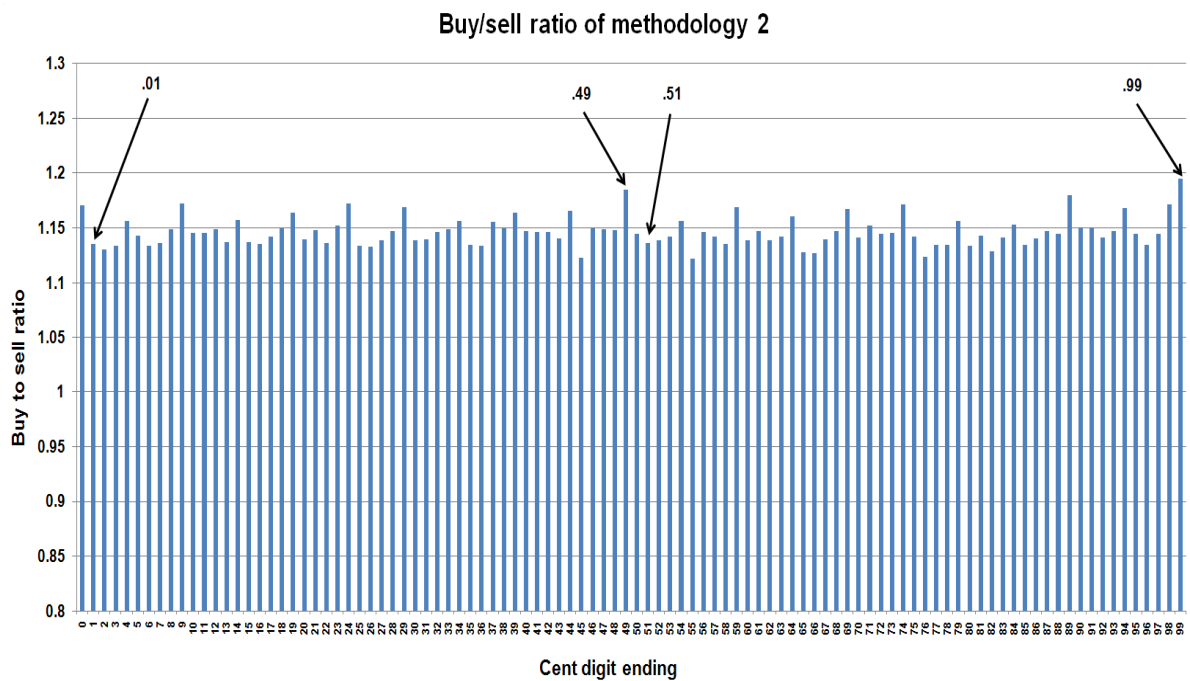
Figure 4: Line graph of buy/sell ratio for average combined cent digits of the NYSE - Methodology 2
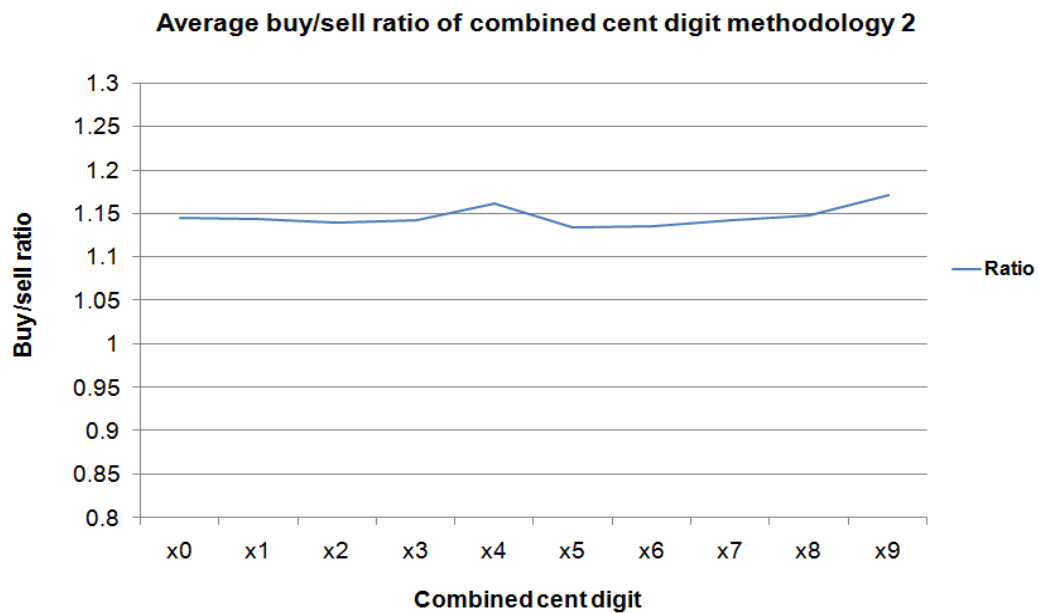
Figure 5: Bar graph of buy/sell ratio for individual cent digits of the NYSE -
Methodology 3
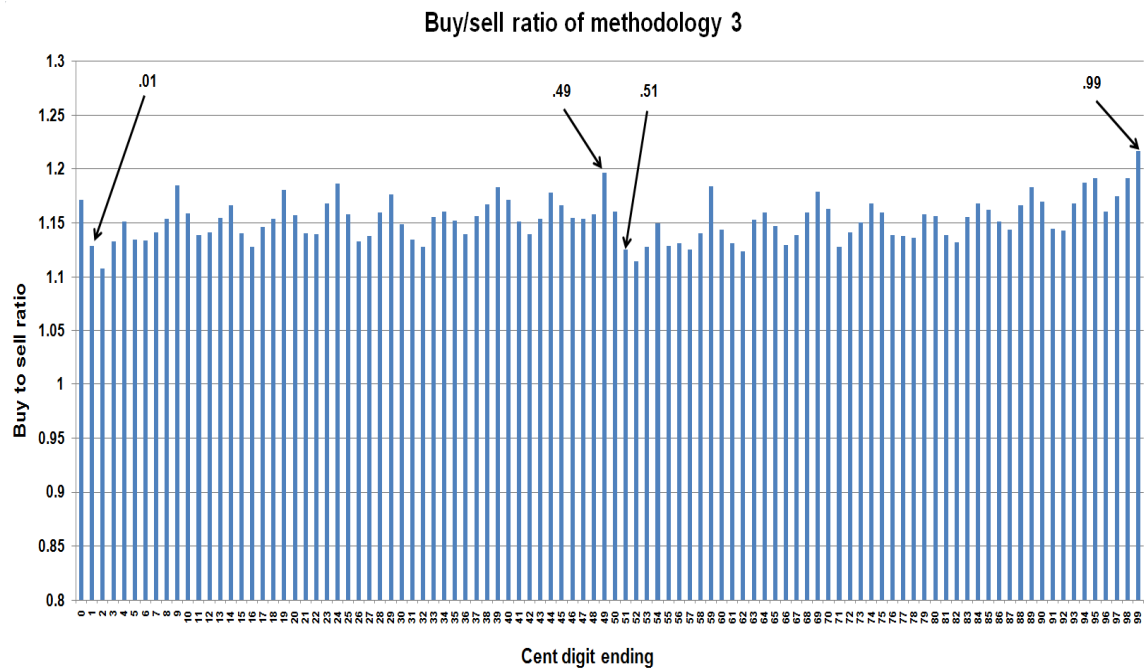
Figure 6: Line graph of buy/sell ratio for average combined cent digits of the
NYSE - Methodology 3
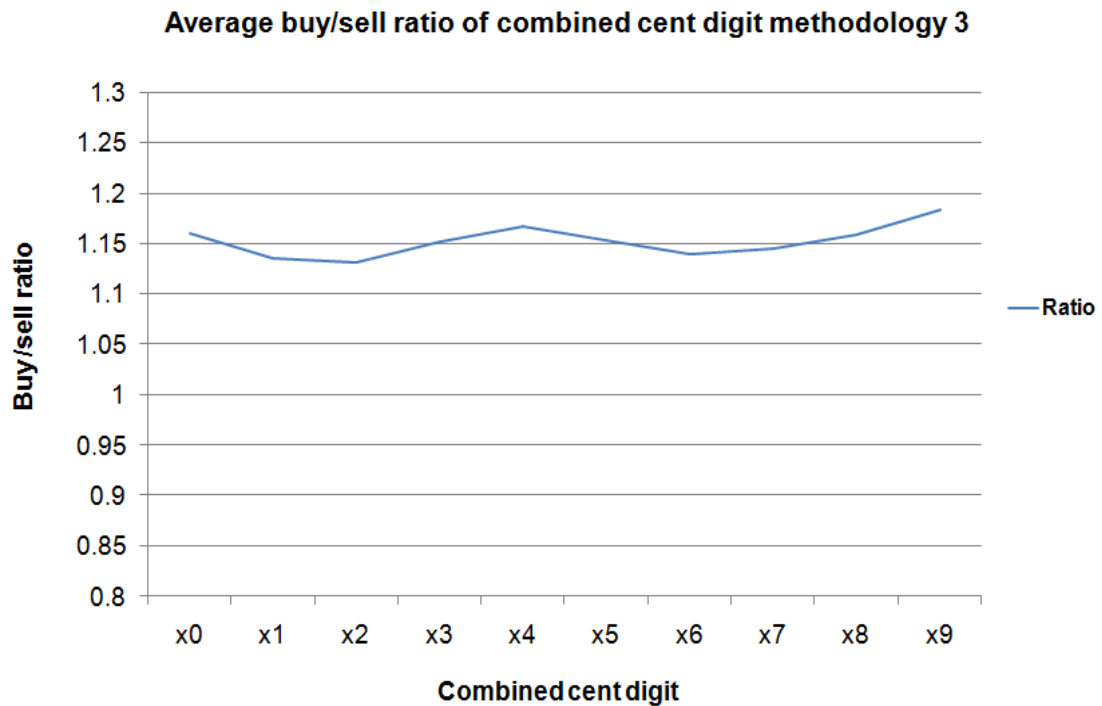
Figure 7: Bar graph of the SSE buy/sell ratio for individual cent digits - Methodology 1


Buy/sell ratio of methodology 1
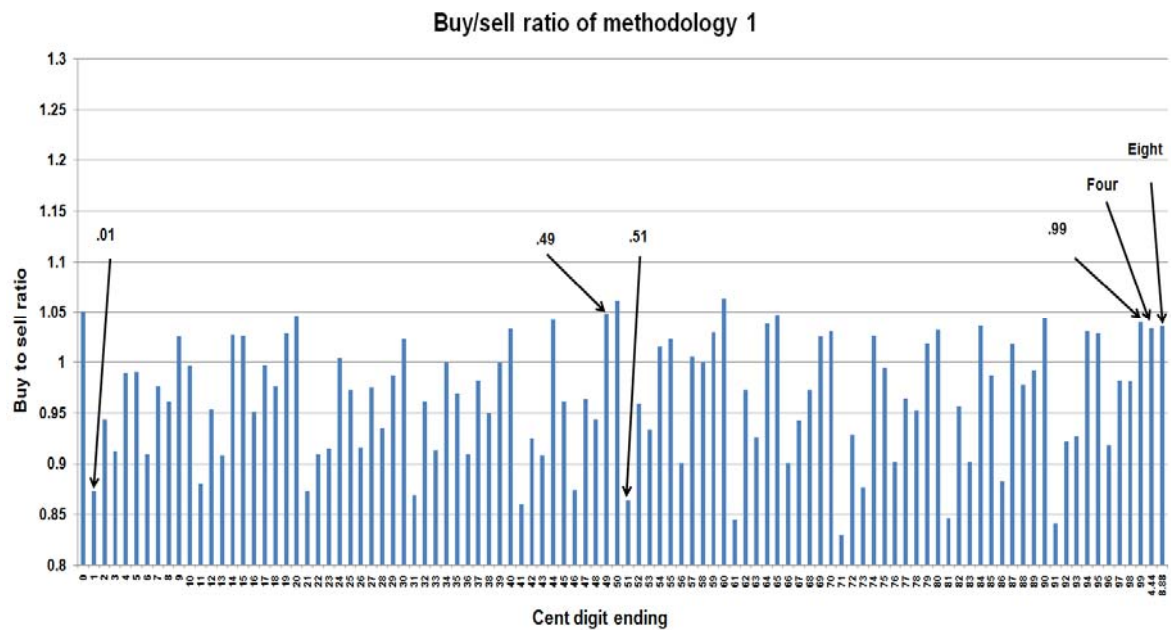
Figure 8: Line graph of combined cent digits in the SSE - Methodology 1

Average buy/sell ratio of combined cent digit methodology 1
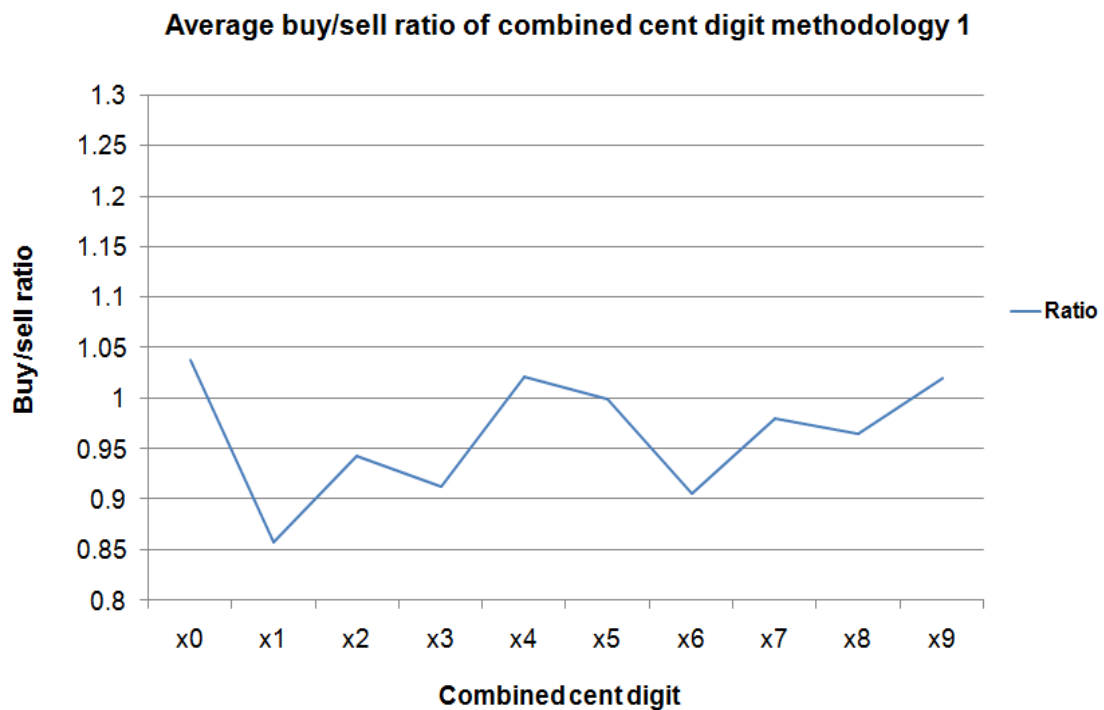
Figure 9: Bar graph of the SSE buy/sell ratio for individual cent digits - Methodology 2
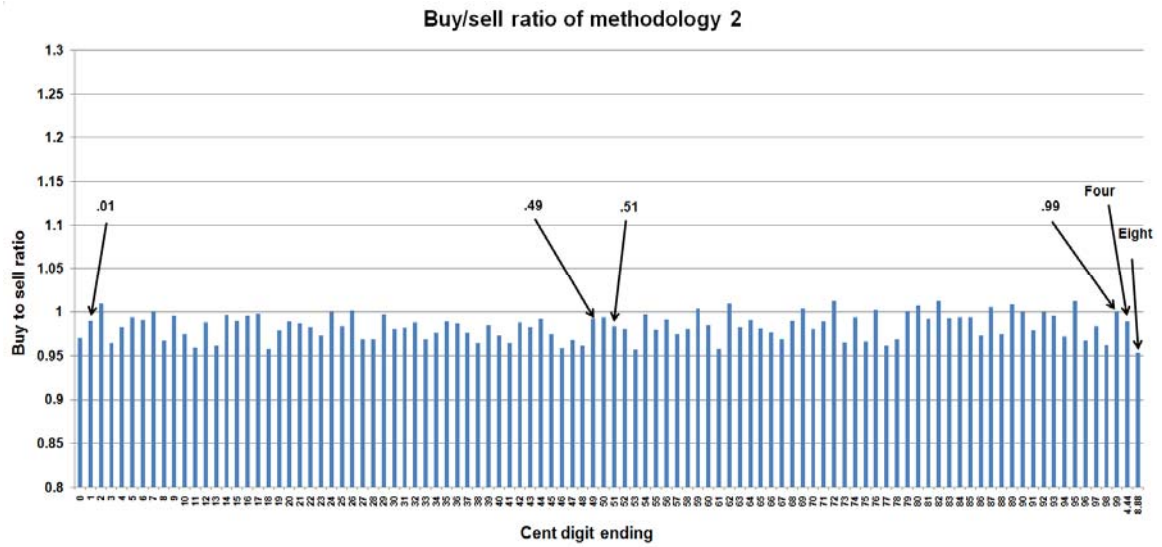
Figure 10: Line graph of combined cent digits in the SSE - Methodology 2
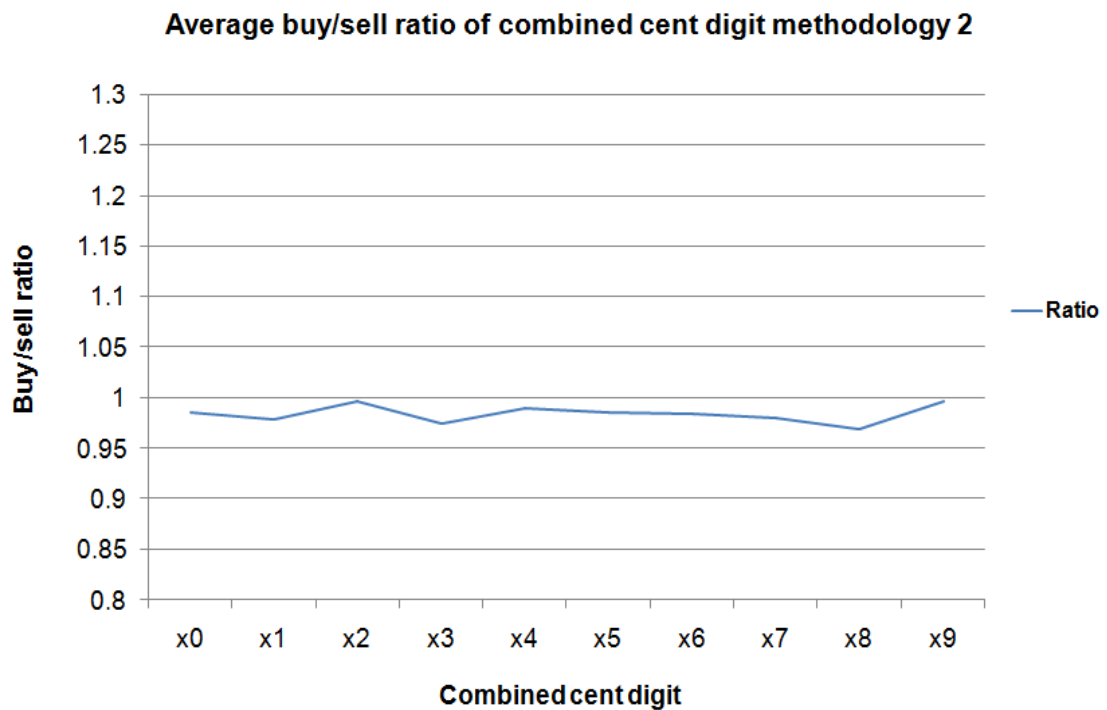
Figure 11: Bar graph of the SSE buy/sell ratio for individual cent digits - Methodology 3
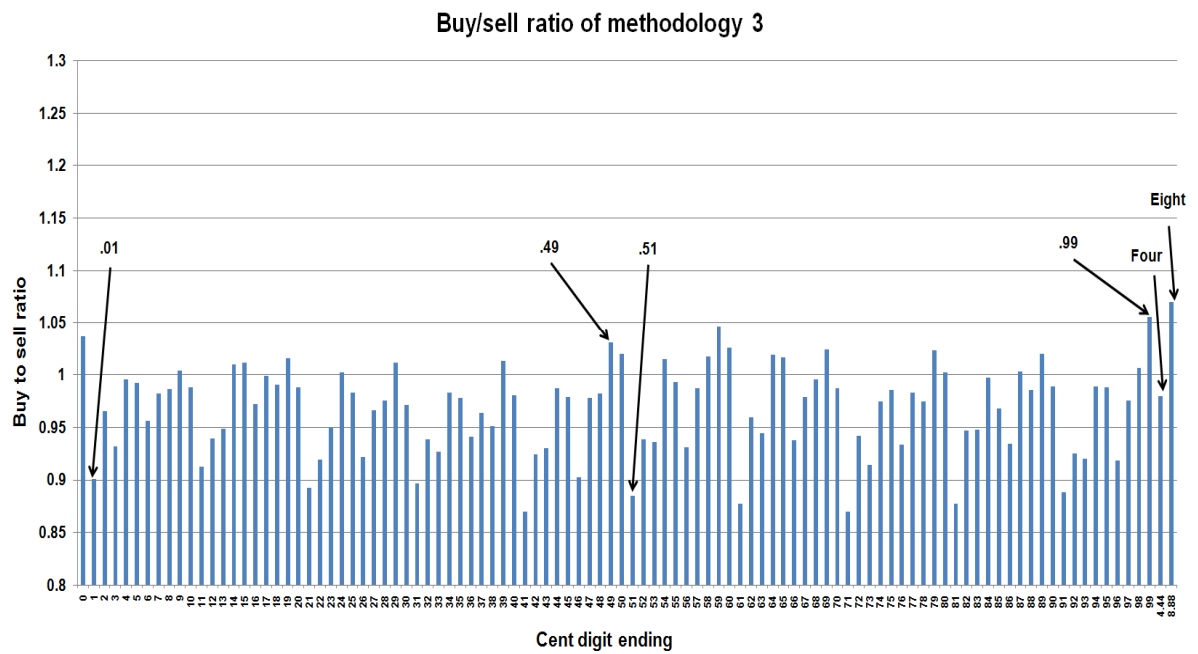
Figure 12: Line graph of combined cent digits in the SSE - Methodology 3
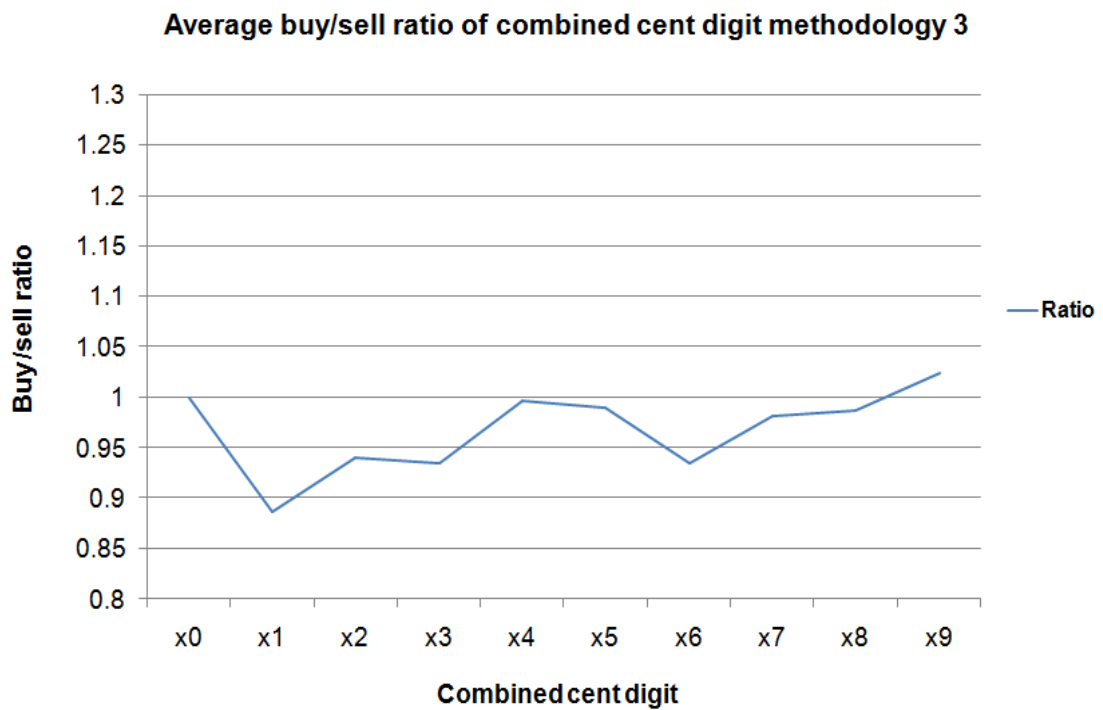
Figure 15: Bar graph of robustness data for buy/sell ratio in individual cent digits for Methodology 1

Buy/sell ratio of methodology 1
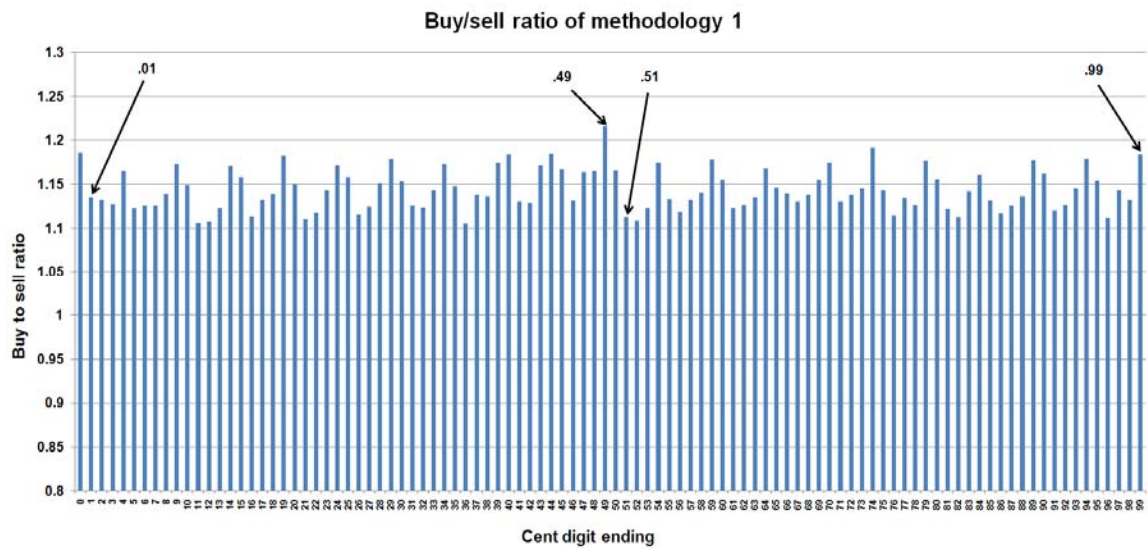
Figure 16: Line graph of robustness data for buy/sell ratio in combined cent digits for Methodology 1

Average buy/sell ratio of combined cent digit methodology 1

93
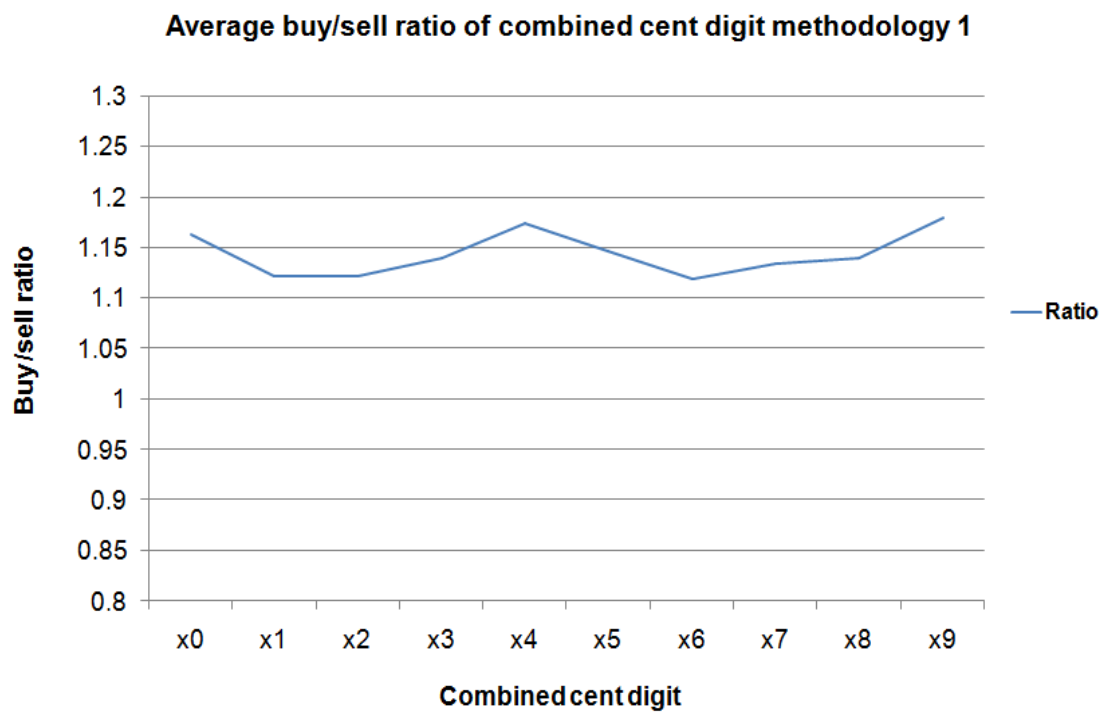
Figure 17: Bar graph of robustness data for buy/sell ratio in individual cent digits for Methodology 2

Buy/sell ratio of methodology 2
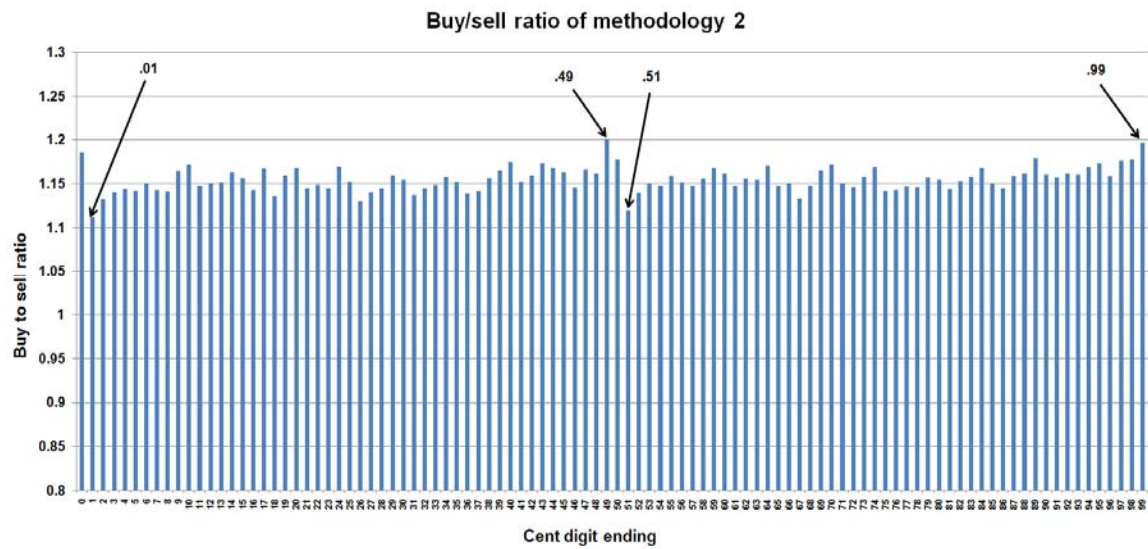
Figure 18: Line graph of robustness data for buy/sell ratio in combined cent digits for Methodology 2


Average buy/sell ratio of combined cent digit methodology 2
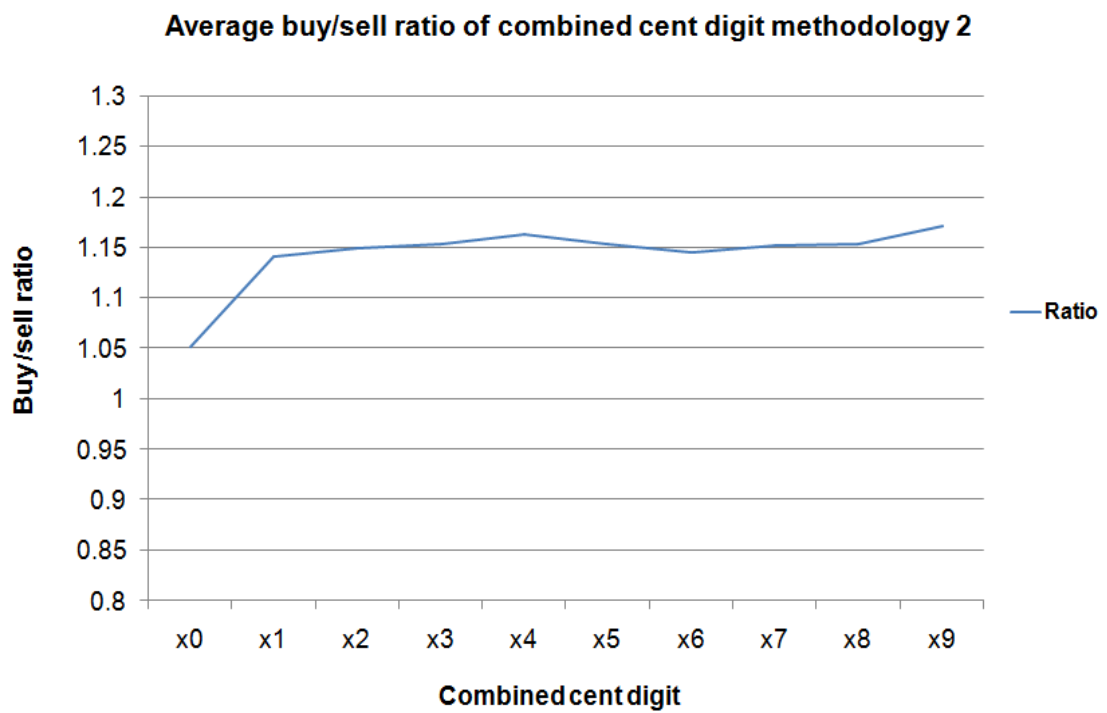
Figure 19: Bar graph of robustness data for buy/sell ratio in individual cent digits for Methodology 3
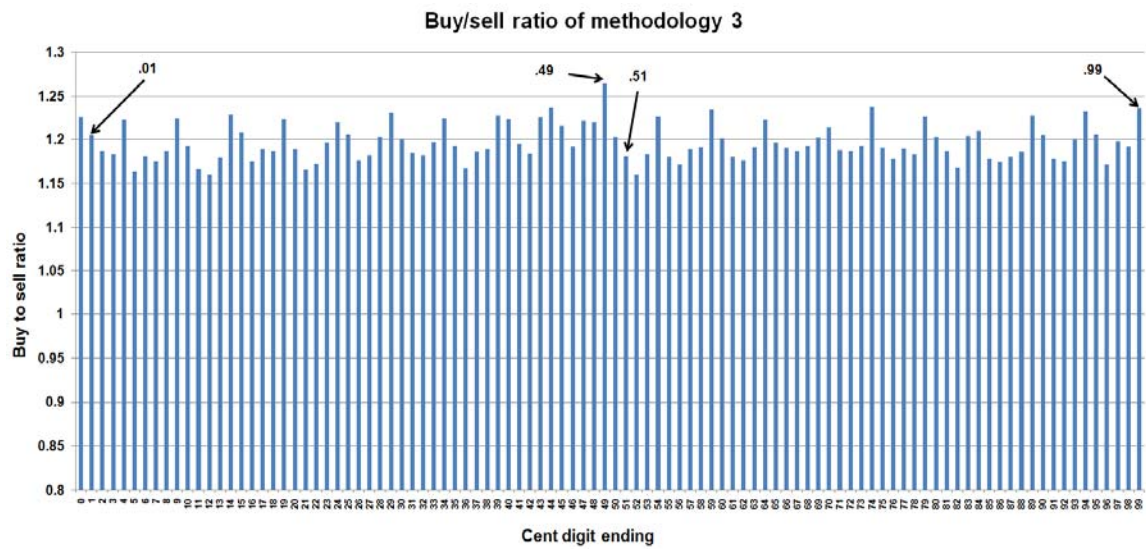
Figure 20: Line graph of robustness data for buy/sell ratio in combined cent digits for Methodology 3
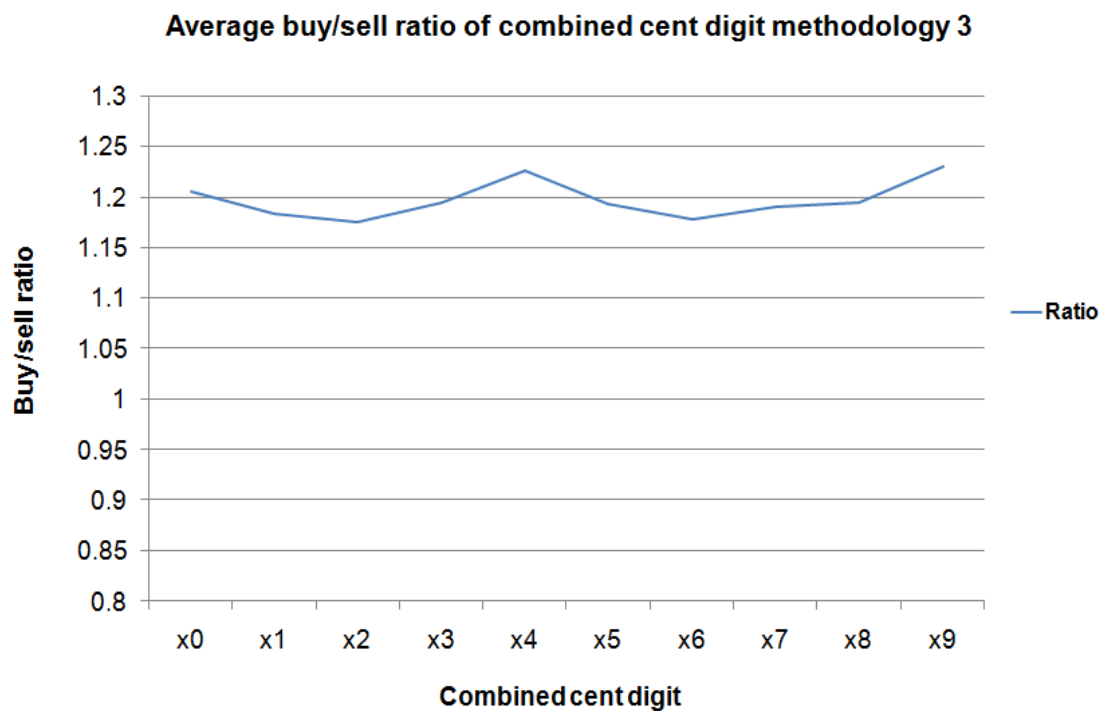
Figure 21: Comparison bar graph of the volume of buy and sell stocks at each combined cent digit ending
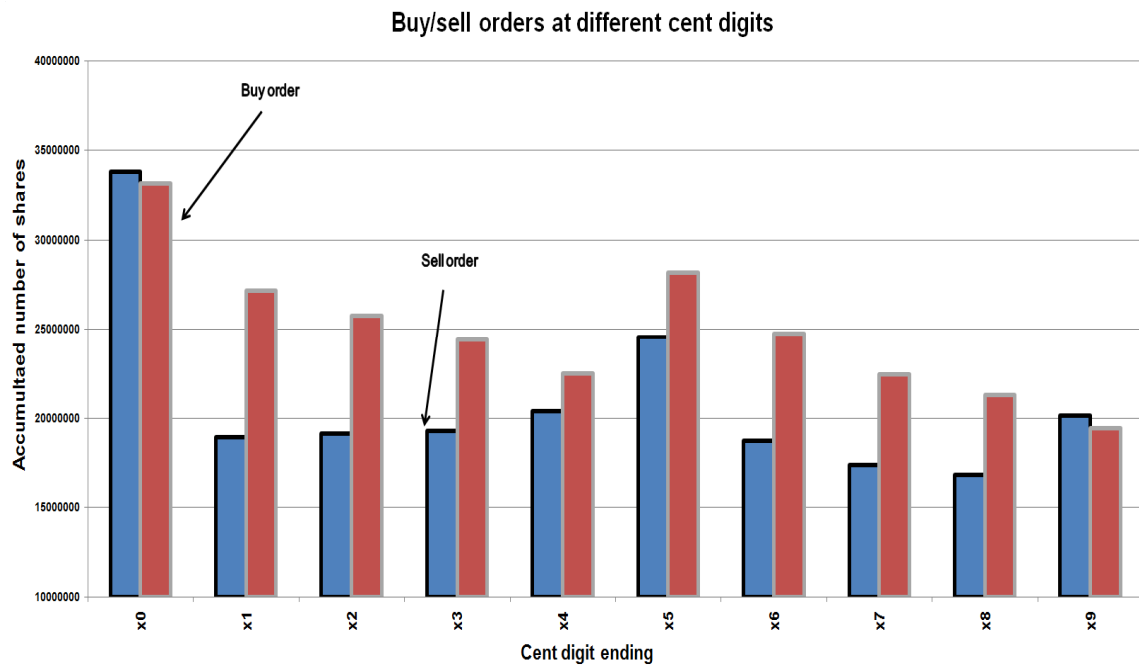
Figure 22: Accumulated numbers of shares ordered for individual cent digit endings
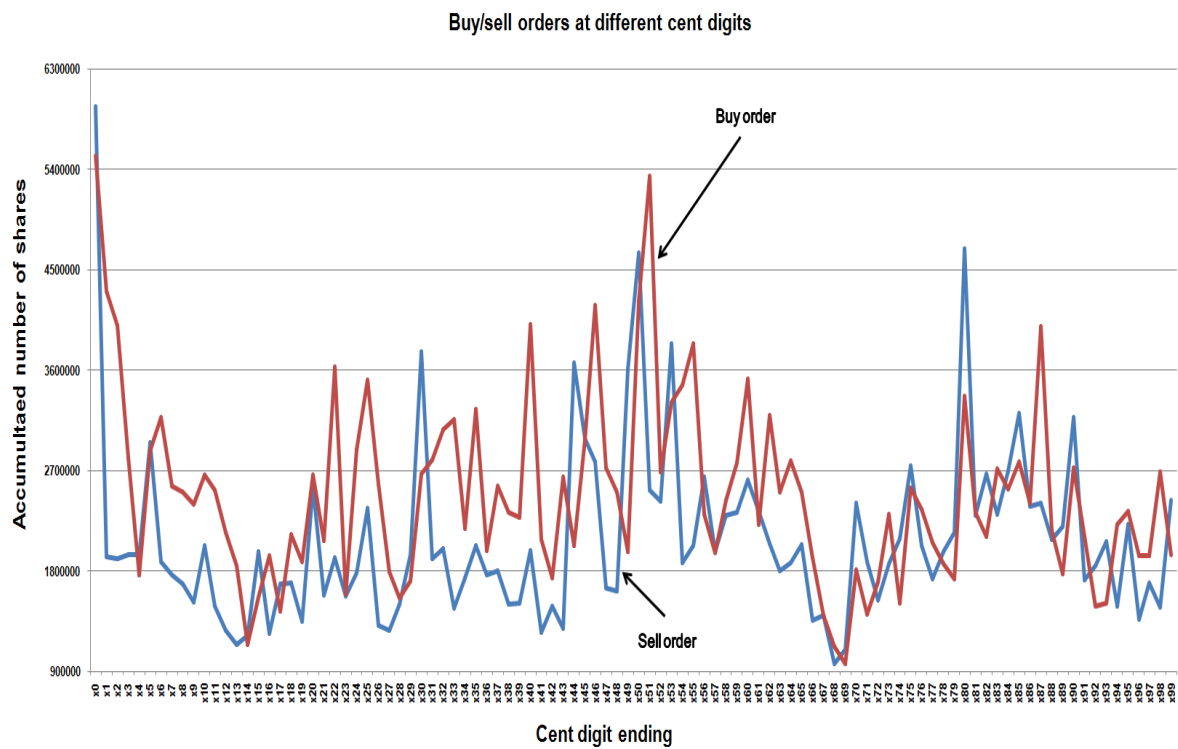
Java program for data extraction

BuySell.java

```java
/**
 * This file is a utility Enum that determines whether a transaction
 * is a buy or sell.
 */
package stock.common;

public enum BuySell
{

    BUY("0"),
    SELL("1");

    private String display;

    BuySell(String display)
    {
        this.display = display;
    }

    public String toString()
    {
        return display;
    }

}
```

Common.java

```java
/**
 * This file contains common utility methods for converting
date/time.
 * It also contains constants that tells which columns are what
fields.
 */
package stock.common;

import stock.utils.FileUtils;
import stock.utils.StringUtils;

import java.io.FilenameFilter;
import java.text.DateFormat;
import java.text.SimpleDateFormat;
import java.util.Date;


public class Common
{

    //---------- Logging ----------//
```

```java
    public final static String LOG4J_XML = "META-INF/log4j.xml";

    //---------- File Filters ----------//

    public   final   static   FilenameFilter   CSV_FILTER   =
FileUtils.createExtensionFilter(".csv");

    //---------- Date/Time Format ----------//

    public final static String DT_FORMAT_WITH_MS    = "dd-MMM-yyyy
HH:mm:ss.SSS";
    public final static String DT_FORMAT_WITHOUT_MS = "dd-MMM-yyyy
HH:mm:ss";
    public  final  static  DateFormat  DT_FORMATTER_WITH_MS    = new
SimpleDateFormat(DT_FORMAT_WITH_MS);
    public  final  static  DateFormat  DT_FORMATTER_WITHOUT_MS = new
SimpleDateFormat(DT_FORMAT_WITHOUT_MS);

    //---------- CSV Constants ----------//

    public final static int COL_DATE = 0;
    public final static int COL_TIME = 1;
    public final static int COL_OPEN = 2;
    public final static int COL_BID  = 2;
    public final static int COL_ASK  = 3;

    public final static String BUY = "0";
    public final static String SELL = "1";

    public final static String[] DIGITS_COMMON  = new String[100];
    public   final   static   String[]   DIGITS_SPECIAL   =   new
String[]{"4.44", "8.88"};
    public   final   static   String[]   DIGITS_ALL              =
combine(DIGITS_COMMON, DIGITS_SPECIAL);

    //---------- Result Headers ----------//

    public final static String[] HEADER_RESULTS = new String[]{"Open
Price", "Bid/Ask  Average", "Next  Open  Price", "Buy[0]/Sell[1]",
"Return"};
    public   final   static   String[]   HEADER_SUMMARY   =   new
String[]{"Digit", "Total Buys", "Total Sells", "Average Return"};

    static
    {
        for(int i = 0; i < 100; i++)
        {
            String digit = getDigit(i);
            DIGITS_COMMON[i] = digit;
            DIGITS_ALL[i] = digit;
        }
    }

    public static String getDigit(int i)
    {
        return (i < 10 ? "0" : "") + i;
    }

    public static String[] combine(String[] array1, String[] array2)
    {
        String[] result = new String[array1.length + array2.length];
```

```java
        System.arraycopy(array1, 0, result, 0, array1.length);
        System.arraycopy(array2,    0,    result,    array1.length,
array2.length);
        return result;
    }

    public static long getTime(String[] line)
    {
        Date date = getDate(line, COL_DATE, COL_TIME);
        return date.getTime();
    }

    public  static  Date  getDate(String[]  line,  int  dateCol,  int
timeCol)
    {
        String dateStr = line[dateCol] + " " + line[timeCol];
        return                    StringUtils.convertToDate(dateStr,
DT_FORMATTER_WITH_MS, DT_FORMATTER_WITHOUT_MS);
    }

}
```

CommandLineUtils.java

```java
/**
 * This file is used to parse command line arguments.
 */
package stock.utils;

import org.apache.commons.cli.*;


public class CommandLineUtils
{

    private final static int TAB_SIZE = 4;
    private final static int LINE_WIDTH = 80;

    public  static  CommandLine  parseArguments(Class  app,  Options
options, String[] args)
    {
        CommandLineParser parser = new GnuParser();

        try
        {
            return parser.parse(options, args);
        }
        catch(ParseException e)
        {
            System.err.println(e.getMessage() + "\n");
            printUsage(app, options, true);
        }

        return null;
    }

    public    static    Option    createOption(String    option,    String
longOption, String description, boolean hasArgs, boolean required)
```

```java
        {
            Option   op   =   new   Option(option,   longOption,   hasArgs,
description);
            op.setRequired(required);
            return op;
        }

    public static void printUsage(Class app, Options options, boolean
error)
        {
            HelpFormatter formatter = new HelpFormatter();
            formatter.setWidth(LINE_WIDTH);
            formatter.setLeftPadding(TAB_SIZE);
            formatter.setDescPadding(TAB_SIZE);
            formatter.printHelp("java " + app.getName(), options, true);
            System.exit(error ? 1 : 0);
        }

}
```

FileUtils.java

```java
/**
 * This  file  contains  utility  methods  that  deals  with  file
input/outputs.
 */
package stock.utils;

import org.apache.log4j.Logger;

import java.io.*;


public class FileUtils
{

    private     final     static     Logger     LOG     =
Logger.getLogger(FileUtils.class);

    public  static  FilenameFilter  createExtensionFilter(final  String
extension)
        {
            return new FilenameFilter()
            {
                public boolean accept(File dir, String name)
                {
                    return name.endsWith(extension);
                }
            };
        }

    public    static    void    ensureExists(File    file)    throws
FileNotFoundException
        {
            if(!file.exists())
                throw new FileNotFoundException(file.getAbsolutePath());
        }
```

```
    public static void closeQuietly(Writer writer)
    {
        try
        {
            writer.close();
        }
        catch(IOException e)
        {
        }
    }

}
```

NumberUtils.java

```java
/**
 * This file contains utility methods that rounds numbers.
 */
package stock.utils;

public class NumberUtils
{

    public static double round(double num, int decimalPlaces)
    {
        int temp = pow(10, decimalPlaces);
        return Math.rint(num * temp) / temp;
    }

    public static int pow(int num, int pow)
    {
        int result = 1;

        for(int i = 0; i < pow; i++)
            result *= num;

        return result;
    }

}
```

ProcessorUtils.java

```java
/**
 * This file contains utility methods that retrieve digits after
 * the decimal point for a given number or string.
 */
package stock.utils;

import java.text.DecimalFormat;
import java.text.NumberFormat;


public class ProcessorUtils
```

```java
{

    private final static NumberFormat NUMBER_FORMAT = new
DecimalFormat("0.00");

    public static String getDigit(String num)
    {
        return getDigit(Double.parseDouble(num));
    }

    public static String getDigit(double num)
    {
        String str = NUMBER_FORMAT.format(num);
        int index = str.indexOf(".");
        return str.substring(index + 1);
    }

}
```

ReflectionUtils.java

```java
/**
 * This class contains utility methods that retrieve
 * resources from the Java classpath.
 */
package stock.utils;

import java.net.URL;


public class ReflectionUtils
{

    public static URL getResourceURL(String resource)
    {
        return getContextClassLoader().getResource(resource);
    }

    private static ClassLoader getContextClassLoader()
    {
        return Thread.currentThread().getContextClassLoader();
    }

}
```

StringUtils.java

```java
/**
 * This class contains utility methods that manipulates Strings,
 * such as date conversion, number conversion.
 */
package stock.utils;

import org.apache.log4j.Logger;

import java.text.DateFormat;
```

```java
import java.text.ParseException;
import java.util.Date;


public class StringUtils
{

    private       final      static      Logger      LOG       =
Logger.getLogger(StringUtils.class);

    public   static   Date   convertToDate(String   str,   DateFormat   ...
formatters)
    {
        for(DateFormat formatter: formatters)
        {
            try
            {
                return formatter.parse(str);
            }
            catch(ParseException e)
            {
                // Ignore exception
            }
        }

        LOG.warn("Cannot parse date string " + str);
        return null;
    }

    public static double convertToDouble(String str)
    {
        return   str   ==   null   ||   str.isEmpty()   ?   0   :
Double.parseDouble(str);
    }

    public static String[] toStringArray(Object ... objects)
    {
        String[] stringArray = new String[objects.length];

        for(int i = 0; i < objects.length; i++)
        {
            Object object = objects[i];
            stringArray[i] = object == null ? "" : object.toString();
        }

        return stringArray;
    }
}
```

DataSplitter.java

```java
/**
 * This file pre-processes CSV data files.
 *
 * It splits opening data file and bid/ask data files into
 * separate CSV files. It also removes duplicate entries,
 * and unused columns.
 *
```

```
 * Two CSV files are created for each firm:
 * 1. Opening prices (with date/time, opening price)
 * 2. Bid/ask prices (with date/time, bid and ask prices)
 */
package stock;

import org.apache.commons.cli.CommandLine;
import org.apache.commons.cli.Options;
import org.apache.commons.csv.CSVParser;
import org.apache.commons.csv.CSVPrinter;
import org.apache.log4j.Logger;
import org.apache.log4j.xml.DOMConfigurator;
import stock.common.Common;
import stock.utils.*;

import java.io.*;
import java.util.*;


public class DataSplitter
{

    private       final    static     Logger       LOG       =
Logger.getLogger(DataSplitter.class);
    private final static Options OPTIONS = new Options();

    private final static int HEADERS    = 1;
    private final static int COL_COMPANY = 0;
    private final static int COL_DATE    = 1;
    private final static int COL_TIME    = 2;
    private final static int COL_OPEN    = 5;
    private final static int COL_BID     = 5;
    private final static int COL_ASK     = 6;
//    private final static int COL_OPEN    = 4;
//    private final static int COL_BID     = 4;
//    private final static int COL_ASK     = 5;

    static
    {

DOMConfigurator.configure(ReflectionUtils.getResourceURL(Common.LOG4J
_XML));

        OPTIONS.addOption(CommandLineUtils.createOption("o",  "open",
"Directory containing opening prices", true, true));
        OPTIONS.addOption(CommandLineUtils.createOption("b",     "bid-
ask", "Directory containing bid/ask prices", true, true));
        OPTIONS.addOption(CommandLineUtils.createOption("s", "split",
"Directory where splitted output is written to", true, true));
    }

    public     static     void     main(String[]     args)     throws
FileNotFoundException
    {
        CommandLine                    command                     =
CommandLineUtils.parseArguments(DataSplitter.class, OPTIONS, args);
        File openDir = new File(command.getOptionValue("o"));
        File bidAskDir = new File(command.getOptionValue("b"));
        File outDir = new File(command.getOptionValue("s"));
        File outOpenDir = new File(outDir, "OPEN");
        File outBidAskDir = new File(outDir, "BID_ASK");
```

```java
        FileUtils.ensureExists(openDir);
        FileUtils.ensureExists(bidAskDir);

        if(!outOpenDir.exists())
            outOpenDir.mkdirs();

        if(!outBidAskDir.exists())
            outBidAskDir.mkdirs();

        LOG.info("Splitting OPEN prices");
        splitData(openDir, outOpenDir, new int[]{COL_DATE, COL_TIME,
COL_OPEN}, true);

        LOG.info("Processing BID/ASK prices");
        splitData(bidAskDir,    outBidAskDir,    new    int[]{COL_DATE,
COL_TIME, COL_BID, COL_ASK}, true);
    }

    private static void splitData(File inDir, File outDir, int[]
cols, boolean skipDuplicates)
    {
        File[] files = inDir.listFiles(Common.CSV_FILTER);
        Map<String, CSVPrinter> printers = new HashMap<String,
CSVPrinter>();
        List<FileWriter> writers = new ArrayList<FileWriter>();
        int total = files.length;
        int current = 0;

        for(File file: files)
        {
            LOG.info("Processing " + file.getAbsolutePath() + " (" +
(++current) + "/" + total + ")");

            try
            {
                Set<Date> dates = new HashSet<Date>();
                FileReader reader = new FileReader(file);
                CSVParser parser = new CSVParser(reader);
                String[] line;

                skipHeaders(parser);

                while((line = parser.getLine()) != null)
                {
                    String companyCode = line[COL_COMPANY];
                    CSVPrinter printer = printers.get(companyCode);

                    if(printer == null)
                    {
                        File out = new File(outDir, companyCode +
".csv");

                        FileWriter writer = new FileWriter(out);
                        printer = new CSVPrinter(writer);

                        printers.put(companyCode, printer);
                        writers.add(writer);
                    }

                    Date date = Common.getDate(line, COL_DATE,
COL_TIME);
```

```java
                    String dateStr = line[COL_DATE];

                    if(!dateStr.endsWith("2005"))
                        break;

                    // Removing duplicate entries with the same
date/time
                    if(!skipDuplicates || dates.add(date))
                    {
                        for(int col : cols)
                            printer.print(line[col]);

                        printer.println();
                    }
                }

                reader.close();
            }
            catch(IOException e)
            {
                LOG.error("Error    splitting    file    " +
file.getAbsolutePath(), e);
            }
        }

        for(FileWriter writer: writers)
            FileUtils.closeQuietly(writer);
    }

    private static void skipHeaders(CSVParser parser) throws
IOException
    {
        for(int i = 0; i < HEADERS; i++)
            parser.getLine();
    }

}
```

DataProcessor.java

```java
/**
 * This is the first version of the data processor.
 * It processes outputs from the DataSplitter.
 * This version uses the PREVIOUS bid/ask prices to calculate the
results.
 */
package stock;

import org.apache.commons.cli.CommandLine;
import org.apache.commons.cli.Options;
import org.apache.commons.csv.CSVParser;
import org.apache.commons.csv.CSVPrinter;
import org.apache.log4j.Logger;
import org.apache.log4j.xml.DOMConfigurator;
import stock.common.Common;
import stock.utils.*;

import java.io.*;
import java.util.*;
```

```java
public class DataProcessor
{

    private        final        static        Logger        LOG        =
Logger.getLogger(DataProcessor.class);
    private final static Options OPTIONS = new Options();

    private final static String SUMMARY    = "SUMMARY";
    private final static String METHOD_1   = "METHOD_1";
    private final static String METHOD_2   = "METHOD_2";

    private final static double[] SPECIAL_NUMBERS = { 4.44, 8.88 };

    private final static Map<String, CSVPrinter> PRINTERS_1    = new
HashMap<String, CSVPrinter>();
    private final static Map<String, CSVPrinter> PRINTERS_2    = new
HashMap<String, CSVPrinter>();
    private final static Map<String, Long>        TOTAL_BUY_1 = new
HashMap<String, Long>();
    private final static Map<String, Long>        TOTAL_BUY_2 = new
HashMap<String, Long>();
    private final static Map<String, Long>        TOTAL_SELL_1 = new
HashMap<String, Long>();
    private final static Map<String, Long>        TOTAL_SELL_2 = new
HashMap<String, Long>();
    private final static Map<String, Double>      TOTAL_RETURN = new
HashMap<String, Double>();

    static
    {

DOMConfigurator.configure(ReflectionUtils.getResourceURL(Common.LOG4J
_XML));

        OPTIONS.addOption(CommandLineUtils.createOption("o",  "open",
"Directory containing opening prices", true, true));
        OPTIONS.addOption(CommandLineUtils.createOption("b",     "bid-
ask", "Directory containing bid/ask prices", true, true));
        OPTIONS.addOption(CommandLineUtils.createOption("r",
"result", "Directory where results are written to", true, true));
    }

    public static void main(String[] args) throws IOException
    {
        CommandLine                    command                    =
CommandLineUtils.parseArguments(DataProcessor.class, OPTIONS, args);
        File openDir = new File(command.getOptionValue("o"));
        File bidAskDir = new File(command.getOptionValue("b"));
        File resultDir = new File(command.getOptionValue("r"));

        FileUtils.ensureExists(openDir);
        FileUtils.ensureExists(bidAskDir);

        List<FileWriter> writers = createPrinters(resultDir);

        LOG.info("Processing data");
        processData(openDir, bidAskDir);

        LOG.info("Writing summary");
```

```java
        writeSummary(PRINTERS_1.get(SUMMARY),                TOTAL_BUY_1,
TOTAL_SELL_1);
        writeSummary(PRINTERS_2.get(SUMMARY),                TOTAL_BUY_2,
TOTAL_SELL_2);

        for(FileWriter writer: writers)
            writer.close();
    }

    private static void writeSummary(CSVPrinter printer, Map<String,
Long> totalBuy, Map<String, Long> totalSell)
    {
        for(int i = 0; i < 100; i++)
        {
            String digit = Common.getDigit(i);
            writeSummary(digit, printer, totalBuy, totalSell);
        }

        for(double num: SPECIAL_NUMBERS)
        {
            String numString = Double.toString(num);
            writeSummary(numString, printer, totalBuy, totalSell);
        }
    }

    private  static  void  writeSummary(String  digit,  CSVPrinter
printer, Map<String, Long> totalBuy, Map<String, Long> totalSell)
    {
        Long buy = totalBuy.get(digit);
        Long sell = totalSell.get(digit);
        Long total = buy + sell;
        Double returnTotal = TOTAL_RETURN.get(digit);
        Double averageReturn = returnTotal / total;

        String[] data = StringUtils.toStringArray(digit, buy, sell,
averageReturn);
        printer.println(data);
    }

    private static void processData(File openDir, File bidAskDir)
    {
        File[] openFiles = openDir.listFiles(Common.CSV_FILTER);
        int total = openFiles.length;
        int i = 0;

        for(File openFile: openFiles)
        {
            String fileName = openFile.getName();
            String company = fileName.substring(0, fileName.length()
- 4);

            File bidAskFile = new File(bidAskDir, fileName);

            LOG.info("Processing " + company + " (" + (++i) + "/" +
total + ")");

            if(!bidAskFile.exists())
            {
                LOG.warn("Cannot find matching Bid/Ask data for " +
company);
                continue;
            }
```

```
            processCompany(company, openFile, bidAskFile);
        }
    }

    private static void processCompany(String company, File openFile,
File bidAskFile)
    {
        try
        {
            FileReader openReader = new FileReader(openFile);
            CSVParser openParser = new CSVParser(openReader);

            FileReader bidAskReader = new FileReader(bidAskFile);
            CSVParser bidAskParser = new CSVParser(bidAskReader);
            TreeMap<Long,         String[]>      bidAskData       =
convertData(bidAskParser);

            String[] currentLine = openParser.getLine();
            String[] nextLine;

            String prevBuySell = null;

            while((nextLine = openParser.getLine()) != null)
            {
                long currentTime = Common.getTime(currentLine);

                String[]     bidAsk     =     lowerBound(bidAskData,
currentTime);

                if(bidAsk == null)
                {
                    currentLine = nextLine;
                    continue;
                }

                // Write results
                double              currentPrice               =
StringUtils.convertToDouble(currentLine[Common.COL_OPEN]);
                double              nextPrice                  =
StringUtils.convertToDouble(nextLine[Common.COL_OPEN]);

                double             prevBid                     =
StringUtils.convertToDouble(bidAsk[Common.COL_BID]);
                double             prevAsk                     =
StringUtils.convertToDouble(bidAsk[Common.COL_ASK]);

                // Ignore invalid entries where prices are 0
                if(currentPrice == 0 || nextPrice == 0 || prevBid ==
0 || prevAsk == 0)
                {
                    currentLine = nextLine;
                    continue;
                }

                double   midPoint   =   NumberUtils.round((prevBid   +
prevAsk) / 2, 4);
                double  returnValue = ((nextPrice - currentPrice) /
currentPrice);

                // Ingore records where open price == mid point
```

109

```java
                if(currentPrice != midPoint)
                {
                    // Strategy 1: if next price > mid point then
it's a BUY otherwise its a SELL
                    String  buySell = currentPrice  >  midPoint  ?
Common.BUY : Common.SELL;
                    recordResult(TOTAL_BUY_1,          TOTAL_SELL_1,
currentPrice, buySell, returnValue);
                }

                // Strategy 2: if current price == next price then
buy/sell is the previous buy/sell
                String  buySell  =  nextPrice  >  currentPrice  ?
Common.BUY : Common.SELL;

                if(currentPrice == nextPrice && prevBuySell != null)
                    buySell = prevBuySell;

                recordResult(TOTAL_BUY_2, TOTAL_SELL_2, currentPrice,
buySell, returnValue);

                currentLine = nextLine;
                prevBuySell = buySell;
            }

            openReader.close();
            bidAskReader.close();
        }
        catch(IOException e)
        {
            LOG.error("Error processing company " + company, e);
        }
    }

    private  static  TreeMap<Long,  String[]>  convertData(CSVParser
parser) throws IOException
    {
        TreeMap<Long, String[]> map = new TreeMap<Long, String[]>();
        String[] line;

        while((line = parser.getLine()) != null)
        {
            long time = Common.getTime(line);
            map.put(time, line);
        }

        return map;
    }

    private  static  void  recordResult(Map<String,  Long>  totalBuy,
Map<String, Long> totalSell, double currentPrice, String buySell,
double returnValue)
    {
        // Increment buy/sell counts
        incrementDigitCount(buySell.equals(Common.BUY)  ?  totalBuy  :
totalSell, currentPrice);

        // Increment return value total
        incrementReturnTotal(currentPrice, returnValue);
    }
```

```java
    private   static   void   incrementDigitCount(Map<String,   Long>
countMap, double currentPrice)
    {
        String digit = ProcessorUtils.getDigit(currentPrice);
        incrementDigitCount(countMap, digit);

        for(double num: SPECIAL_NUMBERS)
        {
            if(currentPrice == num)
            {
                String numString = Double.toString(num);
                incrementDigitCount(countMap, numString);
            }
        }
    }

    private   static   void   incrementDigitCount(Map<String,   Long>
countMap, String key)
    {
        Long count = countMap.get(key);

        if(countMap.get(key) == null)
            count = 1L;
        else
            count++;

        countMap.put(key, count);
    }

    private   static   void   incrementReturnTotal(double   currentPrice,
double returnValue)
    {
        String digit = ProcessorUtils.getDigit(currentPrice);
        incrementReturnTotal(digit, returnValue);

        for(double num: SPECIAL_NUMBERS)
        {
            if(currentPrice == num)
            {
                String numString = Double.toString(num);
                incrementReturnTotal(numString, returnValue);
            }
        }
    }

    private   static   void   incrementReturnTotal(String   key,   double
returnValue)
    {
        Double total = TOTAL_RETURN.get(key);

        if(total == null)
            total = returnValue;
        else
            total += returnValue;

        TOTAL_RETURN.put(key, total);
    }

    private static String[] lowerBound(TreeMap<Long, String[]> data,
long targetTime)
    {
```

```java
        Map.Entry<Long,             String[]>             entry             =
data.floorEntry(targetTime);
        return entry != null ? entry.getValue() : null;
    }

    private  static  List<FileWriter>  createPrinters(File  outDir)
throws IOException
    {
        List<FileWriter> writers = new ArrayList<FileWriter>();

        FileWriter    summary1    =    addPrinter(PRINTERS_1,    outDir,
METHOD_1, "summary.csv", SUMMARY, Common.HEADER_SUMMARY);
        FileWriter    summary2    =    addPrinter(PRINTERS_2,    outDir,
METHOD_2, "summary.csv", SUMMARY, Common.HEADER_SUMMARY);

        writers.add(summary1);
        writers.add(summary2);

        return writers;
    }

    private  static  FileWriter  addPrinter(Map<String,  CSVPrinter>
printers, File outDir, String method, String fileName, String key,
String[] header) throws IOException
    {
        File subDir = new File(outDir, method);

        if(!subDir.exists())
            subDir.mkdirs();

        FileWriter    writer    =    new    FileWriter(new    File(subDir,
fileName));
        CSVPrinter printer = new CSVPrinter(writer);

        printer.println(header);
        printers.put(key, printer);

        return writer;
    }

}
```

Data processor.java

```java
/**
 * This is a slightly modified version of the data processor.
 * It processes outputs from the DataSplitter.
 * This  version  uses  the  NEXT  bid/ask  prices  to  calculate  the
results.
 */
package stock;

import org.apache.commons.cli.CommandLine;
import org.apache.commons.cli.Options;
import org.apache.commons.csv.CSVParser;
import org.apache.commons.csv.CSVPrinter;
import org.apache.log4j.Logger;
import org.apache.log4j.xml.DOMConfigurator;
import stock.common.Common;
```

```java
import stock.utils.*;

import java.io.*;
import java.util.*;


public class DataProcessor
{

    private       final       static       Logger       LOG       =
Logger.getLogger(DataProcessor.class);
    private final static Options OPTIONS = new Options();

    private final static String SUMMARY    = "SUMMARY";
    private final static String METHOD_1   = "METHOD_1";
    private final static String METHOD_2   = "METHOD_2";

    private final static double[] SPECIAL_NUMBERS = { 4.44, 8.88 };

    private final static Map<String, CSVPrinter> PRINTERS_1   = new
HashMap<String, CSVPrinter>();
    private final static Map<String, CSVPrinter> PRINTERS_2   = new
HashMap<String, CSVPrinter>();
    private final static Map<String, Long>       TOTAL_BUY_1  = new
HashMap<String, Long>();
    private final static Map<String, Long>       TOTAL_BUY_2  = new
HashMap<String, Long>();
    private final static Map<String, Long>       TOTAL_SELL_1 = new
HashMap<String, Long>();
    private final static Map<String, Long>       TOTAL_SELL_2 = new
HashMap<String, Long>();
    private final static Map<String, Double>     TOTAL_RETURN = new
HashMap<String, Double>();

    static
    {

DOMConfigurator.configure(ReflectionUtils.getResourceURL(Common.LOG4J
_XML));

        OPTIONS.addOption(CommandLineUtils.createOption("o",  "open",
"Directory containing opening prices", true, true));
        OPTIONS.addOption(CommandLineUtils.createOption("b",    "bid-
ask", "Directory containing bid/ask prices", true, true));
        OPTIONS.addOption(CommandLineUtils.createOption("r",
"result", "Directory where results are written to", true, true));
    }

    public static void main(String[] args) throws IOException
    {
        CommandLine                command                =
CommandLineUtils.parseArguments(DataProcessor.class, OPTIONS, args);
        File openDir = new File(command.getOptionValue("o"));
        File bidAskDir = new File(command.getOptionValue("b"));
        File resultDir = new File(command.getOptionValue("r"));

        FileUtils.ensureExists(openDir);
        FileUtils.ensureExists(bidAskDir);

        List<FileWriter> writers = createPrinters(resultDir);
```

```java
        LOG.info("Processing data");
        processData(openDir, bidAskDir);

        LOG.info("Writing summary");
        writeSummary(PRINTERS_1.get(SUMMARY),              TOTAL_BUY_1,
TOTAL_SELL_1);
        writeSummary(PRINTERS_2.get(SUMMARY),              TOTAL_BUY_2,
TOTAL_SELL_2);

        for(FileWriter writer: writers)
            writer.close();
    }

    private static void writeSummary(CSVPrinter printer, Map<String,
Long> totalBuy, Map<String, Long> totalSell)
    {
        for(int i = 0; i < 100; i++)
        {
            String digit = Common.getDigit(i);
            writeSummary(digit, printer, totalBuy, totalSell);
        }

        for(double num: SPECIAL_NUMBERS)
        {
            String numString = Double.toString(num);
            writeSummary(numString, printer, totalBuy, totalSell);
        }
    }

    private   static   void   writeSummary(String   digit,   CSVPrinter
printer, Map<String, Long> totalBuy, Map<String, Long> totalSell)
    {
        Long buy = totalBuy.get(digit);
        Long sell = totalSell.get(digit);
        Long total = buy + sell;
        Double returnTotal = TOTAL_RETURN.get(digit);
        Double averageReturn = returnTotal / total;

        String[] data = StringUtils.toStringArray(digit, buy, sell,
averageReturn);
        printer.println(data);
    }

    private static void processData(File openDir, File bidAskDir)
    {
        File[] openFiles = openDir.listFiles(Common.CSV_FILTER);
        int total = openFiles.length;
        int i = 0;

        for(File openFile: openFiles)
        {
            String fileName = openFile.getName();
            String company = fileName.substring(0, fileName.length()
- 4);

            File bidAskFile = new File(bidAskDir, fileName);

            LOG.info("Processing " + company + " (" + (++i) + "/" +
total + ")");

            if(!bidAskFile.exists())
            {
```

114

```java
            LOG.warn("Cannot find matching Bid/Ask data for " +
company);
                continue;
            }

            processCompany(company, openFile, bidAskFile);
        }
    }

    private static void processCompany(String company, File openFile,
File bidAskFile)
    {
        try
        {
            FileReader openReader = new FileReader(openFile);
            CSVParser openParser = new CSVParser(openReader);

            FileReader bidAskReader = new FileReader(bidAskFile);
            CSVParser bidAskParser = new CSVParser(bidAskReader);
            TreeMap<Long,        String[]>        bidAskData        =
convertData(bidAskParser);

            String[] currentLine = openParser.getLine();
            String[] nextLine;

            String prevBuySell = null;

            while((nextLine = openParser.getLine()) != null)
            {
                long currentTime = Common.getTime(currentLine);

                String[]      bidAsk      =      upperBound(bidAskData,
currentTime);

                if(bidAsk == null)
                {
                    currentLine = nextLine;
                    continue;
                }

                // Write results
                double                currentPrice                =
StringUtils.convertToDouble(currentLine[Common.COL_OPEN]);
                double                nextPrice                =
StringUtils.convertToDouble(nextLine[Common.COL_OPEN]);

                double                nextBid                =
StringUtils.convertToDouble(bidAsk[Common.COL_BID]);
                double                nextAsk                =
StringUtils.convertToDouble(bidAsk[Common.COL_ASK]);

                // Ignore invalid entries where prices are 0
                if(currentPrice == 0 || nextPrice == 0 || nextBid ==
0 || nextAsk == 0)
                {
                    currentLine = nextLine;
                    continue;
                }

                double   midPoint   =   NumberUtils.round((nextBid   +
nextAsk) / 2, 4);
```

115

```java
                double returnValue = ((nextPrice - currentPrice) /
currentPrice);

                // Ingore records where open price == mid point
                if(currentPrice != midPoint)
                {
                    // Strategy 1: if next price > mid point then
it's a BUY otherwise its a SELL
                    String buySell = currentPrice > midPoint ?
Common.BUY : Common.SELL;
                    recordResult(TOTAL_BUY_1,          TOTAL_SELL_1,
currentPrice, buySell, returnValue);
                }

                // Strategy 2: if current price == next price then
buy/sell is the previous buy/sell
                String buySell = nextPrice > currentPrice ?
Common.BUY : Common.SELL;

                if(currentPrice == nextPrice && prevBuySell != null)
                    buySell = prevBuySell;

                recordResult(TOTAL_BUY_2, TOTAL_SELL_2, currentPrice,
buySell, returnValue);

                currentLine = nextLine;
                prevBuySell = buySell;
            }

            openReader.close();
            bidAskReader.close();
        }
        catch(IOException e)
        {
            LOG.error("Error processing company " + company, e);
        }
    }

    private static TreeMap<Long, String[]> convertData(CSVParser
parser) throws IOException
    {
        TreeMap<Long, String[]> map = new TreeMap<Long, String[]>();
        String[] line;

        while((line = parser.getLine()) != null)
        {
            long time = Common.getTime(line);
            map.put(time, line);
        }

        return map;
    }

    private static void recordResult(Map<String, Long> totalBuy,
Map<String, Long> totalSell, double currentPrice, String buySell,
double returnValue)
    {
        // Increment buy/sell counts
        incrementDigitCount(buySell.equals(Common.BUY) ? totalBuy :
totalSell, currentPrice);
```

```java
        // Increment return value total
        incrementReturnTotal(currentPrice, returnValue);
    }

    private static void incrementDigitCount(Map<String, Long>
countMap, double currentPrice)
    {
        String digit = ProcessorUtils.getDigit(currentPrice);
        incrementDigitCount(countMap, digit);

        for(double num: SPECIAL_NUMBERS)
        {
            if(currentPrice == num)
            {
                String numString = Double.toString(num);
                incrementDigitCount(countMap, numString);
            }
        }
    }

    private static void incrementDigitCount(Map<String, Long>
countMap, String key)
    {
        Long count = countMap.get(key);

        if(countMap.get(key) == null)
            count = 1L;
        else
            count++;

        countMap.put(key, count);
    }

    private static void incrementReturnTotal(double currentPrice,
double returnValue)
    {
        String digit = ProcessorUtils.getDigit(currentPrice);
        incrementReturnTotal(digit, returnValue);

        for(double num: SPECIAL_NUMBERS)
        {
            if(currentPrice == num)
            {
                String numString = Double.toString(num);
                incrementReturnTotal(numString, returnValue);
            }
        }
    }

    private static void incrementReturnTotal(String key, double
returnValue)
    {
        Double total = TOTAL_RETURN.get(key);

        if(total == null)
            total = returnValue;
        else
            total += returnValue;

        TOTAL_RETURN.put(key, total);
    }
```

```java
    private static String[] upperBound(TreeMap<Long, String[]> data,
long targetTime)
    {
        Map.Entry<Long,            String[]>          entry          =
data.ceilingEntry(targetTime);
        return entry != null ? entry.getValue() : null;
    }

    private static String[] lowerBound(TreeMap<Long, String[]> data,
long targetTime)
    {
        Map.Entry<Long,            String[]>          entry          =
data.floorEntry(targetTime);
        return entry != null ? entry.getValue() : null;
    }

    private  static  List<FileWriter>  createPrinters(File  outDir)
throws IOException
    {
        List<FileWriter> writers = new ArrayList<FileWriter>();

        FileWriter    summary1    =    addPrinter(PRINTERS_1,    outDir,
METHOD_1, "summary.csv", SUMMARY, Common.HEADER_SUMMARY);
        FileWriter    summary2    =    addPrinter(PRINTERS_2,    outDir,
METHOD_2, "summary.csv", SUMMARY, Common.HEADER_SUMMARY);

        writers.add(summary1);
        writers.add(summary2);

        return writers;
    }

    private  static  FileWriter  addPrinter(Map<String,  CSVPrinter>
printers, File outDir, String method, String fileName, String key,
String[] header) throws IOException
    {
        File subDir = new File(outDir, method);

        if(!subDir.exists())
            subDir.mkdirs();

        FileWriter    writer    =    new    FileWriter(new    File(subDir,
fileName));
        CSVPrinter printer = new CSVPrinter(writer);

        printer.println(header);
        printers.put(key, printer);

        return writer;
    }

}
```