

## Price clustering on the Australian Stock Exchange

Michael Aitken<sup>a</sup>, Philip Brown<sup>b,\*</sup>, Christine Buckland<sup>b</sup>,  
H.Y. Izan<sup>b</sup>, Terry Walter<sup>a</sup>

<sup>a</sup> *The University of Sydney, Sydney, NSW 2006, Australia*

<sup>b</sup> *The University of Western Australia, Nedlands, WA 6009, Australia*

---

### Abstract

Price clustering is manifest in some prices being observed more frequently than others, when underlying value is uniformly distributed over the range of admissible prices. It is frequently observed in financial markets. We investigate clustering in individual trades effected on the Australian Stock Exchange's wholly computerised, order-driven trading system. We find that price clustering is pervasive, tending to follow an overall pattern somewhat similar to that found in US securities markets. Clustering results from imprecise beliefs ('haziness') about firm value together with the existence of conventional, salient focal points within regions of haziness. Thus we find that clustering increases with the price of the stock (reflecting imprecision in beliefs about firm value) and with surrogates for greater haziness such as higher market-wide volatility, own stock volatility, trade size, and the size of the bid–ask spread. Clustering is lower when price discovery is likely to be more efficient; that is, it decreases with trading frequency, and it is lower for stocks with options traded on them and for stocks that can be sold short.

*JEL classification:* G10; G14

*Keywords:* Price clustering; Stock prices; Australian Stock Exchange

---

---

\* Corresponding author.

## 1. Introduction

Price clustering is pervasive in markets of many kinds. It exists in bid, offer and trade prices in equity markets; in IPO prices, takeover bids and rights issues;<sup>1</sup> and it has been documented in forex and gold markets. It is clearly evident in real estate markets, automobile prices and consumer retailing, although the type of price clustering in automobile and consumer retailing is different from that usually observed in financial markets. Psychological experiments demonstrate more generally that clustering of outcomes at round numbers is a fundamental attribute of human behaviour, whether we are making decisions or simply recording measurements.<sup>2</sup> An understanding of share price clustering is important because, superficially, it is at odds with economic rationality. Certainly it is inconsistent with share prices following a simple random walk process. Clustering of bid and ask (offer) prices also raises important issues about the potential for optimal order placement strategies.

Price clustering is affected by the institutional rules governing tick size. These rules influence the demand for market services and thereby affect the price discovery process. Furthermore, price clustering has serious implications for the measurement of key market metrics such as volatility. It complicates moment estimators, because it exacerbates the effects of tick size (Harris, 1991).

In this paper we document price clustering in individual trades on the Australian Stock Exchange (ASX). To date, most of the literature on share price clustering has employed US data, in particular data from the NYSE specialist market. However, Harris (1991) did find that clustering varies across exchange types. For example, clustering is greater in dealer markets than in public auction markets. The ASX operates a computer based, order-driven market with brokers acting as principals or agents for their clients. The price at which a trade is executed on the ASX results from a type of auction when the market opens, and thereafter it is determined when a market order meets the price of the best opposing limit order. Studying clustering on the ASX is thus a useful comparative exercise.

Section 2 summarises previous studies of price clustering. Our data and research method are outlined in Section 3. Section 4 contains our results, while Section 5 contains our conclusions and suggestions for further work.

---

<sup>1</sup> Buckland (1994).

<sup>2</sup> Rounding fractions or decimals to the nearest whole number (rather than say to the whole number +0.0001) is an efficient measurement convention that is embedded in and preserved by our educational system.

## 2. Previous studies

### 2.1. Equity markets

Osborne (1962) presented the first rigorous empirical evidence of ‘congestion’ in US share prices. Congestion means that “there are price ranges in which a given stock price spends an inordinate amount of time” (p. 369). In the absence of clustering, we would expect to see a uniform distribution over admissible prices.<sup>3</sup> Instead, Osborne found a “pronounced tendency for (closing) prices to cluster on whole numbers, halves, quarters, and odd one-eighths in descending preference” (p. 370).

Niederhoffer (1965) documented clustering of limit orders taken from the order book of a specialist on the NYSE.<sup>4</sup> The ratio of limit order closing prices at the even eighths (0, 2, 4, 6) to those at the odd eighths (1, 3, 5, 7) was 8.8:1, of which prices at whole numbers (0 eighths) constituted 7.7:1. Niederhoffer found clustering in the closing prices of actively and inactively traded shares, in high- and low-priced shares, and in noon as well as closing prices. Higher-priced issues traded mostly at the integers, while stable, lower-priced issues settled at even numbers of eighths.

Niederhoffer argued that the auction market mechanism ensured that price changes show regularity and structure, because of behavioural preferences and specialist trading strategies. He suggested (p. 264) that clustering was the result of the tendency of stock market participants to place their orders at “numbers with which they are accustomed to deal”, such as whole and round numbers. Niederhoffer concluded that such structure and regularity to prices casts serious doubt on the premise that share prices are random. Specialists and floor traders had indicated to him that the congestion of limit orders opens up a lucrative trading technique. The example he cited was a share that recently rose from  $1/8$  to  $7/8$ . There would probably be few buy limits below  $7/8$ , and numerous sell limits one tick higher (at  $8/8$ ). The specialist could sell short at  $7/8$ , hoping to drive the price back to  $1/8$  and make a profit, while feeling relatively safe in the knowledge there should be ample time to cover for a  $1/8$  loss if price were to rise further. A similar trading opportunity might arise if price had recently declined. Niederhoffer speculated that this strategy could explain the Osborne (1962) observation that there were more highs than lows at  $7/8$  and less highs than lows at  $1/8$ .

In a second study, Niederhoffer (1966) used trades data for seven days chosen randomly from the complete record of ticker transactions in 1964 for NYSE

<sup>3</sup> For example we would expect to observe equal frequencies of prices ending in  $1/8$ ,  $2/8$ ,  $3/8$ ,  $4/8$ ,  $5/8$ ,  $6/8$ ,  $7/8$  and whole numbers among prices within the range of the most common US tick size (one-eighth of a dollar).

<sup>4</sup> Niederhoffer’s data related to 837 orders “taken from a specialist’s book in a well-known corporation at 4:00 p.m., January 24, 1964”. He omitted stop orders because they represent a small proportion of trades and market orders because they are executed rapidly.

stocks. Prices were grouped into three strata: (i) 1,000 cases where price was unchanged from the previous trade; (ii) 12,800 cases where it changed by one-eighth; and (iii) 11,000 cases where it changed by more than one-eighth. He found that 58.5% of all trades were at an even eighth, and that clustering at an even eighth was most pronounced in the third stratum. Niederhoffer argued that the clustering he observed was a consequence of more limit and stop orders being placed on specialists' books at even eighths. Price cannot move from or through such a position until all relevant orders have been exhausted. Hence there is the tendency for 'stickiness' at even eighths.

Niederhoffer and Osborne (1966) documented additional properties of dependence in the NYSE ticker prices of six Dow Jones stocks traded in October 1964. While the random walk model states that changes in the price of consecutive transactions are distributed independently, Niederhoffer and Osborne found strong evidence of dependence. For example, after a price rise the odds were about 3:1 that the next price change was a decline; similarly after a decline the odds were about 3:1 in favour of a rise. In addition, after two changes in the same direction, the odds of a continuation in the same direction was almost twice as great as after two changes in opposing directions. While serial dependence in price changes per se is beyond the scope of our present study, we note that their data displayed once again the "stickiness of even eighths". They found (p. 914) that "(r)eversals are relatively more concentrated at integers where stable slow-moving participants offer to buy and sell. There is a concentration of particular types of reversals just above and below these barriers."

A quarter of a century later Harris (1991) gathered evidence from the NYSE to show that stock price clustering had persisted through time and that it conformed to the same hierarchy of degrees of rounding that others had already noted: "integers are more common than halves, halves are more common than odd quarters, odd quarters are more common than odd eighths and other fractions<sup>5</sup> are rarely observed." Harris, building on Ball et al. (1985; discussed below), posited that there are factors that cause the desired price resolution to become more coarse, and hence the extent of clustering to increase in certain circumstances. He suggested that price clustering occurs because traders use a restricted set of prices to simplify their negotiations, which makes them less costly. The existence of a restricted set of discrete prices known to all traders means that negotiation time is reduced, since it limits the number of different prices at which bids and offers will be made. It also limits the amount of information exchanged between traders. Consequently bid and offer prices converge more rapidly, and the time savings alone can be significant.<sup>6</sup>

---

<sup>5</sup> 'Other fractions' meant sixteenths, where they were permitted by the tick size rules.

<sup>6</sup> However, Christie and Schultz (1994, p. 1829) concluded that Harris's "negotiation hypothesis appears incapable of explaining the lack of odd-eighth quotes for the majority of (their NASDAQ) sample".

Cross-sectional variation in price clustering was examined by relating its frequency to individual stock price attributes. The attributes Harris selected were volatility, firm size, transaction frequency, price level and whether the stock is traded primarily on a dealer market. The first three variables reflected the costs of extended negotiations; the price level is a scale variable that was included because “traders are assumed to use discrete price sets based on minimum price variations that are constant fractions of price” (p. 402); and a dealer market dummy variable was included to test Niederhoffer’s limit order explanation of price clustering, relative to Harris’s belief that clustering was due to the incentive to lower negotiation costs and therefore more likely to exist in dealer markets, where limit orders do not exist and dealers build their reputations by acting consistently in repeated trades. Multivariate regressions were fitted across individual stocks, using two different measures for the dependent variable to summarise price clustering in each stock. Clustering was found to increase with the stock’s price level and volatility, and to decrease with firm size and transaction frequency. It was also more prevalent in a dealer’s market.

Whole number price frequencies were documented for all primary market trades in the trading week before and after the October 1987 crash. Price clustering at whole numbers increased at all price levels during the volatile aftermath of the crash, which is noteworthy in the additional sense that *less* clustering normally would be expected after a price fall. Volatility dominated the price level effect in that instance.

Hameed and Terry (1994) investigated the distribution of daily closing prices for 234 stocks that were traded on the Stock Exchange of Singapore Main Board (SES) between January 1980 and July 1994. The SES is an order-driven market with no official market maker for any stock. Prices ending in multiples of ten cents were more common than prices ending in odd multiples of five cents; prices were more likely to end in even cents than odd cents, and least likely to end in fractional cents; and whole dollars occurred more frequently than half dollars, which were more frequent than price multiples of 10 cents. Price clustering increased with a stock’s price level and decreased with trading volume. Unlike Harris, they did not find a consistent relationship between price clustering and stock price volatility. Clustering was found to be inversely related to tick size, especially for higher priced stocks, once the other variables were controlled for.

## 2.2. Other financial markets

Price clustering has also been found in the gold and foreign exchange markets.

Ball et al. (1985) investigated the morning and afternoon fixing prices of gold on the London market from 2 January to 30 April 1981, when the price of gold was unusually volatile. They believed that gold prices are generated under near ideal market conditions. Gold is actively traded in a number of countries; there are numerous types of markets, e.g. physical and derivatives; trading is almost

continuous throughout the world; and competition between markets should limit the commissions and transaction costs in any one market. The widely dispersed regional markets for gold not only increase overall liquidity, but also act as 'sensors', so that information arising anywhere can quickly influence price.

Ball et al. hypothesised that price clustering depends on how well the asset's underlying value is known. If value is not well known, prices will cluster. If value is well known, traders will use a finer set of prices. The maximum precision of pricing (tick size) in the London gold market is USD five cents. They detected clustering in the fixing price, providing clear evidence that the level of rounding is higher than the minimum price tick. The degree of price resolution (rounding) is not constant over time, but rather it changes depending on the amount of information in the market. Rounding increases with the level and volatility of price.

Goodhart and Curcio (1991) examined price clustering in the foreign exchange market, looking at clustering in the final digit.<sup>7</sup> They investigated the same two alternative explanations as Harris (1991): the 'attraction theory'; and Ball et al.'s (1985) price resolution theory, as amplified by Harris. The attraction theory is that the number 0, perhaps because it is more salient, is a stronger attracter than 5,<sup>8</sup> which is stronger than the other even numbers (i.e., excluding 0). Suppose we assume that the final digit of underlying value is uniformly distributed over the integers 0–9. Then according to the theory, 1 and 9 will be strongly attracted to 0, which will be the most commonly observed; 5 will be the next most common, by assumption; it will be more common than 2 and 8, which are even numbers two places removed from the strongest attracter and three places removed from 5; and the least common will be 1 and 9. The relative frequency of 3 will be the same as 7, while 4 and 6 will occur with equal frequency. Whether 3 and 7 are more or less common than 4 and 6 is not readily apparent, because it depends on the 'gravitational pull' of 5 on 4 and 6, relative to the general preference for even over odd numbers. In short the attraction theory, if we can endow it with the status of a theory, predicts the following relationship between the relative frequencies:  $0 > 5 > \{2 = 8\} > \{3 = 7, 4 = 6\} > \{1 = 9\}$ .<sup>9</sup> Harris (1991, p. 35) in fact rejected the attraction theory because he found the frequencies of odd-eighths (in our case, the frequencies of 1, 3, 7 and 9) were approximately the same.

Forex bid and ask prices quoted by foreign exchange dealers and bankers were found to cluster, but clustering did not carry over to the penultimate digit. Goodhart and Curcio concluded that clustering in the final digit depends on the

<sup>7</sup> Forex quotes are to four decimal places.

<sup>8</sup> Put another way, people are more likely to round (say) 129 or 131 to 130 than to round 134 or 136 to 135.

<sup>9</sup> Goodhart and Curcio's prediction was different: they predicted the sequence of relative frequencies would be  $0 > 5 > \{3 = 7\} > \{2 = 8\} > \{4 = 6\} > \{1 = 9\}$ . The sequence depends upon the assumptions one makes about the nature of the attraction.

desired degree of price resolution by traders. However, clustering in the amount of the bid–ask spread appears to be driven by a separate behavioural pattern, consistent with the pure attraction hypothesis.

### 3. Data and method

We examine price clustering at the level of the individual trade. In this section we describe our data sources, establish three measures of price clustering, define our explanatory variables, and indicate their expected relationship to the degree of price clustering that we observe.

#### 3.1. SEATS data

Much of the data used in this analysis is sourced from the complete set of SEATS transactions obtained from the ASX by Aitken, Brown and Walter.<sup>10</sup> A separate database was constructed from this raw data, comprising observations on the ordinary fully paid shares of the 267 listed companies that traded at least five times per day, on average, during the period 3 September 1990 to 3 September 1993 (inclusive). While over 1,000 stocks were listed on the ASX during this period, the method used to select the 267 stocks ensures that the majority of the total market capitalisation and total trade activity on the ASX is captured by the sample.

Only regular trades transacted in the six hours of Normal Trading Mode are considered in this paper. Opening trades were excluded, because they are transacted at an averaged price.<sup>11</sup> They occur at market opening where the price of a bid equals that of an ask, but they also occur where the bid and ask prices overlap (i.e., when the price of at least one bid exceeds that of at least one offer).<sup>12</sup> The averaging process contrasts with the discrete tick size rules that governed admissible prices when the buyer and seller placed their orders. We also excluded any trades that took place after the close of Normal Trading Mode, that is, after 4 p.m. (about 4,000 trades).

We extracted each regular trade's price and volume, and the price of the highest limit order bid and lowest limit order ask immediately before the trade. Additional

<sup>10</sup> See Aitken et al. (1993a) for further information on the SEATS database.

<sup>11</sup> When they overlap, the official opening price is a volume-weighted average of the best bid and best offer prices and is calculated to the nearest tenth of a cent.

<sup>12</sup> The ASX halts trading temporarily after a price-sensitive announcement, so that brokers can adjust their bids and offers for the new information. Bids and asks can also overlap during the information halt, i.e. before trading resumes. Because we could not identify every case where an announcement halt occurred, we removed from the sample all cases where a decimal price was recorded for a trade during the day, since they unequivocally were average prices struck when trading resumed after a halt.

variables were derived from the existing database or obtained externally. The resulting comprehensive database consists of 2,610,400 observations.

### 3.2. Variables

Price clustering increases with uncertainty (haziness) about firm value. Since we can not observe directly the degree of uncertainty about firm value, we use several proxies for it.<sup>13</sup> They are trade size, market-wide volatility, individual stock volatility, firm size, trading frequency, and whether the stock (i) is classified as a resource stock, (ii) has options traded on it, and (iii) may be sold short. We also have a control variable to indicate whether the trade whose price we observe was initiated by the buyer or the seller. Our rationale for choosing each variable is explained below.

*Price clustering.* We use the last digit of price to summarise clustering. We examine all trade prices with a tick size of one cent (i.e., all prices between 10 cents and \$5 from 3 September 1990 to 27 October 1991, and between 10 cents and \$10 from 28 October 1991 to 3 September 1993).<sup>14</sup>

*Stock price.* Assuming the number of issued (outstanding) shares is known, the degree of clustering should be proportional to stock price to a first approximation (Harris, 1991).

*Trade size.* Recognisable information asymmetries initially should prompt greater price clustering. Larger orders are sometimes associated with informed agents (Easley and O'Hara, 1987) and their placement should lead to greater clustering. Furthermore informed traders placing large orders may wish to hide their knowledge by quoting a more clustered price. We acknowledge three factors that weaken the predicted effect. One is the common strategy of large traders to split their orders across brokers to minimise the price impact of their trades, especially when they believe they are better informed. Another is the desire of uninformed, small value traders who, apart from any incentive to quote the clustered prices in order to hide in the crowd, are attracted to round numbers because they more frequently operate in a 'sphere of haziness' about value (Butler and Loomes, 1988). The third is that informed traders improve the efficiency of price discovery, which should reduce price clustering. We measure trade size by the natural logarithm of the dollar value of the trade, to mitigate the potentially undue influence of especially large block trades.

---

<sup>13</sup> Disagreement among analysts' earnings forecasts is an obvious candidate. Unfortunately there is no sufficiently comprehensive analysts' forecast service available in Australia.

<sup>14</sup> About 85% of normal, post-opening trades over the three years we study were within the one cent tick size range. We also examined prices in other tick size ranges and obtained comparable results.



*Market-wide volatility.* Clustering is expected to be greater when value is less certain. Market-wide volatility can be a major source of individual stock volatility. We use the ASX's All Ordinaries Accumulation Index, which we obtained from Equinet, as our proxy for the Australian market. Volatility is calculated on a daily basis during the three year period. It is measured by the absolute value of the change in the closing index value from one day to the next, expressed relative to the previous day's index value. Each daily measure is applied to all trades for the following trading day. By lagging the index one day, any resultant clustering in trade prices is seen as a response by traders to the previous day's market volatility.

*Resources versus non-resources.* Resource stocks are characteristically more volatile (Ball and Brown, 1980) and their price should cluster more strongly. The ASX's Industry Classification code was used to classify each stock.

*Individual stock volatility.* We also measured individual stock volatility directly, by the standard deviation of weekly return estimated over the three years of our study. The effects of price discreteness on return volatility calculations are intensified the shorter the time interval between share prices. Overlapping seven-day price changes (that is, Monday to Monday, Tuesday to Tuesday, etc.) were used to mitigate these effects. A single volatility calculation for each stock is applied to all trades for that stock during the three year period.

*Firm size.* Larger firms tend to release more information to the market place and they are more closely monitored by analysts. Consequently their underlying value is known with greater precision and their prices are less likely to cluster. Firm size is proxied by the natural logarithm of the firm's average market capitalisation over the sample period.<sup>15</sup>

*Liquidity.* Liquidity is associated with efficiency in price discovery; the more liquid the stock the more precisely its value is known, and the less likely its price will cluster. Liquidity is proxied by the natural logarithm of trading frequency, defined as the average number of trades per trading day for that stock over the three year sample period.

*Optioned versus non-optioned stocks.* Active trading in options leads to more efficient pricing of the underlying stock (Damodaran and Subrahmanyam, 1992), perhaps because restrictions on short selling can be overcome and because information based trading is less costly to implement on an options market. Hence the underlying stock's price is less likely to cluster. A dichotomous dummy variable is used to identify the 46 companies that had listed put and call options. The dummy variable is applied to optioned stocks' trades only for the period that their options were listed.<sup>16</sup>

---

<sup>15</sup> The data were sourced from Aitken et al. (1993b).

<sup>16</sup> We are indebted to Elvis Jarnecic for information on option listing dates.

*Short-selling.* Securities approved by the ASX for short selling are like stocks traded on an options market, in that the availability of opportunities for short selling provides incentives to collect information about the stock's value and indirectly to convey that value to the market place. Stocks approved for short-selling are thus expected to have a finer set of prices that are utilised by traders. A dichotomous dummy variable is used to identify the companies that were 'approved securities' for the purposes of short-selling on the ASX.<sup>17</sup>

*Buyer-initiated trades.* It is often found that anomalous returns persist longer for bad news. To control for possible differences between buyer- and seller-initiated trades, a dichotomous dummy variable is created. If the trade took place at the best ask price, then the incoming trader, a buyer, placed a market order and met this ask price. Such a trade is classified as buyer-initiated.<sup>18</sup>

## 4. Results

### 4.1. Univariate relationships

Frequency distributions of the last digit of trade prices are presented in Table 1 across univariate partitions of the database. It shows that clustering, over all trades, is extremely pervasive, with traders having a striking preference for prices with a final digit of 0. Final digits of 5 and the even numbers are also clearly preferred. Note especially the apparent support for the attraction theory in the clustering frequencies:  $0 > 5 > \{2 \approx 8\} > \{3 = 7, 4 \approx 6\} > \{1 \approx 9\}$ . This result contrasts with Harris's (1991) rejection of the attraction hypothesis in the USA.

We now look at each explanatory variable in turn.

*Stock price level.* To demonstrate that clustering increases with the stock's price level, Fig. 1 presents an extract of the frequency distribution for all trades with a tick size of one cent. The value of 10 cents on the Price axis indicates the price range of 11 to 20 cents, as shown on the Final Digit of Price axis. It is obvious from Fig. 1 that clustering increases with price. While clustering is evident in low-priced trades, it becomes more prominent as price rises, especially at a final digit of 0 relative to the other digits.

*Trade size.* Table 1 shows that clustering tends to increase with trade size. For instance, 18.0% of all trades above \$30,000 are transacted at a final digit of 0,

<sup>17</sup> We are indebted to Amaryllis Kua for supplying the data on 'approved securities'.

<sup>18</sup> Although we can determine if the immediate trade was triggered by a buy or sell market order, we recognise that the current trade could be one in a multiple trade game. A major buyer or seller might game by first placing an opposing market order, thereby removing a competing order from the schedule and affecting the quoted current market for the stock.

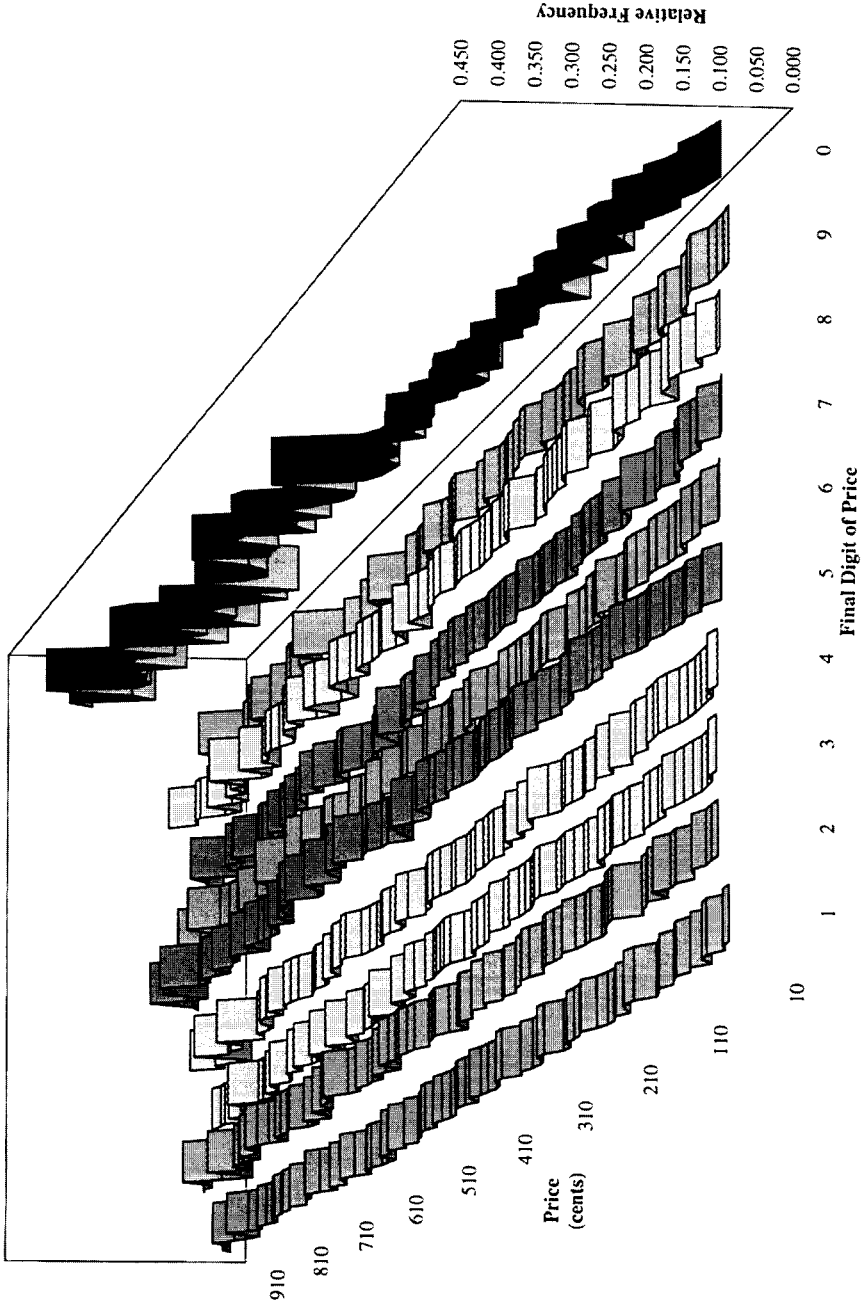


Fig. 1. Relative frequency of final digit of all SEATS trade prices with a tick size of one cent.

Table 1

Clustering of the final digit of price for various partitions of the sample of all trades of 267 stocks that traded on average at least five times a day on the Australian Stock Exchange between 3 September 1990 and 3 September 1993

Partition	Percentage of cases clustered at a final digit of										No. of cases
	0	1	2	3	4	5	6	7	8	9	
All trades	16.5	7.1	10.0	8.4	9.0	13.2	9.2	8.5	10.4	7.7	2,610,400
Trades < \$10K	15.9	7.3	9.9	8.5	9.1	13.0	9.2	8.6	10.4	8.0	1,741,858
Trades \$10K-\$20K	17.6	6.9	9.9	8.2	8.8	13.6	9.0	8.3	10.4	7.4	322,824
Trades \$20K-\$30K	17.0	6.8	10.3	8.4	8.7	13.6	9.2	8.2	10.5	7.4	137,651
Trades > \$30K	18.0	6.5	10.2	8.3	8.8	13.5	9.1	8.1	10.6	6.9	408,067
Mkt volatility < 1%	16.3	7.2	10.1	8.5	9.0	13.1	9.2	8.5	10.4	7.8	2,142,841
Mkt volatility 1%-2%	17.2	6.8	9.7	8.0	8.8	13.7	9.0	8.6	10.7	7.5	430,938
Mkt volatility > 2%	18.5	6.0	9.4	8.1	9.4	13.6	9.0	8.3	10.7	6.9	36,621
Resource stocks	16.0	7.8	10.6	8.9	8.8	13.1	9.1	8.5	10.0	7.2	761,046
Other stocks	16.7	6.8	9.7	8.2	9.0	13.2	9.2	8.5	10.6	7.9	1,849,354
Stock volatility < 4%	16.5	6.8	10.0	8.3	9.1	13.1	9.2	8.4	10.7	8.0	1,537,664
Stock volatility 4%-10%	16.2	7.4	9.9	8.5	8.7	13.3	9.3	8.8	10.5	7.5	823,573
Stock volatility > 10%	17.8	7.8	10.3	9.3	9.1	13.2	8.7	7.9	9.1	6.8	249,163
Firm size < \$36M	16.5	8.2	10.3	9.4	9.6	13.3	8.7	8.1	9.1	6.9	225,520
Firm size \$36M-\$274M	18.4	6.8	10.0	7.9	7.9	14.3	8.9	8.2	10.3	7.4	465,707
Firm size > \$274M	16.0	7.0	10.0	8.5	9.2	12.9	9.3	8.6	10.7	7.9	1,919,173
Tdg freq < 7.5	21.1	6.7	10.0	7.7	8.0	15.8	8.4	7.2	9.2	5.8	238,222
Tdg freq 7.5-18.17	20.0	6.7	10.0	7.8	8.0	14.8	8.7	7.5	9.8	6.8	460,748
Tdg freq > 18.17	15.1	7.2	10.0	8.7	9.3	12.5	9.4	8.9	10.8	8.2	1,911,430
Optioned stocks	14.3	7.5	10.0	8.9	9.5	11.8	9.5	9.1	10.8	8.5	1,368,449
Other stocks	18.9	6.6	10.0	7.9	8.4	14.7	8.8	7.8	10.1	6.9	1,241,951
Short selling stocks	15.6	7.2	10.0	8.5	9.2	12.7	9.4	8.7	10.7	8.1	1,845,095
Other stocks	18.5	6.9	9.9	8.2	8.5	14.4	8.8	7.9	9.9	7.0	765,305
Buyer-initiated	16.9	6.2	9.8	8.2	9.6	13.2	8.5	8.6	10.8	8.3	1,348,599
Seller-initiated	16.1	8.0	10.2	8.7	8.3	13.1	9.9	8.4	10.1	7.1	1,261,801

which is a substantial departure from the 10% expected under a uniform distribution and is a higher percentage than for the other three trade size partitions. Clustering tends to be reflected in larger trades having a higher frequency of prices with a final digit of 0, and a lower frequency of 1 and 9.

*Market-wide volatility.* Clustering appears to increase during periods of higher market volatility, as expected. The effect is pronounced when the index changed the day before by more than 2%. The relationship between market volatility and price clustering is more apparent in the extremes: when volatility is higher there are more trades with a final digit of 0 or 5 and fewer with 1 or 9.

*Resource stocks versus non-resource stocks.* Clustering is expected to be more common among resource stocks, due to their higher volatility. Of the 267 companies in the database, 108 (40%) were classified as resources. The univariate results show the opposite of what is expected, since although clustering is clearly present in both groups, it appears to be less prevalent in resource stocks.

*Individual stock volatility.* Clustering should increase with individual stock volatility. There is a noticeable increase in clustering at the final digit of 0 for stocks that are the most volatile, but other aspects of clustering are less obviously related to stock volatility.

*Firm size.* Clustering is expected to decrease at the individual stock level with increasing firm size. Each firm size partition corresponds to a third of the companies in the sample. As expected, there is less price clustering in the largest 89 companies (market capitalisation > \$274 million). However, the results are not uniform across size partitions, nor do they extend to all aspects of price clustering.

*Liquidity.* Clustering should decrease with increasing liquidity. Liquidity in Table 1 is proxied by the average number of daily trades for each stock. As with firm size, the database has been partitioned three ways according to the stocks' ranked average trading frequency, so that 89 stocks fall into each partition. Clustering decreases markedly with increasing trading frequency, as expected.

*Optioned versus non-optioned stocks.* Optioned stocks are expected to display less clustering in the final digit. That is clearly what we find.

*Short selling.* Stocks that may be sold short are less likely to trade at a clustered price. Again, the results confirm our conjecture.

*Buyer-initiated versus seller-initiated trades.* Some 51.6% of all trades took place at the ask. The results suggest that clustering, at least at the final digit of 0 and 5, is more prevalent in buyer- rather than seller-initiated trades.

#### 4.2. Multivariate analysis

We employ multivariate logistic regression to identify the combined ability of variables to explain clustering. A logistic regression predicts a binary dependent variable (clustered/not clustered) conditional on the values of the explanatory variables. Clustering is measured as a binary variable in three ways. Price clusters if the final digit is 0 rather than 1, 2, 3, 4, 5, 6, 7, 8 and 9; 0 or 5 rather than 1, 2, 3, 4, 6, 7, 8 and 9; and 0, 2, 4, 6 or 8 (evens) rather than 1, 3, 5, 7 and 9 (odds).

First, all potentially explanatory variables are regressed against the dichotomous variable. Then, we fit a more parsimonious regression model that incorporates only five explanatory variables (stock price, market-wide volatility, stock volatility, firm size and trading frequency). These variables, apart from market volatility, were incorporated by Harris (1991) in his aggregated model.

#### 4.2.1. When all the explanatory variables are included

Table 2 reports the results of the logistic regressions for the three definitions of the dependent dichotomous dummy variable that denotes if the price was clustered. In brief, the directional and significance results are as hypothesised for all variables, apart from firm size and resource stocks. These departures coincide with the results from the univariate analysis. Coefficients for these two variables are significant at the 0.0001 level or better if we apply a two-tailed test of significance.

Results for firm size are perhaps surprising because it is a surrogate for factors that are negatively related to clustering.

One way to summarise how well a model explains the data is its predictive accuracy. The percentage of trades classified correctly (as clustered or not clustered) is a relatively high 83.87% in the first logit model (encompassing the first definition of clustering), and declines across the next two definitions. We can compare each model's predictive accuracy with that of two benchmarks, the naive and chance models. The naive model predicts that all prices will belong to the most frequently observed group. Define the observed relative frequency of trades in the clustered group as  $p$ . The naive model's classification accuracy is thus  $\max(p, 1 - p)$ . A weakness is that the naive model predicts no variation in classifying the data: every case is assigned to the same group. The chance model randomly assigns the  $N$  cases to the clustered and unclustered groups such that the

Table 2

Logistic regression coefficients using all independent variables. The dependent variable is whether or not the final digit of price had a particular value. The sample is all trades of 267 stocks that traded on average at least five times a day on the Australian Stock Exchange between 3 September 1990 and 3 September 1993

Variable	Mean	Expected sign	Final digit		
			0	0 or 5	0,2,4,6,8
Price level (\$)	2.94	+	0.0016 *	0.0015 *	0.0007 *
Market volatility (%)	0.58	+	0.0703 *	0.0706 *	0.0282 *
Stock volatility (%)	5.03	+	0.0206 *	0.0176 *	0.0057 *
Firm size (ln \$M)	20.38	-	0.0722	0.0797	0.0258
Trading frequency (ln #)	3.78	-	-0.3111 *	-0.3245 *	-0.1027 *
Trade value (ln \$)	8.70	+	0.0511 *	0.0500 *	0.0205 *
Options-traded stock	0.52	-	-0.2074 *	-0.2222 *	-0.0724 *
Buyer-initiated	0.52	?	0.0609 *	0.0468 *	0.0288 *
Resource stock	0.29	+	-0.0265	-0.0334	-0.0093
Short selling stock	0.71	-	-0.0477 *	-0.0458 *	-0.0118 **
Constant		?	-2.9054 *	-2.1693 *	-0.3136 *
Number of cases			2,610,400	2,610,400	2,610,400
% correctly classified			83.52%	70.92%	55.00%

\* Significant at  $< 0.0001$ .

\*\* Significant at  $< 0.01$ .

Table 3

A comparison of the predictive accuracy of four models of price clustering (% of all prices that were correctly classified). The sample is all trades of 267 stocks that traded on average at least five times a day on the Australian Stock Exchange between 3 September 1990 and 3 September 1993

Final digit	$p^a$	Logistic regression model		Naive model	Chance model
		All variables	Parsimonious		
0	16.49%	83.52%	83.51%	83.51%	72.46%
0 or 5	29.67%	70.92%	70.86%	70.33%	58.27%
0, 2, 4, 6, 8	55.08%	55.00%	55.05%	55.08%	50.52%

<sup>a</sup>  $p$  is the observed relative frequency of trades with a final digit of 0, 0 or 5, etc.

proportion of trades randomly assigned to each group is the same as the observed proportion ( $p$  for the clustered,  $1 - p$  for the unclustered group). Thus the chance model's expected predictive accuracy is  $1 - 2(p - p^2)$ .

Table 3 summarises the predictive accuracy of our logit models, and of the naive and chance models, across the three definitions of clustering. The logistic regression with all variables had the greatest predictive accuracy across the first two definitions of clustering. It equated to the naive model in the third definition. Both logit models outperformed the chance model across each definition of clustering.

#### 4.2.2. Parsimonious model

Regression results for the more parsimonious model, which has only five explanatory variables, were consistent with the first regression model utilising all 11. Again, the influence of firm size was opposite to that expected. The parsimonious model has slightly less predictive content than the first model incorporating all variables, across the three definitions of clustering, as can be seen from Table 3.<sup>19</sup>

#### 4.2.3. Year-by-year models

As a robustness check, three sets of regressions were run, one for each year of trades. The predictive accuracy of the overall model is consistently high over each of the three years. In each year, it outperforms the chance model. However, there were changes in the sign and significance of the variables representing firm size, resource stocks, trades that were buyer- rather than seller-initiated, and stocks that could be sold short.

#### 4.2.4. Other robustness checks

The logit models are not obviously superior to the naive model when fitted to all of the data. It is thus possible that the logit model estimates have converged on

<sup>19</sup> The full results are available from the authors.

the naive model. Further investigation revealed that to be almost so. For instance, the logit model using all explanatory variables correctly classified only 0.2% of cases clustered under the first clustering definition and 99.7% of cases not clustered;<sup>20</sup> these rates may be compared with the corresponding rates of 0% and 100% for the naive model.

Consequently we investigated the case where the naive and chance models have the same accuracy rate: i.e., when the relative frequencies of the clustered and not clustered groups are both 0.5. At the same time we changed our second and third clustering definitions to exclude cases ending in 0 from the second clustering definition and cases ending in 0 or 5 from the third. The outcome is an hierarchy of clustering definitions: 0 v. 1–9; 5 v. {1–4} + {6–9}; and {2, 4, 6, 8} v. {1, 3, 7, 9}. We also considered three time periods: before the change in the price tick rule effective 28 October 1991, after the rule change, and both periods combined.

In brief,<sup>21</sup> the logit models' overall accuracy rate exceeded 50% regardless of the time period and the clustering definition,<sup>22</sup> although the clustered cases were consistently predicted less accurately. The coefficients of the stock price, market volatility, trade size, trading frequency and the option stock dummy variable always had their predicted sign (i.e., for all relevant combinations of model, time period and clustering definition) and there were relatively few inconsistencies in the signs of the other coefficients. Nonetheless, as before, the coefficients of firm size and the dummy variable for a resource stock were almost always the reverse of what we predicted.

We conclude that our findings are relatively robust to time period, tick size range, clustering definition and overall clustering frequency.

## 5. Conclusions

Price clustering is strongly manifest in normal trades on the ASX after the market opens. As expected, we found that clustering increased with the level of the price, with market volatility, own stock volatility, trade size, and the size of the bid–ask spread. It decreased with trading frequency, and also was lower for stocks with options traded on them. Stocks that could be sold short clustered less strongly; when we partitioned our data into three calendar periods of 12 months each, the results were inconclusive in the first period and relatively weak in the second and third periods.

---

<sup>20</sup> The respective accuracy rates for the second definition were 4.4% v. 99.0% or 71% overall; and for the third definition, 97.1% v. 3.4% or 55% overall.

<sup>21</sup> The full results are available from the authors.

<sup>22</sup> The lowest accuracy rate was 52.7% (parsimonious model, third clustering definition, overall time period) and the highest accuracy rate was 57.6% (full model, first clustering definition, second sub-period).



In contrast to Harris (1991), we found support for his attraction hypothesis. Two other findings contrary to what we expected were that larger firms exhibited more not less clustering, and resource stocks were clustered less than industrials, which may call into question the appropriateness of our proxy variables. For instance, our use of a single average market capitalisation metric for the whole three years for each company may have been too crude a proxy to capture firm size effects.

In future work we will extend our study to encompass price clustering separately in market and limit orders, in buy versus sell orders, and in daily highs and lows. We will also investigate any association between price clustering and time-related anomalous returns.

### **Acknowledgements**

The authors are indebted to the Australian Stock Exchange for making the data available, to the Australian Research Council and their respective universities for financial assistance, to the participants in workshops at the Australian Graduate School of Management and the Universities of Adelaide and South Australia, and to our colleagues (especially David Butler and Alex Clarke) at the Universities of Sydney and Western Australia.

### **References**

- Aitken, M., P. Brown and T. Walter, 1993a, The bid/ask schedule under an automated trading regime. Working paper 93-33 (Research Centre in Accounting and Finance, The University of Western Australia, Nedlands).
- Aitken, M., P. Brown and T. Walter, 1993b, Intraday patterns in returns, trading volume, volatility and trading frequency on SEATS, Working paper 93-32 (Research Centre in Accounting and Finance, The University of Western Australia, Nedlands).
- Ball, R. and P. Brown, 1980, Risk and return from investments in the Australian mining industry: January 1958–February 1979, *Australian Journal of Management* 5, 45–66.
- Ball, C.A., W.A. Torous and A.E. Tschoegl, 1985, The degree of price resolution: The case of the gold market, *Journal of Futures Markets* 5, 29–43.
- Buckland, C., 1994, An investigation of the prevalence of price clustering on the Australian Stock Exchange, Unpublished BCom Honours thesis (The University of Western Australia, Nedlands).
- Butler, D.H. and G. Loomes, 1988, Decision difficulty and imprecise preferences, *Acta Psychologica* 68, 183–196.
- Christie, W.G. and P.H. Schultz, 1994, Why do NASDAQ market makers avoid odd-eighth quotes?, *Journal of Finance* 49, 1813–1840.
- Damodaran, A. and M.G. Subrahmanyam, 1992, The effects of derivative securities on the markets for the underlying assets in the United States: A survey, *Financial Markets, Institutions and Instruments* 1, 1–21.
- Easley, D. and M. O'Hara, 1987, Price, trade size and information in securities markets, *Journal of Financial Economics* 19, 69–90.

- Goodhart, C. and R. Curcio 1991, The clustering of bid/ask prices and the spread in the foreign exchange market, Discussion paper no. 110, Financial Markets Group Discussion Paper Series (LSE, London).
- Hameed, A. and E. Terry, 1994, The effect of tick size on price resolution. Unpublished paper presented to the Annual Conference of the Asia Pacific Finance Association, Sydney, 28–30 September 1994.
- Harris, L., 1991, Stock price clustering and discreteness, *Review of Financial Studies* 4, 389–415.
- Niederhoffer, V., 1965, Clustering of stock prices, *Operations Research* 13, 258–265.
- Niederhoffer, V., 1966, A new look at clustering of stock prices, *Journal of Business* 39, 309–313.
- Niederhoffer, V. and M.F.M. Osborne, 1966, Market making and reversal on the stock exchange, *Journal of the American Statistical Association* 61, 897–916.
- Osborne, M.F.M., 1962, Periodic structure in the Brownian motion of stock prices, *Operations Research* 10, 345–379.