4

Occasional Paper Series

Market Cleanliness in the New Zealand Equity Market

OCTOBER 2024

Nam Ngo, Brianne Leef, Sharon Louie, Lynda Sanderson and Stuart Johnson

FMA
FINANCIAL MARKETS AUTHORITY
TE MANA TĀTAI HOKOHOKO

Contents

Foreword by Stuart Johnson, Chief Economist

Market cleanliness goes to the heart of the Financial Markets Conduct Act's purpose – to promote fair, efficient and transparent markets. Both New Zealand businesses looking to raise capital and Kiwi investors, need to be confident that the stock market is safe and free of any large-scale manipulation. A clean stock market also attracts overseas investment, lowering the cost of capital to NZ businesses, supporting firms to manage financial risk and driving innovation. This report shows that over the past 20 years market cleanliness on the New Zealand Equity Market has improved, especially following the introduction of the FMC Act in 2013. This is encouraging for investors, both domestically and internationally, and for Kiwi firms whose stocks are traded on the NZX.

Market Cleanliness is the idea that price sensitive announcements should drive stock price movements only after the announcement is released. Any movement in price before an announcement could suggest some form of information leakage or market manipulation. At the most basic level, if a trader purchases a stock shortly before a positive announcement from the firm, which leads to an increase in the stock price, the trader may potentially have had inside information. If stock prices are moving before an announcement, it could suggest information is leaking and some traders are acting on private information. This damages the integrity of the market. Insider trading and other market manipulative practises extract profits from genuine, law-abiding investors and, in extreme cases where investors shy away from public markets, can increase the cost of capital for firms.

To measure market cleanliness, we use a group of macro or market level metrics which indicate the amount of insider knowledge evidenced on the market. The report examines the number of times prices moved before a price sensitive announcement over the total number of price sensitive announcements. As is the convention internationally, we also use a volume-based measure.

The results show that market cleanliness in the New Zealand Equity Market has improved over the last 20 years. This is evidenced in both the price-based and volume-based measure. While the measures are somewhat volatile, moving from year to year, both metrics show a downward shift over the period (see the graphs on pages 14 and 21 for further detail). This should give investors increased confidence.

Market cleanliness measures capture a macro picture of behaviour on the stock market. At a micro level, stock prices can move ahead of an announcement for a variety of reasons. In 2023, stock prices moved abnormally in advance of a price sensitive announcement in around 1 in 20 cases. This should not raise concerns, as there will always be some cases where prices were trending in one direction ahead of an announcement. Investors should have confidence that these micro situations are already well oversighted by NZ RegCo, the market operator oversight firm. NZ RegCo use a system called SMARTS, a world leading system to monitor market manipulation. Additionally, NZ RegCo and the FMA receive alerts when brokers are concerned about micro level behaviour.

The FMA intends to undertake a review of market cleanliness using the measures outlined in this report and update them on a regular basis. We would like to thank Professor Hai Lin from Victoria University Wellington, a leading expert on market cleanliness who has published widely on the topic. Professor Lin has provided a high-level peer review of our approach, which is included on page 29.

Executive summary

As New Zealand's financial conduct regulator, the Financial Market Authority (FMA) has a statutory requirement to ensure our capital markets are fair, efficient and transparent, to encourage investors to participate and allow businesses to raise funds. Insider trading, among other forms of market abuse, is a threat to the integrity and reputation of markets as it leads to unfair advantages for some investors over others. Our focus is detecting and deterring insider trading and other forms of market abuse. This requires a combination of strong regulation and strict enforcement of the rules, which in turn require high-quality data and statistics.

Improving how we measure and monitor the prevalence of insider trading in the market assists the FMA in our response to market abuse. This report serves that purpose by examining the cleanliness of the New Zealand equity market since 2004. Two main indicators were developed. The Market Cleanliness Statistic (MCS) looks at the abnormality in *stock return* prior to material announcements, given the previous relative performance compared to the market. The Abnormal Trading Volume Ratio (ATVR) assesses whether there are significant differences in the *trading volume* prior to an announcement compared to the usual level.

We found that while both measures are highly volatile year-on-year, there has been a downward shift in the pre-announcement abnormality in both stock returns and trading volume (indicating an overall improvement in market cleanliness) over the past 20 years. Among other regulatory events which took place over the period, the passing of the FMC Act into law in 2013 appears to be correlated with the most significant improvement in market cleanliness indicators. This legislation was soon followed by the first prosecution of insider trading in 2017, though this latter event did not have a clear impact on overall cleanliness measures. A combination of strict regulation and effective enforcement is key to stopping insider trading and other misconduct.

Market volatility is a strong predictor of market cleanliness. Highly volatile periods tend to coincide with more abnormal pre-announcement price movements (APPM) or abnormal trading volumes (ATV). Positive announcements were found to be cleaner than negative ones, which could be due to the absence of a publicly available register of short selling positions in New Zealand, like those found in the UK and Australia. It can also suggest that surveillance and enforcement activities associated with positive announcements may have a stronger deterrent effect on insider trading. Across the market, there is variation in the levels of measured cleanliness across industries and market capitalisations. However, these differences are not strongly significant and are not consistent across the two measures.

The MCS and ATVR appear to be positively correlated with each other, but the relationship is not statistically significant. The two measures capture different aspects of market behaviour, as well as having fundamental differences in methodological approach and sample size. Therefore, it is important to look at a combination of these indicators when analysing the market and sub-markets.

While the two measures can provide useful indications, especially on the long-term trends, they also have limitations. Some of these can be mitigated while others are the byproducts of the statistical approach. For example, it is not possible to attribute all the changes in MCS and ATVR to the prevalence of insider trading. As shown in the analysis, these indicators can move because of genuine market conditions, especially during times of stress. At the same time, as they are both top-down stock-level measures and rely on the power of statistical tests, some actual insider trading activities may not be captured if they were not material enough to move the series. Therefore, it is crucial to focus on longer-term changes rather than year-on-year movements before drawing conclusions.

1. Introduction

The Financial Markets Authority (FMA) has an objective of promoting fair, efficient, and transparent financial markets, including capital markets. We want to ensure these markets reflect genuine supply and demand, in order to preserve their reputation and integrity. This is crucial to encourage investors to participate and allow businesses to raise funds.

Fair, efficient, transparent markets work on the basis that everyone trading has access to the same information. If some people have material information before others and are allowed to take advantage through insider trading, that undermines the fairness of the market and causes harm to others. If this type of misconduct (together with market manipulation) appears to be prevalent, investors may trade more cautiously or pull out of the market altogether. This can increase the cost of capital and have lasting effects on businesses and the wider economy. The detection and deterrence of insider trading requires a combination of strict regulation, effective enforcement of the rules, and high-quality data and statistics.

It is informative to measure and monitor the prevalence of insider trading (and other types of misconduct) at the market level. This adds context to the transaction-level surveillance activities conducted by NZ Reg Co, the market operator oversight firm of the New Zealand Stock Exchange. This report provides market-level context by examining the cleanliness of New Zealand's equity market[1] since 2004 and assessing whether there are signs of possible insider trading. This helps the FMA further understand if current regulations are effective in deterring individuals from insider trading, and if not, what alternative methods we should be proposing. Insider trading laws alone do not prevent insider trading, but efficient enforcement (Bhattacharya & Daouk, 2002) or media coverage on insider trading matters (ASIC, 2024) can discourage such behaviour.

In this report, we explore two measures of market cleanliness. Both measures rely on the difference in how a clean market (with symmetric information) and an unclean market (with information leakage or potential insider trading) would behave. An unexpected and significant movement in the stock return or trading activity *before* a material announcement, in an appropriate direction based on the nature of the announcement, could be indicative of insider trading.

The key difference between the two measures is which aspect of trading behaviour is used to determine abnormality. The Market Cleanliness Statistic (MCS) is a price-based measure, which assesses the abnormality in stock return before an announcement. The second measure is the Abnormal Trading Volume Ratio (ATVR), which assesses the abnormality in volume traded before an announcement compared to the usual level. Both measures rely on the difference in how a clean market (with symmetric information) and an unclean market (with information leakage or potential insider trading) would behave. For the MCS, in a clean market, we expect stock prices to react instantaneously *after* an announcement (e.g. higher price if the news is good). Similarly, for the ATVR, trading activity is also expected to increase *after* an announcement, irrespective of whether it is positive or negative, as people react to the news. These statistics track specific market behaviour and will not represent all insider trading. However, they provide a snapshot that will highlight market inefficiencies.

[1] Note that in this report, when we refer to the New Zealand equity market, we mean the traditional New Zealand Stock Exchange (NZX). There is also the Catalist stock exchange which is smaller and targets small cap firms; this is not part of our analysis.

An unexpected and significant movement in the stock return (in an appropriate direction based on the nature of the announcement), or an increase in the volume of trade *before* a material announcement, could be indicative of insider trading. These methodologies follow and extend previous studies by the FMA (Hensen, 2018) and other regulators including the UK's Financial Conduct Authority (Dubow & Monteiro, 2006; Goldman et al., 2014) and the Australian Securities and Investments Commission (ASIC, 2016; ASIC, 2024). We also assess how these measures change over time and across markets. This analysis has the potential to be turned into a regular release, which can be used to update investors on the integrity of the market.

The rest of the report is structured as follows. Section 2 gives background information on insider trading and how the FMA and other regulators deal with it. Section 3 presents a review of the literature. Section 4 presents the results for the Market Cleanliness Statistic. Section 5 presents the results for the Abnormal Trading Volume Ratio. Section 6 compares the two measures. Section 7 concludes and outlines potential next steps.

2. Background

Insider trading occurs when people with material information[2] conduct trade using this information before it becomes publicly available. Sections 240 - 243 of the 2013 Financial Markets Conduct (FMC) Act prohibit people who hold material information about an issuer that is not generally available to the market (inside information) from trading with that information, disclosing it in certain circumstances, and advising or encouraging other individuals to trade the issuer's shares. Recent insider trading cases the FMA has worked on include Eroad (FMA, 2017), Promisia (FMA, 2019), Pushpay (FMA, 2023), and Heartland Group Holdings (FMA, 2024).

In New Zealand, the NZX Regulation Limited (NZ RegCo)'s surveillance team oversees and detects forms of market misconduct by monitoring movements in stock prices and trading volumes. An in-depth analysis by security, participant or client is completed using technologies and data sources such as the surveillance platform SMARTS, market information from NZX's trading system, and different databases. If insider trading or other market misconduct is detected and needs further investigation, the NZ RegCo surveillance team would refer that to the Participant Compliance team and/or the FMA (NZ RegCo, n.d; NZX Limited, 2024).

In the United Kingdom, the Financial Conduct Authority (FCA) uses SMARTS to analyse transaction reporting data, order book data, benchmark submission and other forms of market data. The market monitoring department regularly connects with the industry and encourages market participants to report suspicious transactions through submitting suspicious transaction and order reports (STORs) (FCA, 2023).

Similarly, the Australian Securities and Investments Commission (ASIC)'s ARTEMIS project, which hunts and finds suspected market misconduct, profitable or suspicious trading patterns won the Australia Public Service Data Analytics and Visualisation Award in 2023 for its sophisticated insider trading surveillance and detection capability (ASIC, 2023). The project combines and uses advanced algorithms to analyse data. Other countries not mentioned in this report have different methods for how they detect and deal with insider trading.

[2] Material information is information that a reasonable person would expect to have a material effect on the price of financial products quoted on a licensed market, if the information was generally available to the market.

3. Literature review

The international literature on market cleanliness has evolved over time, with many studies by both academics and financial regulators. See Table 1 in the Appendix for a short summary of previous studies by the FMA and other international regulators.

The United Kingdom

The Financial Services Authority's (FSA) Occasional Papers 23 (Dubow & Monteiro, 2006) and 25 (Monteiro, Zaman, & Leitterstorf, 2007) were the first research by a financial regulator on market cleanliness using the empirical methods we adopt here. Dubow and Monteiro established the market cleanliness measure by creating a price-based measure to see whether price movements are observed ahead of significant price-sensitive takeover announcements. Occasional Paper 25 provides an update on the measure from the previous paper, including the way abnormal returns are identified and how regulatory changes impact the measure. Many other studies have been published based on the methodologies from these papers.

The next paper conducted by the FCA (the successor to the FSA) is Occasional Paper 4, looking at market cleanliness in UK equity markets between the years 2010 to 2013 (Goldman et al., 2014). This methodology is based on the FSA's Occasional Paper 25. The two changes made in this update include that only takeover events are considered and that a simple model was applied to calculate expected and abnormal returns without considering market volatility.

The FCA continues to publish their findings on market cleanliness in an annual report and has introduced new measures over time. Two new measures were added in 2019 – the Abnormal Trading Volume[3] (ATV) measure and the Potentially Anomalous Trading Ratio[4] (PATR), which are both volume-based measures (FCA, 2019). While the former is based on stock-level data, the latter is based on account-level data, which was made available thanks to recent changes in the reporting regime.

Australia

ASIC Report 487 provides an overview of the cleanliness of Australia's equity markets from the years 2006 to 2015, based on the FCA's price-based measure (ASIC, 2016). This paper includes all material announcements instead of just takeover events and 5-day event windows are used as opposed to the FCA's 2-day event window. ASIC also uses the market capital pricing model (CAPM) to calculate the abnormal returns of stocks. This measure is included in our sensitivity tests in Appendix B.

Two new market cleanliness measures were introduced in this research using trader account level data obtained from ASIC's Market Analysis Intelligence (MAI) surveillance system. This is similar to the FCA's Potentially Anomalous Trading Ratio. It flags trader accounts as suspicious based on timeliness, profitability, and trading patterns. Report 623 is a follow-up project for the years between 2015 to 2018 (ASIC, 2019). The methodology

---

[3] To derive this, apply a T-test to measure if trading volumes in the benchmark window versus the announcement window are significantly different at the 5% significance level.

[4] The PATR requires trade account-level data to identify trading accounts with abnormal anomalous underlying trading behaviour before potentially price sensitive news announcements (PPSNAs), which is compared to the historical trading behaviour.

remains the same as the previous report. The most recent publication, Report 786, finds that Australia's equity market is one of the cleanest compared to other equity markets (ASIC, 2024)

New Zealand

The FMA published the first study on market cleanliness in New Zealand in 2018, with findings relevant for the years 2010 to 2016. The author largely followed the methods by the FCA (Dubow & Monteiro, 2006; Monteiro, Zaman, & Leitterstorf, 2007; Goldman et al., 2014) and ASIC (ASIC, 2016) to produce a price-based market cleanliness measure while modifying some parts to align better with characteristics of New Zealand's equity market.

4. Market Cleanliness Statistic

4.1 Methodology

The price-based Market Cleanliness Statistic (MCS) relies on how stock prices behave differently around a price-sensitive, material announcement in a clean and unclean market, as illustrated in Figure 1. In a "clean", strongly efficient market and given that there is sufficient liquidity, stock prices are expected to increase *instantly* following a positive, material price sensitive announcement (MPSA). This is called trading on good news and is a normal behaviour. Similarly, stock prices should fall following a negative announcement. However, if the market is "unclean" and there is information leakage or potential insider trading, stock prices may exhibit abnormal movements *prior* to the announcement. This behaviour is usually referred to as price run-up.

Figure 1: Market cleanliness statistic – concept

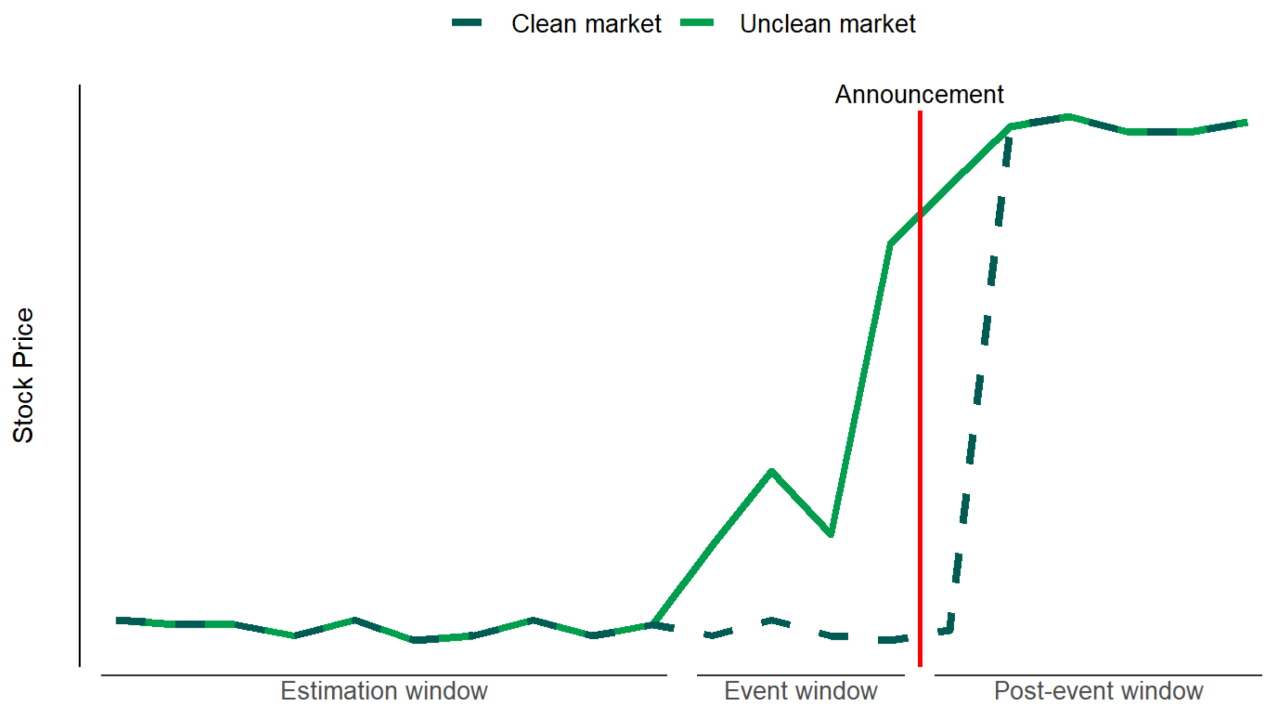


Figure 1 depicts a stylised timeline around a market announcement. The pre-announcement period of interest is called the event window. The analysis also considers the post-event window, which includes the day of the announcement and some days after.

For each announcement, the total daily stock returns including reinvested dividends are calculated. As these are usually small, the following methodologies can be used interchangeably:

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \tag{1}$$

Alternatively, the difference in logs:

$$R_{i,t} = Ln \frac{P_{i,t}}{P_{i,t-1}} \tag{2}$$

where $R_{i,t}$ is the return and $P_{i,t}$ is the adjusted close price of stock i on day t.

Next, linear regression is used to build a predictive model for a stock's expected returns $E(R_{i,t})$ based on the market return during the previous year (equation 3). This is called the market model and is the most common method to estimate expected returns in event studies (MacKinlay, 1995), consistent with the FCA (Dubow and Monteiro, 2006; Goldman et al., 2014). There are 250 trading days in a year and this analysis removes the ten days immediately preceding the day of interest as those returns may be directly impacted by the announcement. The remaining 240 days are called the estimation window.

$$E(R_{i,t}) = \alpha_i + \beta_i R_{m,t} \tag{3}$$

where $E(R_{i,t})$ is the expected return of stock i and $R_{m,t}$ is the return of the market index on day t.

The abnormal return of stock i on day t is the difference between actual returns and expected returns:

$$AR_{i,t} = R_{it} - E(R_{i,t}) \tag{4}$$

Each observed announcement is assessed for materiality by examining how stock returns change in the four days around the announcement date relative to usual, similar to the analysis done by ASIC (ASIC, 2016) and FCA (Dubow & Monteiro, 2006). The 4-day Cumulative Abnormal Return (CAR) for the two days before the announcement, the announcement date, and the day after the announcement is calculated as:

$$CAR_{MPSA} = \sum_{j=-2}^{1} AR_{i,(t+j)} \tag{5}$$

An announcement is considered a material price-sensitive announcement (MPSA) if the CAR is extreme. This is determined using bootstrapping methodology. For each day during the estimation window, the abnormal returns are calculated using the market model in equation (3). Then, a bootstrapped sample of four non-consecutive abnormal returns are taken from the estimation window and added together. That process is repeated 10,000 times to create a sample of "normal" CARs. The CAR in equation (5) is extreme if it is smaller than the sample's 1st percentile or larger than the 99th percentile.

Once the materiality of an announcement is identified, we use a similar approach to analyse the abnormality of price movement. The process of calculating CAR and sampling the estimation window is repeated, but this time only for the event window prior to the announcement. Following ASIC (ASIC, 2016; ASIC, 2024) and FMA (Hensen, 2018), a five-day event window is used for this step:

$$CAR_{APPM} = \sum_{j=-5}^{-1} AR_{i,(t+j)} \tag{6}$$

There are two conditions for determining if the pre-announcement price movement is considered abnormal (APPM):

- The event window CAR is extreme, i.e. less than the 1^{st} percentile or larger than the 99^{th} percentile of the bootstrapped sample.
- The event window CAR matches the direction of the 4-day CAR in equation (5), which was used to determine the materiality of announcements. This means both should be positive (for "good news") or both should be negative (for "bad news").

After the MPSAs and APPMs are identified, the MCS can be calculated as the share of all material announcements which exhibit an abnormal price movement:

$$MCS = \frac{Number\ of\ material\ announcements\ with\ APPMs}{Number\ of\ material\ announcements} \tag{7}$$

An increase in the MCS indicates a deterioration in the cleanliness of the market (which may be because of insider trading). That is, the MCS rises when a larger share of announcements are *preceded* by abnormal price movements. A fall in the MCS would therefore mean an improvement in market cleanliness.

For our analysis, the market model is chosen because it is the most common method for event study in the literature. It is also consistent with the method by the FCA (Dubow and Monteiro, 2006; Goldman et al., 2014). However, other parameter choices like the five-day event window are more consistent with ASIC (2016, 2024) due to the similarities between the New Zealand and Australian markets.

The 1^{st} and 99^{th} percentiles are also consistent with international studies. However, for a small market like New Zealand, these strict criteria can reduce the sample size and increase the volatility of the series. Other potential candidates are the 5^{th} and 95^{th} percentiles. While loosening the significance levels can alleviate the sample size issue, it also increases the risk of misidentifying material announcements and abnormal pre-announcement price movements. Sensitivity tests are used to ensure the results are robust to these choices. The results, which are broadly consistent with the chosen measure, are included in the appendices.

4.2 Data

The securities used in this analysis were those listed on the NZX Main Board in April 2024.[5] The historical stock and index price data were collected from Yahoo Finance for the 2003 to 2024 period. Our analysis includes both listed equities and funds but excludes equity warrants. Listed funds, including ETF, are traded similar to equities, and insider traders can use them to conceal their insider knowledge of a single company (Scannell, 2016; Johnson, 2023; Eglite et al., 2023).

The S&P NZX All index is chosen to represent market return. All price-sensitive announcements from the same period were collected from NZX Company Research.

Some filtering was applied to the announcements:

- Since the stock and index data were at daily frequency, multiple announcements on one day were treated as just one announcement for the associated stock.
- If two announcements from the same stock occurred within 10 days, the later announcement was removed from the analysis because the stock's behaviour would already be affected by the first announcement.
- Any announcement that fell in the first 250 trading days of a stock's data was removed due to an insufficient estimation window prior to the announcement.

4.3 Limitations and mitigations

The Market Cleanliness Statistic (MCS) highlights market inefficiencies and can signal the prevalence of market misconduct, providing a valuable tool for tracking and understanding changes in the market environment. However, a single top-down metric cannot provide a holistic view of market cleanliness. There are three key reasons for this.

Firstly, significant price movements prior to an announcement may be caused by factors other than insider trading. Even in a clean market, speculation may lead to material price changes in the lead-up to an announcement. Periods of market stress and price volatility have been shown to be associated with increases in pre-announcement abnormal price movements (ASIC, 2024). Additionally, where some stocks are traded infrequently or in low volumes, the change from no activity to some activity can lead to a significant abnormal price movement. This is particularly relevant for New Zealand given the small size of our equity market.

The chance of speculatory trading being the sole cause of abnormal price movements is mitigated by our requirement of a directional match between the movement before and after an announcement, as per the FCA's original methodology. However, we cannot rule out that some abnormal price movements were driven by lucky speculation, rather than by investors having access to information in advance of the announcement.

[5] This is chosen to maximise the number of firms or funds with a full set of characteristics, which will be used in the multivariate analysis in section 4.5 and 5.4. This choice may have introduced some bias, especially in the earlier years in which there is an overrepresentation of long-lived firms, which may have more robust information management systems. However, we do not expect this to be a major issue due to the relatively small number of delistings during the period.

Secondly, and conversely, the presence of insider trading will not always result in an APPM. For inside information to lead to a significant movement in the price of a security prior to an announcement, that information would need to be known and acted on either by a large number of people or by a few major shareholders, immediately before the announcement. This leaves the various other scenarios of insider trading unaccounted for. For example, an individual with a small holding may act on private information to their own benefit without significantly affecting the price of the security. Similarly, a major shareholder may become aware of inside knowledge long before an announcement, giving them sufficient time to move their shares in small increments or at an inconspicuous time.

High-level data produces high-level conclusions. A future, more granular analysis using transaction reporting data, such as that available in SMARTs, could capture those unaccounted-for scenarios. For now, the measures' simplicity offers a clear perspective on the overall market's condition and facilitates comparative analysis with other countries.

Finally, as mentioned before, New Zealand has a small equity market. There were only 58 securities included in the first year of this analysis and 177 in the latest year. These theoretically can provide sufficient data for a statistically sound analysis. However, a small sample size makes the statistic noisier compared to larger countries' studies because each announcement has notable weight in the calculation. This can be seen in the following section on results. It is important to keep this in mind when analysing different segments of the market and interpreting the results.

For the headline measure, we have made several parameter and modelling choices. We test the robustness of our measure against these choices, with results available in Appendix B.

4.4 Results

Table 1 and Figure 2 show the price-based stock-level Market Cleanliness Statistic (MCS) in the New Zealand equity market from 2004 to 2024. The 2024 figure is provisional as it only covers the first four months of the year. The series appear to be volatile, with no consecutive increases or decreases. This reinforces the need to focus on the longer-term trend rather than year-on-year movements.

In 2020, there was a sudden and significant deterioration in market cleanliness (with 19 APPMs out of 79 MPSAs), which coincided with the period of extreme market volatility during the COVID-19 pandemic. However, the series reverted to its long-term average of around 7% in the following years. This pattern is not unique to New Zealand. Both the Australian and UK markets showed significant upticks in the number of abnormal price movements in 2020, mostly due to the unusually high level of market volatility (ASIC, 2024; FCA, 2021).

Figure 2: Price-based Market Cleanliness Statistic

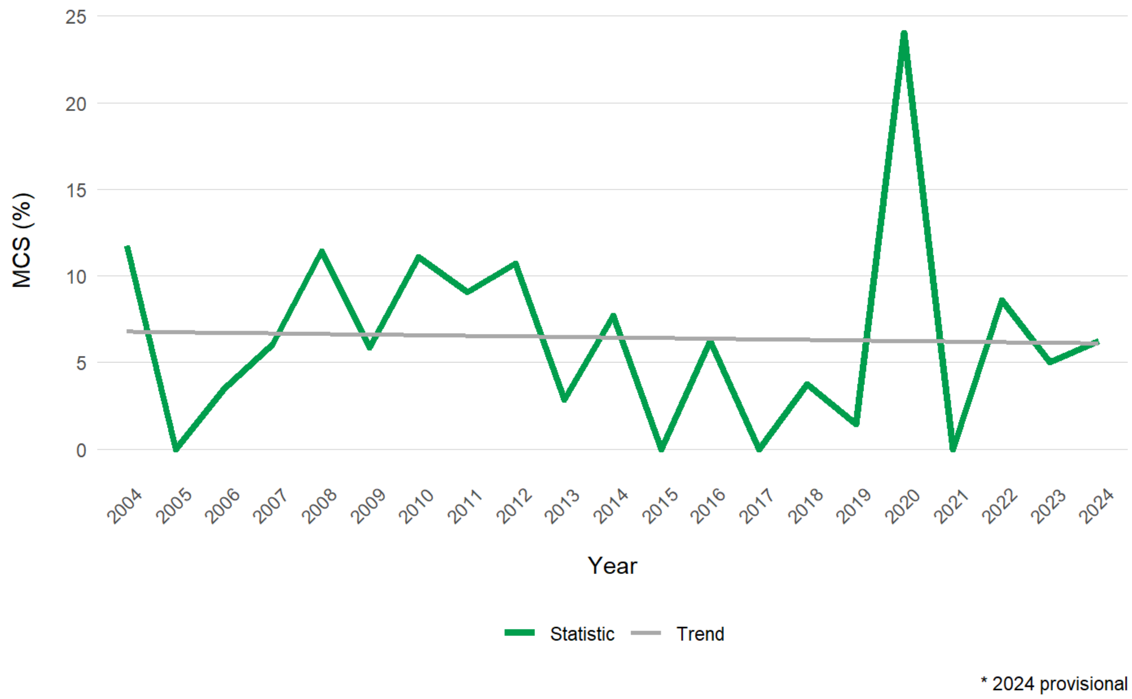


Table 1: Market cleanliness statistic from 2004-2024

Year	MPSA	MPSA with APPM	MCS (%)
2004	17	2	11.76
2005	25	0	0
2006	29	1	3.45
2007	33	2	6.06
2008	35	4	11.43
2009	17	1	5.88
2010	18	2	11.11
2011	33	3	9.09
2012	28	3	10.71
2013	35	1	2.86
2014	26	2	7.69
2015	48	0	0
2016	48	3	6.25
2017	33	0	0
2018	53	2	3.77
2019	69	1	1.45
2020	79	19	24.05
2021	39	0	0
2022	58	5	8.62
2023	60	3	5
2024	16	1	6.25

Source: Authors' calculations.

If we assume 2020 is an outlier and exclude it from the series, we can observe a clearer decline in the MCS statistic, indicating an overall improvement in market cleanliness over the past 20 years (Figure 3). The long-term average reduced from the elevated levels of 8% in the early 2000s to around 4% recently. There were two periods of high abnormality in stock prices: around 2008-2010 during the Global Financial Crisis, and in 2022 while inflationary pressures were rising in most developed economies and there were risks of a deep and prolonged recession.

Figure 3: Market cleanliness statistic (excluding 2020)

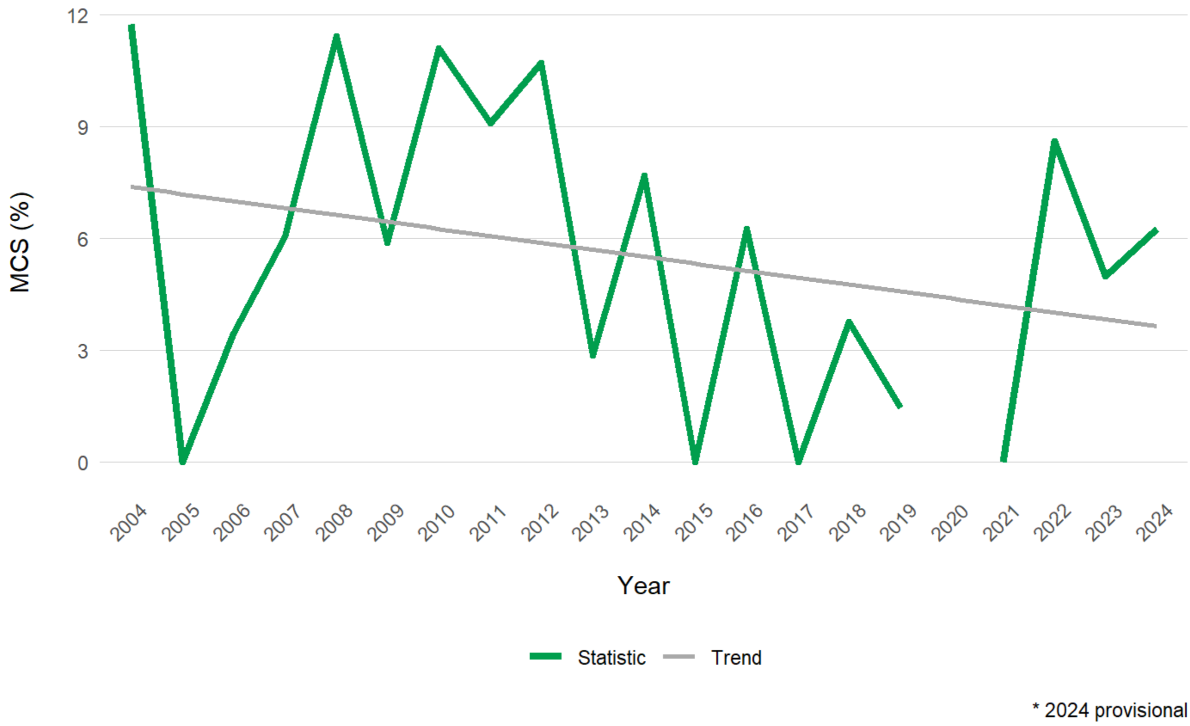
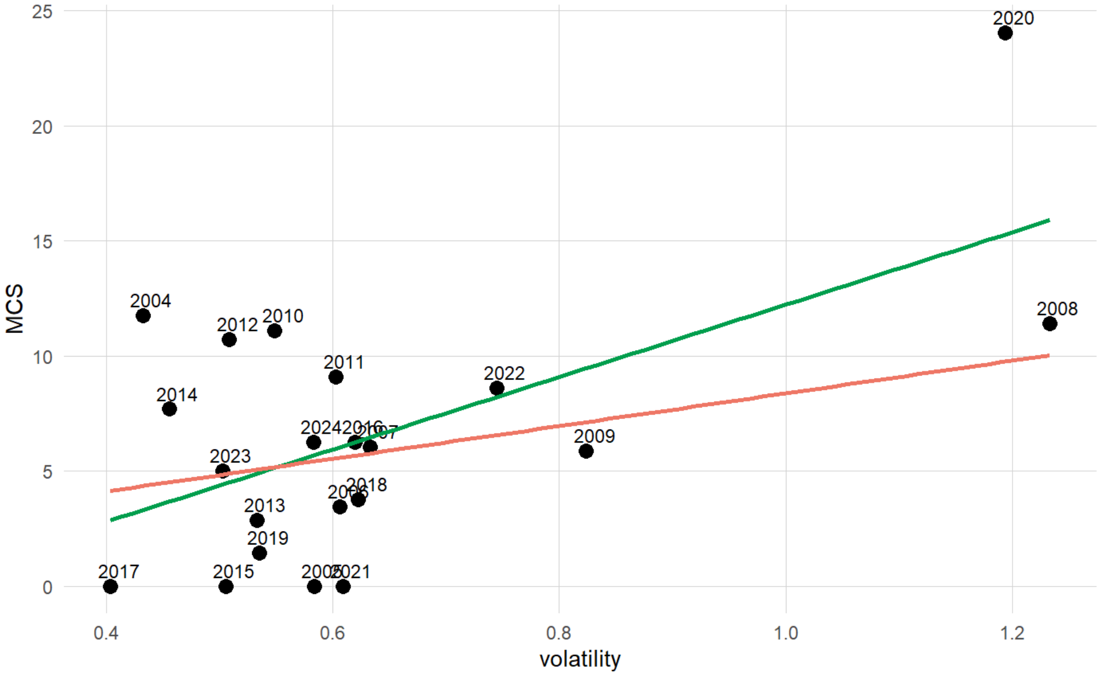


Figure 4 (green line) illustrates the potential correlation between market cleanliness and market volatility, which is measured by the standard deviation of daily stock returns during the same year. It is important to note the role of the 2020 outlier in terms of the prevalence of price abnormality in driving this correlation, as shown by the flatter fitted line (red line) once it is taken out. Overall, the positive correlation between the prevalence of price abnormality and market volatility highlights a key limitation of the MCS. While pre-announcement abnormality can indicate market inefficiencies and potential insider trading, it can also reflect actual market conditions during times of stress (ASIC, 2024).

Figure 4: Market cleanliness and stock market volatility



4.5 Factors associated with abnormality in stock return

In addition to the overall behaviour of measured market cleanliness over time, we are also interested in factors that may help to explain the probability of pre-announcement abnormal stock return (APPM), which forms the basis of the MCS. In this section we use multivariate regressions to examine the observable factors which are associated with the probability of an APPM. The analysis identifies a range of characteristics which are associated with APPMs, while controlling for observable changes in the market environment and the composition of firms over time. This is similar to the approach of the FCA (Goldman et al., 2014) and ASIC (ASIC, 2016).[6]

We estimate a logistic regression equation is as follows:

$$\log\left(\frac{P_{APPM}}{1-P_{APPM}}\right) = f(Market\ factors, Firm\ factors, Announcement\ factors, Time\ factors) \qquad (8)$$

The dependent variable is the log odds ratio of a MPSA being preceded by an APPM. An odds ratio of one indicates that there is a 50-50 chance that an APPM occurs before an announcement. If the odds ratio is greater than one, then a pre-announcement APPM is more likely to happen and vice versa. The explanatory variables include:

[6] Appendix C presents descriptive results from a univariate analysis in which we separately compare the probability of an APPM over different time periods and for groups with different firm characteristics. While the univariate results are simpler, they can be susceptible to small sample size and omitted variable bias and should be treated with caution.

- ***Market volatility*** is measured by the standard deviation of the S&P NZX All index in the 5 days preceding the announcement (the event window). We have chosen a shorter period than ASIC (2014) to reduce the autocorrelation of the variable. Market volatility can affect trading activity and create the possibility of price abnormality, as discussed above.
- ***Trading activity*** is measured by the percentage difference in the average daily volume traded between the 10 days before an announcement and the 20 days before that. This is similar to the concept of the Abnormal Trading Volume Ratio which will be discussed in Section 6. APPM can come with increasing trading volume due to insider trading activity or people trading on speculations.
- ***Industry***. Dummy variables for 10 industries based on the S&P/NZX grouping. The reference group is Consumer Discretionary. This can capture variation in trading behaviour and levels of market cleanliness across industries.
- ***Announcement type*** is a dummy variable which is 1 if the announcement is positive (positive 4-day CAR) and 0 otherwise. This can capture the difference in the probability of APPM between positive and negative announcements.
- ***Market cap***. Dummy variables for 5 quintile groups by the latest market capitalisation. The reference group is Quintile group 1 (smallest). Larger firms may have more resources devoted to disclosure and confidential information management. Stocks of larger firms are usually more liquid, which means insider trading may not be enough to move the price.
- ***Time***. Dummy variables for the periods before and after a regulatory event, plus a dummy for 2020, which is an outlier. The three events of interest are:
    - the establishment of the Financial Markets Authority as a single regulator for insider trading cases (2011);
    - the Financial Markets Conduct Act (FMC Act) was passed into law, which reinforced the regulatory framework relating to disclosure and reporting requirements (2013); and
    - the first prosecution of an insider trading case in New Zealand, relating to the trading in shares of Eroad (2017).

  In each case, the reference period is the period before the event. Similar to Goldman et al. (2014), we include the event year in the before period to account for any adjustment lag in the market. A negative, statistically significant coefficient of the dummy variable for the period after would indicate that the market was cleaner (exhibited fewer APPMs) on average in the years after the event, assuming 2020 is a one-off outlier.

Table 2 presents the results of the logistic regression and Table 3 includes the post-estimation test statistics for the joint significance of the industry, market cap and time dummies. Table 2 reports the odds ratios rather than the raw log odds coefficients as these can be interpreted more easily. For one unit increase in the independent variable, the percentage increase in the odds of an APPM is calculated as:

$$\%\ increase\ in\ the\ odds\ of an\ APPM\ before\ a\ MPSA = (Odd\ ratios\ coefficient - 1) * 100 \qquad (9)$$

Overall, we found that market volatility is positively correlated with the likelihood that an MPSA is preceded by an APPM (less clean). This relationship would make it difficult to use MCS for detecting short-term changes in the level of insider trading and the overall efficiency of the market. Other things being equal, a one unit increase in the standard deviation of the 5-day return increases the odds of an APPM by 56%. Trading volume is positively correlated with pre-announcement price abnormality, but the marginal effect is not statistically significant.

Positive announcements exhibit a higher level of cleanliness than negative ones. Controlling for other factors, a positive announcement is 42% less likely to be preceded by an APPM. This finding is not consistent with what has been found in Australia (ASIC, 2024). According to Dubow and Monteiro (2006) and ASIC (2024), negative material announcements (negative 4-day announcement CAR) can exhibit a lower level of pre-announcement abnormality, or higher level of cleanliness if strict regulations and reporting obligations make it harder to short sell. The difference could, at least to some extent, be attributed to the absence of a publicly available register of short selling positions in New Zealand, like those that exist in Australia (ASIC, 2018) and the UK (FCA, 2024). Even though there are similar requirements for short sellers to disclose their positions to the regulator in all three countries, the register is only made public in Australia (for all positions) and the UK (for large positions that reach 0.5% of the issued share capital of the company). The higher level of cleanliness for positive announcements in New Zealand may also suggest that surveillance and enforcement activities associated with these may have a stronger deterrent effect on insider trading and other types of market misconduct (though it is not clear why the same should not be true in other markets).

Market capitalisation does not have a material impact on the log odds of an APPM. While point estimates are consistent with larger firms experiencing fewer APPMS, none of the coefficients are significant and we fail to reject the null hypothesis of joint significance of all five market cap dummy variables (Table 3). This is consistent with the UK market (Goldman et al., 2014) but not the Australian market (ASIC, 2016), in which announcements by larger firms were found to be cleaner. In terms of industry, Table 3 indicates that the set of industry dummies is only weakly significant in explaining the variation in the dependent variable. Among all industries, Energy and Financials have higher odds of abnormal stock returns prior to announcements.

According to Table 3, including the time dummies does increase the model fit. The likelihood of APPM was significantly higher in 2020, as shown in table 2. An announcement in 2020 is 4 to 5 times more likely to be preceded by an APPM than other years.

Treating 2020 as an outlier, columns 1 to 3 split the observation sample into three bilateral periods, based on the timing of three relevant regulatory events: the establishment of the FMA (2011), the passing of the FMC Act (2013), and the first prosecution of insider trading (2017). Other things being equal, a material announcement in the years from 2014 to 2024 was 51% less likely to be preceded by an APPM compared to the years from 2004 to 2013. While we do not explicitly test for the number or timing of shifts in the series, the smaller and statistically insignificant differences observed for the periods before and after the 2011 and 2017 events are consistent with the 2013 Act (or other changes which took place around this time) being the main driver of the improvement in market cleanliness over this period.

It is also important to note that these interpretations are limited by the number of covariates included in the regression. The odds of having an APPM can also be impacted by other unobservable factors that we cannot control for in our model. While the pseudo Tjur R-squared should be used with caution, the low value (0.116-0.118 across three models) indicates that observable factors explain only a small proportion of the variation in the dependent variable.

Table 2: Logistic regression results - APPM

Predictors	Model 1 *Odds Ratio*	Model 2 *Odds Ratio*	Model 3 *Odds Ratio*
Market volatility	1.55 **	1.56 **	1.58 **
	(0.33)	(0.33)	(0.34)
Trading activity	1.24	1.24	1.21
	(0.26)	(0.26)	(0.25)
Positive announcements	0.59 *	0.59 *	0.58 *
	(0.18)	(0.18)	(0.18)
Industry (Reference group: Consumer Discretionary)			
Consumer Staples	0.75	0.80	0.74
	(0.56)	(0.60)	(0.56)
Energy	8.03 ***	8.74 ***	7.71 ***
	(6.06)	(6.63)	(5.81)
Financials	2.74 *	2.82 *	2.85 *
	(1.46)	(1.51)	(1.53)
Health Care	2.55	2.61	2.48
	(1.62)	(1.66)	(1.58)
Industrials	0.83	0.82	0.84
	(0.51)	(0.50)	(0.52)
Materials	1.72	1.74	1.76
	(1.53)	(1.55)	(1.56)
Real Estate	1.54	1.59	1.49
	(1.13)	(1.17)	(1.10)
Technology	2.12	2.22	2.03
	(1.25)	(1.31)	(1.20)
Telecommunications	0.58	0.56	0.53
	(0.64)	(0.62)	(0.58)
Utilities	0.78	0.82	0.76
	(0.69)	(0.73)	(0.67)
Market capitalisation (Reference group: Quintile group 1)			
Quintile group 2	1.27	1.24	1.30
	(1.07)	(1.04)	(1.09)
Quintile group 3	1.69	1.63	1.73
	(1.13)	(1.10)	(1.16)
Quintile group 4	1.29	1.23	1.33
	(0.86)	(0.83)	(0.89)
Quintile group 5	1.65	1.58	1.70
	(1.05)	(1.01)	(1.07)
Period (Reference group: Before event year)			
Year 2020	4.59 ***	5.14 ***	4.79 ***
	(2.10)	(2.42)	(2.41)
After 2011 (excluding 2020)	0.56		
	(0.21)		
After 2013 (excluding 2020)		0.49 **	
		(0.18)	
After 2017 (excluding 2020)			0.66
			(0.25)
Observations	781	781	781
R^2 Tjur	0.117	0.118	0.116

** p<0.1 ** p<0.05 *** p<0.01*

Table 3: Joint significance tests – Model 2

	Chi-square statistic	p-value
Industry dummies	**17.30**	**0.07**
Market cap dummies	1.02	0.91
Time dummies	**12.03**	**0.00**

5. Abnormal Trading Volume Ratio

Due to the limitations of the price-based market cleanliness statistic as described in section 4.3, it is important to look at a suite of indicators when evaluating the cleanliness of the market. The Abnormal Trading Volume Ratio (ATVR), first developed by the FCA (2019), is a potential candidate. While the MCS focuses on price movements, ATVR focuses on the abnormality in the volume traded. In an unclean market with potential insider trading, we can expect a significant increase in trading volume directly before a price-sensitive announcement, compared to the usual levels, as illustrated in Figure 5.

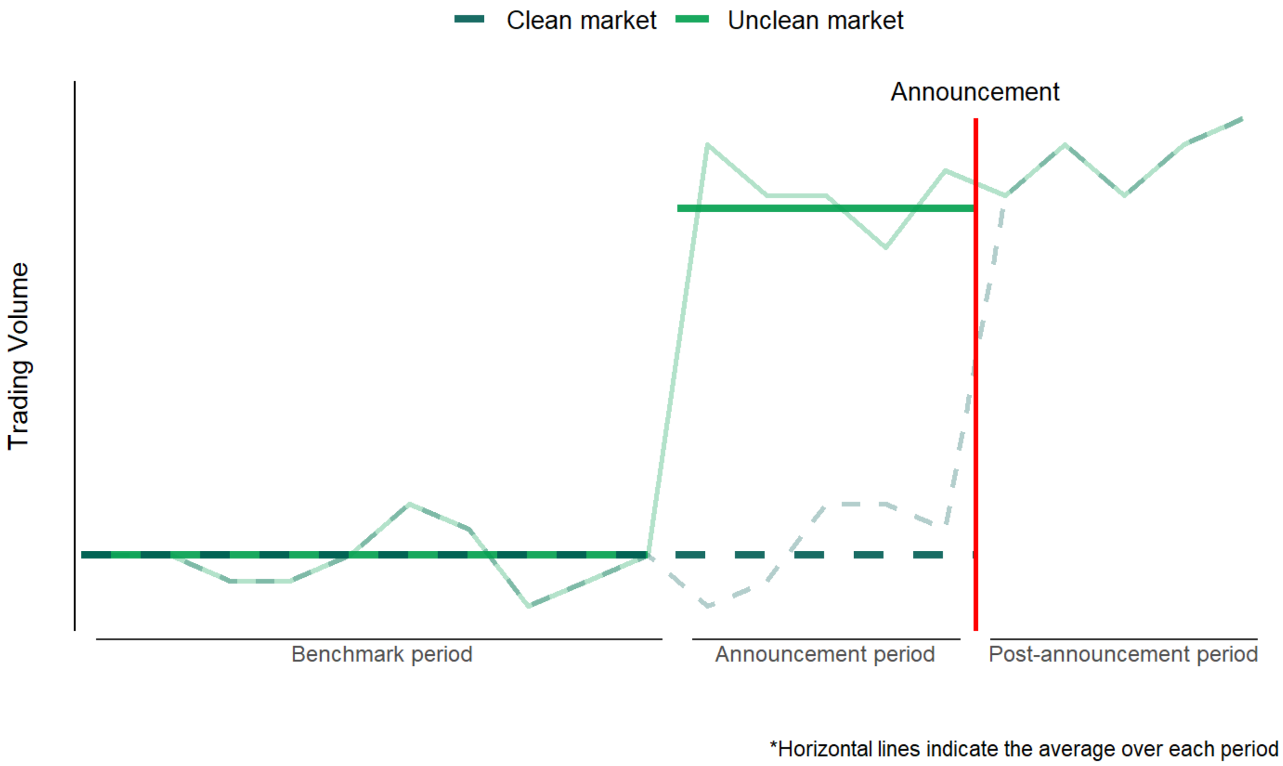
5.1 Methodology

To create the ATVR measure, we look at 30 trading days before the announcement, which is called the observation period. This period is split into two. We compare the volume traded during the 10 days directly preceding an announcement, called the announcement period, versus the 20 days directly preceding the announcement period, called the benchmark period. The statistical test used for these comparisons is the Welch t-test. The null hypothesis is that the average volume traded during the announcement period is not larger than the average volume traded during the benchmark period.

For each announcement, if we can reject the null hypothesis, that announcement is considered to have an ATV. Once all ATVs have been identified, the ATVR can be calculated as:

$$ATVR = \frac{Number\ of\ announcements\ with\ abnormal\ trading\ volume}{Number\ of\ price - sensitive\ announcements} \tag{10}$$

Figure 5: Abnormal Trading Volume Ratio (ATVR) – concept



5.2 Potential issues and mitigations

For the headline ATVR measure, we follow the original methodology (FCA, 2019) in terms of parameter choices and data cleaning process.

The length of the observation period is 30 days, which means an announcement is excluded from the sample if the stock were not traded for at least 30 days before. This is chosen to maximise the number of announcements included while ensuring there are enough observations for the statistical tests.

For a given announcement, the analysis may be less reliable if there are other price-sensitive announcements in either period as the trading volume would already be turbulent. If that is the case, that announcement is removed from the sample in favour of the earlier announcement.

The length of the announcement and benchmark periods are 10 and 20 days, respectively. There is no economic theory underlying these choices. Therefore, we test the robustness of our headline measure against different options for the announcement and benchmark period. See Appendix B for a discussion of the robustness test results.

Although the ATVR looks at potential market misconduct from a different angle to the MCS, it suffers from similar limitations. There are many factors other than insider trading, e.g speculation and rumours, which can lead to increased trading volume before the announcement. Volume traded is also more sensitive to market volatility than stock price. The criteria to determine abnormality in trading volume is also not as strict, which increases the sample size but at the same time limits the robustness of the measure. For these reasons, the ATVR is not as commonly used as the MCS, by other regulators and especially in academic research. Therefore, while the ATVR is a valuable supplement to the MCS it should not be regarded as a standalone measure of market cleanliness.

5.3 Results

Over the past 20 years, the ATVR measure also declined somewhat, consistent with an improvement in market cleanliness (Figure 6). The probability that an announcement is preceded by an ATV has dropped from around 7% to 6%. However, the series is also highly volatile. Therefore, for the ATVR, it is also important to focus on long-term changes rather than year-on-year movements.

Figure 6: Volume-based Abnormal Trading Volume Ratio (ATVR)

Table 4: ATVR over time

Year	Total	ATV	ATVR (%)
2004	27	2	7.41
2005	71	5	7.04
2006	68	4	5.88
2007	99	9	9.09
2008	87	7	8.05
2009	86	6	6.98
2010	81	2	2.47
2011	78	6	7.69
2012	85	8	9.41
2013	95	7	7.37
2014	105	8	7.62
2015	122	6	4.92
2016	138	7	5.07
2017	141	5	3.55
2018	186	9	4.84
2019	258	19	7.36
2020	340	26	7.65
2021	375	25	6.67
2022	385	16	4.16
2023	398	15	3.77
2024	107	10	9.35

5.4 Factors associated with abnormality in trading volume

In addition to the overall decline, we are also interested in factors that may be correlated with the probability of a pre-announcement ATVR, which forms the basis of the ATVR. While we utilise the multivariate approach in this section, we also test our results using the univariate approach for robustness. The results and discussions of that can be found in Appendix C.

Similar to Section 4.5, we estimate a logistic regression with the log odds ratio of an announcement being preceded by an ATV as the dependent variable:

$$\log\left(\frac{P_{ATV}}{1-P_{ATV}}\right) = f(Market\ factors, Firm\ factors, Time\ factors) \tag{11}$$

The explanatory variables include *Market volatility*, *Industry dummies*, and *Time dummies.* All of them are created similarly to Section 4.5, but there are several important differences. First, for the time dummies, we no longer separate the year 2020 as it does not appear to be an outlier for the ATVR measure. Second, the *Trading activity* is no longer part of the set of explanatory variables as it is now the dependent variable. Finally, the dummy for positive announcement is also excluded. There are several reasons for that. Firstly, it is not straightforward to distinguish between positive and negative announcements based only on trading volume. Both good and bad news can lead to heightened trading activities, as people either increase their interest in the equity or look to exit their positions. Secondly, while it is possible to use the cumulative abnormal return, similar to the MCS to identify positive announcements, doing so requires the company's equity to be traded for a year before the announcement, which leads to a reduction in sample size.

Table 5 presents the results. Table 6 includes the statistics for the joint significance test for the industry, market cap and time dummies using model 2.

Overall, market volatility appears to be a strong predictor of the likelihood that an announcement is preceded by an ATV. This is consistent with the results using the likelihood of APPM in Section 4.

Industry dummies are also a predictor of volume abnormality, explaining a significant amount of the overall variation in the ATV statistic (Table 6). The Financials industry has the highest level of market cleanliness according to the estimated coefficients (Table 5). Announcements made by Financials firms are 53% less likely to be preceded by an ATV.

The dummy variables for firm sizes are also jointly significant in predicting the dependent variable. Overall, the odds of having an ATV appear to increase with market capitalisation, though the point estimates are not monotonic. Announcements made by firms within the second largest group in terms of market cap (quintile group 4) are twice as likely to be preceded by an ATV compared to those in the smallest group. Quintile group 5 also appears to have the same increase in odds. However, the coefficient is not statistically significant and definitive conclusion should not be drawn from this.

While odds ratios on the post-event period dummies associated with regulatory events of interest follow the same patterns as those for the APPM model, none of these are statistically significant. This is true for both the individual dummy variables in Table 5 and jointly, as shown in Table 6, in which we fail to reject the null hypothesis of joint insignificance. Therefore, we should be cautious when quantifying the impact of the regulatory changes on market cleanliness.

Table 5: Logistic regression results - ATV

Predictors	**Model 1** *Odds Ratio*	**Model 2** *Odds Ratio*	**Model 3** *Odds Ratio*
Market volatility	**1.66** ***	**1.66** ***	**1.68** ***
	(0.29)	**(0.29)**	**(0.29)**
Industry (Reference group: Consumer Discretionary)			
Consumer Staples	1.40	1.44	1.38
	(0.44)	(0.46)	(0.43)
Energy	0.63	0.62	0.65
	(0.40)	(0.39)	(0.41)
Financials	**0.45** ***	**0.46** ***	**0.44** ***
	(0.13)	**(0.13)**	**(0.13)**
Health Care	0.74	0.75	0.73
	(0.26)	(0.27)	(0.26)
Industrials	0.74	0.74	0.74
	(0.24)	(0.24)	(0.24)
Materials	0.99	0.99	0.98
	(0.51)	(0.51)	(0.51)
Real Estate	0.75	0.75	0.74
	(0.24)	(0.24)	(0.23)
Technology	0.84	0.86	0.83
	(0.35)	(0.35)	(0.34)
Telecommunications	1.48	1.49	1.46
	(0.56)	(0.57)	(0.55)
Utilities	0.82	0.83	0.81
	(0.32)	(0.32)	(0.31)
Market capitalisation (Reference group: Quintile group 1)			
Quintile group 2	1.34	1.33	1.34
	(0.66)	(0.66)	(0.66)
Quintile group 3	1.26	1.23	1.29
	(0.60)	(0.58)	(0.61)
Quintile group 4	**2.13** *	**2.10**	**2.15** *
	(0.97)	**(0.96)**	**(0.99)**
Quintile group 5	1.80	1.75	1.84
	(0.81)	(0.79)	(0.83)
Period (Reference group: Before event year)			
After 2011	0.91		
	(0.17)		
After 2013		0.83	
		(0.14)	
After 2017			1.01
			(0.17)
Observations	3314	3314	3314
R^2 Tjur	0.013	0.013	0.013

** p<0.1 ** p<0.05 *** p<0.01*

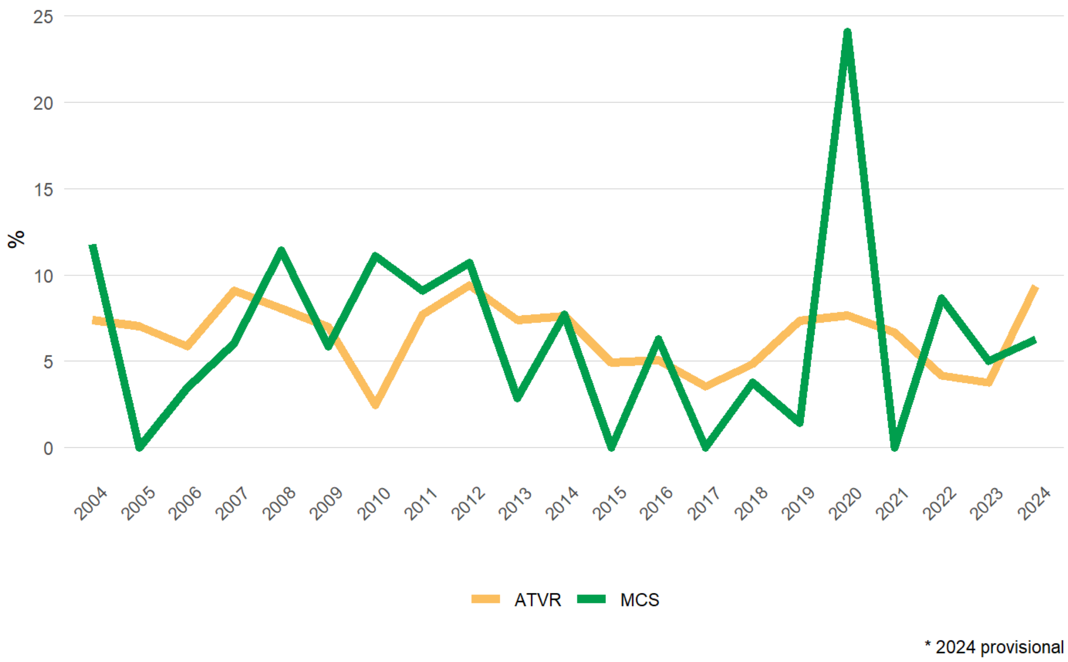
Table 6: Joint significance tests – Model 2

	Chi-square statistic	p-value
Industry dummies	**24.42**	**0.01**
Market cap dummies	**9.11**	**0.06**
Time dummies	0.51	0.47

6. Comparison of the MCS and ATVR

Both the Market Cleanliness Statistic (MCS) and the Abnormal Trading Volume Ratio (ATVR) use abnormality in trading behaviour to understand the prevalence of insider trading. However, we do not expect them to be perfectly correlated. Indeed, over the 2004-2024 period, the pairwise correlation between the two series is only 0.53. Figure 7 illustrates this relationship in which the long-term trend is similar, but the year-on-year changes are much more varied. In the Section 4 logistic regression, the *Trading activity* variable, which is the key input for the ATVR, is positively correlated with the odds of an APPM occurring. However, the coefficient is not statistically significant at any of the conventional levels. There are several fundamental differences between the two measures which lead to this imperfect correlation.

Figure 7: Comparison of the MCS and ATVR



Firstly, the announcements which are included in the analysis of the MCS and ATVR (the denominator in equation 7 in section 4.1 and equation 10 in section 5.1) are different. The MCS focuses only on material announcements, which is determined by the significant movements in stock prices before, during and after the announcement. On the other hand, for the ATVR, only announcements which do not have any other announcement during the preceding 30 days (the observation period) are included in the sample. The two methods also have different liquidity requirements. While the MCS method removes illiquid stocks, which are not traded in at least half of the preceding trading year, the ATVR method requires stocks to be traded in all of the preceding 30 trading days. As a result, in the study period, there are only 316 announcements included in the sample for both measures, out of the 781 announcements in the sample for MCS and 3314 in the sample for ATVR.

Secondly, when insider trading or other market misconduct happens, both the price and the trading volume are expected to exhibit abnormal movements. However, this is not always the case. Table 7 presents the joint probability of APPM (which forms the MCS) and ATV (which forms the ATVR) for the 316 announcements included in both methods. Out of the 32 announcements with at least some abnormalities, only 9.3% (3 announcements) exhibit both price and volume abnormality. Stock prices can display unusual movements without a significant change in the trading volume, especially for thinly traded stocks. On the other hand, in many cases, a surge in trading volume before the announcement may not move the stock price enough to qualify as abnormal in a statistical sense, especially if these reflect speculative trading in the absence of additional information about the outcome of an upcoming announcement. This is more likely to happen for stocks that are historically more volatile.

Table 7: Joint probablity of APPM and ATV

APPM	*ATV* 0	*ATV* 1	***Total***
0	284 89.9 %	21 6.6 %	305 96.5 %
1	8 2.5 %	3 0.9 %	11 3.4 %
Total	292 92.4 %	24 7.6 %	316 100 %

While both the MCS and ATVR can provide some indications of insider trading and other market misconduct, they approach the issue from two related but different angles. The imperfect correlation between the two measures, as shown in this section, suggests it is important to combine them when assessing the cleanliness of the New Zealand equity market. A downward trend in both measures can signal an overall improvement, and an upward trend in both an overall deterioration. If they move in two opposite directions, it is recommended to take a closer look at the underlying drivers – for example, the composition of listed stocks or announcements – before drawing conclusions.

7. Conclusion

This report examines the cleanliness of New Zealand's equity market since 2004 using two measures. The Market Cleanliness Statistic (MCS) looks at the abnormality in *stock returns* prior to the material announcements, given the previous relative performance against the market. The Abnormal Trading Volume Ratio (ATVR) assesses whether there are differences in the *trading volumes* prior to the announcement compared to the usual level.

We found that while both measures are volatile year-on-year, there has been a clear decrease in the abnormality of returns and trading volumes in New Zealand over the past 20 years. This indicates an overall improvement in market cleanliness. Among other regulatory events which took place over the period, the passing of the FMC Act into law in 2013 appears to be most strongly correlated with a downward shift in the statistic. This legislation was soon followed by the first prosecution of insider trading in 2017, though this latter event did not have a clear impact on overall cleanliness. A combination of strict regulation and effective enforcement is key to stopping insider trading and other misconduct.

Market volatility is a strong predictor of cleanliness, i.e. highly volatile periods tended to coincide with more abnormal pre-announcement price movements (APPM) and abnormal trading volumes (ATV) in our data. Positive announcements were also cleaner than negative ones, potentially as New Zealand does not have a publicly available register of short selling positions, like the UK or Australia, or because surveillance and enforcement activities associated with positive announcements may have a stronger deterrent effect on insider trading. At sub-market level, there is some variation in the levels of cleanliness across industries and market capitalisations. However, the differences are not highly significant and are not consistent across the two measures, implying that these differences should be treated with caution.

The MCS and ATVR series appear to be positively correlated with each other, but the relationship is not statistically significant due to the fundamental differences in the approach, methodology design and sample size. Therefore, it is important to look at both indicators when analysing the market and sub-markets.

While the two measures examined in this report can provide some useful indications, especially of the long-term trend, they also have limitations. Some of these can be mitigated while others are the byproduct of the statistical approaches. For example, it is not possible to attribute all the changes in MCS and ATVR to the prevalence of insider trading. As shown in the analysis, these two indicators can move because of genuine market conditions during times of stress. At the same time, as they are both top-down stock-level measures, and rely on the power of statistical tests, it is possible some actual insider trading activities were not material enough to move the series.

In the future, we will consider updating the two measures on a regular basis, along the lines of both the FCA and ASIC. Doing so would provide market participants with a high-level indication of developments in the integrity of New Zealand's equity market and the extent to which potential market misconducts are deterred effectively.

Academic peer review report

Report of "Market Cleanliness in the New Zealand Equity Market"

Hai Lin
School of Economics and Finance, Victoria University of Wellington

This paper discusses the use of two measures to evaluate the cleanliness of the New Zealand stock market. One is based on the abnormality in stock return—the Market Cleanliness Statistic (MCS)—and the other is based on the abnormality in trading volume—the Abnormal Trading Volume Ratio (ATVR). The study finds an overall improvement in the cleanliness of the New Zealand stock market in the last twenty years.

I agree with several points discussed in this paper. First, market cleanliness is critical to providing a fair, efficient, and transparent financial market environment and should be one key focus of financial market regulators. Second, the paper uses stock price response to determine whether an announcement is materially price-sensitive (MPSA). This price reaction approach has also been widely used in academics. Third, there are limits to these measures in that they are based on what is observed, which might arise from other reasons instead of inside trading. The existence of abnormality in trading is not equivalent to inside trading. It provides a case to further look into the corresponding more comprehensive tick-by-tick data. Fourth, return and trading volume should be monitored to give a complete picture.

Several extensions can be made. First, one major issue about the MSC is its small sample size. This can be improved by relaxing the requirement of MPSA. The paper uses the 1st and 99th percentile as the criteria. This criterion can be reduced to 5% and 95%, or 10% and 90%. Reducing the criteria of MPSA will increase the sample size to calculate the MSC, therefore providing more robust results. Second, use the same announcement events to calculate MSC and AVTR. This paper uses different announcement events to calculate the MSC and AVTR, which makes their results incomparable. Alternatively, if the same announcement events such as the MSPA are used, it will provide a coherent and integrated analysis of abnormal trading in the New Zealand stock market. For example, if one MPSA has an abnormal pre-announcement price movement (APPM) and trading volume (ATR), it will make a stronger case for further investigation. Third, other measures can be constructed using high-frequency transaction data. For example, they can be used to calculate order imbalance and other information asymmetry measures. These measures provide evidence of abnormal trading from more comprehensive transaction-level data and supplement the above two measures. Fourth, further explore why positive announcements exhibit a higher level of cleanliness. This finding is inconsistent with the results of other countries and academic literature, which show that positive announcements have a more substantial pre-announcement effect than negative ones due to short sale constraints.

Glossary

Term	Definition
APPM	Abnormal Pre-announcement Price Movement
ASIC	The Australian Securities and Investments Commission, Australia's national corporate regulator.
ATV	Abnormal Trading Volume
ATVR	Abnormal Trading Volume Ratio, a quantification of the market cleanliness concept.
FCA	The Financial Conduct Authority, the UK's financial regulator.
FMA	Financial Markets Authority
Insider trading	Trading in the equity market based on confidential knowledge.
Market cleanliness	The absence of abnormal market behaviour before a material announcement.
MCS	Market Cleanliness Statistic, a quantification of the market cleanliness concept.
Market manipulation	Artificially influencing stock prices to mislead investors.
Market volatility	Periods of unpredictable price movements.
MPSA	Material Price-Sensitive Announcement
NZX	New Zealand Stock Exchange

References

ASIC. (2016). *Review of Australian equity market cleanliness.* Australian Securities and Investments Commission. Retrieved from https://download.asic.gov.au/media/4423818/rep487-published-25-august-2017.pdf

ASIC. (2018). *Short selling Regulatory guide 196*. Retrieved from Australian Securities and Investments Commission: https://asic.gov.au/regulatory-resources/find-a-document/regulatory-guides/rg-196-short-selling/

ASIC. (2019). *Review of Australian Equity Market Cleanliness.* Australian Securities and Investments Commission. Retrieved from https://download.asic.gov.au/media/5218490/rep623-published-31-july-2019.pdf

ASIC. (2023). *ASIC's insider trading detection project wins data award*. Retrieved from Australian Securities and Investments Commission: https://asic.gov.au/about-asic/news-centre/news-items/asic-s-insider-trading-detection-project-wins-data-award/

ASIC. (2024). *Review of Australian equity market cleanliness.* Australian Securities and Investments Commission. Retrieved from https://download.asic.gov.au/media/0jimrh1r/rep787-published-24-july-2024.pdf

Benninga, S. (2014). *Financial Modeling Fourth Edition.* Cambridge: Massachusetts Institute of Technology.

Bhattacharya, U., & Daouk, H. (2002). The World Price of Insider Trading. *Journal of Finance*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=249708

Dubow, B., & Monteiro, N. B. (2006). *Measuring Market Cleanliness.* London: Financial Services Authority. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1019999

Eglite, E., Staermans, D., Patel, V., & Putniņs, T. J. (2023). *Using ETFs to conceal insider trading*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4343579

FCA. (2016). *Short selling*. Retrieved from Financial Conduct Authority: https://www.fca.org.uk/markets/short-selling

FCA. (2019). *Abnormal Trading Volume Ratio*. Retrieved from Financial Conduct Authority: https://www.fca.org.uk/data/abnormal-trading-volume-ratio

FCA. (2020). *Abnormal Trading Volume Ratio*. Retrieved from Financial Conduct Authority: https://www.fca.org.uk/data/abnormal-trading-volume-ratio

FCA. (2023). *Market abuse*. Retrieved from Financial Conduct Authority: https://www.fca.org.uk/markets/market-abuse

FMA. (2017). *FMA files charges for insider trading*. Retrieved from Financial Markets Authority: https://www.fma.govt.nz/news/all-releases/media-releases/fma-files-charges-for-insider-trading/

FMA. (2019). *Former Promisia chairman admits insider trading breaches*. Retrieved from Financial Markets Authority: https://www.fma.govt.nz/news/all-releases/media-releases/former-promisia-chairman/

FMA. (2023). *Insider trading case involving Pushpay shares*. Retrieved from Financial Markets Authority: https://www.fma.govt.nz/about-us/enforcement/cases/insider-trading/

FMA. (2024). *Insider trading case involving Heartland Group Holding Limited shares*. Retrieved from Financial Markets Authority: https://www.fma.govt.nz/about-us/enforcement/cases/insider-trading-case-involving-heartland-group-shares/

Goldman, J., Hunt, S., Minter, P., Suntheim, F., & Zaman, Q. (2014). *Why has the FCAs market cleanliness statistic for takeover announcements decreased since 2009?* London: Financial Conduct Authority. Retrieved from https://www.fca.org.uk/publication/occasional-papers/op-4-market-cleanliness.pdf

Hensen, A. (2018). *New Zealand Equity Market Cleanliness.* Financial Markets Authority.

Johnson, S. (2023). *Insider traders use ETFs to front-run M&A deals, academics say*. Retrieved from Financial Times: https://www.ft.com/content/9ee4e371-f965-4121-a395-cb8b94a5f9e6

MacKinlay, A. C. (1995). Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics, vol. 38, issue 1*, 3-28.

Major, B. (1994). From Social Inequality to Personal Entitlement: the Role of Social Comparisons, Legitimacy Appraisals, and Group Membership. *Advances in Experimental Psychology, 26*, 293-355. Retrieved from https://doi.org/10.1016/S0065-2601(08)60156-2

Monteiro, N., Zaman, Q., & Leitterstorf, S. (2007). *Updated Measurement of Market Cleanliness.* London: Financial Services Authority. Retrieved from https://www.researchgate.net/publication/228282968_Updated_Measurement_of_Market_Cleanliness

Nagata, T. T. (2017). Assessment of informed trading in the Brazilian equity market: Market cleanliness methodology. Retrieved from https://www.pp.u-tokyo.ac.jp/wp-content/uploads/2016/02/51158208.pdf

NZ RegCo. (n.d.). *About NZ RegCo*. Retrieved from NZ RegCo: https://www.nzx.com/regulation/nzregco/about-nzregco

NZX Limited. (2024). *NZX Participant Rules - March 2024*.

Scannell, K. (2016). *SEC investigates ETF use in insider trading*. Retrieved from Financial Times: https://www.ft.com/content/84621418-34a4-11e0-9ebc-00144feabdc0

Appendix A: How different financial regulators measure market cleanliness

Table A1: Market cleanliness methodology – most recent report by financial regulator

	Australia ASIC	**UK FCA**	**NZ FMA**
Year published	2024	2024	2018
Announcement types	All price-sensitive, material announcements	Takeover only	All price-sensitive, material announcements
Time period	2006 – present	2002 – present	2010 – 2016
Comparisons	Market cap Industry Announcement types Regulatory events	Market cap Industry Takeover characteristics Regulatory events	Australia Regulatory events
Methodology	Market Cleanliness Statistic (MCS): Stock-level, price-based Account-based Anomalous trading: Account-level, volume-based	Market Cleanliness Statistic (MCS): Stock-level, price-based Abnormal Trading Volume Ratio (ATVR): Stock level, volume-based Potential Anomalous Trading Ratio: Account-level, volume-based	Market Cleanliness Statistic (MCS): Stock-level, price-based

Appendix B: Robustness tests

Market Cleanliness Statistic

In this Section, we test the MCS against different modelling and parameter choices. Overall, the results are mostly robust, especially the long-term trend.

Expected return methodology

The regression-based market model is used to estimate expected and abnormal returns during the event window. While this is common for event studies (MacKinlay, 1995) and is consistent with the original market cleanliness study by the FCA (Dubow & Monteiro, 2006), there are other methods which can be also be used, each with its own merits and demerits.

First, the Capital Asset Pricing Model (CAPM) uses stock return $R_{i,t}$, market return $R_{m,t}$, and risk-free return $R_{f,t}$ to estimate the expected return $E(R_{i,t})$ (equation 12). The key component is β_i (equation 13), which is calculated by the covariance between stock i and market returns, divided by the variance of stock i returns. β_i measures the relative volatility between stock i and the market, which can drive its expected risks and returns. This analysis used secondary market 10-year NZ Government Bond yields as the risk-free rate. The CAPM model has an advantage of being closely related to economic theory on the risk premium.

$$E(R_{i,t}) = R_{f,t} + \beta_i(R_{m,t} - R_{f,t}) \tag{12}$$

$$\beta_i = \frac{COV_{i,m}}{VAR_m} \tag{13}$$

Another alternative is simply using the mean return of the estimation window $\overline{R_{i,t}}$ as the expected return. This mean-adjusted method was first introduced by Goldman et al. (2014) and is much simpler than the market model or CAPM.

$$E(R_{i,t}) = \overline{R_{i,t}} \tag{14}$$

FMA
FINANCIAL MARKETS AUTHORITY
TE MANA TĀTAI HOKOHOKO

Figure B1: Market cleanliness statistic – by expected return methodology

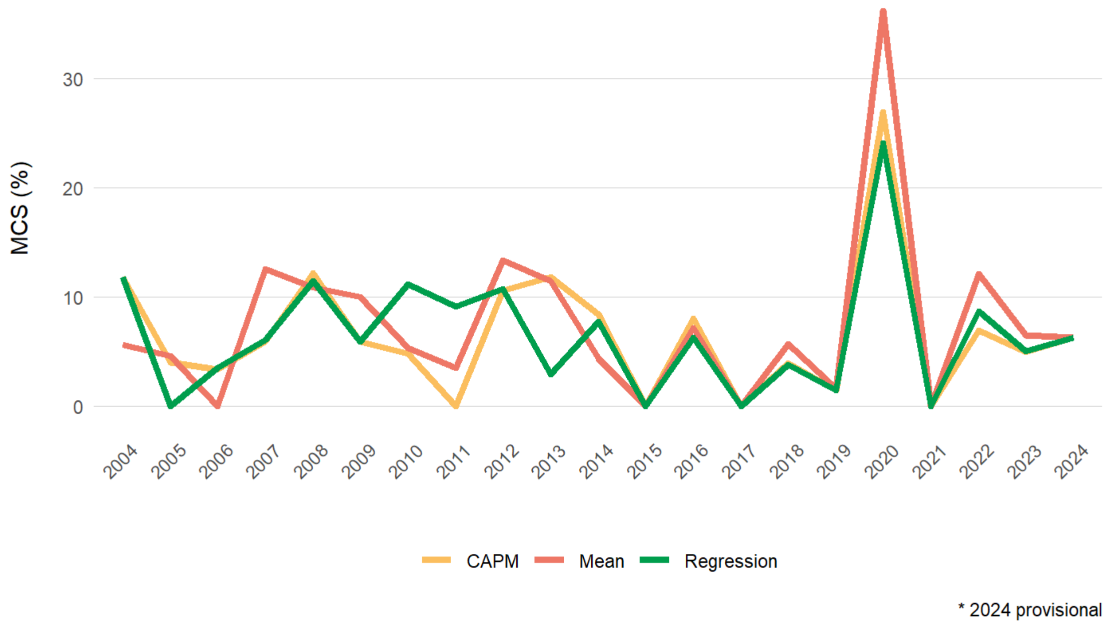


Figure B1 shows that the three methods yield qualitatively similar results over the past 20 years for the New Zealand equity market, despite some minor differences in the levels around 2010-2011. The choice of the expected return methodology becomes irrelevant from 2015 onwards as the three series closely resemble each other.

Liquidity requirements

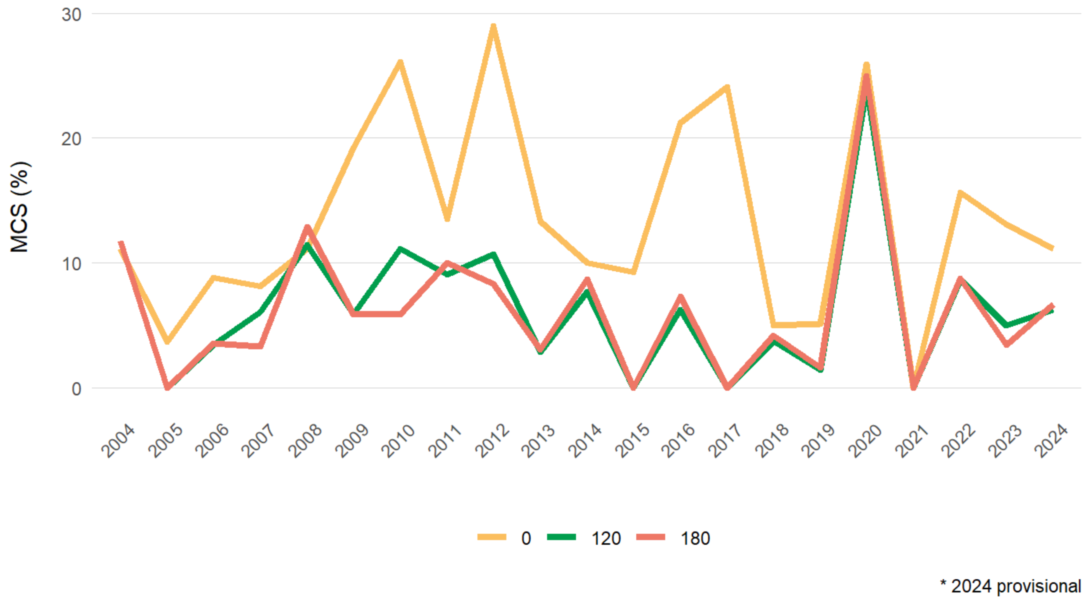
For a stock to be included in the headline measure, it needs to be traded for at least half of the estimation window (120 out of 240 days). This was chosen to strike a balance between including more stocks (which can be an issue with a small market like New Zealand with many thinly traded stocks) and having a sufficiently large sample size to apply the regression-based market model. This is also a common choice in the market model literature (Benninga, 2014).

In this section, we test the robustness of our measure against loosening this requirement (no minimum trading days) and tightening it (stocks needed to be traded for at least three-quarter of the estimation window, or 180 out of 240 days). Using the average liquidity of the New Zealand stock market in the 2023 calendar year, the former adds 8% to the sample while the latter loses around 7%.[7]

Tightening the liquidity requirement in a similar fashion to ASIC (ASIC, 2016; ASIC, 2024) does not materially change the headline MCS (Figure B2). Loosening it, however, significantly increases the level across the time series. This is because some stocks are only traded for a few days during the estimation window, which is not enough for the market-adjusted regression model to yield accurate estimation. This, in turn, leads to an overestimation of the number of abnormal pre-announcement price movements and skews the measure. These results emphasise the importance of the liquidity requirement in estimating market cleanliness statistic.

[7] In 2023, 85% of the stocks on the NZX main board (149 out of 175) were traded for at least three-quarters of the year, 7% were traded for between half and three-quarters, and 8% were traded for less than half of the year.

Figure B2: Market cleanliness statistic – by liquidity requirements

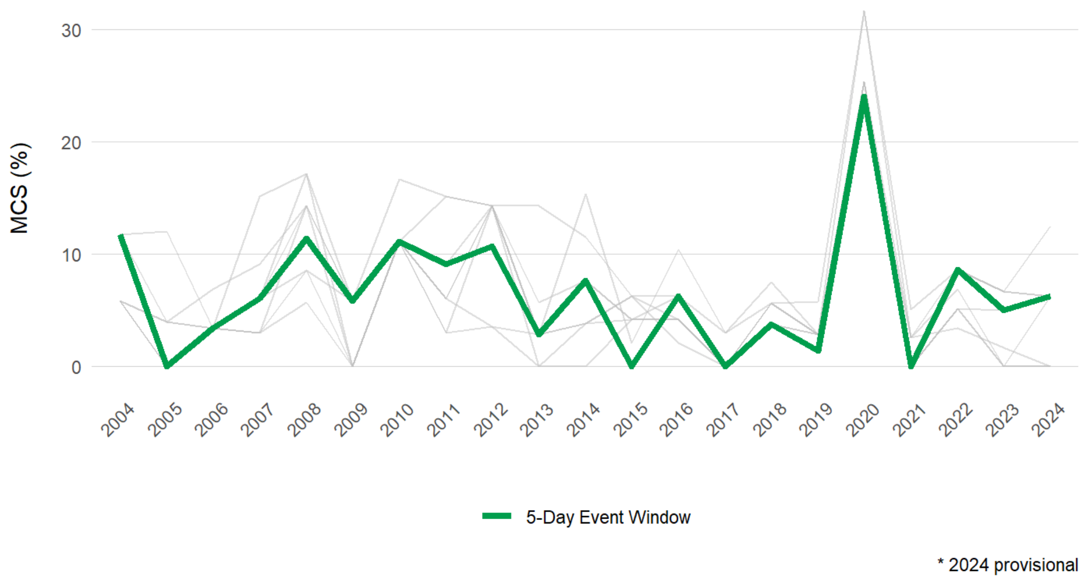


Event window

The headline measure uses a five-day event window before the announcement to determine the abnormality of stock returns, similar to ASIC (ASIC, 2016; ASIC, 2024). The length of the event window can materially impact the final measure. If a longer window is chosen, it would be more difficult to detect statistically significant returns. In contrast, a shorter window may not capture enough trading before the announcement, which can be an issue for a small, relatively illiquid market like New Zealand.

Figure B3 shows the robustness of our measure using an event window ranging from two to eight days. Varying the length of the event window impacts the levels but not the year-on-year changes.

Figure B3: Market cleanliness statistic – by event window



Materiality thresholds

The headline MCS measure uses the strict 1st and 99th percentile criteria to ensure that we can be certain about the materiality of the announcement and the abnormality of the pre-announcement price movement. However, doing so reduces the sample size significantly, which could be an issue for a small market like New Zealand. In this section, we test the robustness of the headline measure by loosening the significance levels to the 5th and 95th percentile. Overall, the results are qualitatively similar. Relaxing the thresholds increases the number of material announcements from 799 to 1799. The actual levels of abnormality are higher compared to the headline measure, as more pre-announcement price movements are classified as abnormal, from 55 to 225. However, the long-term decrease in the probability of price abnormality, or the improvement in measured market cleanliness, remains evident. The 2020 data point remains an outlier, which reflects how unusual the market was during the Covid pandemic, though the outlier is less notable than when the more restrictive threshold is applied.

Figure B4: Market cleanliness statistic – by materiality threshold

Abnormal Trading Volume Ratio

For the headline measure, we have made several parameter and modelling choices. In this section, we test the robustness of our measure to these choices. Overall, the results are mostly robust, especially with regard to the long-term trend.

Liquidity requirements

For the ATVR measure, a stock needs to be traded for at least 30 days before the announcement. While this maximises the sample size of our analysis, it may also fail to exclude illiquid stocks which can bias the results. Figure B4 compares the ATVR with additional liquidity requirements similar to the MCS against the headline measure. We find that the measure is not sensitive to the imposition of additional liquidity requirements and using the 30-day observation period is enough to filter out illiquid stocks.

Figure B4: Abnormal trading volume ratio – by liquidity filter

Period length

In this section, we examine how sensitive the headline measure is to the length of the announcement period and benchmark period. Changing the length of these periods impacts the levels but not the year-on-year changes or the long-term trend.

Figure B5: Abnormal trading volume ratio by announcement periods ranging from 7 to 13 days

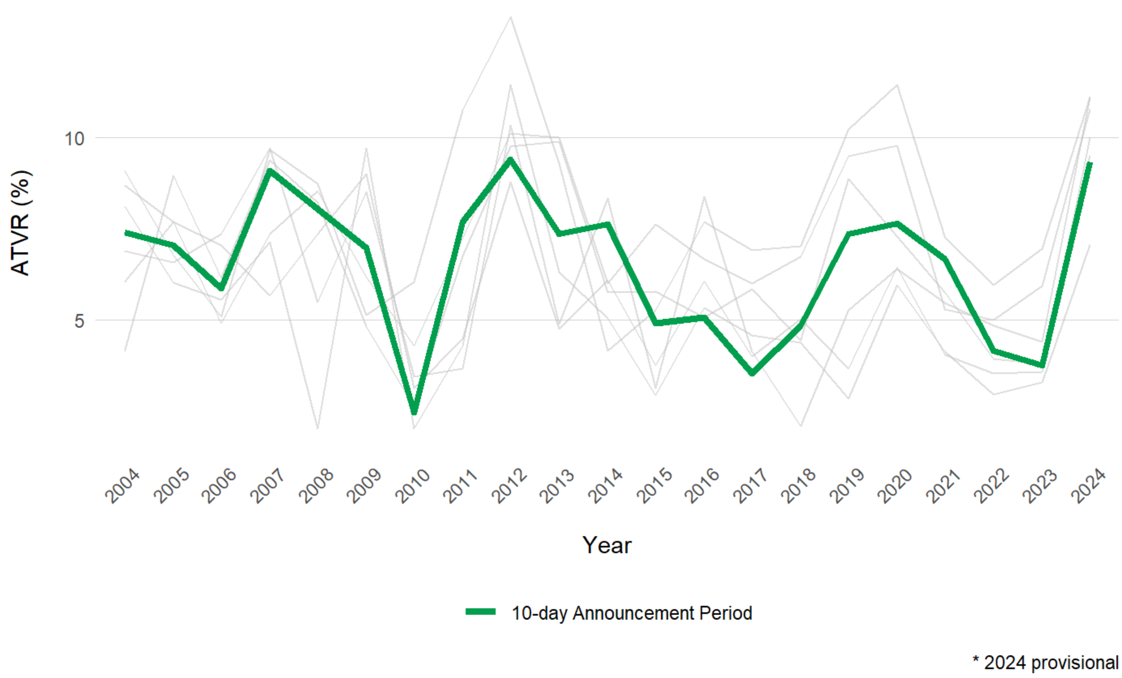
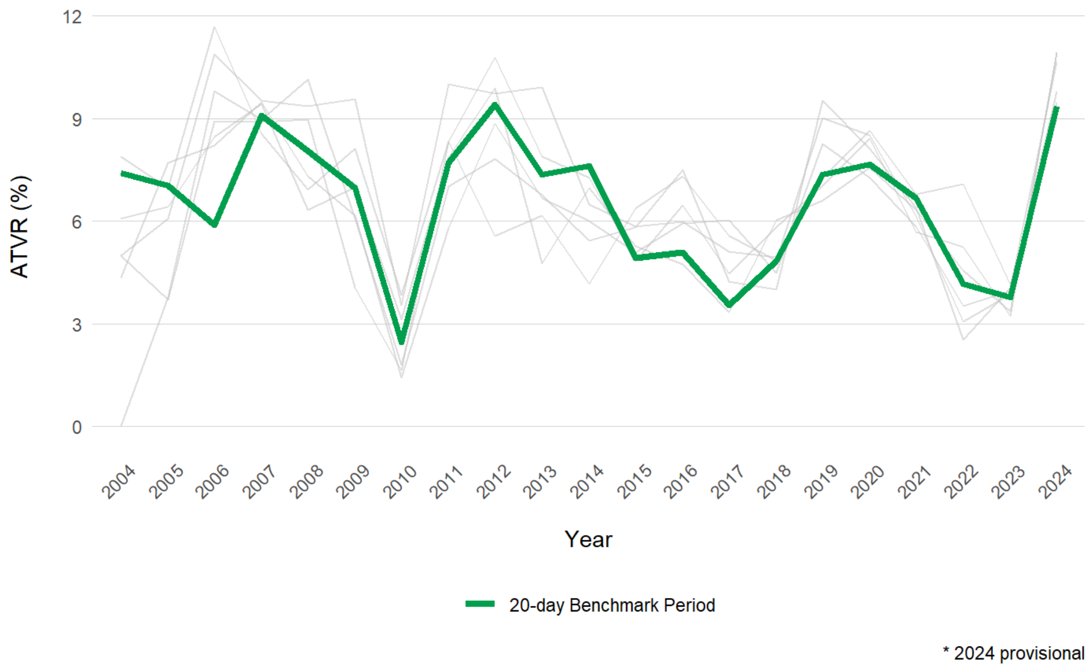


Figure B6: Abnormal trading volume ratio by benchmark periods ranging from 14 to 26 days



Appendix C: Univariate analysis

Market Cleanliness Statistic

This section provides a descriptive account of how the MCS has changed over time and how it differs by firm characteristics. The results are descriptive in the sense that we compare MCS outcomes across different groups, but unlike the multivariate analysis of sections 4 and 5 we do not control for other observable characteristics, such as firm composition and market volatility. These bilateral comparisons should therefore be taken with caution, as they may reflect differences in these unobserved factors rather than being attributable to the variable in question. In most cases, we use the multi-year averages in our analysis instead of the annual numbers to smooth out variation arising from the New Zealand equity market being small with a relatively low number of material announcements over time.

Impact of regulatory events

We first consider how market cleanliness in New Zealand has evolved over time, with references to the three regulatory events discussed in section 4 – the establishment of the FMA (2011), the passing of the FMC Act (2013), and the first prosecution of an insider trading case in New Zealand (2017). Establishing causality is not straightforward as there can be other factors impacting the abnormality of stock returns, including changes in market conditions.

Similar to Goldman et al. (2014) and ASIC (2016), we use a Z-test for difference in proportions to determine whether there was a statistically significant difference in the probability that a MPSA was preceded by an APPM in the periods before and after each regulatory event. Under the key assumption that each MPSA within a group has the same probability of having an APPM and that probability is independent of other MPSAs, the Z-statistic is calculated as:

$$Z = \frac{P_1 - P_2}{\sqrt{\frac{PQ}{n_1} + \frac{PQ}{n_2}}} \tag{15}$$

Here, n_1 and n_2 are the number of MPSA in group 1 (before) and group 2 (after). P_1 and P_2 are the probability of a MPSA being preceded by an APPM. P is the average probability of both groups and $Q = 1 - P$.

We include the event year in the before period to account for potential adjustment lag. We also exclude 2024 as the year is not complete. A statistically significant reduction in the likelihood of an APPM after a regulatory event is consistent with the event have some impact on deterring insider trading and improving market cleanliness. However, we do not empirically test for breaks in the series at these points.

Table C1 presents the results of the univariate tests. For all three events, the differences in probability were not statistically significant. However, the increase after 2017 suggests that the brief spike in 2020 might have skewed the results. Table C2 shows the results without 2020. After excluding 2020, a MPSA is 3.35 percentage points less likely to be preceded by a APPM after 2013 and the improvement is statistically significant at the 5% level. Smaller and non-significant differences for the other two split samples are consistent with these two splits being a mix of the pre- and post-2013 sample periods.

As a robustness check, Table C3 shows the results of the univariate tests using CAPM for estimating abnormal returns. These support the finding of a significant difference in the market cleanliness statistic between the periods prior to and following the passing of the FMC Act.

Comparing across the three splits suggests that the passing of the FMC Act is more likely to have been the driver of the reduction in APPMs over the period than the other two events, consistent with visual examination of the time series shown in Figure 2 which suggests a one-off reduction in the MCS statistic around 2013. At first glance, this is at odds with findings by Bhattacharya & Daouk (2002) that enforcement is more important than just the passing of regulation in improving market cleanliness. This may reflect the credibility of the New Zealand judicial system and of the FMA as a financial regulator. Bhattacharya & Daouk (2002) report that while 87 of the 103 countries in their study had insider trading laws in place, less than half that number had prosecuted any insider trading cases. If market participants in New Zealand had a stronger expectation that the conditions of the FMC Act would be consistently enforced, the distinction between the Act itself and its enforcement may be less relevant here than in some other countries.

Table C1: Univariate tests

Event year	MCS before (%)	MCS after (%)	Difference (p.p)	z-statistic	p-value
2011	7.25	6.77	-0.48	-0.23	0.82
2013	7.04	6.82	-0.21	-0.11	0.91
2017	5.65	8.38	2.73	1.5	0.13

Table C2: Univariate test results, excluding 2020

Event year	MCS before (%)	MCS after (%)	Difference (p.p)	z-statistic	p-value
2011	7.25	4.02	-3.22	-1.79	0.07
2013	**7.04**	**3.69**	**-3.35**	**-1.99**	**0.05**
2017	5.65	3.94	-1.7	-1.02	0.31

Table C3: Univariate test results, excluding 2020 - CAPM

Event year	MCS before (%)	MCS after (%)	Difference (p.p)	z-statistic	p-value
2011	5.77	4.49	-1.28	-0.72	0.47
2013	**6.9**	**3.66**	**-3.24**	**-1.92**	**0.05**
2017	5.74	3.57	-2.17	-1.31	0.19

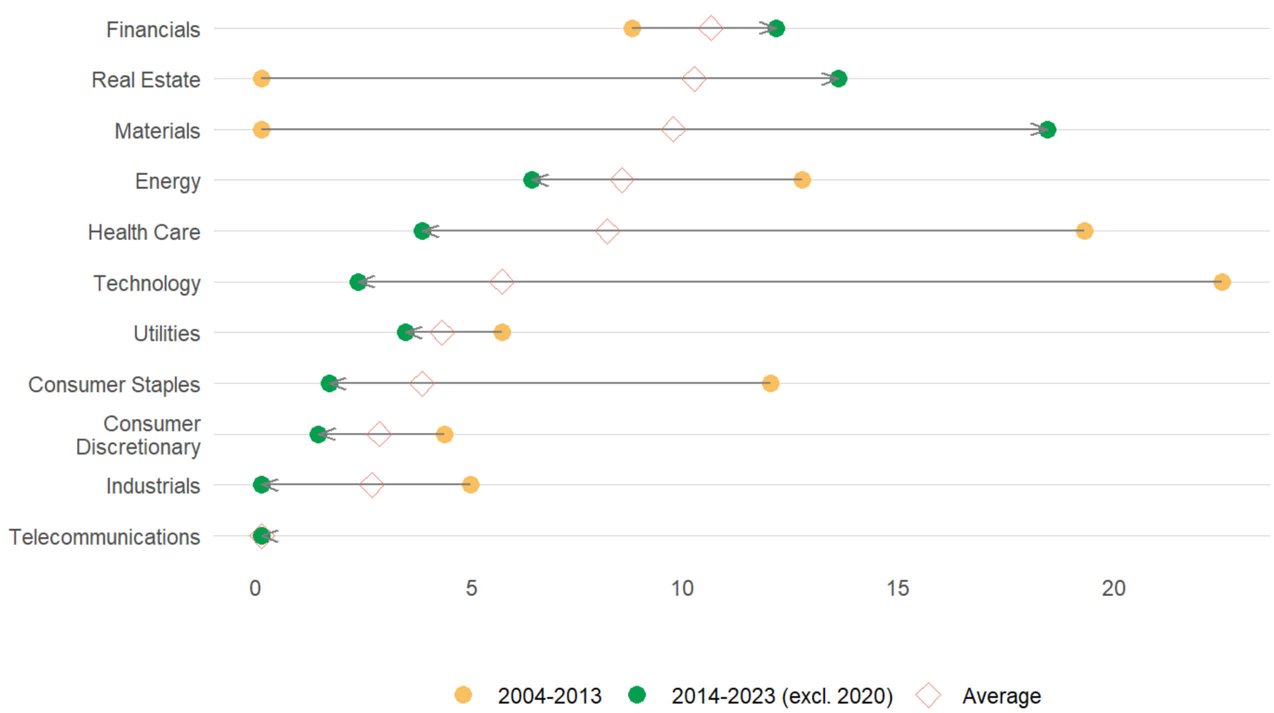
Industry

In this section, we examine the variation in market cleanliness across industries, using the 10-industry grouping by S&P/NZX. We look at both the long-term average and the change in market cleanliness before and after 2013 (excluding the 2020 spike) in Figure C1. These results are presented for completeness, though we note that many of the reported industry results are based on small samples, due to small numbers of both firms and material price-sensitive announcements over the period.

Most industries exhibited improvements in measured market cleanliness in the period after 2013 compared to the pre-FMC period. In contrast, Financials, Real Estate and Materials saw a higher level of APPMs in the post-FMC

Act period, even after disregarding the 2020 spike. These three industries also experienced the highest levels of pre-announcement price abnormality over the past 20 years. Market cleanliness improved the most in Technology, Health Care and Consumer Staples.

Figure C1: Market cleanliness statistic – by industry



Market capitalisation

Market capitalisation can also be a source of variation. In principle, larger companies tend to exhibit a higher level of market cleanliness as they have more resources and procedures in place to comply with disclosure requirements and manage confidential information. Larger companies' stocks also tend to be more liquid, which can limit the impact of insider trading activity on prices.

We divided the stocks on the NZX main board into 5 quintiles, according to their latest market capitalisation. Similar to the industry analysis, we excluded the 2020 spike from the 2014-2024 period. Ideally, we would want to look at market cap quintile at time of announcement. However, the data we have currently is not available at that level. Figure C2 shows the results. All quintile groups show an improvement in market cleanliness over the period, with lower proportions of APPMs in the period from 2014 onwards than in the earlier period. However, the notion that stock from larger companies tend to exhibit higher levels of cleanliness only held true from 2004 to 2013. This is because smaller companies have showed much more significant improvements recently.

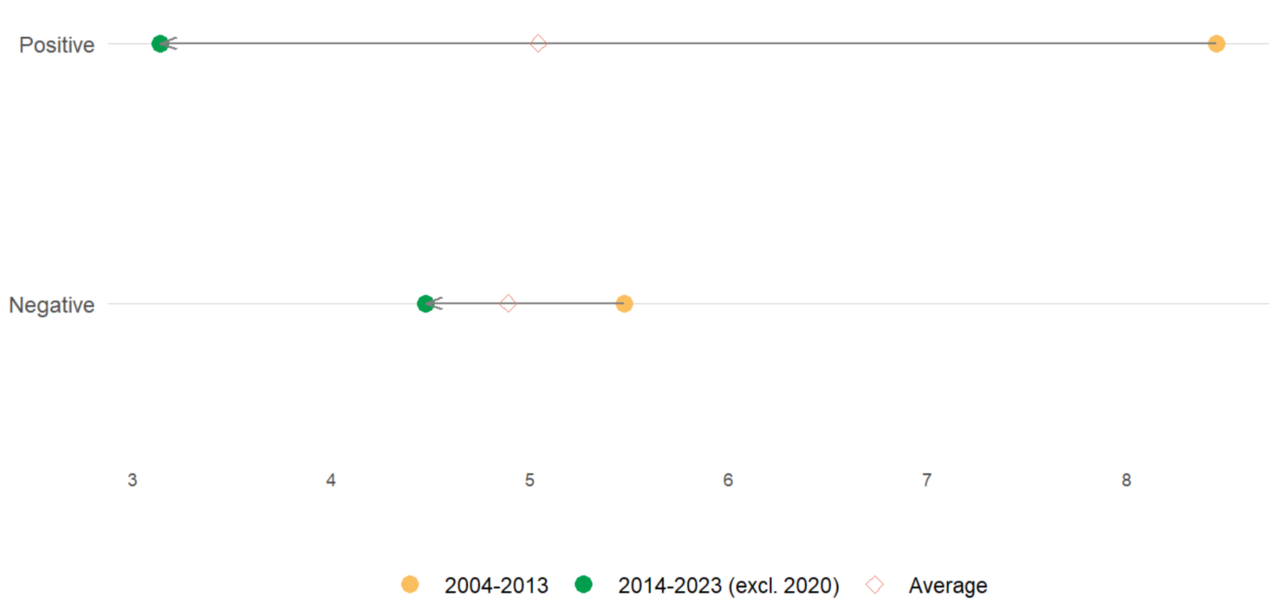
Figure C2: Market cleanliness statistic – by market capitalisation

Announcement type

We also compare the level of market cleanliness between positive and negative announcements. Figure C3 shows the results for the New Zealand market. Similar to the previous section, we also dropped 2020 from our sample. In 2020, the MCS for negative announcements showed a very large increase to 38.6% (17 out of 44 negative announcements were preceded by price abnormality), which could have skewed the overall trend.

Overall, the results are not conclusive, potentially due to the absence of a publicly available register of short selling positions in New Zealand, like in the UK (FCA, 2024) and Australia (ASIC, 2018), or a potential stronger deterrent effect from the surveillance and enforcement activities associated with positive announcements. Positive announcements were more likely to be preceded by price abnormality, or less clean, than negative announcements before 2013. However, after 2013, the cleanliness level for positive announcements has shown a much stronger improvement, dropping to 3% from 9% in the previous period. At the same time, the cleanliness level for negative announcements only slightly improved, dropping to from 5.5% to 4.5%.

Figure C3: Market cleanliness statistic – by announcement type



Abnormal Trading Volume Ratio

This section repeats the analysis above for the ATVR. As with the case of the MCS analysis in the previous section, we use multi-year averages instead of the annual numbers to account for the low number of material announcements in the New Zealand market.

Impact of regulatory events

Table C4 shows that the ATVR was lower in the 2010s and 2020s than it had been in the 2000s. Consistent with the univariate analysis of MCS in the previous section and the multivariate analysis in the main text, the change in market cleanliness appears to have occurred around the time of the FMC Act in 2013. The probability of a pre-announcement ATV was 1.65 percentage points lower in the post-2013 period than in the earlier period, significant at the 10% level.

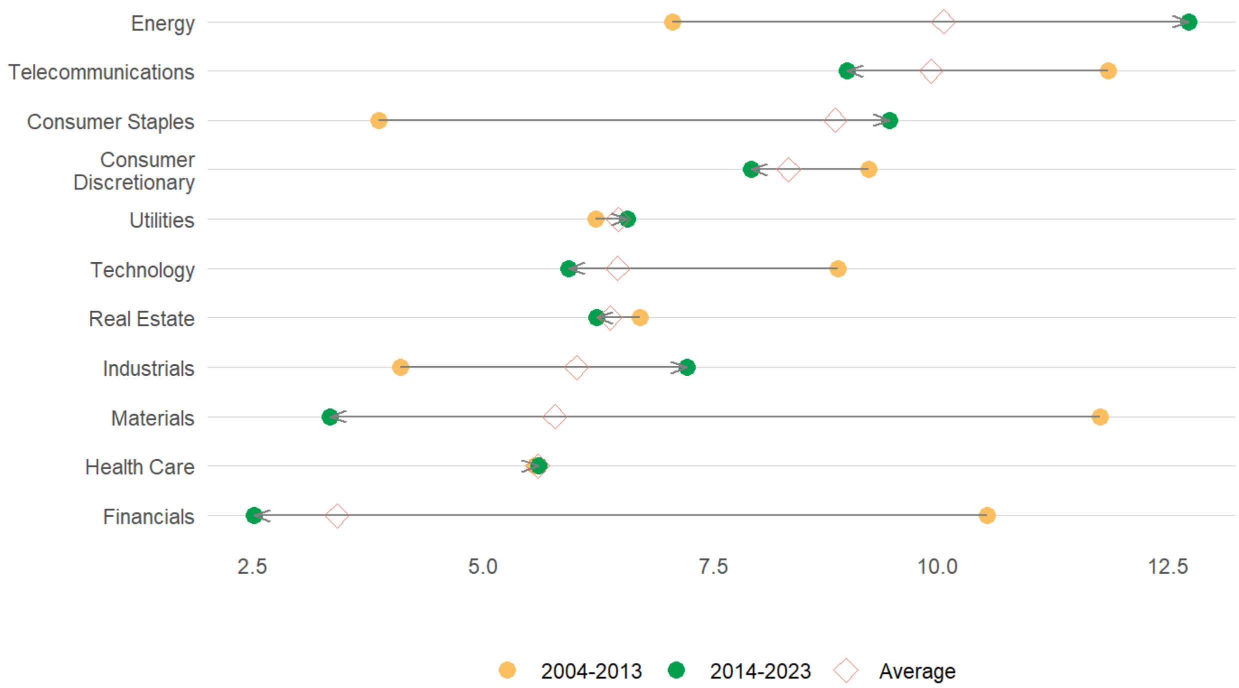
Table C4: Univariate tests - ATV

Event year	ATVR before (%)	ATVR after (%)	Difference (p.p)	z-statistic	p-value
2011	6.87	5.75	-1.12	-1.05	0.3
2013	**7.21**	**5.56**	**-1.65**	**-1.7**	**0.09**
2017	6.39	5.66	-0.73	-0.85	0.39

Industry

Figure C4 shows the variation in the ATVR across industries. Over the longer term, Energy, Telecommunications and Consumer Staples had the highest average levels of ATV. On the other end of the spectrum, Materials, Health Care and Financials have the lowest levels of ATV, or the highest levels of market cleanliness. These results are quite different from the MCS results, in which Financials is one of the three industries with the highest levels of stock return abnormality. These differences suggest that it is important to look at a suite of indicators for market cleanliness, not only for the headline measure but also for the analysis of market segments.

Figure C4: Abnormal trading volume ratio – by industry



Market capitalisation

In theory, larger companies with more liquid stock should exhibit higher levels of market cleanliness. However, comparison of the ATVR by market cap suggests that in New Zealand trading volume abnormality is more prominent in larger companies (Quintile group 4 and 5). Quintile group 1 (smallest) and Quintile group 3 (middle) showed the largest improvement in market cleanliness. The reason for this is unclear.

Figure C5: Abnormal trading volume ratio – by market capitalisation