The Integrity of Closing Prices

Ryan J. Davies
Babson College

September 10, 2018

Prepared as a chapter for “The Handbook of Financial Market Manipulation and Fraud” (Wiley)
Edited by Carol Alexander and Douglas Cumming

Why closing prices matter

The integrity of closing prices matters. Closing prices are used to determine mutual fund net asset value (NAV), option and futures final settlement values, variation margin, mark-to-market and daily settlement for futures, forwards, and other derivative contracts. Closing prices are also used for benchmarking institutional and broker trade execution and for determining fund manager remuneration. Some crossing networks use closing prices as the benchmark price. Closing prices are also used in the pricing of seasoned equity offerings and in the pricing of mergers and acquisitions. Exchange listing requirements typically specify minimum closing prices over a given period. Closing index values and index rebalancing decisions are based on closing prices of constituent stocks. Some momentum-based strategies and technical analysis-based strategies are influenced by closing prices, particularly when those prices exceed or fall below key threshold values. Finally, there is some evidence that end-of-day manipulation can reorient corporate managers to become more short-term focused and can worsen certain traditional corporate finance outcomes. For example, firms that experience end-of-day manipulation appear to file less patents subsequently (see Cumming, Ji, and Peter (2016)), and are more likely to have mergers and acquisition deals withdrawn (see Cumming, Ji, Johan, and Tarsalewska (2016)).

Given the importance of closing prices, it is essential to understand how to detect financial market misconduct that reduces closing price integrity, and design markets and regulation that deter such behaviour. Detection can be challenging and only a small fraction of closing price manipulation is prosecuted. Comerton-Forde and Putniņš (2014) estimate that one percent of closing prices are manipulated, using an approach whereby they jointly estimate the probability of being detected and the

---

1 Associate Professor of Finance and Chair of the Finance Division. Ryan Davies is a Business Advisor for Volos Portfolio Solutions, LLC., which provides software for option strategy optimization. 231 Forest St., Babson Park MA 02457. Tel: 781-239-5345. Email: rdavies@babson.edu.
process by which violations occur. Often researchers can only infer whether manipulation occurred based on changing in underlying market conditions. For example, Comerton-Forde and Putnīņš (2011) examine the effects of closing price manipulation and find that manipulation is associated with large increases in day-end returns, return reversals, trading activity and bid-ask spreads.

In this chapter, I examine three related research themes related to the potential manipulation of closing prices. First, I examine the incentives for mutual fund and hedge fund managers to inflate the closing price of their existing holdings by trading near the close of key performance measurement periods, such as the end of the quarter or the end of the year. Second, I examine how other financial intermediaries reacted to a specific well-documented instance in which a portfolio manager used “bang-the-close” trades to influence the settlement prices of commodity futures contracts. Third, I examine potential manipulation of closing stock prices on equity option expiration dates. The chapter concludes with a discussion of regulatory implications and recent market design issues.

**Painting the tape and portfolio pumping**

“Painting the tape” or “leaning for the tape” is often used to describe when mutual fund or hedge fund managers purchase securities that they already hold with the intention of inflating the end-of-quarter price to improve their relative performance. This behaviour is also referred to as portfolio pumping and net asset value (NAV) inflation. These trades occur near the market close with the goal of triggering a short-term price impact that increases the closing prices used to calculate end-of-quarter reported NAV. The short-term price impact typically unwinds at the opening of the next day, and as such, fund managers are effectively increasing the returns of the ending quarter at the expense of the next quarter. Quarterly returns are reported to fund investors in their quarterly statements and often are used to determine fund performance rankings. These returns are used by rating agencies and regulators, and used by retirement investment committees with a fiduciary responsibility to monitor and evaluate available fund choices. Fund manager remuneration is also often determined, in part, by end-of-year fund return performance.

---

2 In other contexts, the term “painting the tape” is used to refer to a series of transactions designed to give the impression of activity or price movement.

3 Painting the tape is not to be confused with “window dressing”, which is characterized by purchasing recent “winners” and selling recent “losers” prior to the end of quarters, so that reported mutual fund holdings look more favourable to investors. Lynch, et al. (2014) use trade data to distinguish between window dressing, institutional tax-loss selling, and risk shifting at year-ends. Sias and Starks (1997), Agarwal, Gay, and Ling (2014), Lakonishok, et al. (1991), and Ng and Wang (2004) also study window dressing and end-of-year risk adjustments.
The empirical evidence that portfolio pumping occurs is quite strong. Carhart, et al. (2002) find that 80% of funds beat the S&P on the last trading day of the year (62% for other quarter-end dates), but only 37% (40% other quarters) do so on the first trading day of a new year. They find that the difference is greater for small-cap funds, and that end-of-year performance is more pronounced for better historical performers. Bernhardt and Davies (2005) argue that Carhart, et al. (2002) may actually be underestimating the effect. Using net flows to mutual funds as an explanatory variable in a time-series regression framework, Bernhardt and Davies (2005) provide evidence that “painting the tape” may be so widespread that its effect is reflected in the overall index. Specifically, daily returns of the equally weighted index on the last trading day of a quarter greatly exceed the daily returns on the first trading day of the succeeding quarter, and this return difference rises with the share of total equity held by mutual funds. As such, empirical studies such as Carhart, et al. (2002) that benchmark price inflation relative to the index are likely to underestimate the true impact of painting the tape. As further evidence of painting the tape (as well as window dressing), Brown, et al. (2017) find that active funds underperform the market and other passive benchmarks only in the first month of a quarter.

End-of-quarter and end-of-year return patterns may be driven by other factors as well, such as the impact of tax loss selling and window dressing. The pattern may extend to end-of-months, as McConnell and Xu (2008) find a turn-of-the-month returns pattern has persisted in U.S. markets over the period 1926-2005, and find similar patterns in 31 out of 35 countries examined. They argue that these patterns cannot be explained by the performance-flow incentives of fund managers alone.

Direct evidence of portfolio pumping is obtained from institutional trade data. Gallagher, Gardner, and Swan (2007) use daily trading data of a set of active fund managers in Australia to provide direct evidence of ‘painting the tape’ at the end of quarters and show that such gaming behaviour is more likely to occur in smaller, illiquid stocks, momentum stocks, and stocks in which the fund is overweight. Hu, McLean, Pontiff, and Wang (2014) examine institutional trade data provided for U.S.-based funds using a specific institutional broker (Abel Noser Solutions) and argue that year-end price inflation may be caused more by depressed selling, rather than excessive buying. Conditional on buying at year-end, they find that institutions tend to buy stocks in which they already have large positions.

Portfolio pumping may help explain some of the other puzzling empirical anomalies for mutual funds and hedge funds. For instance, Bernhardt and Davies (2009) develop a model of portfolio pumping that incorporates three well-known observations: (i) past fund performance influences subsequent net fund flows; (ii) fund manager compensation rises with total assets under management;
and (iii) trading has short-term price impacts. These observations provide fund managers the incentive to mark up their holdings at quarter-end through aggressive trading of stocks they already hold. Fund managers’ incentives to distort investment of new cash inflows toward stocks in which the fund has larger positions leads to the empirically observed short-run persistence and long-run reversal in fund performance. The intuition is that funds with greater past performance will have more cash inflows, providing them with more ability to engage in end-of-quarter trading, which leads to greater reported performance and greater inflows the next quarter. In the model, the benefits of portfolio pumping are not sustainable long-term, since each subsequent quarter, the fund will start with a greater deficit as the price impact of the end-of-quarter trades decay, requiring the fund to inflate subsequent returns by even more. Eventually, the fund will not be able to “paint the tape” enough to overcome the performance deficit, leading to a reversal in performance. The model provides a possible reason for why funds appear to exhibit clairvoyant stock selection and yet under-perform long term.

Mutual fund incentives to “paint the tape” are driven by the tournament-like investment behaviour in the fund management industry, with cash inflows going to winners. Sirri and Tufano (1998) examine the flow of funds into mutual funds in response to performance and show that flows go disproportionately into winners (i.e., the performance-flow relationship is convex). Berk and Green (2004) develop a model to show that the observed flow-performance relationship can arise endogenously in rational markets, even when performance is not persistent and active managers, with varying skill levels, underperform passive benchmarks on average.

In the model of Bernhardt and Davies (2009), the response of fund flow to past fund performance is exogenous. Given the convex performance-flow relationship, mutual funds with the best recent performance are the most likely to have both the incentive and the capacity (new funds) to engage in painting the tape behaviour. This feature is consistent with the patterns observed in U.S. markets by Carhart, et al. (2002) and Duong and Meschke (2016), and in Korean markets by Lee, Baek, and Park (2014). In these markets, portfolio pumping appears motivated by the flow-performance relationship. In contrast, Shackleton, et al. (2017) argue that end-of-quarter price inflation is larger for the worst-performing fund managers in China because of the threat of replacement and dismissal for poor performance. Similarly, Gallagher, et al. (2009) focuses on institutional, wholesale funds in Australia that lack a strong performance-flow relationship. These institutional funds have professional trustees that will dismiss poorly performing fund managers, and as such, poor-performing managers appear more likely to engage in gaming trades, presumably to protect their jobs.
Hedge fund managers may have an additional incentive to inflate end-of-year portfolio values since their performance incentive fees are directly tied to fund values at certain key measurement points. Ben-David, et al. (2013) show that hedge funds also engage in portfolio pumping trades, particularly for illiquid stocks. Agarwal, et al. (2011) show that hedge funds tend to have higher December returns if they have high incentives and more opportunities to inflate returns. They argue that a hedge fund’s incentive to inflate its returns depends on: (i) the asymmetric call-option-like payoff of its incentive fees; (ii) its flow-performance sensitivity and current performance ranking; and (iii) its penalties for poor performance, including lockup and restriction periods. Similarly, they argue a hedge fund has more opportunities to inflate returns when its returns are more volatile, thereby allowing it to hide returns management, and when it has higher exposure to liquidity risk, thereby allowing it to paint the tape more effectively.

Mutual fund and hedge fund managers are financial intermediaries acting on behalf of their investors. By engaging in NAV inflation, they impose direct and indirect costs on investors, largely to increase their own compensation. In the next section, I examine a specific example of how financial intermediaries reacted to a long-term scheme to inflate commodity futures settlement prices using “bang-the-close” trades.

“Bang-the-close” manipulation: The response of financial intermediaries

Atanasov, Davies, and Merrick (2015) study the response of NYMEX floor traders to an alleged manipulation, whereby a hedge fund portfolio manager (PM) submitted large “bang-the-close” buy orders for platinum and palladium futures contracts over a 7-month period (November 2007 to May 2008). Detailed court records allow them to examine whether the floor traders, acting as financial intermediaries, mitigated or magnified the alleged manipulative scheme. In this case, court transcripts outline the PM’s desire to purchase these contracts at the highest possible prices, thereby triggering “buy signals” through new highs in the futures daily settlement prices. The PM had a major financial motive to increase these prices as his fund held large net long positions in platinum and palladium – worth almost a billion dollars at the peak of this period. Higher fund returns would result in higher compensation for the PM and would also attract fund inflows, leading to higher fund management fees.

---

At the time of the alleged manipulation, platinum and palladium futures contracts were traded via open outcry on the NYMEX floor and via the Globex electronic limit order book. While most trading had migrated to the electronic platform, about 20% of trading volume still occurred on the floor. The platinum open outcry session ran from 8:20am to 1:05pm, and the palladium open outcry session ran from 8:30am to 1:00pm. Electronic trading for both platinum and palladium occurred in a near 24-hour session from 6pm to 5:15pm the next day. The daily settlement price for the futures contracts was based on the volume-weighted average price of all NYMEX floor and Globex transactions during a two-minute closing period (1:03pm – 1:05pm for platinum; 12:58pm – 1pm for palladium).

Theoretical models predict that tacit collusion can persist in a transparent market with repeated interaction by a small number of similar participants and high barriers to entry (see, for example, Ivaldi, et al. (2003)). At the time of the alleged manipulation, there were less than a dozen NYMEX floor traders remaining in these markets. These traders interacted with each other on a daily basis in a highly visible manner. Market frictions limited the ability of non-floor participants to interact with floor trades. Under these conditions, Atanasov, et al. (2015) find evidence that, in response to the bang-the-close trades, the floor traders appear to have extracted significant rents through non-competitive pricing and behaved in a manner consistent with tacit (implicit) collusion.

The PM directed his orders to the NYMEX floor rather than the more liquid, parallel electronic market. Over time, these trades became predictable and contained no fundamental market information. Irrespective of the PM’s motives, competition among NYMEX floor traders should have limited the price impacts of his repetitive bang-the-close trades to a normal “size and immediacy” mark-up. Empirical evidence suggests otherwise. The PM trades had price impact both because the trades received higher prices on the floor and because floor trading enabled him to trade in higher quantities (and thereby influence the weighted settlement price by more). Importantly, the floor traders had hand-held devices that allowed them to trade directly on the electronic order book, allowing them to

---

5 Friedman (1971) shows that collusion may be a possible equilibrium outcome if the future costs to each player of deserting the equilibrium exceed the immediate gains. Dutta and Madhavan (1997) develop a game-theoretic model to show that, when order flow access is restricted, implicit collusion can arise even from non-cooperative behaviour among dealers. In experimental settings, Cason (2000) shows that dealers can tacitly collude to widen spreads, even without any direct communication, and Kluger and Wyatt (2002) find that the ability for dealers to internalize order flow allows them to coordinate on a less competitive equilibrium. Back and Zender (1993) show that collusive strategies can be self-reinforcing in uniform-price divisible-goods auctions. Khwaja and Mian (2005) find instances in which stock brokers in Pakistan, acting as financial intermediaries, colluded to trade among themselves to raise prices artificially in a “pump and dump” price manipulation scheme. These brokers traded for their own advantage at the expense of outside investors.
unwind positions at the walk-the-book price with very little risk. Floor trading rules allowed the floor traders to specify the quantity they wished to trade, so that they did not need to sell the entire position requested by the PM. The PM directed his futures commission merchant’s floor broker to execute his orders at very close to (possibly even slightly after) the official market close. As a consequence, the floor traders who were counterparty to the PM’s trades needed to unwind their short positions on the electronic book after the closing settlement period. Observed intraday trading patterns were consistent with this hypothesis, with trade volumes in the two-minute window immediately after the floor close higher during the alleged manipulation period than during the non-manipulation period.

Atanasov, et al. (2015) develop counterfactual futures contract pricing benchmarks to separate the share of estimated damages caused by the non-competitive behaviour of floor traders from the share caused by the direct impacts of the PM’s bang-the-close trades. These benchmarks are constructed using data from trades and orders on the parallel Globex electronic platform. Through this approach, damages can be attributed properly across the multiple market participants.

The decomposition of the impact of the bang-the-close trades is illustrated by considering three cases. In the first case, they consider the proportion of the PM’s order that could have been executed immediately on the Globex electronic limit order book by walking-the-book. This “smoking gun” evidence compares prices for the PM’s exchange floor-filled orders to the concurrent volume-weighted average price for tradeable depth on the Globex platform. The difference between the walk-the-book price and the price obtained on the floor reflects non-competitive pricing on the floor. In the second case, they hold constant the mark-up obtained on the floor but now consider the impact of the ability of the PM to trade larger volumes on the floor than available on the electronic limit order book. Specifically, $\Theta_{\text{PM}} > \Theta_{\text{WTB}}$, where $\Theta_{\text{PM}}$ is defined as the PM’s trade volume on floor during the closing period as a share of the total trade volume on both the floor and Globex during the closing period; and $\Theta_{\text{WTB}}$ is defined as the available walk-the-book depth at the close divided by the sum of the closing period volume on Globex, the available walk-the-book depth at the close, and the closing period floor trades not involving the PM. Finally, in the third case, they examine the excess price impact on the residual portion of the PM volume, over and above the mark-up on the quantity of the PM’s trades immediately available on the electronic limit order book. Figure 1 illustrates this final decomposition.
Figure 1: Decomposition of Settlement Price Artificiality when the number of contracts traded for the PM on the floor exceeds the available walk-the-book depth on Globex, $\Theta_{PM}^r > \Theta_{WTB}^r$, and there exists an additional mark-up for the contracts that cannot be immediately executed on Globex (Residual Mark-Up $> 0$). Now, the residual settlement price artificiality can be attributed to both the ability of the PM to execute larger volume on the floor than Globex and to the additional price impact of this larger volume. The figure also assumes that, excluding trades involving the PM as counterparty, the average price of floor trades executed during the 2-minute closing period is equal to the VWAP of Globex trades during the period (the Globex Settlement Price).

The PM submitted large trades at the close almost every day. To investigate how the floor brokers reacted over time, the alleged manipulation period was split in halves. In a competitive market, one would expect that the floor brokers and other traders on the floor would reduce the price impact of these trades over time. The opposite was observed: the PM’s bang-the-close trades had greater price impacts and received worse prices in the second half of the alleged manipulation period. During the second half of the alleged manipulation period, floor traders executed the PM’s platinum futures contract buy orders at prices that were 40-80 ticks above competitive benchmarks. This pattern is consistent with tacit (implicit) collusion. The observed impact is much larger than the conjecture of Pirrong (1996) that the effect on prices from cooperation by floor traders (or locals) in a typical open outcry market may be limited to one or two ticks.

Atanasov, et al. (2015) estimate the potential profits of the floor brokers using two approaches. The first approach estimates profits based on the floor brokers executing only the immediately available
walk-the-book portion of the PM’s trades. Under these assumptions, the floor traders are estimated to have achieved a profit of $1 million that is close to risk-free. The second approach estimates profits assuming that the floor brokers executed the volume exceeding the walk-the-book portion by participating in half of the volume of buyer-initiated trades on Globex post-close. Under these assumptions, the floor traders are estimated to have achieved a profit of roughly $6 million with minimal risk. Regulators and exchange operators often presume that competitive forces will limit the ability of prices to deviate from fair values. But this example highlights that the inflated contract settlement prices resulting from an alleged manipulative scheme can be facilitated and magnified by the actions of other financial intermediaries. As such, these parties also share part of the blame for the resulting damages.

Unlike “painting the tape” which is intended to have a short-term price impact, “banging the close” manipulation is often intended to trigger longer-term follow-on impacts. While Atanasov, et al. (2015) are able to decompose the short-term reaction of other market participants to the bang-the-close orders, it is more challenging to determine the long-term second-order effects from the manipulation. Court records suggest that the PM intended for his trades to cause the price to cross key technical levels, triggering trades by other market participants that use technical analysis and/or momentum-based strategies. Anecdotal evidence suggests that the rise in platinum and palladium prices during this period may have been largely driven by other macro trends and not the PM’s trades, but additional analysis would be needed to attribute these effects.

The fragmentation of order flow across two alternative trading channels enabled the PM’s trades to have a greater price impact. Off-the-floor futures market participants did not, or could not, place limit orders on the floor to take advantage of the mispricing. This observation provides guidance for the design and regulation of today’s markets, in which fragmentation can occur through both barriers to access and differences in latency.

At the time, the settlement price mechanism for platinum and palladium futures contracts used a volume-weighted average price across the two markets. This mechanism appears to have enabled the alleged manipulation to occur. Subsequent to this case, the mechanism was changed such that only Globex trades would be used in the determination of the settlement price.

With the rise of electronic trading of commodities futures, the settlement price mechanisms for other products have also changed. For instance, the method for determining the settlement price for
corn futures on the Chicago Board of Trade changed in June 2012. Prior to the change, the settlement price was based solely on a weighted average of floor trade prices during the last minute of floor trading. After the change, the price of both floor trades and electronic trades, occurring during the last minute of floor trading, were included in the calculation. Onur and Reiffen (2018) examine how trading behaviour, and the incentives to bang-the-close, changed in response to the new “blended settlement” methodology for corn futures. They provide some evidence that manipulation decreased after reducing the importance of floor trading, as measured in terms of less price reversals and in terms of price movements more closely resembling a random walk.

Bang-the-close trades have been studied in other contexts as well. In the context of foreign exchange “fix” orders, Saakvitne (2016) develops an equilibrium model of trading behaviour and benchmark prices. This model shows that bang-the-close trading can arise naturally as a solution to an optimal order execution problem for dealers that receive fix orders that are benchmarked against a price at a certain time. Similarly, Baldauf, et al. (2018) develop a model of an agency broker executing a large trade for a client, where performance is benchmarked to a reference price, such as the closing price. Their model shows that the broker has the incentive to time their trades over the day in a manner that impacts the reference price. These papers highlight how principal-agent problems can arise when closing prices are used as reference prices to evaluate broker performance.

In the next section, I discuss how trading patterns in a security that underlies a soon-to-expire derivative contract can also raise concerns about potential manipulation. The issue is complex, since the hedging trades of financial intermediaries, such as market makers, may be difficult to disentangle from intentional manipulation.

**Stock price pinning on option expiration dates**

In option markets, “pinning” occurs when the underlying stock price closes at, or near to, an option strike price on an option expiration date. Pinning causes both call and put options at this strike price to expire worthless, which has led some market participants to believe that pinning is often caused by the manipulation of stock prices. Ni, Pearson, and Poteshman (2005) show that closing stock prices do tend to cluster at option strike prices on option expiration dates. They are, however, largely

---

6 Edmunds, et al. (1984) provides an early study of stock pinning on option expiration. More generally, the interaction between stock prices and options has been studied by Klemkosky (1978), Sorescu (2000), and others. The issue of pinning is related to a growing, much larger literature that shows that non-fundamental (or technical) factors have become more important in determining security prices.
unable to resolve whether the pinning is caused by deliberate manipulation or as a natural by-product of market maker delta hedging. The importance and frequency of stock pinning has likely increased over time, corresponding to the dramatic increase in exchange-traded option trading volumes.

Before proceeding, consider the following explanation for why delta hedging can lead to stock pinning. Delta-hedging agents, such as market makers, must frequently rebalance their portfolios in order to maintain their delta hedge. As the price of the underlying asset changes, the delta of an option position changes. Recall that the rate of change in delta with respect to the price of the underlying asset is called gamma. The gamma of a long (bought) call or put option position is always positive. As the underlying share price rises, agents with a written option contract will need to purchase additional shares to maintain their hedge. This action corresponds to buying more shares to hedge a written call option position and reducing the number of shares sold short to hedge a written put option position. A high gamma indicates that delta is highly sensitive to the price of the underlying asset.

As expiration approaches, the delta of an out-of-the-money call option approaches zero and the delta of an in-the-money call option approaches one. When a call option is trading close to at-the-money, its delta will tend to remain close to 0.50 until the final day of trading; as the final closing price on expiration becomes clearer, its delta will converge rapidly toward either zero (out-of-the-money) or one (in-the-money). A similar pattern exists for put options, with the delta of a near at-the-money put option converging rapidly toward either minus one (in-the-money) or zero (out-of-the-money) at expiration. In both cases, a short-life option trading near at-the-money will have a high gamma and will require active rebalancing to remain hedged. Furthermore, the delta of an in-the-money (out-of-the-money) call option and the delta of an out-of-the-money (in-the-money) put option both increase (decrease) as expiration approaches.

Stock price pinning can occur as a natural by-product of delta hedging on expiration dates when two conditions occur: (i) only a subset of agents actively delta hedge their option positions; and (ii) this group of agents has a net bought option position (i.e., their positions have a net positive gamma). To see why, first suppose that the stock is trading below the nearest strike price. As time passes, as long as the share price remains below the strike price, these delta-hedging (positive gamma) agents will rebalance their hedge by purchasing additional shares. If their share purchases have sufficient price impact, their purchases will drive the share price toward the strike. Next, suppose that the stock is trading above the nearest strike price. Now delta-hedging (positive gamma) agents will rebalance their hedge by selling shares as expiration approaches; again, the price impact of these sales will tend to drive the share price towards the strike price.
Because optimal delta hedging is a more complex trading strategy that can incur substantial transaction costs, it is reasonable to suppose that only market makers and some sophisticated institutional investors will typically engage in it. Prior research (see, for example, Lakonishok, et al. (2007)) finds that the most common option position for retail traders is a net written call option, which suggests that other, more sophisticated, investors are likely to have an offsetting net bought call option position. If so, the tendency of delta-hedging agents to have a net bought option position could lead to a greater tendency for a stock to be pinned.

The unconditional likelihood that delta-hedging leads to a particular stock pinning at a strike price depends on many factors. These factors include: (i) the coarseness of the option strike price grid; (ii) the volatility of the stock price; (iii) the average trade volume and open interest in the options (both in an absolute sense and relative to equity trade volume); (iv) the average net gamma position of market makers and institutional traders in nearby and distant contracts; and (v) the price impact of equity trades. At a given time, the conditional likelihood of pinning also depends on the distance from current stock price to the closest strike price, which acts like a “magnet”.

Krishnan and Nelken (2001), Avellaneda and Lipkin (2003), and Avellaneda, Kasyan, and Lipkin (2012) develop mathematical models of stock pinning. The key feature of these models is how the price impact of trading is determined. Krishnan and Nelken (2001) model the stock dynamics assuming that the \textit{ex-ante} likelihood of pinning is determined by a random variable (formally, they use a Brownian-bridge framework). Avellaneda and Lipkin (2003) model the effect of option market maker hedging trades on the underlying stock price. They allow pinning to be determined endogenously by the hedging trades, based on linear price impact function, incorporating an exogenous price elasticity of demand. Avellaneda, Kasyan, and Lipkin (2012) extend the model of Avellaneda and Lipkin (2003) to incorporate non-linear price impact functions. Ultimately, these models highlight the importance of the magnet effect of strike prices. A reduction in the coarseness of the strike price grid increases the likelihood that a stock will trade at prices close enough to a strike price to be pinned, while, at the same time, the strength of each magnet is likely reduced by spreading open interest across more strike prices.

In a series of regulatory decisions, options exchanges have been allowed to increase the number of stocks with options at a finer strike grid. More strikes appears to make pinning more likely since stock prices will be more likely to be in the range of a strike price near expiration. Does doubling the number of strikes cause the likelihood of pinning to double? It depends on how it impacts the strength of the magnets. More strike prices increase the likelihood that the nearest strike is closer to the current price
(increasing the magnet strength), but at the same time, more strike prices may lead to less outstanding contracts at each strike (reducing the magnet strength). The optimal coarseness of the strike price grid for exchange-traded options is an important decision for regulators and exchange operators.

Delta hedging does not eliminate all risk. It is difficult for market makers to hedge their positions in options that are close to at-the-money near expiration. Small changes in the price of the underlying stock can cause the delta of these options to go from almost one to almost zero very quickly (i.e., their gamma is very high). Some traders may elect to exercise their options when it is not expected (particularly during the last few days prior to expiration), leaving the other side of the trade with an unexpected delivery. A market maker with a written option position faces the risk that the option is exercised unexpectedly after the market closes on the last day of trading prior to expiration. Since after-hours equity trading, particularly on Friday, is often highly illiquid, the market maker may be unable to close the unexpected position over the weekend, exposing them to considerable price risk. This risk suggests that optimal delta-hedging trades might continue after-hours on expiration dates. As such, after-hours trading patterns may depend on the deadline for option exercise decisions and their settlement. Weekly options typically expire on Friday evening, whereas serial options typically expire on Saturday. While this distinction appears to have received little attention, the later expiration time may increase the possibility of a late, unfavourable delivery.

Determining whether the effect is due to market maker delta hedging or intentional manipulation depends critically on understanding the net position of market participants that use delta hedging. Ni, et al. (2005) studied the period 1996 to 2002, during which time U.S. equity option trading occurred only on the Chicago Board Options Exchange (CBOE). Their dataset included complete information on option market maker positions. Unfortunately, market maker positions are much harder to identify today. As of July 2018, equity options are publicly traded on 16 exchanges in the U.S., each with their own open-close data. Not only does the open-close data not typically include confidential market maker position information, the fragmentation of order flow, along with the possibility that offsetting market maker positions could be taken on another venue, makes it difficult to develop proxies for market maker positions. It is also complicated by a blurring of who is a market maker in option markets, as sophisticated high frequency traders also engage in market making activities, in part to due to the market-taker fee structure on many option exchanges.

Stivers and Sun (2013) examine stock return patterns during option expiration weeks for S&P 100 constituent stocks. They argue that delta-hedging rebalancing by market makers leads to
predictable weekly patterns in returns. Specifically, the tendency for market makers to have a net long position in call options leads them to have short equity positions. These short positions are unwound as expiration approaches, which causes higher weekly returns during option expiration weeks. To examine the role of high frequency trading in pinning, Figueiredo, Mishra, and Jain (2017) use Nasdaq Totalview-ITCH data to examine intraday trading behaviour near the market close on option expiration dates. Non-HFT limit order submissions appear to be clustered near round number strike prices, whereas the HFTs appear to submit fleeting orders that are inside or outside the round number prices.

A closely related widespread belief among traders is that stocks are manipulated such that they close at the “max pain” price on option expiration dates. Max pain is commonly defined as the stock price that causes the most outstanding written options to expire “out-of-the-money” (i.e., to be worthless). Despite being widely discussed within the trading community, the academic literature has been largely silent on max pain and other forms of manipulation focused on option expiration. This silence appears to be related to the difficulty in identifying manipulation and measuring its effects.

To the extent that market participants believe in stock pinning and max pain, these beliefs can become self-fulfilling as traders begin to trade in anticipation of the stock closing at the nearest strike or max pain price. With the introduction of weekly options, additional strikes, and an expansion in the number of stocks with listed options, these option expiration date effects could occur on more days and for more securities. Interestingly, the magnet effect of a strike price could do some of the manipulator’s work for them – the manipulator would simply need to push the stock price in the direction of the desired strike and the magnet effect would do the rest.

Conclusion: Lessons for the regulation and design of financial markets

This chapter has examined three main research themes related to the manipulation of closing prices. In all three cases, the economic consequences of closing price manipulation are potentially very large. Artificially high mutual fund NAV results in a transfer of wealth from new investors to existing investors in the fund. Improper price signals can distort management decisions and lead to an inefficient allocation of resources. Manipulation of the settlement prices for derivatives can reduce their effectiveness as a hedging instrument.

Zhang, Chen, and Cai (2015) examines the impact of introduction of weekly options in 2010 on stock price pinning. They find that pinning tends to occur less frequently with weekly options, due in part because of their lower trading volumes during the period studied (June 2010 – December 2012).
The issues raised in this chapter have implications for the design of markets and for the design of fund manager performance-based contracts. A natural question is whether the adoption of a certain market closing mechanism can prevent or reduce closing price manipulation. Unfortunately, there does not appear to be an easy solution. Some exchanges have adopted closing call auctions to increase transparency and liquidity at the close. These call auctions, however, may also be manipulated. There is a complex interaction between the closing call auction and trading of the limit order book immediately prior to the close. A determined manipulator can find ways to abuse indicative quotes, indicative order imbalances, and specialized order types (e.g., market-on-close).

Cordi, Foley, and Putniņš (2015) study the introduction of closing batch mechanisms in 20 markets. They divide the mechanisms broadly into those with closing call auctions that occur after continuous trade ends, and those with on-close facilities that operate in parallel to continuous trading, with limitations on new orders based on imbalances during the pre-close period. Both mechanisms appear to improve closing price efficiency relative to using the last trade. Cordi, et al. argue that the on-close facility lowers trading costs and improves price stability and informational efficiency. Call auctions cause spreads to increase because order flow is drawn away from the continuous market. In general, neither mechanism has a significant effect on the prevalence of closing price manipulation. Overall, they argue for a closing batch mechanism that has randomized closing times and extensions if volatility thresholds are breached. As well, they argue the mechanism should prevent traders from altering their orders during the pre-close time and should not display indicative closing prices.

Another natural question is whether the reporting of mutual fund performance, and the associated incentive structure for fund managers and other financial intermediaries, should be changed. Again, the answer is not clear. Most mutual fund investors already have easy access to daily mutual fund prices. Limited attention causes them to focus disproportionately on quarterly returns and their behavioural biases cause them to chase returns by investing in past top performers. Net asset values could be determined based on a weighted average of trade prices over a window of time (e.g., last five minutes of trading day), but these prices would also be subject to the potential for manipulation. Clearly, the compensation structure for hedge fund performance, combined with limited transparency,

---

is also fertile ground for potential manipulation. Hedge fund investors, however, are typically sophisticated investors who should be able to perform their own oversight.

Goetzmann, et al. (2007) propose that a “manipulation-proof” measure for evaluating a fund manager’s performance that accounts for non-normal returns distributions and avoids rewarding “information-free” trading. This measure resembles the average of a power utility function and is based on the fund’s return history. By focusing on the entire return history, rather than the fund value at a particular date, this measure should reduce incentives for portfolio pumping. While this “manipulation-proof” measure is highly cited in the academic literature, anecdotal evidence suggests that it has not been widely adopted in the finance industry.

Both opening and closing prices can be subject to manipulation. In general, closing price manipulation is typically trade-based, whereas opening price manipulation is typically quote or order-based. As after-hours trading becomes more common and trading becomes increasingly fragmented, the distinction between opening and closing mechanisms has become blurred. As such, lessons from the literature on market opening mechanisms are relevant to this discussion of the integrity of closing prices. Biais, Hillion, and Spatt (1999), Cao, Ghysels, and Hatheway (2000), Davies (2003), and Lescourret (2017) study price discovery and “gaming behaviour” during preopening sessions. Bellia, et al. (2017) find that high frequency traders are able to extract valuable information during the preopening of Paris Euronext, allowing them to profit at the opening auction. Biais, Bisière, and Pouget (2014) use an experimental setting to examine how price discovery at the market opening depends on whether preopening orders are binding and whether the preopening time is deterministic or random. They highlight some of the trade-offs in market design, such as noting that a random opening time may reduce manipulation, while increasing the cost of valuable “preplay communication”. It is difficult to know, however, the extent to which their conclusions would apply outside of their highly stylized experimental setting.

Hauser, Kamara, and Shurki (2012) show that the introduction of a random market opening time on the Tel-Aviv Stock Exchange improved price discovery and reduced excess volatility, especially on option expiration dates. Hauser, Kedar-Levy, and Milo (2018) study the parallel preopening sessions for stock and options on the Tel-Aviv Stock Exchange. As a proxy for mispricing at the opening, they

---

9 Other market openings, such as those after trading halts, or in the case of initial public offerings, may also be subject to manipulation but are beyond the scope of this chapter. See, for instance, Kuk, Liu, and Pham (2016) for a discussion of the preopening periods for IPOs.
measure the percentage gap between the at-the-money option-implied index and the indicative index. They find evidence that this gap is significantly larger on option expiration dates, which combined with subsequent reversals during trading hours, is suggestive of manipulation of these dates.

From a regulatory and legal perspective, it can be difficult to distinguish between legitimate trading behaviour and illegal fraud and manipulation. Typically, enforcement requires more than one-off trading because everyone has an Alternative Plausible Explanation (“APE”). Hence, enforcement of regulation is only possible (or more likely) with a clear pattern of repeat behaviour. End-of-day manipulation may be easier to detect and enforce since it often occurs in a regular pattern (e.g., option expiration, end-of-quarter). That said, legitimate explanations for some of these patterns are still possible. Mutual funds have the incentive to trade at the close to reduce tracking error and to reduce their trading costs. Market makers, faced with inventory constraints, may rationally respond to repetitive manipulative trades by increasing their price impact. Option market makers must delta hedge to manage their risk exposure.

Davies and Sirri (2018) summarize many of the economic and regulatory issues facing U.S. trading markets today. They highlight that a particular regulatory challenge is determining how different trading venues interact with each other, with regards to order protection rules, best execution requirements, market linkages, market access fees and minimum tick sizes, and pre- and post-trade transparency. As the earlier example of “bang-the-close” trades in platinum and palladium futures contracts highlighted, market frictions can enable manipulative trades to persist and have greater impact. A manipulator could elect to trade in venues with lower liquidity (and higher price elasticity) to maximize the impact of their trades on closing prices. Today, traditional market making, performed by specialists or other dedicated human market makers with affirmative obligations, has been largely replaced by high frequency trading. Fleeting quotes mean that trade execution costs can only be predicted on a probabilistic basis. In such an environment, reference prices, such as closing prices, may be difficult to establish and manipulation of these prices can be difficult to detect and enforce.

Traditionally, derivative markets and equity markets have operated under different sets of rules, governed by different regulatory organizations. Markets have become increasingly interconnected, and as such, regulatory environments need to be designed to detect and prevent financial market misconduct that occurs across different markets and financial products.
There is some evidence that regulation can curtail manipulation and make the effect of manipulation on markets less severe.\textsuperscript{10} In Europe, the 2003 Market Abuse Directive (MAD)\textsuperscript{11} sought to implement an EU-wide market abuse regime. The subsequent LIBOR benchmark manipulation scandal and the development of new trading platforms highlighted the need to update MAD with more robust legislation. This need led to publishing in 2014 of the Market Abuse Regulation (MAR)\textsuperscript{12} and the Directive on Criminal Sanctions for Market Abuse (CSMAD)\textsuperscript{13}, which, in combination, are known as MAD II. These new regulations were implemented into national law by the EU Member States in July 2016. Notably, the UK and Denmark opted out of the CSMAD. These regulations are intended to be consistent with the new Markets in Financial Instruments Directive (MiFID II). Among other things, MAR introduced offences related to the manipulation of benchmarks and commodities, as well as cross-market manipulation and the manipulation of opening and closing prices. The expanded scope of these regulations highlight the importance of closing and settlement prices to the proper functioning of financial markets.

From their study of alleged manipulation of commodity futures markets, Atanasov, Davies, and Merrick (2015b) highlight six regulatory implications:

1. The mechanism used to determine the settlement (or closing) price can magnify the impact of the bang-the-close trades.
2. Tacit collusion can occur in an environment with frequent, repeated interaction by a small number of similar participants in a transparent market with barriers to entry.
3. Market access plays a role in preventing market manipulation.
4. Participants may not be fully forthcoming about possible misconduct.
5. Timely, accountable enforcement of exchange rules can be lacking.
6. Counterparties to manipulative trades can be responsible for a significant share of the total artificiality in prices caused by the manipulation scheme.

\textsuperscript{10} For instance, Cumming, Johan, and Li (2011) find that exchange trading rules for market manipulation, insider trading, and broker-agency conflict can have significant impacts on liquidity.
These findings have implications for the determination of LIBOR, foreign exchange rates, and other interest rate benchmarks. For example, in the $5.3 trillion / day foreign exchange (FX) market, the WM/Reuters currency benchmarks are set at regular times by interbank trades occurring in a 1-minute interval (2-minutes for less active currencies). These interbank trades represent only a subset of FX trades and are conducted by a small group of participants, in a repeated process that is not easily accessible by outsiders. Dealers in this market, acting in the role of a financial intermediary, are alleged to have colluded to move these currency rates in anticipation of the settlement of customer orders. Many of the economic incentives of these FX dealers are similar to those of the floor traders studied by Atanasov, et al. (2015).

Barclays Bank, UBS, Rabobank, RBS, Deutsche Bank, JP Morgan, and other major money center banks, have been fined, or are under investigation, for their role in apparent collusion to set daily LIBOR fixings over an extended period of time. Total fines and settlements have exceeded $9 billion to-date and are expected to rise further.14 Gandhi, et al. (2017) estimate manipulation of LIBOR could have increased the market value of panel banks by over $22 billion. Abrantes-Metz, et al. (2012) find statistical evidence of patterns in LIBOR rate submissions that appear inconsistent with those expected to occur under conditions of market competition. In response to the LIBOR scandal, the oversight of the LIBOR rate setting process was transferred from the British Banking Association (BBA) to the Financial Services Authority (FSA). Under the new system, the individual submissions of panel banks are published with a three month lag, thereby reducing the immediate transparency of submissions and the ability of participants to enforce a collusive equilibrium.

In contrast to some of the changes to LIBOR, Aspris, Foley, Gratton, and O’Neill (2017) examine changes to the London ‘fix’ intended to reduce potential manipulative conduct by increasing transparency. They find that increased transparency improved market quality for the gold and silver fix, in terms of less information leakage, higher quoted depth, and lower transaction costs and volatility. They did not find similar effects for the less liquid platinum and palladium contracts. These results are interesting in light of models that suggest that transparency may increase implicit collusion, as it allows a small number of participants to enforce a collusive equilibrium in a repetitive setting, as evidence in Atanasov, et al. (2015) supported.

Equity markets and option markets have become increasingly dependent on each other. In the past, option prices were assumed to be \textit{derivative} from the underlying equity prices. Today, option markets have grown to sufficient size that the “tail is wagging the dog”: trade in options can and does influence underlying equity prices. For some widely-held stocks (e.g., Alphabet Inc.), the average daily trading in option contracts can represent a number of shares exceeding 25% or more of the average daily trade volume of the underlying equity.\footnote{For example, in July 2018 the average daily option volume in Apple Inc. (AAPL) options on the CBOE alone (just one of several U.S.-based option trading venues) was 63,435 contracts. These contracts represent 6.3 million shares (100 shares / contract), which is about a quarter of AAPL’s average daily trade volume of about 24 million.} Furthermore, for many of these names, the total open interest across option contracts, in terms of shares represented, \textit{far exceeds} the typical daily trade volume in the shares.

The potential gains from manipulation of the underlying equity on option expiration are substantial. If a market maker or another trader has a large outstanding option position, the cost of moving the underlying stock price could be much smaller than the resulting gains from the options.\footnote{Jarrow, Fung, and Tsai (2018) provide some evidence consistent with cross-market manipulation between index futures and index options in Taiwan.} In part, this phenomenon arises from the embedded leverage in options. All else equal, this type of manipulation should be easier when the stock is relatively illiquid (trades have high price impact), but the option has low implied volatility (options are cheaper, allowing for a larger option position).

The embedded leverage in options also manifests itself in products tied to option volatility. Because of the manner in which these products are settled, they are also subject to possible manipulation. Griffin and Shams (2018) find price and volume patterns in S&P 500 Index (SPX) options consistent with strategic trading with the apparent goal of influencing the settlement value of the CBOE Volatility Index (VIX). Their analysis leverages the feature that only some options are used in the VIX calculation and that some options have a higher and discontinuous influence on VIX. These “influential” options are more likely to exhibit abnormal trading near VIX settlement, consistent with market manipulation. Empirical evidence does not support alternative explanations such as hedging and coordinated liquidity trading. Overall, these results highlight the importance of the settlement procedure and how manipulation may be more likely to occur when price elasticities differ across different, interconnected contracts.
New concerns related to VIX and settlement prices have arisen because of large losses incurred by investors in XIV, the VelocityShares Daily Inverse VIX Short-Term ETN (exchange-traded note). As an ETN, XIV was a debt instrument issued by Credit Suisse, with a value inversely related to the S&P 500 VIX Short-Term Futures Index. The XIV prospectus disclosed that Credit Suisse expected to hedge its exposure to XIV by trading VIX futures (and other related derivatives) and noted that such hedging activities might adversely affect the level of the underlying index. At the close of equity trading at 4pm on February 5th 2018, XIV was worth $99 per share. After the market close, VIX futures spiked dramatically, as much as 96%. Between 4:10pm and 5:09pm, the Intraday Indicative Value of XIV was not updated from its last reported value of approximately $24.70. At 5:10pm, the Intraday Indicative Value of XIV was updated to $4.22 per share, reflecting the large spike in the VIX futures. Overall, this spike led to a drop of 97% in the value of the XIV from its prior closing value on February 1st, when it had been worth approximately $1.6 billion. This fall was sufficiently large that it triggered an acceleration event, enabling Credit Suisse to announce the next day its intention to close the ETN to prevent further losses. Regardless of whether Credit Suisse is found to be legally at fault for these losses, this case raises many concerns: (i) the demise of XIV was driven by a single spike in prices, which appears extreme relative to prices before and after; (ii) the hedging program of Credit Suisse was not aligned with the interests of the XIV investors and likely magnified the losses sustained by those investors; (iii) the XIV product and its hedging program may not have been well-understood by investors; and (iv) during the market dislocation, Intraday Indicative Values of XIV were not updated every 15 seconds, as was typical (Janus Index and Calculation Services LLC was responsible for providing estimates of the notes value).

The settlement process for options (and other derivatives) impacts the ease with which manipulation can occur. For instance, manipulation incentives may be different if an option contract is cash settled or has physical delivery at expiration. With cash settlement, the manipulator only needs to influence the price of the underlying security at expiration, at which point the manipulator settles in cash the value of the option based on this price. In contrast, with physical delivery, the manipulator may receive (or deliver) the underlying asset – the unwinding of the resulting position will cause the price to move against the manipulative trades, reducing the value of the manipulation. This phenomenon is sometimes known as the challenge of “burying the corpse” after the manipulation.

---

Regulation of trade-based manipulation tends to focus on the intent of the trader. The examples highlighted in this chapter highlight that it can be very difficult to identify intent. Automated rule-based trading algorithms that do not have direct human intervention magnify these challenges. Trading programs based on artificial intelligence and machine learning could execute trades near the market close that appear manipulative but arise from “innocent” optimization routines. Interestingly, Aitken, Cumming, and Zhan (2015) provide empirical evidence that the presence of high frequency trading (HFT) has helped mitigate the frequency and severity of suspected end-of-day price manipulation. Of course, high frequency trading firms also have the capacity to engage in manipulation. In 2012, the high frequency trading firm Optiver Holding reached a $14 million settlement with the CFTC in relation to allegations that Optiver, two of its subsidiaries, and three employees attempted manipulation of crude oil, heating oil, and gasoline futures on the NYMEX. The traders were alleged to have used bang-the-close trades in at least 19 instances during March 2007 to manipulate the settlement prices of crude and distillate oil futures contracts.19 The settlement prevented the company from trading oil futures in the three minutes prior to the market close for two years.

To the extent that HFT may be a “force for good”, it is also important to note that high frequency and low latency trading, combined with market fragmentation, have dramatically increased the sheer quantity of reported trades and quotes, and as such, it has become increasingly challenging in practical terms to detect and infer the intent of potential manipulative trades. To enforce trading rules in such an environment, regulators need to invest in computer surveillance technology and update regulations to reflect these new realities. The integrity of closing prices depends on it.

References


Atanasov, V., Davies, R.J., Merrick, J.J. (2015a). Financial intermediaries in the midst of market manipulation: Did they protect the fool or help the knave?, *Journal of Corporate Finance* 34, 210-234.


