

"This is the peer reviewed version of the following article: Accounting and Finance, 2015, 55 (3), pp. 717 - 748, which has been published in final form at <http://dx.doi.org/10.1111/acfi.12067>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving."

A Model of Emulation Funds

Zhe Chen ^{a, b, †}, F. Douglas Foster ^c, David R. Gallagher ^{a, b}, Adrian D. Lee ^c

12 September 2013

^a Macquarie Graduate School of Management, Sydney, NSW, 2109, Australia

^b Capital Markets CRC Limited, Sydney, NSW, 2000, Australia

^c UTS Business School, University of Technology, Sydney, NSW, 2007, Australia

JEL classification: G23

Keywords: multi-manager, fund-of-funds, transaction costs, emulation funds

*The authors gratefully acknowledge the assistance and support of an anonymous pension fund sponsor. The Capital Markets CRC Limited also sponsored this research.

[†]Corresponding author. E-mail address: zchen@cmcrc.com (Z. Chen).

A Model of Emulation Funds

Abstract

Emulation funds are a potentially cost-effective way for multi-manager funds to improve their investment performance by delaying and netting trade signals from underlying managers. We develop a model to represent the expected sources of differential performance in an emulation fund relative to its underlying multi-manager portfolio. The model formalises the expected interaction between potential savings and opportunity costs, and allows us to observe complexities in the emulation process that are hidden without a benchmark. Finally, the functional representation of the model allows sensitivity analysis of the emulation fund to key parameters, and enables us to determine theoretically optimal lag periods.

1. Introduction

In this study we develop and analyse a model for the process by which an emulation fund generates differential gross returns relative to its target portfolio. The model provides a number of analytical advantages over simulation alone. These include formalising the expected interactions between cost and benefit drivers in the emulation process, identifying complexities that are not consistent with expected model outcomes, enabling the forecast of emulation fund performance based on forward looking expectations rather than historical data, and allowing numerical optimisation of the exogenously determined emulation parameters.

An emulation fund is a multi-manager investment strategy relying on delegated portfolio management services, whereby the fund-of-funds uses the trade signals of their constituent fund managers to coordinate a separate portfolio that tracks the holdings of the underlying actively managed portfolio. A key feature of emulation funds is that trade signals are followed on a delayed basis to prevent competition for market volume with the underlying active fund managers. Commercial emulation funds rebalance using weekly or fortnightly snapshots of the aggregated underlying portfolio. However, we use a continual rolling window in this paper to smooth discrepancies between using particular start dates for a fixed window. The lag enables opposing market signals, within the delay period, to be offset against each other, and results in potential transaction cost savings. However, the altered timing of the trade signals introduces potential opportunity costs arising from adverse market price movements in the delay period.

Large institutional investors (e.g. pension funds and sovereign wealth funds) typically employ multi-manager investment organisations to manage their assets. The delegation of stock-picking decisions to multiple fund managers in a decentralised framework (i.e. constituent fund managers do not actively communicate with one another) also helps to mitigate manager-specific risk (Sharpe,

1981, Elton and Gruber, 2004, Brands and Gallagher, 2004). However, the use of active fund managers may introduce an additional level of management fees (Brown, Goetzmann and Liang, 2004, Ang, Rhodes-Kropf and Zhao, 2008), and ostensibly results in some degree of trading redundancy whereby constituent fund managers execute trades on opposite sides of the market within a short period of time. Hence, transaction costs are incurred with little change in net exposure to the underlying stock in the multi-manager portfolio. Emulation strategies seek to address both these issues by (1) allowing the multi-manager investor to negotiate low additional fees with the underlying fund manager to use their trade signals indirectly¹, and (2) internally offsetting opposing order signals. Emulation funds are also purported to deliver capital gains tax savings through their reduced turnover relative to the underlying multi-manager portfolio. However, due to the idiosyncratic nature of each investment service provider's tax structure, we leave this issue to be researched in a later study. Given the recent growth and interest in emulation products, the topic of emulation funds is increasingly relevant to large multi-manager funds.

Commercial interest in emulation funds as a low cost method of extending investment capacity has been primary driven by their pursuit in reducing transaction costs (and taxes), which are both well-observed and significant. For example, studies on brokerage commissions have found these costs to range from 0.11% of trade value in the US and 0.15% in Europe (Goldstein et al., 2009). In Australia, brokerage rates have fallen significantly over the past fifteen years. In our representative sample of active Australian equity fund manager trades, we observe a steady decline in commissions paid per trade from 0.36% in 1996 to 0.16% in 2010. On-market trading also incurs substantial price impacts, which have been found to be as high as 0.46% (Domowitz, Glen and Madhavan, 2001). Conceptually, emulation strategies are appealing as they are meant to reduce on-

¹ Informal discussion with investment executives suggests that reusing a fund manager's trade signals typically costs around 0.10% of funds under management per year.

market trading, thereby lowering transaction costs. However, there is little academic research that investigates the effect of opportunity costs associated with altered trade timing.

While it is possible to simulate emulation funds ex-post based on historical trade flow (Chen, Gallagher, Foster and Lee, 2012), there are a number of limitations with this approach. The primary concern of this paper is to address the “black box” characteristic of simulation based analysis, which relies solely on historical data. By introducing a structured framework, we are able to decompose the process of fund emulation, which leads to further avenues of enquiry and a deeper understanding of emulation fund mechanics. Further, the model provides a benchmark for empirically simulated results, and enables us to identify and quantify otherwise unexpected sources of emulation fund volatility. The model also allows us to determine the proportion of trades that can be internally offset as a function of the delay period and the trade frequency of the underlying security. Finally, the model can be used to optimise the lag period between receiving a trade signal from the underlying fund and implementing that signal in the emulation fund.

We use the model to show that the potential costs and savings of an emulation fund are strongly related to the proportion of trades that can be internally crossed (the offset ratio). Further, the main determinants of this offset ratio are the chosen delay period and the frequency of trading of underlying stock holdings. We find that the model has the most explanatory power with respect to the most heavily traded stock groups: large-cap stocks, moderate growth and style neutral stocks, and stocks with relatively neutral past period historical returns. Incidentally, for the multi-fund manager data we used to simulate an emulation fund, we find that stock prices tend to immediately increase following buy trade signals and immediately decrease after sell signals. Hence, short timing delays in exploiting fund manager trade signals result in a significant opportunity cost, which in the majority of instances outweighs the commission and price impact savings from internal

crossings.² The model reflects our expectation that both the lag period and a security's expected trade frequency have a positive (but marginally diminishing) relation to a security's expected offset ratio, and this is confirmed by the empirical analysis. We also use the model to consider potential savings, relative to potential costs, to derive necessary conditions of an optimal lag period.

2. Background

Emulation funds are a new and largely unexplored form of multi-manager investment. They share risk-moderating and transaction cost-saving characteristics with other multi-manager portfolio structures, and through examining these alternative structures, we can gain some insight into the mechanisms that drive emulation fund costs and savings. However, because emulation funds have a unique system for altering the timing of trades, there remains a gap in understanding that this paper seeks to address.

The key advantage of multi-manager investment is diversification of manager-specific risk. Indeed, Lhabitant and Learned (2002) and Brands and Gallagher (2005) show that significant reductions in manager-specific risk can be achieved with five to ten different fund managers. Beyond this, adding more underlying funds may be detrimental to the skewness and kurtosis of portfolio returns (Brands and Gallagher, 2005) and may decrease potential alpha (Gallagher and Gardner, 2005). However, there has been little work on the level of trading redundancy that multi-manager frameworks exhibit.

There is also interest in determining whether multi-managers generate value beyond that created by their constituent managers. Emulation funds represent a plausible example of this approach, as they

² Note that this does not take into consideration the management fees paid to active fund managers, which is not covered in this study. In addition, tax implications are also beyond the scope of this research.

take the trade signals of the underlying managers to exogenously create capacity with much lower fees and potential transaction cost savings. Multi-managers charge additional fees that represent the cost of maintaining an additional level of management and reporting activities. Brown, Goetzmann and Liang (2004) find that within the fund-of-hedge-funds environment, individual hedge funds tend to dominate fund-of-funds in terms of both after-fee return and Sharpe ratio. However, Ang, Rhodes-Kropf and Zhao (2008) argue that the fees charged by funds of hedge funds are justified because the correct benchmark should be the direct returns of an uninformed investor in the underlying hedge fund pool, rather than the constituent funds. This takes into consideration the value of fund research services that multi-managers provide. Nevertheless, emulation funds are advantageous in this context as the fees paid to the underlying managers in using their signals are usually much lower than the fees paid for direct management³, and hence espouse a new dimension to cutting the costs of multi-manager investing.

Concerns about trading redundancy are a significant driver of research in centralised multi-manager investment strategies, and is a key motivator for the development of emulation funds. Redundant trades occur when independently managed constituent managers execute trades on the same underlying security, but on opposite sides of the market almost coincidentally. These trades incur transaction costs but may provide very little additional net return to the overall portfolio. Approaches designed to address this problem can be broadly classified into three categories: inventory funds, explicit forecast models, and paper portfolios.

³ From discussion with investment personnel, it appears that active management fees for direct funds management range between 30 and 70 basis points of funds under management per annum. It was suggested that active fund managers typically charged 10 basis points for the use of their trade data in emulation strategies.

An inventory fund acts as a 'buffer' between the trades of the individual managers and the market. Orders entered by the constituent managers are executed against the inventory fund, which periodically rebalances by routing orders to the market. As in emulation funds, offsetting trades within a predetermined period of time are netted. A variation of this is to structure the inventory fund so that it tracks a market index within a specified band and rebalances when active manager trades push its index tracking beyond an acceptable tolerance range. These market index funds essentially contribute a passive investment component to the overall portfolio, and may take the place of other passive market investments. In either case, inventory funds delay the execution of trade signals from the underlying managers. However, this delay is typically much longer than is the case for emulation funds. Ferguson (1978) presents a case against market index inventory funds by arguing that if managers are knowledgeable, the offsetting mechanism will remove many valuable trades. On the other hand, if managers are not knowledgeable, a simple passive index fund should outperform a portfolio of active managers net of fees. This argument is also pertinent to the case of emulation funds. Wagner and Zipkin (1978) examine real and simulated inventory funds using US data and show cost savings of 0.8% of assets under management over a six month period (assuming total transaction costs of 1.5%). However, transaction costs have fallen significantly since the 1970s and there appears to be a notable lack of recent research.

A seminal paper by Rosenberg (1977) proposes a more active approach to centrally managed multi-manager structures, and specifies that constituent managers provide explicit numerical forecasts on stock returns. A central fund manager then trades based on a consensus of these forecasts. At that time, the primary drawback of the model was that many fund managers were highly qualitative and thus could not readily provide detailed numerical data. Sharpe (1981) recognised these difficulties and discusses a number of other issues involved in constructing and managing a portfolio of funds. DiBartolomeo's (1999) extension on Rosenberg's (1977) model introduces a paper portfolio system

where individual fund managers submit their trades to the central manager, who then uses these trades to generate the numerical forecasts required for Rosenberg's model.

Both Rosenberg's (1977) and DiBartolomeo's (1999) approaches require considerable cooperation between the constituent fund managers and the central manager. However, this may not be feasible. In a commercial setting, fund managers are often reluctant to disclose alpha forecasts for both privacy and economically-sensitive (i.e. intellectual property) reasons. Elton and Gruber (2004) address this problem by assuming only partial information sharing between the underlying managers and the central decision maker — namely, that each fund manager will only share information about the portfolio, and not forecasts about individual securities. Nevertheless, these centralised portfolio approaches require a fundamental shift in information flows.

Emulation funds differ from Rosenberg's (1977) and DiBartolomeo's (1999) approaches in two ways. First, the abovementioned structures are designed to replace the existing multi-manager framework. In contrast, emulation funds merely provide additional capacity while maintaining the incumbent arrangements with the underlying fund managers. Second, emulation funds only use trade signals after the underlying fund manager has exploited them. Hence, an emulation fund does not require cooperation from the underlying fund managers beyond contractually attaining their permission to use their trade signals. Emulation funds represent a compromise between traditional decentralised multi-manager investing and the centralised investment structures proposed by Rosenberg (1977), DiBartolomeo (1999), and Elton and Gruber (2004). Emulation funds are a simplification of these past approaches: they extend investment volume without changing the existing management arrangements, and the consequences are purely incremental. However, concerns regarding the extent to which the anticipated performance drivers in emulation funds (i.e. potential transaction cost savings and opportunity costs of delayed trade timing) affect their overall

investment outcome remain unresolved. Our research attempts to fill this gap in the existing literature.

Chen et al. (2012) provides a case analysis of emulation funds. The authors simulate an emulation fund using trade level data and show that a hypothetical emulation fund on average underperforms its target fund. This is primarily due to the opportunity cost of delayed trading outweighing the benefits of reduced transaction costs. We build on the Chen et al. (2012) study in a number of ways. Our model enables us to develop performance expectations of an emulation fund, which can then be compared to realized performance. Hence, can link assumptions about trading signal patterns to performance and investigate potential sources of risk. In addition, our model makes transparent the mechanism through which emulation funds accrue differential returns to its tracking fund. This allows projections of expected costs and savings that are tied to beliefs about patterns in future share returns. Finally, we can use the model to analyse the sensitivity of emulation performance to key assumptions and parameters.

3. The Model

An emulation algorithm takes as inputs trade signals from the target fund managers. These specify the timing, side (i.e. to buy or sell) and volume for each trade signal. The internal logic of the emulation algorithm then computes the lags and offsets, and produces a sequence of trade signals to be acted on by the emulation fund. For simplicity, we assume that the emulation fund receives trade signals from the underlying fund managers at particular frequencies, represented by a uniform distribution (we find that this assumption does not significantly impact the modelled rate of offsetting trade signals). If acted upon, these trade signals become purchase or sell orders that change their own directly managed holdings of the emulation fund. Trade efficiencies associated with emulation funds may arise when conflicting trade signals (i.e. a buy signal and a sell signal)

arrive from different underlying target fund managers within a short period of each other — a period of time less than the selected delay period. When this happens, the conflicting signals are internally offset, so that only the net trade signal is executed by the emulation fund at the end of the selected delay period. Figure 1 provides an example of the offsetting process. The offset ratio (OR) describes the proportion of trade signals (weighted by the nominal value of each signal) that is internally offset in the emulation fund. Section 3.1.1 describes this process in greater detail.

The offset ratio (OR) influences a number of cost and saving drivers within the emulation structure. The four fundamental performance factors that account for the relative return difference between an emulation fund and its underlying tracking fund are: Brokerage commissions (C), price impact (PI), the opportunity costs associated with internal crossing (X) and those associated with delayed execution (D). The model represents the interaction between these factors and an offset ratio (OR) parameter, which refers to the portion of the order that is internally crossed. We expect reductions in commission and price impact, as well as the opportunity costs of internally crossing trades (which reduces the crystallisation of trade gains or losses) to be positively associated with the OR while the opportunity cost of delayed execution is negatively associated. The model is specified on a per-security basis.

The expected profit ($\pi_{i,m,L}$) from an emulation strategy, implemented on a particular security, isolated for a particular market side, and using a given lag period is:

$$\begin{aligned} E(\pi_{i,m,L}) = & E(OR_{i,m,L}) \left(C_{i,m} + E(PI_{i,m}) - E(X_{i,m,L}) \right) \\ & - \left(1 - E(OR_{i,m,L}) \right) E(D_{i,m,L}) \end{aligned} \tag{1}$$

Here, i is a unique security identifier, m denotes a specific market side (either buys or sells) and L denotes the lag period. Hence, the profit in the strategy arises from the proportion of total transaction costs $(C_{i,m} + E(PI_{i,m}))$ that can be internally offset $(E(OR_{i,m,L}))$, minus the proportion of all crystallised gains $(E(X_{i,m,L}))$ that are eliminated with internal offsetting, and the opportunity cost of executing the proportion of trade volume that is not offset $(1 - E(OR_{i,m,L}))$ and executed on a delayed basis $(E(D_{i,m,L}))$.

Expression (1) reflects our expectation that as more trade signals are internally offset (i.e. $OR_{i,m,L}$ increases), we realise a greater proportion of total potential reductions in commission ($C_{i,m}$) and price impact ($PI_{i,m}$). At the same time, the opportunity cost of internally offsetting crossed trades ($X_{i,m,L}$) at an intermediate price, rather than executing at differing market prices, will also increase. Conversely, the $OR_{i,m,L}$ negatively affects the delayed execution opportunity cost ($D_{i,m,L}$) because a smaller proportion of total trade signals require delayed on-market execution if more of these signals are internally offset. We explore each of these parameters in further detail in the following sections.

3.1 Offset Ratio

The offset ratio $OR_{i,m,L}$ is the proportion of trade signal volume on security i and market-side m that can be eliminated for on-market trading through an L -day delay period. If the timing and volume of each trade is known explicitly (e.g. ex-post analysis), this ratio can be computed through an iterative application of the offset function. That is, for trade signal x in a security-specific trade sequence, we identify the specific percentage volume of the signal that can be offset:

$$OR_{i,m,L_x} = \frac{\min(V_x, \sum_{y=1}^{N_{L_x}} V_{x+y}^-)}{V_x} \quad (2)$$

V_x	Trade volume of trade signal x
V_{x+y}^-	Trade volume of a transaction signal subsequent to x , on the opposite market side and executed by a different manager
N_{L_x}	The number of same security trades on the opposite market side within transaction x 's offset window.

In Equation (2), the left hand side denotes the offset ratio associated with x , a trade signal on security i , market side m and within an emulation framework with lag period L . The numerator in the right hand side is the entire volume of trade signal x if it is completely offset by subsequent crossed volume, or as much volume as can be crossed within the L day lag window. This is expressed as a proportion of the entire volume of x . Note that in practical application, an iterative procedure is applied to prevent double offsetting.⁴ Thus, the V_{x+y}^- term is updated to the post-offset volume after each offset event.

Where the expected trade frequency is known but individual trades cannot be inferred, we use an algebraic formulation of $OR_{i,m,L}$ incorporating the expected trade frequency and the lag period in the emulation structure. This has the advantage of being a continuous differentiable function with respect to key emulation parameters. Hence, the model is not restricted by the availability of historical data and can be used in scenario analysis. The offset ratio function gives the proportion of trade signal volume that can be crossed, which depends on (a) the collision rate $CR_{i,m,L}$ (the

⁴ For example, in a buy-sell-buy trade sequence of equal volumes, we must ensure that the sell does not offset against both buys, as this would incorrectly inflate our offset ratio.

probability that a buy (sell) trade signal will occur within L days of another manager's sell (buy) trade signal on the same security), and (b) the overlap ratio VR (the expected proportion of trading that we expect to offset if two trades are within L days of each other).

We need to make a number of assumptions about the nature of the projected trade signals on market side m . In the simplest case, we assume that the timing of all trades on the opposite market side m^- follow a uniform distribution \mathcal{U}_i , both buy and sell side trade signals lie in the same range (i.e. the domain of \mathcal{U}_i) and all funds that buy security s also sell it, and vice versa. Hence, the probability of a trade signal lying within L days of subsequent offsetting signal is expressed as:

$$E(CR_{i,m,L}) = 1 - \lim_{P \rightarrow \infty} \left(\frac{P-L}{P} \right)^{2 \cdot E(tf) \cdot P \cdot \left(\frac{f-1}{f} \right)} \quad (3)$$

$CR_{i,m,L}$	The collision rate of signals for security i on the market side m
P	Period over which trade frequency is measured
L	Trade delay period
tf	Trade frequency on the opposite market side, expressed as the number of trade signals over period P
f	Total number of funds that traded in security i .

Appendix 2 provides the formal derivation of expression (3). Note that in this specification, the collision rate is not static, given an expected trade frequency tf , but is asymptotic to the true collision rate for large values of P .

In practice, there are a number of effects that a uniformly distributed pattern of trade signal arrival does not take into consideration. First, institutional trades may be auto-correlated if fund managers trade on short-lived information or herd (Hirshleifer, Subrahmanyam and Titman, 1994, Kothari and Warner, 2001, Brown, Wei and Wermers, 2007). Hence, institutional trade imbalance⁵ is expected to be biased towards the traded side (i.e. positive institutional trade imbalance following institutional buying and negative institutional trade imbalance following institutional selling) in the period following the trade. Second, total institutional investor trading activity in general may be higher in the period following other institutional trade, leading to non-uniformity in absolute trade frequency. Lastly, in securities where the trade frequency is high, one trade signal may offset against multiple subsequent signals on the opposite market side. We do not take multiple-signal offsetting into consideration in our basic model, but we do investigate how this might cause divergences between the model and the simulated results in section 3.2.

Once we have modelled the probability of two signals being crossed, we now estimate the expected proportion of trade volume that overlaps when two trade signals are crossed. This is twice the overlapping volume divided by the total volume of the two trade signals:

$$VR = \frac{2 \cdot \min(V_1, V_2)}{V_1 + V_2} \quad (4)$$

The overlap ratio depends on the distribution of typical trade sizes. Kyle and Obizhaeva (2011) note that after adjusting for trade activity, order size distributions across stocks in different volume and volatility groups closely resemble a log-normal. In this study, we will not calculate the order

⁵ Defined by $\frac{V_b - V_s}{V_b + V_s}$ where V_b is the volume of buys executed by mutual funds and V_s is the volume of sells executed by fund managers on any particular day.

statistics of lognormal distributions as this is a non-trivial problem⁶. Instead, we use repeated two-signal sampling of a 1-year segment of our data to establish an approximate overlap ratio (VR) of 0.5 (i.e. given two trades drawn at random, the smaller volume is half the size of the greater volume). We acknowledge that this is parametrically insensitive to variations in the mean and standard deviation of trade volume distributions across securities but, as shown in section 3.2, the overall overlap ratio does not seem to vary significantly with time.

Combining the collision rate with the overlap ratio, the expected offset ratio is given as:

$$E(OR_{i,m,L}) = 0.5 \cdot \left(1 - \lim_{p \rightarrow \infty} \left(\frac{P-L}{P} \right)^{2 \cdot E(tf) \cdot P \cdot \left(\frac{f-1}{f} \right)} \right) \quad (5)$$

Figure 3 illustrates model forecasts of offset ratios as a function of hypothetical (but realistic) values of the delay period (L) and expected trade frequency ($E(tf)$). This demonstrates the need for longer (shorter) delay periods for less (more) frequently traded stocks to achieve a substantial internal crossing. We assume the number of underlying funds (f) supplying trade signals to the emulation fund is constant. Since we do not account for a signal offsetting against multiple signals on the opposite market side (i.e. one large buy trade offsets against multiple smaller sell signals), we expect some model prediction error of the offset ratio for larger values of lag days and higher expected trade densities. We provide an example of how the model can be used to infer the $OR_{s,m,L}$, given the expected number of trades on either market side over a hypothetical period of trading and specified lag period.

⁶ For example, see Nadarajah, S. 2008. Explicit Expressions for Moments of Log Normal Order Statistics. *Economic Quality Control*, 23, 267-279.

Assume in the coming 100-day period that we expect 20 buy side trade signals and 25 sell side trade signals from four separate fund managers about security x . Let the lag period be 5 days. The buy side offset ratio, computed by using $L = 5$, $n = 25$, $P = 100$, is:

$$E(OR_{s, \text{buy}, 5}) = 0.5 \cdot \left(1 - \lim_{p \rightarrow \infty} \left(\frac{P-5}{P} \right)^{2 \cdot \frac{25}{100} \cdot P \cdot \left(\frac{4-1}{4} \right)} \right) = 0.423$$

Similarly, the offset ratio for sell side signals, where we use $n_{\bar{m}} = 20$, is:

$$E(OR_{s, \text{sell}, 5}) = 0.5 \cdot \left(1 - \lim_{p \rightarrow \infty} \left(\frac{P-5}{P} \right)^{2 \cdot \frac{20}{100} \cdot P \cdot \left(\frac{4-1}{4} \right)} \right) = 0.388$$

In this example, the modelled offset ratios on the buy and sell side are 42.3% and 38.8% respectively. Hence, using Equation 5, we can deduce an expected level of offsetting from the trade frequency, lag period, and the number of funds contributing trade signals to the emulation portfolio.

3.2 Brokerage Commission

Brokerage costs are usually computed as a fixed percentage of trade value, though this may vary between security classes and brokers. Where the commission rate is not explicitly known, it may be proxied by weighted average commission rates estimated from historical data. The model assumes that the commission structure is proportional.

3.3 Price Impact

We use prior-day close prices to benchmark trade prices and determine historical price impact. This is used as an indication of the expected price impact on trades. The price impact calculation follows

from Comerton-Forde et al. (2005), Chiyachantana et al. (2004), Keim and Madhavan (1996) and Chan and Lakonishok (1995). Chiyachantana et al. (2004) refers to this as the decision price measure and uses an approximation of the market price when the trade decision is made by the fund manager. The expected price impact on trade signal x is a hypothetical price impact should it be executed on-market.

The expected price impact of trade signal x is calculated as a weighted aggregate of the individual price impacts of the constituent trades within the underlying trade package:

$$E(PI_x) = \begin{cases} \ln \left(\frac{\sum_{x' \in x} w_{x'} P_{x'}}{P_x \sum_{x' \in x} w_{x'}} \right) + \varepsilon & \text{for buy trade signals} \\ -\ln \left(\frac{\sum_{x' \in x} w_{x'} P_{x'}}{P_x \sum_{x' \in x} w_{x'}} \right) + \varepsilon & \text{for sell trade signals} \end{cases} \quad (6)$$

- $w_{x'}$ Trade value (price \times volume) of individual trade x' within the trade package
- $P_{x'}$ Trade price of individual trade x' within the trade package x
- P_x Closing price on day prior to initiating trade day.

The numerator within the natural log function represents the total cash flow (less commissions) involved in the trade package (if we assume trade signal volume as the weighting factor), while the denominator is the hypothetical cost of the package if it had been fully executed at the decision price. Our approach differs slightly from the original Chiyachantana et al. (2004) measure in that we do not subtract the market price movement in the trade package period. This was done to fully preserve the differential returns of the emulation fund relative to the tracking fund. The previous day close price was also chosen as the benchmark price since both our performance metrics (close-to-close returns and characteristics-based alpha) are based on closing prices.

In an ex-ante context, the expected price impact on a hypothetical trade signal is taken as the expected price impact (weighted by trade value) of historic on-market trading in the stock on the particular market side:

$$E(PI_{i,m}) = \frac{\sum_{x \in i,m} w_x E(PI_x)}{\sum_{x \in i,m} w_x} \quad (7)$$

3.4 Internal Crossing Opportunity Cost

When crossed trade signals are offset in the emulation fund, we forgo the crystallisation of gains or losses that arise through execution at differing market prices. The internal crossing opportunity cost ($X_{i,m,L}$) is therefore the opportunity cost of eliminating exposure to market price shifts, and is measured as the log difference in the benchmark prices on buy and sell signal days. We use a benchmark price (in this case, the previous-day close price) rather than the actual trade price since the price impact is computed separately.

If we assume that fund managers have short-term trade timing ability, then post-trade prices on securities will move favourably in the direction of the trade (i.e. prices will increase after institutional purchases and decrease after institutional sales). While individual price paths following trading are uncertain, we consider a “characteristic” price path function $\rho_{i,m}(t)$ that incorporates the ability of managers to anticipate favourable stock price movements. Empirically, this is defined by the trade-value weighted mean price movement following trading on a particular stock (Equation 8).

$$\rho_{i,m}(t) = \frac{\sum_{x \in i,m} w_x \left(\frac{P_{x+t} - P_x}{P_x} \right)}{\sum_{x \in i,m} w_x} \quad (8)$$

If we assume, as in Equation (3), that trade signals on the opposite market side are uniformly distributed following an initial institutional trade signal, then the expected price movement between the initial trade and a subsequent offsetting trade within L lag days is given by:

$$X_{i,m,L} = \frac{\sum_{t \in [1,L]} \rho_{i,m}(t)}{L} \quad (9)$$

Note that this equation is unique to the market side of the initial trade (m), and is equivalent to the average return on a security in the L days after that security has been traded.

3.5 Delayed Execution Opportunity Costs

An inherent feature of emulation funds is the requirement to delay the exploitation of trade signals by a predetermined number of days. This is to ensure that on-market trading performed by the emulation fund does not compete with the underlying active funds for market volume. When a trade cannot be fully offset within the delay period, they must be executed at a subsequent market price to ensure tracking of the target fund. This leads to an opportunity cost in altering the timing of the trade, given by $\rho_{i,buy}(L)$ for buy side signals and $-\rho_{i,sell}(L)$ for sell side signals. When the performance of buy-side signals is aggregated with the performance of sell-side signals in the emulation fund, the absolute values of $\rho_{i,buy}(L)$ and $\rho_{i,sell}(L)$ are not important for our analysis. Rather, it is the difference between them (i.e. $\rho_{i,buy}(L) - \rho_{i,sell}(L)$) which is of main significance. If this value is positive from 0, then we expect greater opportunity costs associated with delayed trade execution. If this value is negative, then delayed trade execution would correct fund manager

mistiming. Convergence to 0 in the long-term would indicate that fund managers neither add nor subtract long-term value.

4. Data, Calibration and Simulations

The aim of this section is to assess the accuracy of the model and demonstrate its strengths and weaknesses across different types of stocks. Because expectations data are not readily available, we use observed historical trading data as a proxy for expectations:

$$E(OR_{s,m,L}) = \overline{OR}_{s,m,L} + \varepsilon_1 \quad (10)$$

$$E(PI_{s,m,L}) = \overline{PI}_{s,m,L} + \varepsilon_2 \quad (11)$$

$$E(X_{s,m,L}) = \overline{X}_{s,m,L} + \varepsilon_3 \quad (12)$$

$$E(D_{s,m,L}) = \overline{D}_{s,m,L} + \varepsilon_4 \quad (13)$$

This gives an empirical form of:

$$\begin{aligned} E(\pi_{s,m,L}) = & (\overline{OR}_{s,m,L} + \varepsilon_1) \left(C_{s,m,L} + (\overline{PI}_{s,m,L} + \varepsilon_2) + (\overline{X}_{s,m,L} + \varepsilon_3) \right) \\ & + \left(1 - (\overline{OR}_{s,m,L} + \varepsilon_1) \right) (\overline{D}_{s,m,L} + \varepsilon_4) \end{aligned} \quad (14)$$

In the following sections, we generate model estimates for each of the performance drivers using our sample of multi-manager data, and compare these empirical results to those computed by a simulated emulation fund that uses actual trades as hypothetical trade signals. The simulation algorithm iteratively applies Equation (2) to compute the specific offset ratio on individual offsetting events, and uses this to determine commission and price impact savings. Both of these factors are assumed to, on average, scale linearly with reductions in on-market traded values on an

individual trade signal basis. The opportunity cost of internal crossing is determined as the sell signal price minus the buy signal price – by offsetting these trades, the emulation portfolio foregoes crystallization of any gains or losses generated by trading two opposed signals on market at different prices. Trade signals that are not fully offset are executed at a delayed point in time; the difference between benchmark prices on the original signal date and the hypothetical delayed execution date are used to compute the delayed execution cost. The model provides projected estimates specific to each security and market side; we present these in an aggregated context to provide an overall perspective on the model’s performance. For robustness, we compare the accuracy of the model results to the simulation on a variety of security sub-portfolios partitioned by size, style and return characteristics.

4.1 Data

The data used to compare the model with a simulated emulation fund comes from the Australian Equities component of a major Australian superannuation fund. Contained within the data set are daily aggregated transactions that include the trade price, volume, fund manager identifier and broker identifier over the period 2005 to 2009 (inclusive). The multi-fund portfolio includes all major fund styles, and no emulation strategy is currently utilised by the superannuation fund. Statistics for this data set are summarised in Figure 4. This institutional trades data set is supplemented by stock price data from the SIRCA Australian Equities Tick History (AETH) database, market capitalisation and dividend payment data from the SIRCA Share Price and Price Relative (SPPR) database, and earnings data from Aspect Huntley.

4.2 Simulation Method

The simulation is executed with an algorithm that iterates through a list of trade signals from the tracking portfolio fund managers (i.e. their daily trades). The algorithm delays these trade signals

for a specified lag period, over which time the signal may be partially or fully offset against opposing signals from other constituent funds. Outstanding volume at the end of the lag window is executed on the following day at an inferred market price.

4.2.2 Transaction Cost Savings

The two transaction costs we look at are brokerage and price impact. Brokerage commissions are known explicitly from the daily trade data – hence, commission savings are calculated pro rata according to the offset ratio. The price impact measure we use follows from the widely used open-to-trade method in prior literature (e.g. Comerton-Forde et al. (2005), Chiyachantana et al. (2004) and Keim and Madhavan (1996)). This is defined as:

$$PI_{\text{open}} = \sum_i^N w_i \left(\frac{P_i - OP_1}{OP_1} - \frac{M_i - M_1}{M_1} \right) \quad (15)$$

Trade packages have been deconstructed into their N constituent trades. w_i is the volume weighting of each trade in the package, P_i is the trade price of trade i in the package, OP_1 is the opening on the first day of the trade package, and M_i is the value of the market index on the day of trade i , M_1 is the value of the market index on the day of the first trade in the package. The simulation uses the underlying assumption that price impact scales down linearly with the offset ratio.

4.2.3 Timing-related Opportunity Costs

There are two opportunity costs associated with altering the timing of trades in the emulation fund relative to the tracking portfolio. When two trade signals on the same underlying security are offset in the emulation fund, we forgo the crystallisation of either a gain or a loss from the difference in

their on-market trade prices. Where a gain was made in the tracking portfolio, each unit of volume that is offset incurs an opportunity cost associated with internal crossing (x) equivalent to:

$$x_i = x_j = \frac{P_b - P_s}{2} \quad (16)$$

P_b is the benchmark price of the day of the buy order and P_s is the benchmark price on the day of the sell order. The total cost of internal crossing attributable to each of the crossed trades is:

$$X_i = X_j = x_i \cdot \text{argmin}(V_i, V_j) \quad (17)$$

Here, $\text{argmin}(V_i, V_j)$ represents the crossed volume between two opposing trade signals with volumes V_i and V_j .

The second opportunity cost arises from the delayed execution of trade signals from the tracking portfolio that have not been internally offset in the emulation fund. These result from the price movement in the lag period between when the trade is executed in the tracking portfolio and when it is emulated. The simulated delayed execution opportunity cost (D) is hence:

$$D_i = \begin{cases} P_{i+n} - P_i & \text{if buy} \\ P_i - P_{i+n} & \text{if sell} \end{cases} \quad (18)$$

4.3 Offset Ratio Modelling

We first examine the model's projected offset ratio compared to the measured offset ratio in the simulated fund. The model is calibrated on the same data as those used for the simulation, but uses

only aggregated statistics concerning the trades (i.e. trade frequency) while the simulation processes every single trade signal. Figure 5 shows a comparison between the offset ratio (OR_L) predicted by the model and that derived through simulation on the same data over 5, 10 and 20 days. The model appears to be quite accurate over the shorter delay windows, but underestimates the projected OR_L more significantly over 20 days. There are two short-term deviations from our assumption about normally distributed trade signals that could possibly affect the projected offset ratio: the skewed trade imbalance following institutional trades and the elevated levels of institutional buying and selling. While we do observe these effects,⁷ the fact that the model-projected offset value over 5 days closely matches the offset value observed through simulation indicates that the rate of offsetting is not greatly affected by the combination of these two factors.

At longer delay periods, the model underestimates the simulated offset ratio more substantially. We believe this is primarily caused by the model ignoring crossing events where a trade signal offsets against multiple signals on the opposite market side. As expected, this effect occurs more commonly as the lag period (L) increases. Overall, the model's offset ratio projections appear to be very accurate for emulation strategies with a 5-day lag period and do a reasonably good job of modelling OR_L for 10-day lags. Over longer lag windows, a more sophisticated model of multiple trade offsetting needs to be considered. However, since current commercial emulation strategies primarily operate with short 5 to 10 day lags, our simple 2-trade crossing model appears to be sufficient in this context.

⁷ In the 5 days after a fund manager trades, the institutional buy (sell) side trade imbalance, excluding other trades by that manager, is 7.71% (5.55%) away from the long-term equilibrium imbalance in the buy (sell) direction. In the same 5 day period, institutional sell (buy) activity decreased (increased) by 0.62% (4.98%) following an institutional buy (sell).

4.4 Performance Drivers Modelling

Next, we compare the model predictions for each of the performance drivers against the results from our simulated emulation fund. Panels A, B and C of Figure 6 show a side-by-side comparison of each of the four performance drivers for 5, 10 and 20 lag days. The modelled commission savings turn out to be slightly lower than the true savings, which reflects the model's underestimation of the simulated offset ratio. Conversely, the model significantly overestimates both the price impact savings and the internal crossing opportunity costs. The former observation suggests that in the simulated emulation fund, trade signals with relatively low price impact are more likely to be internally offset. This may be due to the presence of contrarian trading— e.g. if a fund manager buys when other managers are selling, then those buy signals would incur relatively low or even negative price impacts. The emulated signals would also have a greater probability of being offset, by definition, since there is elevated institutional supply of that stock. This view is supported by the presence of both growth and value managers in the data sample. The latter observation, that the modelled internal crossing opportunity costs are greater than those seen in the simulation, in conjunction with the observed tendency for $\rho_{i,buy}(t) - \rho_{i,sell}(t) > 0$ for small values of t , suggests that offsetting events tend to occur close to each other rather than spread out through the lag period. Finally, the delayed execution opportunity cost appears to be underestimated by the model over 5 days, closely estimated over 10 days and overestimated over 20 days. Since the modelled offset ratio underestimates the simulated offset ratio, this implies that with a corrected OR, the model would generally tend to slightly underestimate the true delayed execution opportunity cost. This suggests that stocks that are less likely to be offset (e.g. small-cap stocks with lower trade densities) incur greater than expected opportunity costs associated with delayed execution of trade signals, and confirms previous observations that fund managers exhibit greater stock selection ability in small stocks (e.g. Chen et al. (2010)). We also observe that opportunity costs associated

with delayed trading seem to decrease with longer lag periods. This implies that $\rho_{i,buy}(t) - \rho_{i,sell}(t)$ converges for larger values of t , and indicates that longer lag periods would result in a more profitable emulation strategy.

The model demonstrates that the opportunity costs of delayed trading are at least as significant as the transaction cost savings. While trading efficiency and transaction cost reduction are often touted as the selling points of an emulation fund, the actual cost savings from reduced turnover is bounded by the maximum value of market impact and brokerage. On the other hand, both the opportunity cost of internal crossing and the opportunity cost of delayed execution are dependent on the structure of $\rho_{i,buy}(t)$ and $\rho_{i,sell}(t)$, which are potentially unbounded. Empirically, the effects of these post-trade price movement functions are much more significant than the transaction cost savings and, at least over short lag periods, adverse to overall performance.

In general, the model shows some discrepancy in predicted and actual factor performance as computed by the simulated emulation algorithm, particularly in price impact reduction and internal offsetting opportunity cost. However, the model effectively projects both commission savings and the delayed execution opportunity cost (which has the most significant influence on overall emulation performance).

4.5 Model Robustness

We test the model against simulated emulation sub-portfolios constructed along market capitalisation, book-to-market and historical returns characteristics. Results are summarised in Figure 7. For 5 and 10-day lags, the model forecasts within 0.30% of simulated outcomes for all stocks ranked between 1 and 300 in size (Figure 7, Panel A). Stocks ranked below 300 exhibit large

errors in the forecasted performance relative to simulation, and are not modelled reliably. Over a 20-day lag, the model forecasts simulated emulation performance within 0.15% for stocks ranked down to 200 in size. However, accuracy significantly drops for smaller stocks (forecasting errors of 1.27% and 1.68% for stocks ranked in 201 to 300 and 300+ respectively). Along the value-growth dimension, the model appears to be more accurate forecasting emulation performance on stocks in the high market-to-book (i.e. growth) end of the spectrum compared to the low market-to-book (i.e. value) end (Figure 7, Panel B). We also note that over the 20-day lag period, the model tends to underestimate performance, relative to the simulation for growth stocks, and overestimate performance (and much more significantly so) for value stocks. The model is also less effective for forecasting performance in stocks with very negative prior-year returns, and to a lesser extent, those with very positive prior-year returns (Figure 7, Panel C). A comprehensive breakdown of the model's predictions against those simulated directly from the data, along with analyses across market capitalization, book-to-market ratio and momentum categories, can be found in an online supplement.⁸

The forecasting accuracy of the model is also a function of the lag length employed. There appears to be a lower margin of error in the offset ratio prediction (Figure 5) for shorter lags compared to longer ones. This may be a product of the model's implicit assumption that offsetting only occurs in independent, opposed signal pairs (i.e. a buy and a sell). In reality, the offsetting process happens continuously where one signal (e.g. a buy) may be offset against multiple opposing signals (e.g. several sells). As the lag period is increased, the frequency of these complex, multi-signal interactions also increases, which leads to the model underestimating the simulated offset ratio. However, the model's overall predictive accuracy actually increases as the lag period lengthens up to 20 days. This is a product of the model producing greater underestimation bias on commission

⁸ https://dl.dropboxusercontent.com/u/37092/ChenGallagherFosterLee2012_Supp.docx.

savings and internal crossing opportunity costs as the lag period lengthens, but an opposite effect in the errors associated with the delayed execution opportunity cost – the model underestimates this cost with 5 and 10 days but overestimates it with the 20-day lag. The net result is that errors cancel each other out with the 20-day lag window. The errors between the model and the simulated emulation fund are 0.08%, 0.07% and 0% for 5, 10 and 20 lag days respectively. We suggest that this model should be applied with caution for emulation portfolios with lag periods longer than 20-days, since we expect the error on the predicted offset ratio to be exacerbated by the longer lag.

5. Model Applications

Having established that the model provides a fair representation of the differential performance that an emulation fund generates relative to its tracking fund (in the stocks that constitute the majority of trading), we look at applications of this model to a range of scenarios. In section 5.1, we assume that the short-term price evolution of a stock subsequent to purchase is not significantly different from the price movement of a stock subsequent to sale. Hence, we only look at savings generated by the internal crossing mechanism. In section 5.2, we add a price movement model based on empirical observations about how prices evolve subsequent to being traded by mutual fund managers.

5.1 Offset Ratio Optimisation

We first assume that purchased stocks exhibit short-term price movement patterns that are not significantly different to sold stocks after they are traded by mutual fund investors. This is generally the assumption implicit in the justification of commercial emulation products. Under this condition, the expected value of timing opportunity costs (i.e. $E(X_{i,m,L})$ and $E(D_{i,m,L})$) is zero. Equation (1) hence simplifies to:

$$E(\pi_{i,m,L}) = E(OR_{i,m,L}) (C_{i,m} + E(PI_{i,m})) \quad (19)$$

That is, the profit from the emulation strategy arises purely as a linear proportion of the offset ratio. The offset ratio, in turn, is a function of the predetermined lag period and the expected trade frequency (equation (5)). Partial derivatives of equation (5) with respect to the lag period and the expected trade frequency suggest that the expected payoff from emulation increases with longer lags and higher trade densities, but these exhibit rapidly diminishing marginal returns.

The lag period parameter can be directly modified by the emulation fund provider, and hence determining the sensitivity of the offset ratio, $E(OR_{i,m,L})$, to changes in the lag period, L , is of prime importance (Equation (20)).

$$\frac{\partial E(OR_{i,m,L})}{\partial L} = \lim_{P \rightarrow \infty} \left(2 \cdot E(tf) \cdot \left(\frac{f-1}{f} \right) \cdot \left(\frac{P-L}{P} \right)^{2 \cdot E(tf) \cdot P \cdot \left(\frac{f-1}{f} \right) - 1} \right) \quad (20)$$

We note that $\frac{\partial E(OR_{i,m,L})}{\partial L}$ is strictly positive – hence the offset ratio (and marginal profit from emulation) monotonically increases with the lag period. However, Figure 8 Panel A shows that the marginal benefit of increasing the lag rapidly diminishes with the size of the lag, particularly for stocks with high trade frequency.

A similar pattern occurs when we examine the partial derivative of $E(OR_{i,m,L})$ with respect to $E(tf)$ (Equation (21)).

$$\frac{\partial E(OR_{i,m,L})}{\partial E(tf)} = -1 \cdot \lim_{P \rightarrow \infty} \left(P \cdot \left(\frac{f-1}{f} \right) \cdot \log \left(\frac{P-L}{P} \right) \cdot \left(\frac{P-L}{P} \right)^{2 \cdot E(tf) \cdot P \cdot \left(\frac{f-1}{f} \right)} \right) \quad (21)$$

Again, the function is monotonically increasing with respect to the expected trade frequency, but marginal improvements to the expected offset ratio rapidly diminish with a larger flow of trades. While multi-fund managers cannot directly control the trade volume of each constituent fund, the issue of trade frequency becomes relevant when the multi-fund manager decides to include additional funds in the investment strategy. The inclusion of additional underlying funds not only increases the expected trade frequency, but also the value of f – the number of funds. However, the effect of increasing f is deemed to be insignificant for most large multi-fund managers since $\frac{f-1}{f}$ diminishes rapidly as f increases. In the base case where the post-trade performance of buys and sells are homogenous, increasing the lag period or adding additional fund managers to the tracking portfolio always improves the return of the emulation fund relative to the tracking portfolio.

5.2 Lag Period Optimisation with Heterogeneous Post-Trade Price Movement

The simulation shows that the characteristic price movement following fund manager buys is different from that following sells. Historical data suggests that mutual funds tend to buy stocks that outperform sold stocks in the short term subsequent to trading. To model this, we introduce the post-signal price movement function ($\rho_{i,m}(L)$), which describes the evolution of stock prices immediately after they are traded by a mutual fund. We can substitute $\frac{\rho_{i,m}(L)}{L}$ for $E(X_{i,m,L})$ if we assume that offsetting trades are uniformly distributed across time, and we can also substitute $\rho_{i,m}(L)$ for $E(D_{i,m,L})$, since delayed execution occurs at the end of the lag period. Hence, the marginal payoff of emulation becomes:

$$\begin{aligned}
& E(\pi_{i,m,L}) \\
&= \begin{cases} E(OR_{i,m,L}) \left(C_{i,m} + E(PI_{i,m}) - \frac{P_{i,m}(L)}{L} \right) - (1 - E(OR_{i,m,L})) \rho_{i,m}(L) & \text{if } m \text{ is buy} \\ E(OR_{i,m,L}) \left(C_{i,m} + E(PI_{i,m}) + \frac{P_{i,m}(L)}{L} \right) + (1 - E(OR_{i,m,L})) \rho_{i,m}(L) & \text{if } m \text{ is sell} \end{cases} \quad (22)
\end{aligned}$$

In the context of maximising relative returns in the emulation fund, the optimal lag period is specific to each security and depends on the expected trade frequency of that security and the structure of its post-signal price movement function ($\rho_{i,m}(L)$). Combining the buy and sell sides, the payoff from emulating a particular stock i is given by:

$$\begin{aligned}
E(\pi_{i,L}) &= E(OR_{i,L}) \left(2C_i + E(PI_{i,buy}) + E(PI_{i,sell}) - \frac{P_{i,sell}(L) - P_{i,buy}(L)}{L} \right) \\
&\quad + (1 - E(OR_{i,L})) (\rho_{i,sell}(L) - \rho_{i,buy}(L)) \quad (23)
\end{aligned}$$

Here, we assume that both the volume of trading and the brokerage commissions are symmetric between the buys and sells. Note that $P_{i,m}(L)$ is the integral of $\rho_{i,m}(L)$ with respect to L . In the context of emulation portfolio optimisation, the ideal lag period depends on the characteristic price movement following mutual fund buys compared to that following sells. Figure 9 provides an application of Equation (23) on the “average” stock (with weighted mean characteristics of all stocks). The underlying data suggests that buys outperform sells in the short term, but this effect reverses between 50 and 100 days post-trade. A proposed emulation portfolio for this particular stock is therefore expected to generate negative marginal returns using a short lag period, but positive marginal returns with a very long lag period (in our example 100 trade days). Given the large degree of variation occurs between the post-trade price movements of different stocks, the

model would be useful in determining whether a long lag is optimal (as in the example provided), a short lag is optimal (e.g. if buys tend to outperform sells), or some intermediate lag period (e.g. if the buy-minus-sell returns difference initially diverges after trade and then subsequently converges).

We note that it may also be desirable to minimise the tracking error between the emulation fund and the target portfolio; this would make shorter lag periods more attractive. The idiosyncratic risk tolerance of the multi-manager fund manager is not incorporated into our model.

6. Conclusions

Our model formalises the interaction between potential transaction cost savings and the opportunity costs associated with internal offsetting and delayed trade timing. This provides both predictive and explanatory advantages over simulation results. From a forecasting point of view, we can estimate the reduction in turnover based purely on expected trade signal frequency and a predefined lag period. This, in turn, allows us to estimate performance of a proposed emulation fund based on historical observations about transaction costs and post-trade price movements. The model — in conjunction with simulation — enables us to evaluate hypotheses about the underlying statistical processes that drive emulation funds, and hence reveal where assumptions may break down.

Importantly, the model highlights aspects of the emulation process which we may not have thought to investigate in a pure simulation setting. For example, the observed price impact saving determined in simulation is much lower than that predicted by the model — indicating that low price impact trade signals are more likely to be internally offset in the emulation fund. Merely observing the level of price impact saving itself would not have allowed us to reach a similar conclusion. Further, having the model as a benchmark enables us to determine that infrequently traded stocks have a greater than expected opportunity cost to delayed trading, and hence are likely

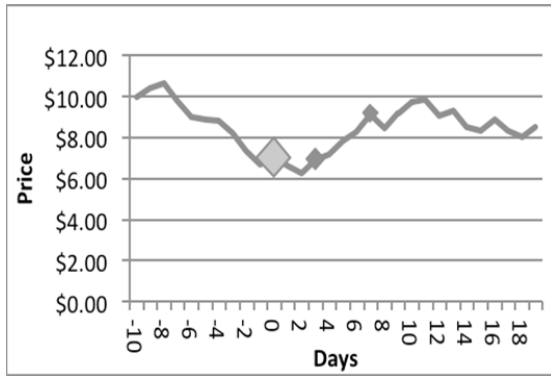
to be more skilfully traded. This suggests that these stocks should be emulated with shorter lags or excluded from the emulation fund altogether.

When tested against simulation, we show that the model does reasonably well at forecasting both the offset ratio and overall performance for shorter lag periods (i.e. less than 10 lag days). Hence, by examining the functional representation of the model, we can actually determine the sensitivity of the offset ratio to changes in lag period and trade frequency. In particular, we observe that the marginal benefit to increasing the lag period beyond 7 days is extremely low, even for securities with very low trade signal frequencies in our data sample. From a practitioner's point of view, this allows us to compare the potential opportunity costs determined by the post-trade price movement function (i.e. $\rho_{i,m}(L)$) to the potential savings arising from higher levels of internal trade signal offsetting. Where the post-trade price adjustment functions of individual stocks can be reasonably estimated, we can also use the model to infer the optimal lag structure of the emulation fund on a stock-by-stock basis.

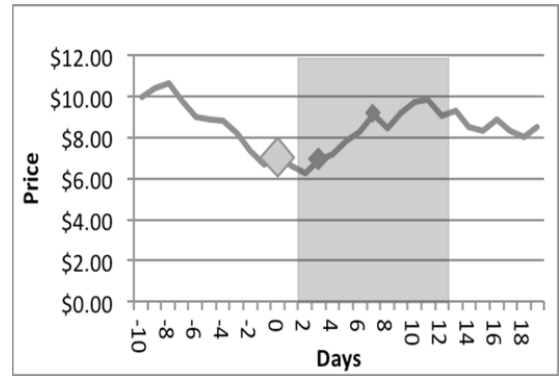
Finally, we note that emulation outcomes also depend on factors that have not been addressed here. These include the tax consequences of administering the emulation fund, active management fees, and implementation issues such as brokerage, risk and cash flow management in the emulation fund. These topics would provide rich avenues for future research.

Appendices

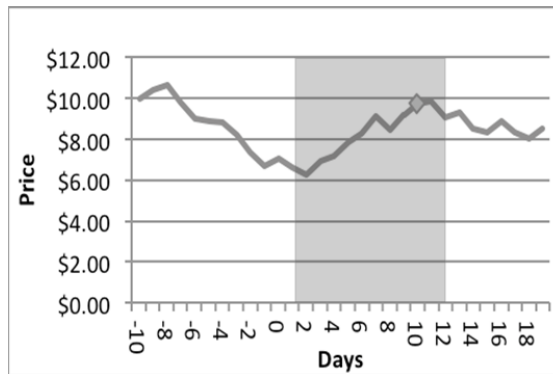
Figure 1: Example of the trade signal offsetting process (from Chen et al. (2012)).



- A.** Assume the following trade sequence:
1. Manager X issues buy signal for x100 on day 0
 2. Manager Y issues sell signal for x30 on day 3
 3. Manager Z issues sell signal for x20 on day 7



- B.** Offset subsequent sell signals (x30 and x20) against the initial buy signal, which in this case is lagged for 10 days.



- C.** Execute residual buy volume on-market at the end of the lag delay period (x50).

Appendix 2: Explanation and derivation of the collision rate equation

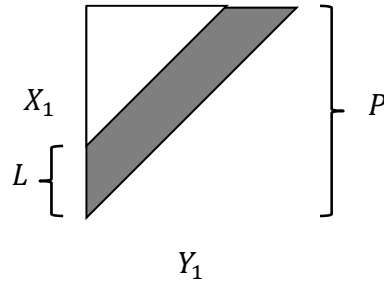
Assume sell side trade signals are drawn from a uniform distribution with range P . P represents the total period of observation from which trades can be drawn:

$$Y_1 \sim \mathcal{U}_s(0, P)$$

Further assume that a buy side signal (X_1) is drawn from the same range and distribution as sell side signals. The collision rate is the probability that X_1 is within distance L of Y_1 for $L \geq 0$:

$$\begin{aligned} CR_{X_1, Y_1} &= P(|X_1 - Y_1| \leq L) \\ &= P(X_1 - Y_1 \leq L \mid X_1 \geq Y_1) \cdot P(X_1 \geq Y_1) + P(Y_1 - X_1 \leq L \mid Y_1 \geq X_1) \cdot P(Y_1 \geq X_1) \end{aligned}$$

Since the two sides are symmetrical, we can simply solve for $P(X_1 - Y_1 \leq L \mid X_1 \geq Y_1)$. The figure below graphically characterises the distributions of X_1 and Y_1 on orthogonal axes, and the grey area represents the region where $0 \leq X_1 - Y_1 \leq L$.



Hence:

$$\begin{aligned} P(X_1 - Y_1 \leq L \mid X_1 \geq Y_1) &= \frac{P^2 - (P - L)^2}{P^2} = \frac{L(2P - L)}{P^2} \\ \therefore CR_{X_1, Y_1} &= \frac{L(2P - L)}{P^2} \end{aligned}$$

We take the n^{th} power of the base probability, where n is the number of opposed trade signals. Since signals issued by the same fund manager are excluded from offsetting, we scale the number of trade signals on the opposite market side by $\frac{f-1}{f}$, where f is the number of managers that participated in trading the particular security.

Hence, for n signals on the opposite market side:

$$CR_{X_1, Y_1 \dots Y_n} = 1 - \left(1 - \frac{L(2P - L)}{P^2}\right)^{n \left(\frac{f-1}{f}\right)} = 1 - \left(\frac{P - L}{P}\right)^{2 \cdot n \left(\frac{f-1}{f}\right)}$$

Finally, we formulate this equation with n as a stochastic variable with respect to an expected trade frequency td (i.e. $td \sim n/P$). Since the trade frequency assumes that P is unbounded, we must take the limit of P as it approaches infinity. Indeed, the function is non-stationary when taking small values of P , but asymptotic as P increases:

$$CR_{L, td} = 1 - \lim_{P \rightarrow \infty} \left(\frac{P - L}{P}\right)^{2 \cdot E(td) \cdot P \cdot \left(\frac{f-1}{f}\right)}$$

For practical purposes, we can achieve an arbitrary degree of accuracy for the projected collision rate by using a sufficiently large value of P .

Figure 3: Plot of the expected offset ratio function with realistic value ranges for the lag period and expected trade frequency. The expected offset ratio function is given by:

$$E(OR_{i, m, L}) = 0.5 \cdot \left(1 - \lim_{p \rightarrow \infty} \left(1 - \frac{L(2P - L)}{P^2}\right)^{E(td) \cdot P \cdot \left(\frac{f-1}{f}\right)}\right)$$

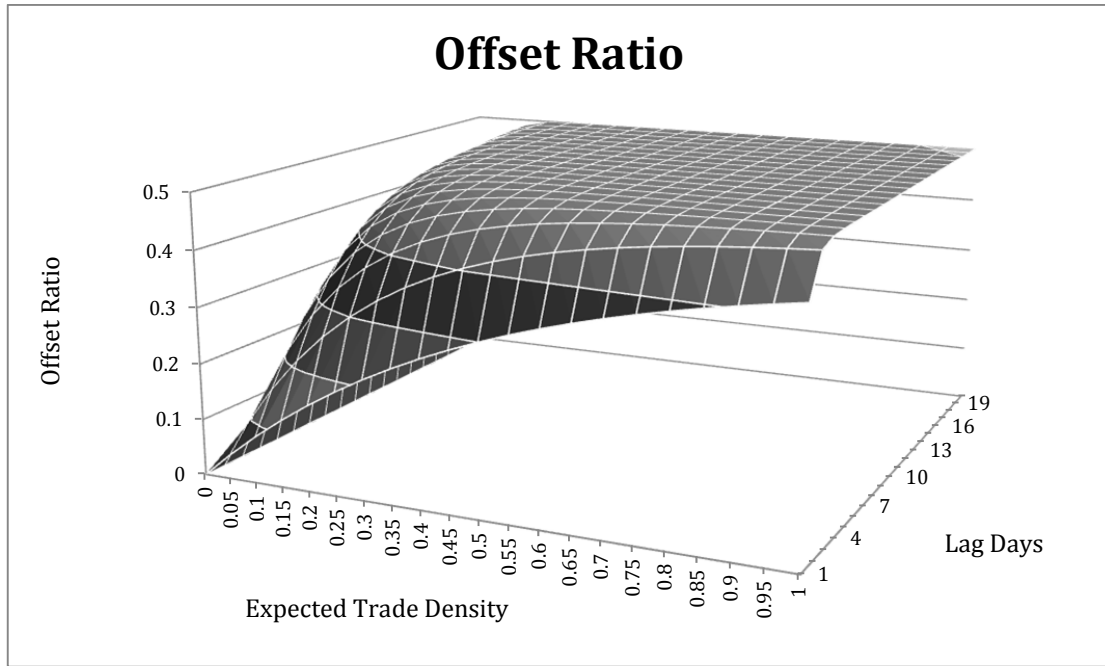


Figure 4: Descriptive statistics of the dataset. Trade packages are determined from individual trades by aggregating successive trades executed by a single manager on one side of the market, and unbroken by a trade by that manager on the opposite market side or a trading gap of more than five days. Data for this table is entirely sourced from our sample multi-manager.

	2005	2006	2007	2008	2009	All Years
Multi-manager Composition						
Enhanced Passive	4	4	4	4	4	4
Growth	1	1	1	2	2	2
Long/Short	1	1	1	2	2	2
Style-Neutral	1	1	2	3	3	3
Value	3	3	4	5	5	5
Total	10	10	12	16	16	16
Total trades						
Buy	4,137	5,669	5,746	8,272	9,097	32,921
Sell	3,640	4,892	6,112	7,630	8,650	30,924
Total trade packages						
Buy (% in parentheses)	1,392	1,980	2089	2691	3073	11,225
Sell (% in parentheses)	1,431	1,753	2188	3048	3528	11,948
Total Fund Value (\$m)	6,782.2	8,153.4	9,084.3	8,010.7	8,351.7	
Annual Turnover (%)	37.88	41.14	38.58	50.83	46.71	
Unique Securities Traded	155	165	215	220	248	332

Figure 5: Comparison of aggregated offset ratios predicted by the model and those observed in the ex-post simulated emulation fund. The offset ratio represents the proportion of traded value that can be internally crossed (and hence removed from on-market trading) in the emulation fund relative to the target fund. The x-axis represents the three lag periods that were used for both the model and the simulated fund.

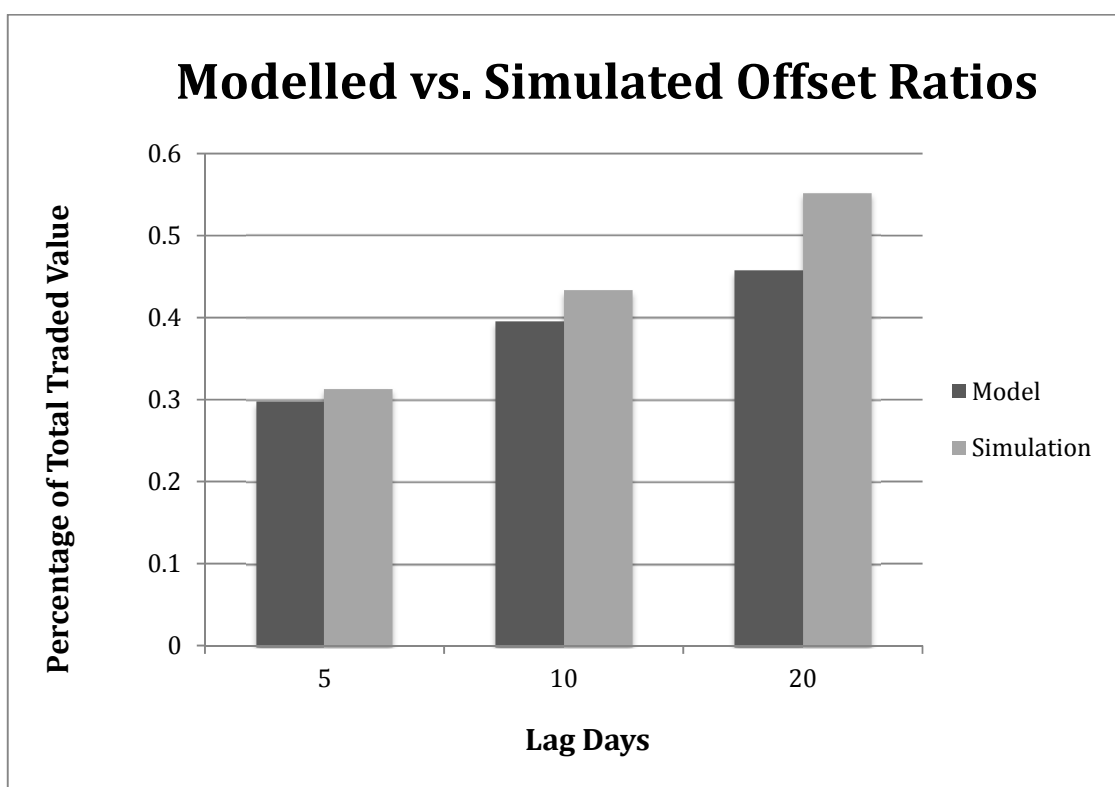


Figure 6: Individual performance drivers as forecasted by the model compared to those observed in the simulated emulation fund. Results are aggregated on a trade-value weighted basis across the observation period. We present results for emulation funds with 5, 10 and 20-day lag periods. Note that commissions and price impacts are presented as savings, which create positive cash flows to the emulation fund, while internal crossing and delayed execution represent costs. The total relative

returns represent the aggregated costs and benefits, with positive values representing excess performance and vice versa. Performance is measured as a proportion of total traded value – i.e. if the commission saving is reported as 0.05% in the simulation, it means that the total benefit arising from commission savings is 0.05% of the total traded value (buys plus sells).

Lag Days	5			10			20		
	Model	Simulation	Model - Simulation	Model	Simulation	Model - Simulation	Model	Simulation	Model - Simulation
Commission Savings	0.04%	0.05%	-0.01%	0.05%	0.06%	-0.01%	0.06%	0.08%	-0.02%
Price Impact Savings	0.04%	0.02%	0.02%	0.05%	0.01%	0.03%	0.05%	0.02%	0.03%
Internal Crossing Opportunity Cost	0.12%	0.08%	0.04%	0.18%	0.13%	0.05%	0.22%	0.15%	0.07%
Delayed Trading Opportunity Cost	0.39%	0.45%	-0.05%	0.36%	0.36%	-0.01%	0.30%	0.28%	0.02%
Total Impact	-0.43%	-0.47%	0.03%	-0.44%	-0.42%	-0.02%	-0.41%	-0.33%	-0.08%

Figure 7: Summary of model accuracy relative to a simulated emulation portfolio using subportfolios of different stock size (Panel A), market-to-book ratio (Panel B) and prior 1-year return (Panel C). Positive outcomes represent where the model/simulation outperforms the target fund; negative outcomes represent underperformance.

Panel A

Size Rank	Lag Days								
	5			10			20		
	Forecast	Actual	Forecast - Actual	Forecast	Actual	Forecast - Actual	Forecast	Actual	Forecast - Actual
1 - 20	-0.30%	-0.50%	0.19%	-0.29%	-0.47%	0.18%	-0.21%	-0.31%	0.11%
21 - 50	-0.41%	-0.51%	0.10%	-0.38%	-0.42%	0.03%	-0.27%	-0.24%	-0.03%
51 - 100	-0.53%	-0.37%	-0.16%	-0.46%	-0.39%	-0.07%	-0.46%	-0.38%	-0.08%
101 - 200	-0.53%	-0.38%	-0.15%	-0.47%	-0.25%	-0.22%	-0.57%	-0.42%	-0.15%
201 - 300	-0.98%	-1.18%	0.20%	-0.84%	-0.56%	-0.28%	-0.90%	0.38%	-1.27%
301 +	-0.60%	-3.58%	2.98%	-0.49%	-4.92%	4.43%	-1.11%	-2.79%	1.68%

Panel B

	Lag Days								
	5			10			20		
Size Rank	Forecast	Actual	Forecast - Actual	Forecast	Actual	Forecast - Actual	Forecast	Actual	Forecast - Actual
Highest MtB*	-0.19%	0.28%	-0.47%	-0.13%	0.43%	-0.56%	-0.37%	0.60%	-0.96%
2	-0.56%	-0.69%	0.13%	-0.61%	-0.56%	-0.04%	-0.71%	-0.51%	-0.19%
3	-0.50%	-0.50%	0.00%	-0.29%	-0.20%	-0.09%	-0.18%	0.11%	-0.28%
4	-0.39%	-0.33%	-0.07%	-0.29%	-0.17%	-0.12%	-0.24%	0.01%	-0.25%
5	-0.49%	-0.53%	0.04%	-0.46%	-0.39%	-0.07%	-0.26%	-0.60%	0.33%
6	-0.26%	-0.30%	0.04%	-0.32%	-0.30%	-0.03%	-0.30%	-0.06%	-0.24%
7	-0.45%	-1.36%	0.91%	-0.52%	-1.55%	1.04%	-0.59%	-1.76%	1.17%
8	-0.63%	-0.69%	0.06%	-0.51%	-0.51%	-0.01%	-0.58%	-1.06%	0.48%
9	-0.24%	0.64%	-0.88%	-0.39%	0.19%	-0.57%	-0.76%	-1.20%	0.45%
Lowest MtB*	-0.34%	-4.21%	3.87%	0.14%	-5.43%	5.57%	-0.74%	-5.99%	5.25%

* Market-to-book ratio

Panel C

	Lag Days								
	5			10			20		
Size Rank	Forecast	Actual	Forecast - Actual	Forecast	Actual	Forecast - Actual	Forecast	Actual	Forecast - Actual
Lowest Pr1yr*	-0.24%	-2.11%	1.87%	0.11%	-2.46%	2.57%	0.23%	-2.81%	3.04%
2	-0.41%	0.54%	-0.95%	-0.25%	0.10%	-0.35%	-0.17%	0.56%	-0.72%
3	-0.39%	-0.40%	0.01%	-0.35%	-0.35%	-0.01%	-0.25%	-0.10%	-0.15%
4	-0.34%	-0.27%	-0.08%	-0.33%	-0.47%	0.14%	-0.32%	-0.51%	0.19%
5	-0.38%	-0.56%	0.19%	-0.41%	-0.42%	0.02%	-0.35%	-0.63%	0.27%
6	-0.31%	-0.67%	0.36%	-0.27%	-0.38%	0.11%	-0.20%	-0.50%	0.30%
7	-0.38%	-0.56%	0.18%	-0.35%	-0.32%	-0.03%	-0.32%	-0.32%	0.00%
8	-0.42%	-0.50%	0.08%	-0.39%	-0.34%	-0.05%	-0.31%	-0.15%	-0.16%
9	-0.57%	-0.51%	-0.06%	-0.53%	-0.82%	0.30%	-0.58%	-0.39%	-0.19%
Highest Pr1yr*	-0.73%	-0.36%	-0.37%	-0.67%	-0.20%	-0.46%	-0.75%	1.01%	-1.77%

* Prior 1-year return

Figure 8: Partial derivatives of the expected offset ratio equation, given by:

$$E(OR_{i,m,L}) = 0.5 \cdot \left(1 - \lim_{p \rightarrow \infty} \left(\frac{P-L}{P} \right)^{2 \cdot E(td) \cdot P \cdot \left(\frac{f-1}{f} \right)} \right)$$

Panel A plots the partial derivative of this with respect to the lag days L , with realistic value ranges for the lag period and expected trade frequency:

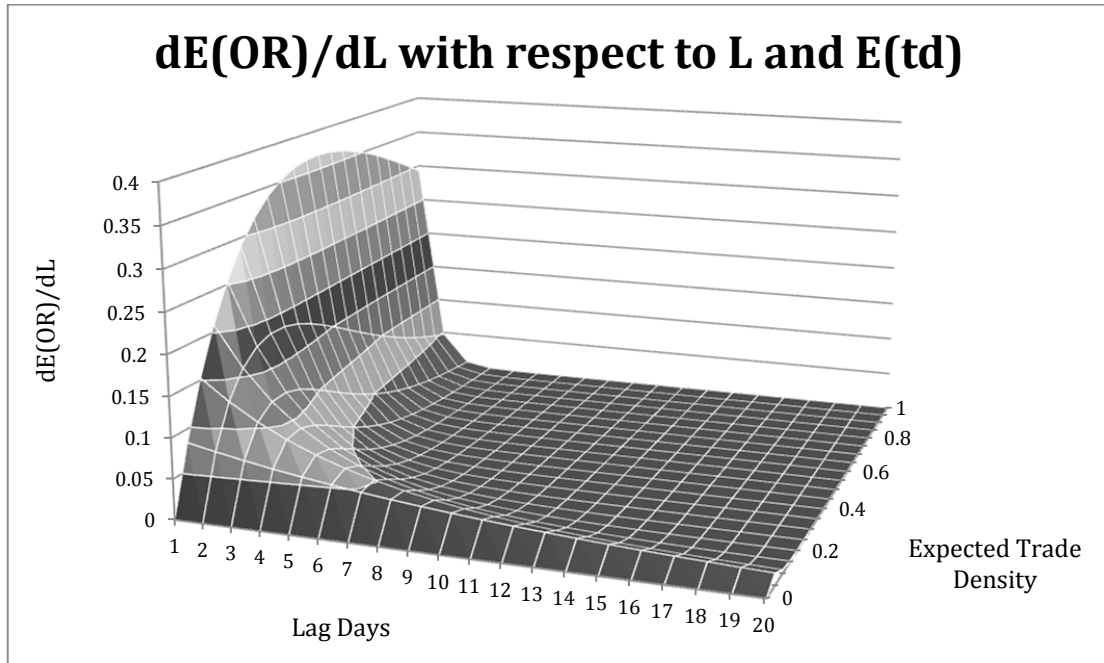
$$\frac{\partial E(OR_{i,m,L})}{\partial L} = \lim_{P \rightarrow \infty} \left(2 \cdot E(tf) \cdot \left(\frac{f-1}{f} \right) \cdot \left(\frac{P-L}{P} \right)^{2 \cdot E(tf) \cdot P \cdot \left(\frac{f-1}{f} \right) - 1} \right)$$

Panel B plots the partial derivative of the expected offset ratio equation with respect to the expected trade frequency $E(tf)$, for realistic value ranges for lag period and expected trade frequency:

$$\frac{\partial E(OR_{i,m,L})}{\partial E(tf)} = -1 \cdot \lim_{P \rightarrow \infty} \left(P \cdot \left(\frac{f-1}{f} \right) \cdot \log \left(\frac{P-L}{P} \right) \cdot \left(\frac{P-L}{P} \right)^{2 \cdot E(tf) \cdot P \cdot \left(\frac{f-1}{f} \right)} \right)$$

We use $f = 12$ for the purposes of creating the graphs below as this was the mean number of individual fund managers that traded each stock on a trade-value weighted basis.

Panel A



Panel B

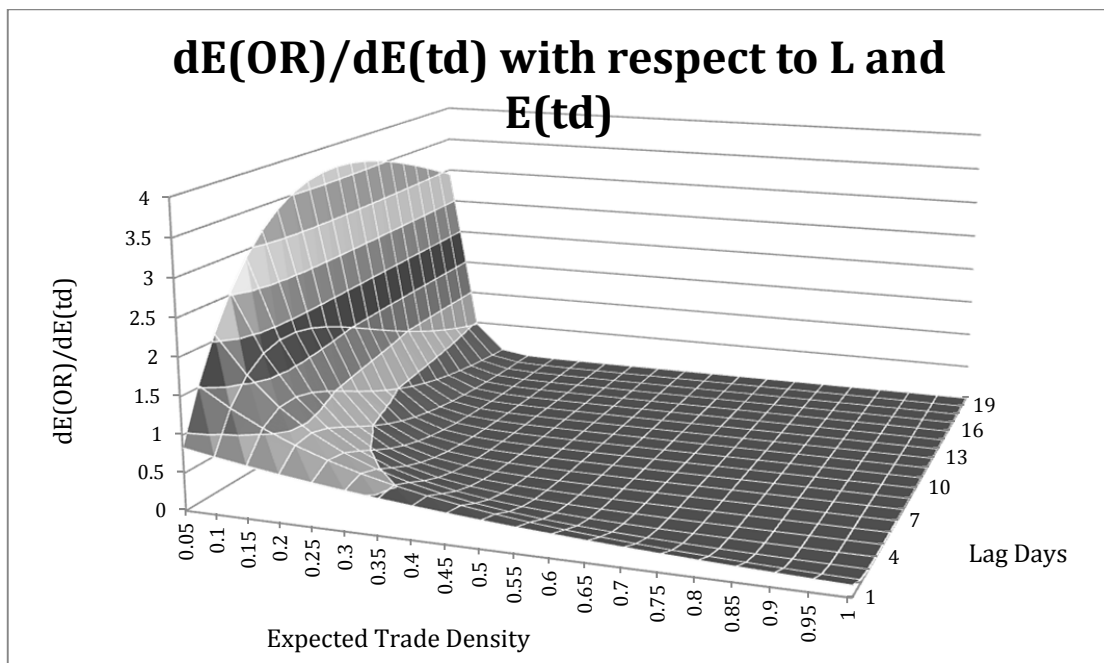


Figure 9: Application of the model for lag period optimisation. The model is specified as:

$$E(\pi_{i,L}) = E(OR_{i,L}) \left(2C_i + E(PI_{i,buy}) + E(PI_{i,sell}) + \frac{P_{i,sell}(L) - P_{i,buy}(L)}{L} \right) \\ + (1 - E(OR_{i,L})) (\rho_{i,sell}(L) - \rho_{i,buy}(L))$$

The expected offset ratio is given by the following equation:

$$E(OR_{i,L}) = 0.5 \cdot \left(1 - \lim_{p \rightarrow \infty} \left(\frac{P - L}{P} \right)^{2 \cdot E(tf) \cdot P \cdot \left(\frac{f-1}{f} \right)} \right)$$

Input variables are based on observations in the mutual fund trade dataset to provide realistic estimates of emulation performance.

L	5	10	20	50	100
$E(td_i)$	0.05	0.05	0.05	0.05	0.05
f	8	8	8	8	8
$E(OR_{i,L})$	17.72%	29.16%	41.31%	49.37%	49.99%
C_i	0.15%	0.15%	0.15%	0.15%	0.15%
$E(PI_{i,buy})$	0.10%	0.10%	0.10%	0.10%	0.10%
$E(PI_{i,sell})$	0.02%	0.02%	0.02%	0.02%	0.02%
$\rho_{i,buy}(L)$	0.30%	0.52%	0.93%	2.60%	6.62%
$\rho_{i,sell}(L)$	0.06%	0.19%	0.58%	2.57%	7.53%
$\frac{P_{i,buy}(L)}{L}$	0.15%	0.28%	0.50%	1.26%	2.94%
$\frac{P_{i,sell}(L)}{L}$	0.03%	0.08%	0.23%	1.04%	3.04%
$E(\pi_{i,L})$	-0.14%	-0.17%	-0.14%	0.08%	0.72%

References

- Ang, A., Rhodes-Kropf, M. & Zhao, R. 2008. Do Funds-of-Funds Deserve Their Fees-on-Fees? *NBER Working Paper*, 1-32.
- Brands, S. & Gallagher, D. R. 2004. Risk and return properties of fund-of-funds. *JASSA*, 33-40.
- Brands, S. & Gallagher, D. R. 2005. Portfolio selection, diversification and fund-of-funds: a note. *Accounting & Finance*, 45, 185-197.
- Brown, N. C., Wei, K. D. & Wermers, R. 2007. Analyst Recommendations, Mutual Fund Herding, and Overreaction in Stock Prices. *Working Paper*, 1-70.
- Brown, S., Goetzmann, W. & Liang, B. 2004. Fees-on-Fees in Funds-of-Funds. *Journal of Investment Management*, 24, 39-56.
- Chan, L. K. C. & Lakonishok, J. 1995. The behavior of stock prices around institutional trades. *Journal of Finance*, 50, 1147-1174.
- Chen, C., Comerton-Forde, C., Gallagher, D. R. & Walter, T. S. 2010. Investment Manager Skill in Small-Cap Equities. *Australian Journal of Management*, 35, 23-49.
- Chen, Z., Foster, F. D., Gallagher, D. R. & Lee, A. D. 2012. Does portfolio emulation outperform its target funds? *Australian Journal of Management*, Forthcoming, 1-39.
- Chiyachantana, C. N., Jain, P. K., Jiang, C. & Wood, R. A. 2004. International evidence on institutional trading behavior and price impact. *Journal of Finance*, 59, 869-898.
- Comerton-Forde, C., Fernandez, C., Frino, A. & Oetomo, T. 2005. How broker ability affects institutional trading costs. *Accounting & Finance*, 45, 351-374.
- Dibartolomeo, D. 1999. A radical proposal for the operation of multi-manager investment funds. *Northfield Information Services*, 1-15.
- Domowitz, I., Glen, J. & Madhavan, A. 2001. Liquidity, Volatility and Equity Trading Costs Across Countries and Over Time. *International Finance*, 4, 221-255.
- Elton, E. J. & Gruber, M. J. 2004. Optimum Centralized Portfolio Construction with Decentralized Portfolio Management. *Journal of Financial and Quantitative Analysis*, 39, 481-495.
- Ferguson, R. 1978. Do Market Inventory Funds Really Make Sense? *Financial Analysts Journal*, 38-45.
- Gallagher, D. & Gardner, P. 2005. Portfolio design and challenges inherent in multiple manager structures. *JASSA*, 20-25.
- Goldstein, M. A., Irvine, P. J., Kandel, E. & Wiener, Z. 2009. Brokerage Commissions and Institutional Trading Patterns. *Review of Financial Studies*, 22, 5175-5212.
- Hirshleifer, D., Subrahmanyam, A. & Titman, S. 1994. Security Analysis and Trading Patterns When Some Investors Receive Information Before Others. *Journal of Finance*, 49, 1665-1698.
- Keim, D. & Madhavan, A. 1996. The Upstairs Market for Large Block Transactions: Analysis and Measurement of Price Effects. *Review of Financial Studies*, 9, 1-36.
- Kothari, S. P. & Warner, J. B. 2001. Evaluating Mutual Fund Performance. *Journal of Finance*, LVI, 1985-2010.
- Kyle, A. S. & Obizhaeva, A. A. 2011. Market Microstructure Invariants. *Working Paper*.
- Lhabitant, F. S. & Learned, M. 2002. Hedge Fund Diversification: How Much is Enough? *SSRN Electronic Journal*, 1-44.

- Nadarajah, S. 2008. Explicit Expressions for Moments of Log Normal Order Statistics. *Economic Quality Control*, 23, 267-279.
- Rosenberg, B. 1977. Institutional Investment with Multiple Portfolio Managers. *Proceedings of the Seminar on the Analysis of Security Prices*, 55-160.
- Sharpe, W. F. 1981. Decentralized Investment Management. *Journal of Finance*, 36, 217-234.
- Wagner, W. H. & Zipkin, C. A. 1978. Can Inventory Index Funds Improve Active Equity Performance? *Financial Analysts Journal*, 34, 34-76.