Chapter 3

Factors Influencing Transaction Costs

Introduction

In a perfectly competitive and complete market, traders can engage in trading instantly at an equilibrium price for any quantity. In that environment, when executing a trade, a trader's cost of trading encompasses two components: (1) the value of the stock or asset (*cost of investment*), a payment to the seller; and (2) a fixed charge, such as commission, a payment to the broker or intermediary who facilitates the trade. The second component is typically known as *transaction cost*.

The performance on trading depends on the proceeds from the sale of the stock, net of transaction cost. In an intensely competitive financial market in which every trader is seeking alpha, transaction cost, no matter how small, can dictate performance when comparing with a benchmark or peer traders. Empirical evidence reveals that transaction costs can range from as low as 30 bps to as high as 300 bps, depending on the size of the trade and the liquidity of the stock (Wagner and Glass [2001]). This evidence clearly indicates that transaction cost can affect investment performance and warrants careful management.

A seminal paper by Perold [1988] was the first to address the importance of transaction cost and gives a clear understanding and insights of transaction cost. He provided evidence that true transaction cost is not merely the commission to be paid to a broker for facilitating the trade; the reality involves the cost of trading and the cost of not trading, which can magnify the true transaction cost. According to Perold, many factors that have not been incorporated in the past can adversely contribute to a trader's transaction cost, and the performance of traders and institutional managers may have been overstated. These factors, such as liquidity of the market (if the market is illiquid, a trader may not be able to trade all shares), ability to trade at one price (if the market moves fast, a trader may not be able to trade all shares at that price), and presence of informed traders (who trade with information that the trader does not have, known as *asymmetric information trading*), can influence the true transaction cost. Perold [1988] called this *implementation shortfall* (IS). IS comprises (1) all fixed charges such as commission and fees, and (2) the amount as a result of these factors, known as *opportunity cost*.

Wagner and Edwards [1993] also showed that the naïve definition of transaction cost (commission only) is far from reality and is merely the obvious lip of the execution cost structure. The scope of transaction cost, according to Wagner and Edwards [1993], also should be broadened to cover the entire process of implementing a trade, which starts from the moment a trader decides to trade to the moment the trade actually is executed. During this process, a trader can endure price impact in soliciting and making a trade, timing cost that causes any price variation before trading can be done, and inability to complete total investment that results in opportunity cost for the trader.

Wagner [1991] and Wagner and Edwards [1993] closely followed the framework of Perold [1988] and provided an expanded IS by clearly identifying various components that can play an important role in transaction cost. Later, a study by Kissell

[2006] revisited this expanded IS and further regrouped and categorized components in an *expanded IS* to provide better and clearer classification of transaction cost. In another work, Kissell [2014] provided a list of 10 components of transaction cost by considering factors identified in the expanded IS of Wagner and Edwards [1993] and Kissell [2006]. These components are commission, fees, taxes, rebates, spreads, delay cost, price appreciation, market impact, timing risk, and opportunity cost. The first four items are considered fixed cost (i.e., they are beyond the control of a trader and must be paid), and the remaining six items are invisible and result from implementation strategy.

In this research, further review of all three mathematical frameworks of IS and improving the analysis by providing further classification on opportunity cost that previous studies have not addressed. The model used by Kissell [2006] and attempt to further classify components of opportunity cost is closely followed. The study discovered two additional components that are not addressed by Kissell [2006] or others. These two subcomponents are, however, additional venues for traders to manage carefully in order to control transaction costs while trading. By further identifying these factors, the study contribute to the existing literature with a view that traders can now detect and manage these factors properly to reduce the IS of trading. As a result, along with stock picking ability, traders will be able to implement trading strategy by appropriately executing trades through controlling these two new factors that add to opportunity cost. In addition, using a numerical experiment, it is also shown that this model calculates ISs that are equal to all three existing models. To our knowledge, we are the first to provide such calculations and to identify factors that contribute to the IS.

By being able to detect and control these factors, traders will be able to reduce or eliminate these costs while executing trades.

Understanding and managing transaction cost is highly critical to portfolio performance in an environment in which more than half of mutual funds have consistently underperformed the S&P 500. Sophisticated investors, mutual fund managers, and hedge fund managers who have superior analytical ability in stock selection and follow optimal asset allocation techniques also must pay close attention to IS identified in the literature. IS is relevant especially for short-term traders (or day traders) and dynamic traders because these traders engage in trading quite frequently in a short period of time. Transaction cost from frequent trading can add up to a large amount and essentially eat up the profits if proper care is not taken to avoid or contain any the factors identified in the literature. Amihud and Mendelson [2013] emphasized that in today's decimalization era of high-frequency trading, transaction cost is an essential element of market microstructure that cannot be removed but can and should be managed. Transaction cost is also considered in recent research to evaluate the effectiveness of dynamic portfolio strategies (Kim and Viens [2012]) and optimal portfolio decisions (Garleanu and Pedersen [2013]). Understanding and measuring transaction cost accurately is undoubtedly an important issue for any trader or portfolio manager and has been a subject of research among academicians and practitioners.

Perold [1988] defined transaction cost as IS and used paper return (PR) and real return (RR) to essentially measure the degree to which a trader is unable to exploit his or her stock selection skill, which greatly depends on the trader's implementation strategy. The trader's implementation strategy, on the other hand, depends on how quickly the trader wants to trade at any price (in which trade can take place instantaneously) or at a fair price (in which a trader may have to wait a while for full or partial execution), which may contribute to a trader's inability to completely execute his or her trade.

A trader's inability to execute stock selection skills due to implementation strategy is known as *opportunity cost*. IS, as a result, consists of not only execution cost (which relates to executed transaction and measures all the obvious costs such as brokerage commissions and transfer tax) but also opportunity cost (which relates to the transactions a trader fails to execute and simply measures the paper performance of the buy and sell that could not be executed). Opportunity cost is the key contribution of Perold [1988] and has received greater attention in the subsequent research.

Wagner [1991] and Wagner and Edwards [1993] expanded IS by providing three classifications of opportunity cost. The motivation of this further classification is to give traders an in-depth understanding of opportunity cost components they may face and how to minimize the cost as much as possible. These three categories of opportunity cost are (1) delay-related cost, (2) trading-related cost, and (3) opportunity cost (other costs, not due to price impact and timing but due to incomplete execution). It should be noted here that Wagner and Edwards [1993] named these three categories as (1) price impact cost, (2) timing cost, and (3) opportunity cost. Price impact is affected by order size, market depth, liquidity, trade urgency, and broker skill. Timing cost depends on a trader's willingness to reduce price impact by waiting and trading slowly, which can result in larger price change by the time a trade is completed.

Opportunity cost depends on the amount of unexecuted trade. The seminal paper by Kissell [2006] further analyzed expanded IS by making slightly different, with better groupings of the cost elements proposed by Wagner [1991] and Wagner and Edwards [1993]. Kissell [2006] identified three components and named them a little differently: He renamed price impact as trading-related cost and timing cost as investment-related cost. Additionally, whereas Wagner and Edwards [1993] incorporated trading-related delay cost and operation-related delay cost as part of timing cost (or delay-related cost), Kissell [2006] removed operation-related delay cost from investment-related cost and categorized it as a part of the opportunity cost because it is related to the unexecuted components of shares. As a result, opportunity cost, according to Kissell [2006], is made of two subcomponents: investment-related opportunity cost and trading-related opportunity cost. This is the key contribution of Kissell [2006].

The work of Kissell [2006] is further expanded and provide additional classifications to the opportunity cost component, keeping the remaining two components the same as Kissell's [2006]. The study extend the opportunity cost component into three subcategories to provide better understanding of the opportunity cost that plays a role in the transaction cost. These components are (1) investment- or operation-related opportunity cost (the same as the operation-related delay cost of Kissell [2006]), (2) first trading–related opportunity cost, and (3) residual trading–related opportunity cost.

Using a numerical example, it is proved in this research that this model provides the same result for IS as that found by Perold [1988], Wagner and Edwards [1993], and Kissell [2006]. It is also proven mathematically that all three models are identical in measuring IS. It is believed that this research is timely and very important because of the increase in electronic order flow and algorithmic trading and the dominant presence of high-frequency traders in today's financial market. This further classification of opportunity cost may help a portfolio manager or trader to focus and avoid the cost in his or her investment decisions and improve portfolio performance. Exhibit 3.01 (in Appendix 3) shows the evolution of IS.

Modeling Transaction Cost

Following Kissell [2006], in this section, the study provides the mathematical framework of IS. Considering unexecuted trade as part of the model, the study first present the model of Perold [1988] and its expansions proposed by Wagner [1991], Wagner and Edwards [1993], and Kissell [2006]. Finally, The classification of expanded IS, following and expanding on the classifications in Kissell [2006] is presented here.

Transaction costs have two components as mentioned earlier. There are some fixed components in transaction costs that are unavoidable, such as commission, fees, and taxes. These components are also known as *visible costs* or *fixed costs*. Variable and invisible components in transaction costs depend on factors that influence the implementation strategy of a trader and are manageable through implementation strategies. These transaction costs are known as *nontransparent costs*, such as market impact costs, timing cost, and opportunity cost due to un-executed trade. In the next section, a brief definition of each component of IS is given. A detailed definition was given by Kissell [2014].

The study begins with the IS framework proposed by Perold [1988] to measure transaction costs, which encompasses both visible and nontransparent costs (cost of trading and not trading). According to the IS methodology used by Perold [1988], once stocks are selected, one should implement a decision (take a position). A position taken in *paper* (i.e., a hypothetical trade with no real investment made) offers what is known as PR, and an actual position (i.e., a real transaction completed) offers what is known as portfolio RR. IS is then defined as the difference between PR and RR. Algebraically it can be written as

$$\text{IS} = \text{PR} - \text{RR} \tag{3.01}$$

Let us now present it mathematically. Consider that a trader wishes to buy X number of shares that are currently trading at P_d dollars per share and at P_N at the closing of trading day. The PR in dollars is calculated as follows:

$$\text{PR} = X \times (P_N - P_d) \tag{3.02}$$

Where $X \times P_N$ is the ending value of the portfolio and $X \times P_d$ is the beginning value of the portfolio. This PR is calculated as if the trader has (hypothetically) initiated the transaction at the beginning and closing price. Please recall that, in practice, the real transaction price can be anything but the beginning and ending price as price moves by seconds, depending on the trader's decision to trade and the time gap in initiating the transaction to execution. As a result, the portfolio RR can be different from the (hypothetical) PR. If the trader engages in trading shares, depending on the available number of shares at each trading point, the trader may accumulate his or her total shares in multiple trades at different prices. Putting this together, the portfolio RR can be estimated as follows:

$$RR = X \times P_N - (\sum x_j \times P_j - \text{Fixed}) \quad (3.03)$$

Where $\sum x_j$ is the number of shares purchased at the *j*th transaction at corresponding price P_j, and the sum of x_j is equal to X. The fixed fee represents the total of commission, tax, clearing, and settlement charges that are fixed and must be paid to the intermediaries, government, and other parties involved in facilitating the transaction. Please note that X and x_j could be positive (for long position) for cash investment or negative (for short position) for cash redemption. Now the IS can be calculated as follows:

$$\text{IS} = \underbrace{(X \times P_N - X \times P_d)}_{\text{Paper return}} - \Big(X \times \underbrace{P_N - (\sum x_j \times P_j - \text{Fixed})\Big)}_{\text{Real portfolio return}}$$

$$= \sum x_j P_j - X \times P_d + \text{Fixed} \quad (3.04)$$

If the entire desired trade cannot be completely executed, a portion of cash remains not invested in stocks, which creates an opportunity cost due to inability to invest in stocks. Consider that at time t_d, a trader has $X \times P_d$ amount of cash that he or she wishes to invest; this can be viewed as the beginning portfolio and or cash value. If the actual transaction value is defined as $\sum x_j \times P_j$, the idle cash amount (not invested), due to the unexecuted order, is the dollar amount that is the difference between cash available and cash invested; that is, $X \times P_d - \sum x_j \times P_j$. As a result, the ending value of the portfolio has a component of idle cash along with stocks. These values are defined as follows:

Beginning portfolio value = $\underbrace{X \times P_d)}_{\text{Available cash}}$

$$\text{Ending portfolio value} = \underbrace{(\sum x_j \times P_N)}_{\text{Stock value}} + \underbrace{X \times P_d - (\sum x_j \times P_j)}_{\text{Idle cash}} \qquad (3.05)$$

Portfolio RR is then calculated as follows:

$$\text{RR} = \text{Ending portfolio value} - \text{Available cash}$$

$$= \underbrace{(\sum x_j \times P_N + X \times P_d - (\sum x_j \times P_j) - X \times P_d - \text{Fixed}}$$

$$= \underbrace{\sum x_j \times P_N - \sum x_j \times P_j - \text{Fixed}} \qquad (3.06)$$

IS with unexecuted order can now be calculated as follows:

$$\text{IS} = \underbrace{(X \times P_N - X \times P_d)}_{\text{Paper return}} - (\sum \underbrace{x_j \times P_N - (\sum x_j \times P_j - \text{Fixed})}_{\text{Realportfolio return}} \qquad (3.07)$$

Rearranging Equation (3.07), it is in the following form:

$$\text{IS} = (X - \sum x_j) \times P_N - X \times P_d + \sum x_j P_j + \text{Fixed} \qquad (3.08)$$

where $(X - \sum x_j)$ is the number of shares unexecuted. The number of initial shares, X, can now be written as $X = \underbrace{(X - \sum x_j) + \sum x_j}$. Replacing the definition of X into Equation (3.08), IS can be redefined as follows:

$$IS = (X - \sum x_j) \times P_N - ((X - \sum x_j) + \sum x_j) \times P_d + \sum x_j P_j + \text{Fixed}$$

$$= (X - \sum x_j) \times P_N - ((X - \sum x_j) P_d - \sum x_j \times P_d + \sum x_j P_j + \text{Fixed} \quad (3.09)$$

$$= (\sum x_j \times P_j - \sum x_j P_d + (X - \sum x_j) \times (P_N - P_d) + \text{Fixed}$$

$$= X\left(P_n - P_d\right) - \sum x_j \left(P_n - P_j\right) + \text{Fixed}$$

The Equation (3.09) is the IS defined by Perold [1988], which distinguishes between execution cost and opportunity cost of the order and can be expressed as follows:

$$IS = \underbrace{\sum x_j \times P_j - \sum x_j \times P_d}_{\text{Execution cost}} + \underbrace{(X - \sum x_j) \times (P_N - P_d)}_{\text{Opportunity cost}} + \text{Fixed} \quad (3.10)$$

In Equation (3.10), opportunity cost is defined as the portion of loss or gain from the price difference of the opening and closing price multiplied by the portion of unexecuted order. Wagner [1991], Wagner and Edwards [1993], and Kissell [2006] expanded that execution cost component of the IS. As indicated earlier, Wagner [1991] and Wagner and Edwards [1993] expanded the execution cost component into two parts: price impact (delay-related trading cost) and timing cost (trading-related cost). Kissell [2006] named them investment-related cost and trading-related cost. The study focuses and continue the model development based on Kissell's [2006] because present model further expands that of Kissell [2006]. The investment horizon, according to Kissell

[2006], is the time period from the investment decision t_d to the time that trade begins t_o, and the trading horizon is the time period from the commencement of trading at t_o to the end of the trading, t_n. Accordingly, if P_t represents the midpoint of the bid–ask spread at the time the order is entered to the market (i.e., arrival price), the price change over the period t_d to t_n can be written in terms of P_0 as follows:

$$(P_N - P_d) = (P_N - P_0) + (P_0 - P_d) \tag{3.11}$$

By replacing ($P_N - P_d$) in Equation (10) with the extension shown in Equation (3.11), Equation (3.10) can now be rearranged as follows:

$$\text{IS} = (\sum x_j \times P_j - \sum x_j P_d + (X - \sum x_j) \times (P_N - P_0) + (X - \sum x_j) \times (P_0 - P_d) + \text{Fixed}$$

$$= X(P_N - P_0) + \sum x_j \times P_j - \sum x_j P_d + (X - \sum x_j) \times (P_N - P_0) + \text{Fixed} \tag{3.12}$$

Equation (3.12) can be rearranged—a simple mathematical manipulation and regrouping of items—to develop the expanded form of IS, proposed by Wagner and Edwards [1993] and used by Kissell [2006], as follows:

$$\text{IS} = \underbrace{X(P_0 - P_d)}_{\text{Investment related}} + \underbrace{\sum x_j \times P_j - \sum x_j \times P_d}_{\text{Trading related}} + \underbrace{(X - \sum x_j)(P_N - P_0)}_{\text{Opportunity cost}} + \text{Fixed} \tag{3.13}$$

As shown in Equation (3.13), Kissell [2006] expanded IS (total transaction or execution cost) into four components that sum up to total transaction costs. The first component is identified as investment-related costs, which arise during the investment decision phase. This part of the IS constitutes the period of time from the investment decision to the time the order is released to the market for transaction. Any gap between these two time periods is called a *delay in implementing investment decision*. This delay-related cost can arise because of lack of communication between portfolio manager and trader or because of a delay in selecting an appropriate broker, algorithm, or algorithmic parameter. If manager and trader can work closely to determine the strategies most consistent with the investment objective, the delay-related cost can be reduced or eliminated.

The second component is trading-related costs that comprise the largest subset of transaction costs and include those that arise during implementation of the investment decision (the time period from the start of trading to the end of trading). The largest trading-related costs are market impact and timing risk. Market impact is highest when using an aggressive trading strategy and lowest with a passive strategy. Timing risk, on the other hand, is highest with a passive strategy and lowest with an aggressive strategy. There is a clear need for an optimum trade-off that can minimize overall costs from these two components. Other trading-related costs are price appreciation cost and spread cost. The higher and faster the price movement and spreads, the larger the trading-related costs. Market liquidity and competition for market making play a role in these two components, over which managers and traders have no control.

The third component is the opportunity cost, which is a part of the transaction cost that represents the foregone profit or loss resulting from being unable to trade the entire order. It is measured as the number of unexecuted shares multiplied by the price difference over the period of the order. The opportunity cost has two components, investment-related opportunity cost and trading-related opportunity cost.

The fourth component comprises fees that constitute all fixed charges related to commission, rebate, and taxes. The expanded IS of Kissell [2006] now can be written as follows with its subcomponents in each of the four categories:

IS = Investment related + Trading related + Opportunity cost + Fixed

Or,

$$\text{IS} = \text{IRC} + \{\text{TC} = \text{SC} + \text{MIC} + \text{PAC} + \text{TRC}\} + \{\text{OC} = \text{IROC} + \text{TROC}\} + \{\text{Fixed} = C + F + R + T\} \quad (3.14)$$

where IRC is investment-related cost, TC is trading-related cost, SC is spread cost, MIC is market impact cost, PAC is price appreciation cost, TRC is timing risk cost, OC is opportunity cost, IROC is investment-related opportunity cost, TROC is trading-related opportunity cost, C is commission, F is fees, R is rebates, and T is tax.

Kissell [2006] differed from Wagner [1991] and Wagner and Edwards [1993] slightly in classification of components and in their inclusion. The expanded IS model of Wagner [1991] (and Wagner and Edwards [1993]) also has four components:

$$\text{IS} = \underbrace{X(P_0 - P_d)}_{\text{Delay related}} + \underbrace{\sum x_j \times P_j - \sum x_j \times P_d}_{\text{Trading related}} + \underbrace{(X - \sum x_j)(P_N - P_0)}_{\text{Opportunity cost}} + \text{Fixed} \quad (3.15)$$

Trading-related cost (also called *price impact* by Wagner and Edwards [1993]) is defined as the difference between the price at which trade is revealed to the broker and the execution price (i.e., the price at which trade actually takes place). Kissell [2006] also defined it as trading-related cost.

The delay-related cost of Wagner [1991] (also called timing cost by Wagner and Edwards [1993]) is identified as investment-related costs by Kissell [2006], with a deviation. Delay-related cost is defined as price movements between the initial submission of an order to the trader and the exposure of the order to the broker. Wagner [1991] and Wagner and Edwards [1993] incorporated two factors in timing- or delay-related cost: trading-related delay cost (Kissell [2006] only considers this as investment-related cost) and operation-related delay cost. Kissell [2006], however, did not include operation-related delay cost as part of investment cost and rather treated it as part of the opportunity cost because it is related to unexecuted shares. As a result, opportunity cost according to Kissell [2006] has one additional item (operation-related delay cost) that is not present in the definitions used by Wagner [1991] and Wagner and Edwards [1993]. Putting it formally, the concept used by Wagner [1991] and Wagner and Edwards [1993] can be presented as follows:

IS = Delay-related cost + Trading-related cost + Opportunity cost + Fixed

or,

$$IS = \begin{Bmatrix} DC \text{ or } TC \\ = TRDC + ORDC \end{Bmatrix} + \begin{Bmatrix} TC \text{ or } PI \\ = SC + MIC + PAC + TRC \end{Bmatrix} + OC + Fixed \qquad (3.16)$$

Where DC is delay cost (also known as TC, timing cost); TRDC is trading-related delay cost; ORDC is operation-related delay cost; TC is trading-related cost (also known as PI, price impact); and SC, MIC, PAC, TRC, and OC are defined as earlier.

In this research, the opportunity cost component is further investigated and classified. It is done in order to better understand elements of opportunity costs to enable traders to notice and address them and thus minimize overall transaction cost. Accordingly, a subdivision is done of the trading-related opportunity cost of Kissell

[2006] into two categories: (1) first trading–related opportunity cost and (2) residual trading–related opportunity cost. By appropriately setting the price for the first trade, if traders can execute all shares they will be better able to reduce these costs. Kissell's [2006] trading-related opportunity cost in this model is included as a third component. With this extension the study present the model of IS with its subcomponents as follows:

IS = Investment related + Trading related + Opportunity cost + Fixed

or,

$$\mathrm{IS} = \mathrm{IRC} + \{\mathrm{TC} = \mathrm{SC} + \mathrm{MIC} + \mathrm{PAC} + \mathrm{TRC}\} + \{\mathrm{OC} = \mathrm{IROC} + \mathrm{FTROC} + \mathrm{RTROC}\} + \left\{ \begin{matrix} \mathrm{Fixed} \\ = C + F + R + T \end{matrix} \right\} \quad (3.17)$$

Where FTRC is first trading–related opportunity cost and RTROC is residual trading–related opportunity cost. All other variables are the same as those in Kissell [2006].

The investment-related and trading-related costs is defined in the same way as Kissell [2006]. In the opportunity cost component, the study extend based on Kissell's [2006] work, as stated earlier. Finally, fixed, in our model, takes the definition used by Kissell [2006]. This extension of opportunity costs contributes to the existing literature and is the main contribution of our research.

Through numerical examples as experimental research, the study develop different scenarios and offer calculations of IS (or transaction costs) for the models of Perold [1988], Wagner and Edwards [1993], Kissell [2006], and that used in this study. If, through various classifications of opportunity cost, the same result as that given by Wagner and Edwards [1993] and Kissell [2006] is arrived, that means this research is interesting and makes a case for traders to pay attention to each of the categories of opportunity cost.

When transaction cost for this extended model is calculated, the result shows that one can reach the same transaction cost as Wagner and Edwards [1993] and Kissell [2006]. Exhibit 3.01 (in Appendix 3) displays the IS models based on Perold [1988], Wagner and Edwards [1993], Kissell [2006], and the present model. Before presenting this experiment in calculating IS, it would be useful first to provide the definition of each cost component of opportunity costs.

Definition of Cost Components of Opportunity Costs

The opportunity cost is the cost associated with an inability to complete an order. Most often, opportunity cost accompanies limit order–based strategies, but it can also be present in market-order execution. The inability to fulfill an order can be due to one of several factors: (1) the market price has never crossed the limit price; (2) the market does not have sufficient liquidity (demand or supply) to fulfill the order at the desired price; or (3) the price moves away so quickly that fulfilling the order would render the transaction unprofitable, and the transaction is canceled as a result. The opportunity cost is measured as the profit expected to be generated if the order were executed. Kissell [2006] provided two subgroups for opportunity cost, namely, investment-related opportunity cost and trading-related opportunity cost. The study classify opportunity cost into three subgroups: investment-related opportunity cost, first trading–related opportunity cost, and residual trading–related opportunity cost. The sum of the last two opportunity cost components is equal to the trading-related opportunity cost of Kissell

[2006], which is defined as the difference between the closing price of the day and the order release price for the unexecuted shares.

Investment- or operation-related opportunity cost

This cost is the difference between the order release price and decision price for the unexecuted shares, whereas the decision price is the price of the share in the market at the time a decision is made to purchase and the order release price is the price at which the order is placed in the market. The gap between these two prices can be positive, which would raise the opportunity cost for the trader, or vice versa. Kissell [2006] also defined it as investment- or operation-related opportunity cost. Trading-related opportunity cost may arise because of inadequate liquidity conditions and/or substantial adverse price movements.

First trading–related opportunity cost

First trading–related opportunity cost is one of our innovations. It is the difference between the first traded price and the order release price for the unexecuted number of shares. A first traded price that is higher than the order release price would increase the opportunity cost for the trader, and vice versa. Assume that a trader wishes to trade 100,000 shares and places the order when the price is at P_0 (say $101.00). Consider that trader is able to execute at price P_1 (say $102.00) for 20,000 shares. At the end of the day, consider that the trader is able to trade only 90,000 shares. To determine the opportunity cost for the unexecuted shares—10,000 shares—the study identified two additional subcomponents: first trading–related and residual trading–related opportunity costs. The first trading–related opportunity cost is calculated by multiplying the price difference between these two prices ($P_1 - P_0$ = \$102 – \$101) by the number of unexecuted shares (10,000), or 10,000 × (\$102 – \$101) = \$10,000. This is an extended subcomponent in our definition of opportunity cost that no other earlier researchers have identified.

Residual trading–related opportunity cost

Residual trading–related opportunity cost is our second innovation in estimating opportunity cost. It is the difference between the closing price of the day and the first executed trading price for the unexecuted shares. To extend the example given for first trading–related opportunity cost, consider that a trader is able to trade 40,000 shares at P_2 (\$103), and without any other trade, the closing price of the day, P_N, is \$103.50. To determine the residual trading–related opportunity cost for the unexecuted shares—10,000 shares—then multiply the price difference between first traded price and the closing price of the day, ($P_N - P_1$ = \$103.5 – \$102) by the number of unexecuted shares (10,000), or 10,000 × (\$103.5 – \$102) = \$15,000. As can be observed, if the closing price is higher than the first executed trading price, the opportunity cost is higher for the unexecuted shares.

Again, it is beyond the scope of this study to pursue this calculation, and it remains a topic for future research. It should be noted that for a sequence of multiple trades, further extension based on second, third, and other trades is possible to add other sequential trading–related opportunity costs in a similar fashion. The residual trading–related opportunity cost is the difference between the closing price and the last executed trading price, which is also beyond the scope of this research.

Numerical Experiment of Transaction Costs under Various Models

The following numerical example is used to measure the different components of transaction cost (IS) of all three previous models and our proposed model. Through this numerical experiment the study prove that all three models produce the same IS. This is the first time, to the best of knowledge, that such a comparative analysis has been conducted, and it provides a perspective on different components of IS for traders to consider while trading.

Example Case

Consider a manager who decides to buy 100,000 shares (X) of ABC. The share of ABC trades at the beginning of trade at \$100.00/share (P_s) and is currently trading at \$100.00/share (P_d). However, by the time the trader chooses a broker and submits the order to the market, the stock price has increased to \$101.00 (P_0). The broker has executed 20,000 shares at \$102.00/share (P_1), 40,000 shares at \$103.00/share (P_2), and 30,000 shares at \$104.00/share (P_3). The ending price (P_n) of the stock for the day is \$103.50/share. The average daily volume (ADV) of the stock ABC is 1,000,000 shares. These volume data are necessary to convert cost into basis points. Assume the National Best Bid and Offer prices as given in Exhibit 3.02 (in Appendix 3).

Here, best bid and best offer is needed to estimate the midpoint price P_m. P_b is the best offer price in a sequence of time j in which j = 1, 2, 3,…, K. Also, P_m is the midpoint price in a sequence of time j. Further assume that the expected prices, based on price trend (P_p), are \$101.25, \$102.15, and \$103.65 for the time t_1, t_2, and t_3, respectively. Also assume the events occur in the timing given in Exhibit 3.03 (in Appendix 3).

Computation of IS proposed by Perold [1988]

$$\text{IS} = X\left(P_n - P_d\right) - \left[\sum x_j\left(P_n - P_j\right) - \text{Fixed}\right] \tag{3.18}$$

where X = 100,000; P_n = \$103.50; P_d = \$100.00; and P_j = \$102.00, \$103.00, and \$104.00 are executed prices for trade 1, 2, and 3 respectively. The number of shares traded is 20,000, 40,000, and 30,000 for trade 1, 2, and 3 respectively.

$$X\left(P_n - P_d\right) = \text{PR}$$

$$= 100{,}000(\$103.50 - \$100.00) = \$350{,}000. \tag{3.19}$$

$$\sum x_j\left(P_n - P_j\right) - \text{Fixed} = \text{Portfolio RR}$$

$$= 20{,}000(\$103.50 - \$102.00) + 40{,}000(\$103.50 - \$103.00) +$$

$$30{,}000(\$103.50 - \$104.00) - \text{Fixed}$$

$$= \$35{,}000.00 - \text{Fixed} \tag{3.20}$$

$$\text{IS} = X\left(P_n - P_d\right) - \left[\sum x_j\left(P_n - P_j\right) - \text{Fixed}\right]$$

$$= \$350{,}000 - (\$35{,}000 - \text{Fixed}) = \$315{,}000 + \text{Fixed, or } 315 \text{ bps} + \text{Fixed} \tag{3.21}$$

To convert the value into basis points calculation of total order value is required. Here total order value = $X \times P_d$ = 100,000 × \$100.00 = \$10,000,000. So, IS (in bps) = (\$315,000/\$10,000,000) × 10^4 + Fixed = 315 bps + Fixed.

Computation of Expanded Implementation IS Proposed by Wagner and Edwards [1993]

$$\text{IS} = \underbrace{X(P_0 - P_d)}_{\text{Delay related}} + \underbrace{\sum x_j * P_j - \sum x_j * P_d}_{\text{Trading related}} + \underbrace{(X - \sum x_j)(P_n - P_0)}_{\text{Opportunity cost}} + \text{Fixed} \qquad (3.22)$$

The extended form is given as follows:

$$\text{IS} = \left\{ \begin{array}{l} \sum x_j (P_0 - P_d) \\ + (X - \sum x_j)(P_0 - P_d) \end{array} \right\} + \left[\sum x_j \left\{ \begin{array}{l} (P_{bj} - P_{mj}) \\ + (P_j - P_{bj}) \\ + (P_{pj} - P_0) \\ + (P_{mj} - P_{pj}) \end{array} \right\} \right] + \left(X - \sum x_j\right)(P_n - P_0) + \text{Fixed } (F) \qquad (3.23)$$

Here, $\sum x_j (P_0 - P_d)$ = Trading-related delay cost and $(X - \sum x_j)(P_0 - P_d)$ = Operation-related delay cost. These two costs sum up to total delay-related cost [$X(P_0 - P_d)$]. Trading-related cost [$\sum x_j \times P_j - \sum x_j \times P_d$] has four subcomponents: (1) spread cost, $\sum x_j (P_{bj} - P_{mj})$; (2) market impact cost, $\sum x_j (P_j - P_{bj})$; (3) price appreciation cost, $\sum x_j (P_{pj} - P_0)$; and (4) timing risk cost, $\sum x_j (P_{mj} - P_{pj})$. Finally, $\left(X - \sum x_j\right)(P_n - P_0)$ is the opportunity cost.

$$\text{Trading-related delay costs} = (P_0 - P_d)\sum_{j=1}^{k}(x_j)$$

$$= (\$101.00 - \$100.00) \times 90{,}000 = \$90{,}000 \qquad (3.24)$$

Order value = $X \times P_d$ = 100,000 × \$100.00 = \$10,000,000.00. (3.25)

Thus, the trading-related delay cost (in bps) = (\$90,000/\$10,000,000) × 10^4

= 90 bps. (3.26)

Operation-related delay costs $= (P_0 - P_d)[X - \sum_{j=1}^{k}(x_j)]$

$= (\$101.00 – \$100.00) \times 10{,}000 = \$10{,}000$ (3.27)

So, the operation-related delay cost (in bps) = ($10,000/$10,000,000) × 10^4

= 10 bps. (3.28)

Total delay-related cost = Trading-related delay costs + Operation-related delay costs

= $90,000.00 + $10,000.00 = $100,000.00 (3.29)

Trading-related costs = Spread cost + Market impact cost + Price appreciation cost +

Timing risk cost where spread cost $= \sum_{j=1}^{k}[x_j(P_{bj} - P_{mj})]$

= 20,000($101.50 – $101.35) + 40,000($102.70 – $102.40) +

30,000($103.80 – $103.70)

= $18,000.00 or 18 bps (3.30)

and

Market impact cost $= \sum_{j=1}^{k}[x_j(P_j - P_{bj})]$

= 20,000($102.00 – $101.50) + 40,000 ($103.00 – $102.70)

+30,000 ($104.00 – $103.80)

= $28,000.00 or 28 bps. (3.31)

Mathematically, market impact cost can be separated into two components, as stated earlier, and can be presented as follows:

$MI_\$ = 0.95 I_\$ \eta^{-1} + 0.05 I_\$$; So, $I_\$ = MI_\$ / (0.95\eta^{-1} + 0.05)$; and $\eta = (X + 0.5ADV) / X$ (3.32) where $I_\$$ is instantaneous market impact cost; and η is market participation.

Here, η = (100,000 + 0.5 × 1,000,000)/100,000 = 6; and $I_\$$ = \$28000 / (0.95 × 6^{-1} + 0.05) = \$134,400.

Thus, temporary market impact cost = $0.95 I_\$ \eta^{-1}$ = 0.95 × \$134400 × 6^{-1} = \$21,280 = 21.28 bps or 21 bps, and Permanent market impact cost = 0.05 $I_\$$ = 0.05 × \$134,400 = \$6,720 = 6.72 bps or 7 bps.

$$\text{Price appreciation cost} = \sum_{j=1}^{k} [x_j (P_{pj} - P_0)]$$

= 20,000(\$101.25 – \$101.00) + 40,000(\$102.15 – \$101.00) + 30,000(\$103.65 – \$101.00)

= \$130,500.00 or 130.5 bps. (3.33)

$$\text{Timing risk cost} = \sum_{j=1}^{k} [x_j (P_{mj} - P_{pj})]$$

= 20,000(\$101.35 – \$101.25) + 40,000(\$102.40 – \$102.15) + 30,000(\$103.70 – \$103.65)

= \$13,500.00 or 13.5 bps. (3.34)

Total Trading-related cost = \$18,000.00 + \$28,000.00 + \$130,500.00 + \$13,500.00 = \$190,000.00.

$$\text{Opportunity cost} = \left(X - \sum x_j\right)(P_n - P_0)$$

= (100,000 – 90,000)(\$103.50 – \$101.00) = \$25,000.00 or 25 bps. (3.35)

IS = Delay-related cost + Trading-related cost + Opportunity cost + Fixed

= \$100,000.00 + \$190,000.00 + \$25,000.00 + Fixed

= \$315,000.00 + Fixed or 315 bps + Fixed (3.36)

Computation of Expanded IS proposed by Kissell [2006]

IS = Investment-related cost + Trading-related cost + Opportunity cost + Fixed

$$IS = \sum x_j (P_0 - P_d) + \left[\sum x_j \left\{ \begin{array}{l} (P_{bj} - P_{mj}) \\ + (P_j - P_{bj}) \\ + (P_{pj} - P_0) \\ + (P_{mj} - P_{pj}) \end{array} \right\} \right] + (X - \sum x_j) \left\{ \begin{array}{l} (P_n - P_0) \\ + (P_0 - P_d) \end{array} \right\} + F \qquad (3.37)$$

Here, $\sum x_j (P_0 - P_d)$ = Investment-related cost. It should be noted that trading-related cost and its subcomponents are identical to those of Wagner and Edwards [1993].

Opportunity cost has two subcomponents: $(X - \sum x_j)(P_n - P_0)$ = Trading-related opportunity cost, and $(X - \sum x_j)(P_0 - P_d)$ = Investment-related opportunity cost.

Investment-related cost = $\sum x_j (P_0 - P_d)$

= 90,000 × ($101.00 – $100.00) = $90,000 or 90 bps. (3.38)

Trading-related cost = $190,000.00 or 190 bps (the same as in Wagner and Edwards [1993]).

Opportunity cost = Trading-related opportunity cost + Investment-related opportunity cost

Trading-related opportunity cost = $(X - \sum x_j)(P_n - P_0)$

= (100,000 – 90,000)($103.50 – $101.00) = $25,000.00 or 25 bps. (3.39)

Investment-related opportunity cost = $(X - \sum x_j)(P_0 - P_d)$

= ($101.00 – $100.00) × (100,000 – 90,000) = $10,000.00 or 10 bps. (3.40)

Opportunity cost = $25,000.00 + $10,000.00 = $35,000.00 or 35 bps.

IS = $90,000.00 + $190,000.00 + $35,000.00 + Fixed

= $315,000.00 + Fixed or, 315 bps + Fixed. (3.41)

Computation of Expanded IS proposed by the study

IS = Investment-related cost + Trading-related cost + Opportunity cost + Fixed

$$IS = \sum x_j (P_0 - P_d) + \left[\sum x_j \left\{ \begin{array}{l} (P_{bj} - P_{mj}) \\ + (P_j - P_0) \\ + (P_{pj} - P_0) \\ + (P_{mj} - P_{pj}) \end{array} \right\} \right] + \left[\begin{array}{l} (X - \sum x_j)(P_0 - P_d) \\ + (X - \sum x_j)(P_1 - P_0) \\ + (X - \sum x_j)(P_n - P_1) \end{array} \right] + F \qquad (3.42)$$

Here, investment-related cost is $\sum x_j (P_0 - P_d)$, the same as in Kissell [2006], and trading-related subcomponents are identical to those of Wagner and Edwards [1993] and Kissell [2006].

The opportunity cost has three subcomponents that are defined as

$(X - \sum x_j)(P_0 - P_d)$ = Investment- or operation-related opportunity cost;

$(X - \sum x_j)(P_1 - P_0)$ = First trading–related opportunity cost; and

$(X - \sum x_j)(P_n - P_1)$ = Residual trading–related opportunity cost.

Investment-related cost = $90,000.00 or 90 bps (same as in Kissell [2006])

Trading-related cost = $190,000.00 or 190 bps (same as in Wagner and Edwards [1993] and Kissell [2006]).

Opportunity cost = Investment- or operation-related opportunity cost + First trading–related opportunity cost + Residual trading–related opportunity cost.

Investment- or operation-related opportunity cost = $(X - \sum x_j)(P_0 - P_d)$

$$= (\$101.00 – \$100.00) \times 10{,}000 = \$10{,}000.00 \qquad (3.43)$$

First trading–related opportunity cost = $(X - \sum x_j)(P_1 - P_0)$

$$= (\$102.00 – \$101.00) \times 10{,}000 = \$10{,}000.00 \qquad (3.44)$$

Here, P_0 is the price at which the order is released, and it is beyond the control of the trader. P_1 is the possible price of the next trade (if the order is not fully executed at the original price).

Residual trading–related opportunity cost = $(X - \sum x_j)(P_n - P_1)$

$$= (\$103.50 – \$102.00) \times 10{,}000 = \$15{,}000.00 \qquad (3.45)$$

Opportunity cost = \$10,000.00 + \$10,000.00 + \$15,000.00 = \$35,000.00 or 35 bps.

IS = \$90,000.00 + \$190,000.00 + \$35,000.00 + Fixed

= \$315,000.00 + Fixed or 315 bps + Fixed (3.46)

From this numerical experiment, summarized in Exhibit 3.04 (in Appendix 3), the study claim that if a trader can control the FTROC and RTROC or either of the two, s/he can reduce the opportunity cost by 25 bps or 15 bps or 10 bps. It is worth pointing out here that *transaction cost management* (TCM) is the key to reducing first trading– and residual trading–related opportunity costs, which requires a balance with market impact cost. With TCM, a trader will be able to assess the growing order size in the market and implement his or her position within an indicated price range. By observing real-time market price trend and order flow information from a sequence of trades, a trader can predict any unfavorable trading situations.

Reduction of the opportunity cost can be attained by using this pre-trade analysis to approve orders sized properly for present market situations. For large orders, the investor must decide whether to risk market impact or spread the trade over several days, risking exposure to a price change. Setting price limits efficiently (so that the order is filled or almost filled) will ensure that the FTROC- and RTROC-related opportunity cost will be as low as possible for the trader. If the price trend is higher, opportunity cost will be minimal from setting the price aggressively in the first trade so that all or almost all trading volume is executed, and residual trading cost will tend to zero. The opportunity cost, on the other hand, can be reduced by setting a passive pricing strategy and executing less in the first trade and more on the later trade if the price trend is downward.

In addition to setting the price limit, good communication between the broker (who executes the order) and trader is vital to permit swift execution and to alter any implementation strategy (if necessary). A trader must decide how to balance the costs of trading with the opportunity cost. It should be noted that there is a trade-off between trading rapidly and being persistent. Executing a full trade rapidly drops opportunity costs but increases probable market impact costs. Slower implementation, on the other hand, lowers market impact but increases opportunity cost because the full order may not be executed or the trade may be filled later at a disadvantageous (average) price.

Exhibit 3.04 (in Appendix 3) shows the calculation for each of the four models to give a clear picture of the cost. Any experienced trader can imagine that it is extremely challenging to accurately predict the market and make a choice between market impact and price impact; as such, accurately modeling transaction costs in a back test is also challenging. Although it is virtually impossible to account for every eventuality, it is worthwhile to explore multiple back tests assuming different turnover levels to better understand the balance between transaction and opportunity costs so that in future trade a trader is able to reduce the first trade–related and residual trade–related opportunity costs as much as possible.

Conclusion

Transaction costs play a very significant role in affecting investment performance. Above and beyond the fixed charges (such as fees, taxes, and commission), other hidden or variable factors may adversely increase trading costs for traders. These factors, like delay cost, price appreciation cost, market impact cost, timing risk cost, and opportunity cost, can add to costs of transaction and reduce portfolio return.

The study by Perold [1988] was the first to formally identify this as IS, which traders and fund managers must consider while analyzing portfolio performance. Once security selection is done, prompt action in taking a position, avoiding price run in executing transaction, and ability to complete total transaction are equally important to performing well. Partial trade or adverse price effect due to front run may lower the return of investment. Although various factors are beyond the control of the trader, such as ability to trade at the right price and complete the total transaction, traders must be prompt in executing investment decisions and should review carefully those factors that are part of opportunity costs, as identified in the literature and extended in our research, so they can avoid or reduce these costs of investing.

Although these costs may not be such a significant factor for long-term traders who wish to buy and hold assets for a long time, they are highly important for day traders and dynamic portfolio managers. The work of Wagner and Edwards [1993] extended the transaction cost by incorporating various factors, such as price impact, timing cost, opportunity cost, and commission, that add up to total transaction costs, which can significantly affect portfolio performance. The model proposed by Kissell [2006] further modified cost elements that Wagner and Edwards [1993] incorporated.

The study closely follow and extend the Kissell [2006] model by further identifying trading-related costs that consist of opportunity costs. Our research contributes to the existing literature by extending the opportunity cost subcomponents that traders or fund managers can identify and, by taking an appropriate trading implementation strategy, reduce or eliminate from total transaction cost. Our innovative sub classification of opportunity costs is one of the key contributions of this research because these subcomponents have not been discovered in previous literature.

To validate this proposed framework, the study provide a single numerical experiment of calculating IS for all three existing models along with our model and show that all four models provide the same IS. Traders and fund managers now have two additional factors they can identify that adversely affect IS; they can control these factors by identifying their sources and creating implementation strategies to remove or reduce them while executing trades. As a result of lower IS, portfolio or fund performance can be improved. TCM is the key to reducing first trading– and residual trading–related opportunity costs, which require a balance with market impact and price impact cost. After thorough back testing, market price trend, and pretrade analysis, setting price limits efficiently (so that the order is filled or almost filled) will ensure the first trading–related opportunity cost and residual trading–related opportunity cost are as low as possible for the trader. If the price trend is higher, opportunity cost will be at a minimum from setting price aggressively in the first trade so that all or almost all trading volume is executed (so that residual trading cost will tend to zero). The opportunity cost, on the other hand, can be reduced by setting a passive pricing strategy and executing less in the first trade and more on the later trade if the price trend is downward.

In addition to setting the price limit, good communication between the broker (who executes the order) and trader is vital to permit swift execution and to alter any implementation strategy (if necessary). Readers can pursue future research focusing on testing portfolio performance with IS for day traders and dynamic portfolio managers using a benchmark portfolio for comparative evaluation. This research can provide further insight into whether there is any relationship between higher performance and lower IS via carefully controlling extended components of opportunity cost. In addition, a further extension of TCA should incorporate both buy and sell trading for all sequential trades (not only buy-side analysis, as has been done in previous research) and further test the relationship between IS and trading performance.